TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

CHAO JIANG

SAMPLING FREQUENCY OPTIMIZATION AND TRAINING MODEL SELECTION FOR PHYSICAL ACTIVITY CLASSIFICATION WITH SINGLE TRIAXIAL ACCELEROMETER

Master of Science thesis

Examiner: prof. Ilkka Korhonen
Examiner and topic approved by the
Faculty Council of the Faculty of
Natural Sciences
on 3rd June 2015

ABSTRACT

CHAO JIANG: Sampling frequency optimization and training model selection for physical activity classification with single triaxial accelerometer
Tampere University of Technology
Master of Science Thesis, 52 pages, 0 Appendix page
September 2015
Master's Degree Programme in Science and Bioengineering
Major: Information Technology for Health and Biology
Examiner: Professor Ilkka Korhonen

Keywords: sampling frequency optimization, physical activity, acceleration, activity classification, decision tree, impersonal model, hybrid model, personal model

Ambulatory monitoring system with accelerometers can provide a reliable, continuous, unsupervised and objective monitoring of human physical activities. The system can in many cases recognize the type of activity being performed, and calculate the duration and intensity. This kind of information can be utilized to help people to follow up their physical activities and remind people to be more active, because physical inactivity can cause some health problems. However, especially for mobile devices continuous sampling, signal processing and activity recognition rapidly depletes the system's energy, which is a critically constrained resource.

In this thesis work, several methods for reducing energy consumption in physical activity recognition were reviewed and discussed, i.e., 1) reducing the number of sensors used; 2) selecting low power sensors; 3) reducing the number of axes; 4) decreasing the sampling frequency; 5) adopting an adaptive sampling strategy. In this thesis, a single tri-axial accelerometer was utilized for sensing the accelerations, and sampling frequency was optimized in order to lower the energy consumption. The physical activity recognition was performed with different sampling frequencies and training strategies, with the target to reach good classification accuracies and low energy consumption.

Based on the obtained classification results, several conclusions were drawn. Firstly, personal models did not always achieve better classification accuracies over impersonal and hybrid models. However, personal models performed much better for some activities, e.g., biking, lying, and rowing. Secondly, there was no uniform optimal sampling frequency for all activities. Sampling frequencies no larger than 10 Hz were enough to classify all activities.

To further optimize the energy consumption, adaptive sampling rate logic was designed and implemented. It adaptively used 1 Hz when sampling the accelerations from lying activity and 10 Hz for other activities. The results showed it worked effectively and efficiently.

PREFACE

This thesis work was carried out between January 2015 and June 2015 in the Department of Signal Processing at Tampere University of Technology (TUT).

I would like to thank Antti Särelä, D.Sc. Hannu Nieminen and Prof. Ilkka Korhonen for their guidance, supports, comments and suggestions during the thesis work. A lot of thanks go to VTT for providing the raw data source for the work.

Also I would like to thank my mentor Timo Valli and my friends Guangliang Liu, Dinglin Cai, and Mohammed Al-Mosawi for their encouragement and sharing their working, studying and life experience with me. Finally, special thanks go to my family for their support and love.

Tampere, 17.09.2015

Chao Jiang

CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

AccX	Up-down chest acceleration
AccY	Left-right chest acceleration
AccZ	Back-forth chest acceleration
Palantir2003	Palantir Context Data Library 2003
Palantir2004	Palantir Context Data Library 2004
TUT	Tampere University of Technology
CDT	Custom decision tree
HMM	Hidden Markov model
SVM	Support vector machine
ANN	Artificial neural network

1. INTRODUCTION

Physical inactivity can cause several health problems. To encourage people to be more physically active, wearable devices performing physical activity recognition are becoming popular. They can recognize some daily activities with good accuracies. However, their energy consumption is a painful obstacle towards the adoption of continuous sensing driven applications, e.g., continuously performing physical activity recognition with the embedded accelerometer sensors.

The objective of the thesis work was to find suitable sampling frequencies and a suitable training strategy to reach an acceptable tradeoff between the energy overhead and the classification accuracy. Generally, to obtain better classification accuracies, higher sampling frequencies and/or more complex algorithms need to be used. These will definitely increase the energy consumption. Thus, lower sampling frequencies and simpler algorithms would save power, provided that the classification accuracy stays high enough.

This thesis is structured as follows. Chapter 2 introduces the background of the thesis work, including the purpose of the physical activity recognition, its related research and power consumption problems faced by the wearable devices. Chapter 3 provides an overview of the Palantir project, which provided the raw data source utilized in this thesis. Chapter 4 shows the whole process of acceleration data analysis, starting with data preprocessing, followed by feature extraction, then describes the implemented custom decision tree logic, and ends with the implemented feature selection. It also introduces adaptive sampling rate logic which uses 1 Hz as the sampling frequency only for activity “lying”. Chapter 5 demonstrates and discusses the classification results for each training model, each activity and each sampling frequency. Chapter 6 concludes the thesis work.

2. BACKGROUND

Physical activity recognition system records and recognizes physical activities. With the help of this system, the information of the type of activity that is performed, the duration and the intensity, can be extracted. This kind of information can be used to remind and promote people to be more physically active. With enough physical activity, people can maintain and promote their health. However, continuous sensing and classification easily creates energy consumption problems, which limits the use of the techniques especially in mobile devices.

2.1 Physical inactivity

Physical inactivity (lack of physical activity) is a global public health problem. It has been identified as one of the ten leading risk factors for global mortality. Moreover, physical inactivity is estimated to be the main cause for approximately 21-25% of breast and colon cancers, 27 % of diabetes and approximately 30% of ischaemic heart disease burden (World Health Organization 2009). People who are insufficiently active have a 20% to 30% increased risk of death compared to people who are sufficiently active (World Health Organization 2015). Globally, around 23% of adults were not active enough, and more than 80% of the adolescent population were insufficiently physically active in 2010 (World Health Organization 2015). The current levels of physical inactivity are partly due to insufficient participation in physical activity during leisure time and an increase in sedentary behavior during occupational and domestic activities. Likewise, an increase in the use of “passive” modes of transportation also contributes to insufficient physical activity (World Health Organization 2015).

Physical activity is defined as any bodily movement produced by skeletal muscles, which requires energy expenditure (Caspersen 1985, p. 126). Physical activity in daily life can be categorized into occupational, sports, condition, household, or other activities. The term “physical activity” should not be mistaken with “exercise”. Exercise is a subcategory of physical activity that is planned, structured, repetitive, and purposeful in the sense that the improvement or maintenance of one or more components of physical fitness is the objective. Physical fitness is a set of attributes that people have or achieve that relates to the ability to perform physical activity. Physical activity includes exercise as well as other activities which involve bodily movement and are done as part of playing, working, active transportation, house chores and recreational activities. So physical activity can be done almost anywhere and does not necessarily require equipment. Carrying groceries, wood, books, or children are good complementary physical activities, as is climbing the stairs instead of using the elevator. It is not imperative to go to a gym, pool or other special sports facility to be physically active. It is found that regular moderate intensity physical activity, e.g., walking, cycling, or participating in sports, has significant benefits for health (World Health Organization 2015).

Both moderate and vigorous intensity physical activity brings health benefits. The intensity of different forms of physical activity varies between people. In order to be beneficial for cardiorespiratory health, all activity should be performed in bouts of at least 10 minutes duration (World Health Organization 2015). Regular physical activity of moderate intensity, e.g., walking, cycling or doing sports, has significant benefits for health. At all ages, the benefits of being physically active outweigh potential harm, for example through accidents. Some physical activity is better than doing none. By becoming more active throughout the day in relatively simple ways, people can quite easily achieve the recommended activity levels. Regular and adequate levels of physical activity helps improve muscular and cardiorespiratory fitness; improve bone and functional health; reduce the risk of hypertension, coronary heart disease, stroke, diabetes, breast and colon cancer, and depression; reduce the risk of falls as well as hip or vertebral fractures; and are fundamental to energy balance and weight control (World Health Organization 2015).

Actually it only takes 30 minutes daily of moderate-intensity physical activity to improve and maintain the health for adults, and 60 minutes of moderate to vigorous-intensity physical activity daily for children and adolescents. By following the minimum recommendation, many health benefits can be obtained when compared with completely inactive people. The basic level of physical activity can be achieved by everyday activity like walking at work, shopping, gardening, cleaning, etc. Although achieving the minimum recommendation of physical activity brings many health benefits, even more health benefits can be achieved by taking part in a more vigorous and wider spectrum of physical activities. For example, endurance-enhancing activities and activities maintaining flexibility and muscular strength bring benefits that are not achieved with basic activity (Pate 1995). Endurance can be enhanced, e.g., with energetic walking, jogging, cycling and rowing. Activities maintaining functions of the musculoskeletal system are, e.g., ball games, gym, and dancing. Thus, in addition to daily energy expenditure, activity types play an important role in overall wellbeing and health. This fact indicates the need of recognition of activity types.

Increasing physical activity is a societal, not just an individual problem. Wearable devices can not only serve as tools for people to manage their physical activity data, but enable people communicate with each other over the Internet. In such case, a physical activity recognition system can be built on a society level. This fact, together with the big data techniques, will create more both health and economic benefits.

2.2 Research on physical activity recognition

Generally, the physical activity recognition covers the whole processing of sensing and data analysis. This section provides overviews of placements of accelerometers, and classification techniques used to recognize different activities. Specially, it gives a brief description of the decision tree algorithm, because the custom decision tree works as the classification algorithm in this thesis work.

2.2.1 Accelerometer and its placement

In the field of physical activity recognition, the inertial sensor system can provide a low-cost, effective and privacy-aware solution. The most widely used inertial sensors are accelerometers and gyroscopes. Gyroscope sensors measure the angular velocity by using the tendency of vibration in the same plane as an object vibrates (Avci 2010). Accelerometers can sense applied acceleration acting along a sensitive axis in a real-time fashion which can be used to calculate the rate and intensity of body movement in up to three planes (Godfrey 2008). Specifically, accelerometers measure both static (e.g. gravity) and dynamic (e.g. vibration) acceleration, and three accelerometers can be incorporated into a single tri-axial accelerometer providing information on three-dimensional movements (Culhane 2005, p. 556).

In general, the current draw of modern accelerometers is several μA, which is much lower when compared with the one of gyroscopes, which is typically several mA. Moreover, in (Pärkkä 2007), the results showed that accelerometers outperformed gyroscopes in estimating the intensity of physical activity, with accelerometers and gyroscopes attached to ankle, wrist and hip. Based on this, this section focuses on the research using accelerometers.

Accelerometers have been widely used for the human movement detection, both in the clinical settings and in free-living environments, of which some example studies are (Veltink 1996; Foerster 1999; Bao 2004; Mathie 2004b; Aminian 2004; Ravi 2005; Yang 2010).

M. J. Mathie et al. (2004b, p. R2) summarized some desirable advantages of accelerometer in the monitoring of human physical activity: firstly, they respond to both frequency and intensity of human movement; secondly, some types of accelerometers can be used to measure tilt as well as body movement; thirdly, enhancements in integrated microelectromechanical systems technology have enabled the manufacture of miniaturized, low cost accelerometers, which still demonstrate a high degree of quality and reliability in measurement, with little variation over time. These advantages have enabled the development of lightweight, small, low cost and portable systems which can be worn by a free-living subject without impeding movement, drawing too much awareness, and causing much discomfort, but still providing objective and reliable monitoring. Thus, accelerometry has been adopted as a suitable tool for long-term monitoring of free-living subjects in a wide range of applications, including classification of movements, assessment of physical activity level, estimation of metabolic energy expenditure, and assessment of balance, gait and sit-stand transfers (Mathie 2004b, p. R16). For example, F. Foerster et al. (1999, p. 571) found that accelerometry was highly reliable in the detection of posture and motion.

The placement of the accelerometer on the body is an important aspect in the measurement of body movement. Normally, it is the part of the body whose movement is being studied, because the generated accelerations during human movement vary across the body and depend on the activity being performed. Some researchers used several accelerometers and attached them to different parts of the body, for example (Foerster 1999; Olguin 2006; He 2007). This is called the multi-site attachment strategy. The corresponding one is single-site attachment. The single placement can be the pelvic (Ravi 2005), chest (Godfrey 2011; Lee 2011), wrist (Yang 2008; Siirtola 2009; Zhang 2012), hip (Staudenmayer 2009), waist (Mathie 2004a; Long 2009), thigh (Kwapisz 2011), and ankle (Ermes 2008a).

In (Mannini 2013), it is mentioned that the multi-site attachment was preferable to detecting a larger variety and finer complexity of activities because both upper and lower body motion can be captured independently. But fixing multiple sensors to the body might pose restrictions on physical activities performed in daily life context, and would be cumbersome to use (Long 2009). In addition, this strategy will cost more, be less comfortable, and require a higher dimension and a more complex feature set (Yang 2008). Thus the single-site attachment is more practical for continuous, long-term monitoring, as the simplicity and ease-of-use of the single instrument facilities compliance and minimizes cost (Mathie 2004a). However, detailed classification cannot be achieved in this approach because it uses only one sensor and only the acceleration at the position of this unique sensor is available, which cannot reflect the movement of the whole body (He 2007). For example, in (Staudenmayer 2009), they claimed that the approach of using a single hip-mounted uniaxial accelerometer had the inherent limitation that different activities could produce very similar accelerometer signals.

There are design tradeoffs between the number of accelerometers used, the cost, the usability and the transferability of an ambulatory monitoring system, but the design choices are determined to a large extent by the purpose and duration of the monitoring (Mathie 2004b). In short-term, supervised monitoring situations, the multi-site attachment is preferred because it allows the collection of greater quantities of information, leading to very accurate classification of a specific subset of activities for a certain application. However, in long-term, unsupervised monitoring environments, the single-site will be the choice, because of the much less complexity and cost of the system. For example, in (Mathie 2004b), a single, waist-mounted accelerometry system was developed for long-term, ambulatory monitoring of a range of different parameters of human movement in an unsupervised setting.

The recognition of basic physical activities, e.g. resting, walking, running and cycling, is a well-researched area, and shows that the recognition of these activities is possible even with just one tri-axial accelerometer (Reiss 2011). For example, D. O. Olguin and A. S. Pentland (2006) found that it was possible to recognize some of the most common activities using a single accelerometer on the chest with average classification accuracy around 60%. There were 8 activities involved in their study, including sitting down, running, squat, walking, standing, crawling, lying down on the chest, and standing with hand movements. M. W. Lee et al. (2011) also used a single chest mounted tri-axial accelerometer in their study. They developed a portable real-time personal life log system for human activity recognition and exercise information generation. A validation of the system was performed with 6 activities, i.e. lying, standing, walking, going-upstairs, going-downstairs, and driving. They obtained average classification accuracies of 94.43% and 96.61% for subject-independent and subject-dependent recognition respectively.

About the acceptance of placements, D. O. Olguin and A. S. Pentland had lengthy discussions with three user groups, i.e. office workers, hospital patients and army soldiers, and found that there was broad consensus about the acceptability of the chest and hip locations (Olguin 2006). This fact, together with the feasibility of the single-site attachment in common daily activity recognition, indicates that the decision of choosing single chest mounted tri-axial accelerometer in the thesis work is valid.

2.2.2 Overview of classification techniques

A lot of different classifiers have been employed in physical activity recognition. Those classifiers could be divided as base-level classifiers and meta-level classifiers, with the former ones being widely used. Some common base-level classifiers include decision tables, decision trees, k-nearest neighbors, HMM, SVM, ANN, and naïve Bayes. The meta-level classifiers can be clustered into three frameworks, i.e. voting, stacking, and cascading (Ravi 2005). The comparison of these two types of classifiers on different activity recognition problems has been made in Ravi (2005) and Reiss (2014). N. Ravi et al. found that meta-level classifiers generally performed better than base-level classifiers, and plurality voting combining multiple base-level classifiers achieved the best performance. A. Reiss and D. Stricker (2014) compared the performances of 4 base-level classifiers, i.e. C4.5 decision tree, k-nearest neighbor, SVM and naïve Bayes, and two different voting strategies (boosting and bagging). Their results showed that the boosted decision tree obtained the best accuracy 90.65%, but with very huge size and high computational cost.

The hidden Markov model (HMM) has been widely employed to represent and learn the sequential and temporal characteristics in human physical activity sequences, e.g. (Yamato 1992; Brand 1997; Duong 2005; Olguin 2006; He 2007). When using HMM in human activity classification, the human activity series are considered as a Markov process and activities as states. J. He et al. (2007) used HMM to estimate the hidden states series from the acceleration-related data series. There were 8 hidden states in their study, which were divided into 2 groups: transition states and stable states; the stable states included lying, sitting, and standing; the transition states included standing to sitting, sitting to lying, sitting to standing, lying to sitting, and falling (He 2007, p. 3194). They achieved very good classification results with an overall accuracy of 95.82%. J. Yamato et al. performed the human action recognition from a set of time-sequential images, and got recognition rates over 90% for some sports scenes. M. Brand et al. (1997) proposed the coupled HMM (CHMM) for complex action recognition. They used a self-calibrating stereo blob tracker to obtain 3D hand tracking data for three T'ai Chi gestures involving arm-motions: the left single whip, the left cobra, and the left brush knee (Brand 1997, p. 996). Their results showed that CHMMs were more reliable than conventional HMMs, and yielded higher likelihood models with better discriminatory power in fewer epochs and these models often ran faster than comparable HMMs in a modified Viterbi algorithm (Brand 1997, p. 999). In (Duong 2005), the Switching Hidden Semi-Markov Model (S-HSMM) method was introduced and the discrete Coxian distribution was used, to exploit both the inherent hierarchical organization of the activities and their typical duration. According to their results, the S-HSMM performed better than existing models including the flat HSMM and the hierarchical hidden Markov model in both classification and abnormality detection tasks, alleviating the need for pre-segmented training data; the discrete Coxian duration model yielded better computation time and generalization error than the classic explicit duration model (Duong 2005).

The support vector machine (SVM) was adopted in physical activity recognition in (He 2009; Sun 2010; Anguita 2012; Mannini 2013). In (He 2009), a tri-axial accelerometer ADXL330 was put into the subject's trousers pocket to measure the accelerations from four daily activities, i.e. running, still, jumping and walking. They decomposed the multi-class activity classification problem into several binary class problems with One-versus-One Strategy, and got average 97.51% accuracy. D. Anguita et al. (2012) proposed a novel hardware-friendly approach for multiclass classification, i.e., Multiclass Hardware Friendly Support Vector Machine (MC_HF_SVM), to reduce energy consumption and computing power. Their results showed a significant improvement in terms of computational costs while maintaining similar accuracy when compared with the traditional SVM (Anguita 2012).

The artificial neural network (ANN) method served as the classification algorithm in (Yang 2008; Staudenmayer 2009; De Vries 2011). J. Y. Yang et al. introduced a systematic design approach for constructing neural classifiers for human physical activity recognition. Their design approach consisted of two steps: in the first step, a divide-and-conquer strategy was applied to separate dynamic activities from static activities; in the second step, these two different types of activities were classified separately. Three neural networks were actually built, including one for pre-classification of all activities into dynamic and static activities, one for the classification of static activities, and one for dynamic activities. Finally, they achieved an overall recognition accuracy of 95% for eight daily activities, i.e. standing, sitting, walking, running, vacuuming, scrubbing, brushing teeth, and working at a computer. J. Staudenmayer et al. (2009) adopted the ANN to estimate physical activity energy and identify physical activity type from an accelerometer. Activity types were divided into four categories, i.e. low-level activities, locomotion, vigorous sports, and household activities/other activities. They found that the ANN correctly classified activity type 88.8% of the time (Staudenmayer 2009, p. 1300). In (De Vries 2011), S. I. De Vries et al. drew a conclusion that relatively simple ANN models perform well in identifying the type but not the speed of the activity of adults from accelerometer data.

X. Long et al. implemented the naïve Bayes classifier in their study, but made a modification of its input features, by applying the principal components analysis to remove the correlation among features vectors and reduce the feature vector dimension. To evaluate the performance of the modified naïve Bayes classifier, they applied it in the classification of five activities including walking, running, cycling, driving and sports, and also built a decision tree classifier and took the classification results obtained by the decision tree classifier as the reference. The comparison between these two sets of results showed that their modified naïve Bayes classifier obtained a comparable result as the decision tree did. But, as they claimed, the Bayes classifier had the advantage to be more extensible, requiring little effort in classifier retraining and software update upon further expansion or modification of the target activities, e.g. when new features or activities need to be incorporated (Long 2009).

J. Yang performed the physical motion recognition using mobile phones with built-in accelerometer sensors. The task involved six common activities, i.e. sitting, standing, walking, running, driving, and bicycling. Four static classifiers, i.e. C4.5 decision trees, naïve Bayes, k-nearest neighbor and the SVM, were evaluated. The results showed that decision tree achieved the best performance; a well-pruned decision tree with simple time domain features and less over-fitting on the training data could provide a usable model for inferencing a physical activity diary, refined by a similarity match from k-means clustering results and smoothed by an HMM-based Viterbi algorithm (Yang 2009, p. 1). L. Bao and S. S. Intille implemented decision table, instance-based learning, C4.5 decision tree, and naïve Bayes classifiers in their study, and found that decision tree classifiers achieved the best performance recognizing 20 daily activities (Bao 2004). U. Maurer et al. evaluated and compared C4.5 decision trees, k-nearest neighbor, naïve-Bayes and the Bayes net classification algorithms; they claimed the decision tree classifier could provide a good balance between accuracy and computational complexity (Maurer 2006). In (Pärkkä 2006), the custom decision tree (CDT), automatic decision tree and the ANN were employed in physical activity recognition. The custom decision tree achieved the best recognition accuracy for more than half of the activities, but automatic decision tree obtained the best overall classification accuracy. The ANN got the same overall accuracy as the CDT did, but its recognition result of running was very poor (22%).

In (Ermes 2008a; Reiss 2011), only the CDT was employed in activity recognition task. Specifically, M. Ermes et al. built a binary decision tree to classify 5 activities, i.e. lying, sitting and standing, walking, running, and cycling. They got overall average classification accuracy as 94%. A Reiss and D. Stricker also adopted the binary tree strategy, and obtained an overall performance as 86.80% for 7 activities, including lying, cycling, sitting and standing, running, Nordic walking, walking, and one default activity named as ‘other’. The ‘other’ activity involved ironing, vacuum cleaning, playing soccer, rope jumping, ascending stairs, and descending stairs. The reason for this combination was that the CDT could not classify these activities with good accuracies. For this point, they clearly mentioned that for more complex classification problems, more advanced classification techniques were needed, they also claimed the main drawback of custom decision trees was that they did not necessarily provide the most suitable classifier for a given set of features, so other techniques had to be investigated to exhaustively explore the solution space (Reiss 2014, p. 107). But the CDT has obvious advantages including simple implementation, low computation requirements, and a good understanding of the classifier’s structure. In practice, the binary decision tree algorithms with simple threshold rules can be implemented as an if-else structure (Ermes 2008a, p. 4453), which will reduce the computational complexity further.

2.2.3 Decision tree techniques

Decision tree method is one of the most popular inductive inference algorithms. Normally, it approximates discrete-valued target functions and represents the learned function with a built decision tree.

Generally, decision trees can be divided into two categories, i.e. automatic decision trees and custom decision trees. Here, the “custom” means the structure, leaves and testing nodes of the tree are defined by users before training or building the tree, while, the “automatic” indicates the tree is built automatically without user-defined order of testing nodes and user-defined leaf nodes. This difference indicates that automatic decision tree algorithms differ from custom decision tree algorithms in the training stage. Specifically, the automatic decision tree method requires more work on determining the node question and the tree size. Normally, there are two methods for automatically building a decision tree. One is based on a greedy strategy and the other on the so-called grow-then-prune strategy. The greedy method starts with a single node and recursively performs the split based on some question until some stop criterion is met. There are some issues related to this method. For example, the size of the resulting tree is usually very large. So the generated hypothesis is overly complex with high VC-dimensions and thus could cause overfitting problems. This problem could be avoided by using the grow-then-prune strategy. It starts with growing a very large-size tree until the tree fully fits the training sample or until no more than a very small number of points are left at each leaf. Then, the built tree is pruned back to minimize a user-defined objective function, which helps to reduce the tree size and to avoid overfitting problem. But for the CDT, the tree size, the order of testing nodes, leaves, and node questions are defined by users before the training stage. So an additional prior knowledge is needed when using the CDT. The splitting rule plays a key role in the construction of decision trees. A splitting rule is a prescription for deciding which variable, or combination of variables, should be used at a node to divide the data samples into subgroups, and for deciding how the values that the variable takes should be partitioned (Webb 2011). It can be selected based on the node impurity function used. For example, if the entropy is employed to measure the node impurity, the entropy gain can work as the splitting rule.

In practice, the binary decision tree algorithm is commonly used. It can be explained via a binary tree model, consisting of a root node, some binary decision nodes, left and right branches, and leaf nodes. The root node is the first testing node but without parent node. It, together with other testing nodes, performs a binary testing to decide how and where the query example flows. The leaf nodes store the class label information. In the prediction stage, a query example starts at the root node of the built tree and goes down the tree until a leaf is reached, by flowing along the right branch of a node when the decision to the node question is positive, and along the left branch otherwise. The query example is associated with the label stored in the reached leaf.

In general, decision trees can also be viewed as a representation of a disjunction of conjunction of constraints on the attribute values of instances. Each path from the tree root to a leaf corresponds to a conjunction of attribute tests and the tree itself to a disjunction of these conjunctions. So the learned tree can be re-presented as sets of if-else rules to improve human readability, and be converted into sets of if-else logic in implementation to reduce the complexity of the classification process and the computational workload. For custom decision trees, they can be implemented with if-else logic before training, while, for automatic decision trees, they can only be converted into if-else logic after training.

The C4.5 decision tree, mentioned in the section 2.2.2, is one type of automatic decision trees. It performs differently from the CDT in the learning stage, but shares almost the same prediction stage. Specifically, the C4.5 adopts a divide-and-conquer approach to growing decision trees and the post-pruning algorithm to reduce the tree size to avoid overfitting. As described in section 2.2.2, C4.5 decision trees and CDT were both successfully adopted in the physical activity recognition. But the CDT is a simpler algorithm compared with the C4.5. So the CDT is preferred and adopted in the thesis work.

2.3 Energy consumption problem

Ambulatory activity monitoring using accelerometers is a reliable technique, providing continuous, unsupervised and objective monitoring of mobility (Culhane 2005, p. 558). However, continuous sensing with mobile devices rapidly depletes the wearable system's energy, which is a critically constrained resource (Krause 2005). Basically, the energy consumption is caused by the sensor and by the analysis. The sensor consumes energy in the sampling process which is directly affected by the sampling rate. In the analysis process, the computational load determines the energy consumption, which is affected by the number of samples it needs to process and the complexity of the analysis. A low sampling frequency reduces the number of samples needed to be processed on a second-by-second basis, and a simple analysis method the complexity of the analysis. As described in the section 2.2.1, the current draw of modern accelerometers is very low, so now the energy is mainly consumed in the analysis process. To lower the energy consumption, it is important to reduce the computational load.

2.3.1 Academic research

Some studies have been performed on the topic of reducing the energy consumption, for example (Godfrey 2011; Zhang 2012; Yan 2012; Yurur 2013). In former section 2.2.1, two attachment strategies have been introduced and analyzed. Generally, the single-site attachment beats the multiple-site attachment in terms of lowering energy consumption, because of less number of sensors used. But, some devices incorporated high-power sensors, such as gyroscopes, thereby still causing high energy consumption. For example, B. Najafi et al. (2003, p. 711) used only one kinematic device attached to the chest in detection of body postures and periods of walking in the elderly, but the kinematic device was composed of one miniature piezoelectric gyroscope (4.5 mA) and two miniature accelerometers (ADXL202, <0.6 mA) (Najafi 2003, p. 712). Based on (Najafi 2003), A. Godfrey et al. (2011, p. 1128) replaced the device with single tri-axial accelerometer which consisted of two bi-axial Analog Devices ADXL210 (<0.6 mA). Classification results were similar with the ones in (Najafi 2003), but the use of a low-power sensor and simplified classification algorithm reduced the energy consumption.

S. Zhang et al. (2012) noticed that the number of sensors and the sampling rate were closely associated with the size and performance (e.g., data storage and processing capacity, reliability of monitoring, and battery life), when using accelerometry devices. So they performed a study to achieve a clear understanding of the effect of different sampling frequencies and number of axes on energy consumption. The aim of the study was to determine whether reduced axes and sampling frequency still provided acceptable classification accuracies. Results showed that when using the wrist-worn GENEA (a triaxial acceleration sensor), high accuracy for classification of sedentary activities, household activities, walking, and running could be achieved using data sampled at 10-20 Hz on a single axis, capturing forward/backward motion. This finding, as they claimed, was important because a lower sampling rate combined with fewer axes of measurement in the monitor would prolong the period of use the accelerometer on a single battery charge and reduce the processing time and power required to detect and classify activities.

Z. Yan et al. (2012, pp. 19) found empirically that the major difference in energy consumption arose from a choice of using none or any of the domain-specific features, because the exclusion or inclusion of a specific frequency or time domain feature had only a marginal effect. In their study, they measured the energy consumption on the smartphone when only time-domain features were used, and measured again when including the frequency-domain features. They found the total energy overhead in continuous physical activity recognition clearly increased with sampling frequency, and the additional energy overhead incurred by including frequency-domain features was a non-linear and logarithmic function of the sampling frequency (Yan 2012). In addition, A. Krause et al. (2005) developed a context-aware eWatch as a sensing and notification platform. They measured the battery lifetime for a time domain analysis and for a frequency domain analysis, separately. The results showed a slight increase because there was no FFT calculation in the time domain analysis.

Some researchers have attempted to adaptively utilize the optimal sampling frequency for each activity instead of adopting one uniform low sampling frequency for all activities to further cut down the energy consumption. Z. Yan et al. (2012, p. 17) introduced an activity-adaptive approach for continuous activity recognition, where the choice of both the accelerometer sampling frequency and the classification features were adapted in real-time according to the activity being performed by the subject. The reason for adaptively changing the values of these two independent parameters in the accelerometer-based activity recognition process was that they jointly affected a tradeoff between two mutually-conflicting objectives, i.e. increasing classification accuracy and lowering energy consumption. Generally, a higher sampling frequency and a richer set of features helps to achieve better classification accuracies, but decrease of the sampling frequency, duty cycle and/or the set of features is needed to reduce the energy consumption (Yan 2012, p. 17). They first investigated how the choice of accelerometer sampling frequency and classification features affected, separately for each activity, the tradeoff between the energy overhead and classification accuracy. They found that such tradeoff was activity specific. Based on the finding, they developed an activity-sensitive algorithm, i.e. A3R, short for Adaptive Accelerometer-based Activity Recognition. This algorithm could run continuously for physical activity recognition, with the sampling frequency and the classification feature set being adaptively switched to the corresponding optimal values in real-time according to the detected activity. They evaluated its performance and found it worked well in reducing the energy consumption. For example, they obtained an overall energy saving of 20% to 25% when running it on the Android phones. But they did not evaluate the classification accuracies after adopting this adaptive strategy.

In this thesis work, single low power tri-axial accelerometer and simple analysis method were used. Based on the literature review, the author focused on studying the effect of the decrease of the sampling frequency to find low sampling frequencies still providing good classification results, and studying the use of an adaptive sampling strategy to further reduce the energy consumption.

2.3.2 Modern commercial solutions

One of the main difficulties with commercial wearables is how to provide enough energy for the electronics to run over a reasonable amount of time without making the battery too large or the device bulky. The reason is that people care the battery life a lot when they buy wearable bands and/or smartwatches.

Apple watch employs GPS to track cycling, and an accelerometer to steps and calories burned. But it cannot automatically recognize activities, and users are required to tape the workout type when they start performing some activity. It has an up to 18-hour battery life. Jawbone UP3 adopts a tri-axis accelerometer to count steps. It claims that users can log their different workout types to get the full picture of their activity. Its battery type is lithium-ion polymer and the battery life is up to 7 days. Fitbit surge can track steps, distance, calories burned, floors climbed and active minutes. But it requires a separate Strava workout app to help recording running, cycling, cross training and other workouts. Related sensors include GPS, 3-axis accelerometers, and a 3-axis gyroscope. Its battery type is lithium-polymer and it has an up-to 7-day battery life. Microsoft band can track steps, calories burned, sleep time, running and biking by using a 3-axis accelerometer, GPS, and a gyrometer. It has dual lithium-ion polymer batteries and provides a 48-hours battery life of normal use. Mi band adopts a military-grade accelerometer to track steps, calories burned and sleep time. It has one ultra-thin lithium polymer battery with a 30-day battery life.

In conclusion, modern commercial products adopt low power sensors and chips, and powerful battery to reduce the energy consumption problem. But advanced hardware increases the cost and thus the price of the product. Besides, with single tri-axis accelerometer, modern commercial products only track steps, calories burned and sleep time. It seems that recognizing several activities is a challenging task to them. For example, GPS is additionally employed to record running and biking. But using more sensors will increase the energy consumption and the cost. Employing single tri-axis accelerometer, low sampling frequency and simple analysis methods would improve the battery life without additional economic cost. Considering the wearables of today even increasing the battery life by a few days would be a huge improvement. So the thesis work is of great interest.

3. PALANTIR PROJECT

This thesis work utilizes the data set collected earlier in the Palantir project, which was performed by the VTT (Technical Research Center of Finland) in 2003 and 2004. Its objective was to study how and which meaningful contexts it would be possible to recognize and classify from the data collected with wearable sensors. Several publications (Pärkkä 2004; Pärkkä 2006; Ermes 2008b) have been issued based on the Palantir project results.

3.1 Overview of the Palantir project

The Palantir project aimed to identify which contexts could be realistically recognized and which sensors and data processing methods were the most useful ones for this purpose. The context refers to any information that describes the surroundings or situation, e.g., the type of the physical activity or a social situation. Different devices can then use the context information in different ways, e.g., for automatically keeping a digital diary, for automatically adapting the user profile, for automatically recommending a service or information for the user, etc.

The approach of the Palantir project to context sensitivity was data-oriented and empirical. A wearable data collection system was developed to allow realistic context data collection with several wearable sensors, from several different realistic contexts by several volunteers. Generally, volunteers carried sensors, and followed a scenario to perform several daily activities. The context information was marked down via an annotation application on a personal digital assistance (PDA). An example of the annotation application is shown in Figure 1. The data collected were stored in a large context data library, i.e., Palantir Context Data Library.

3.2 Key results from the Palantir project

During the Palantir project, the data was collected in a large context data library, which consisted of two parts, i.e., Palantir Context Data Library 2003 (Palantir2003) and Palantir Context Data library 2004 (Palantir2004). Palantir2003 contained 2-hour recordings measured from 16 volunteers. Palantir2004 contained 6-hour recordings measured with 12 volunteers.

J. Pärkkä et al. first did an experimental study with Palantir2003 in 2006. They extracted features from the raw data with 1 Hz sampling rate. The extracted features are roughly depicted in Table 1. Feature selection was performed based on the combination of a prior knowledge and the distribution bar graphs analysis (Pärkkä 2006). Based on this selection method, six features were selected (in order): 1) peak frequency of the up-down chest acceleration; 2) median of the up-down chest acceleration; 3) peak power of the up-down chest acceleration; 4) variance of the back-forth chest acceleration; 5) sum of variances of three-dimensional wrist accelerations; 6) power ratio of frequency bands 1-1.5 Hz and 0.2-5 Hz measured from the left-right magnetometer on chest (Pärkkä 2006, p. 122). Three different classification methods were utilized in the study: the custom decision tree, automatic decision tree and artificial neural network. Their results showed that several daily activities, i.e., lying, rowing, biking, sitting and standing, running, Nordic walking, and walking, could be automatically recognized with good accuracies by using the chest accelerations, wrist accelerations and chest compass. The more detailed classification results are shown in Table 2.

The Palantir2004 was first studied by M. Ermes et al. in 2008. In their study, the extracted features were similar to the ones in (Pärkkä 2006). Instead of the distribution bar graphs they utilized the receiver operator characteristic (ROC) curve, in the feature selection. The extracted features were (in order): 1) peak frequency of the up-down hip acceleration; 2) range of the up-down hip acceleration; 3) mean of the up-down hip acceleration; 4) peak frequency of the horizontal wrist acceleration; 5) sum of variances of the three-dimensional hip acceleration; 6) spectral entropy of the up-down hip acceleration; 7) speed measured from the GPS (Ermes 2008b, p. 22). Four classification algorithms were studied, i.e., custom decision tree, automatic decision tree, artificial neural network, and a hybrid method of custom decision tree and artificial neural network. The activities involved lying, sitting and standing, walking, running, cycling with an exercise bike, rowing, playing football, Nordic walking, and cycling with a regular bike. Their results showed that these activities could be well recognized in controlled and/or uncontrolled conditions. The controlled conditions referred to the supervised periods with exact scenario and accurate supervisor-made annotations, while the uncontrolled conditions referred to unsupervised periods with subject-made annotations.

In Palantir2003, the chest accelerations and chest compass were both originally sampled at 200 Hz, and the wrist accelerations 40 Hz. In 2004, they reduced the sampling frequencies, and sampled both the hip and wrist accelerations with 20 Hz. With all these in consideration, the author of this thesis decided to use the Palantir2003 as the raw data source. More information about the Palantir2003 data set structure will be introduced in the section 3.3.

Figure 1. *Annotation application on PDA. Checkboxes on the left were used to expand and collapse between the title line and full view. Radio buttons were used to mark the active context information. The asterisk was used to mark the context value as "other" (Pärkkä 2006, p.121).*

Table 1. *Features extracted in (Pärkkä 2006).*

	Time-domain features	**Frequency-domain features**
body position; humidity; blood oxygen saturation SaO2; skin resistance; skin temperature; environmental temperature.	mean, variance, median, skew, kurtosis, 25% percentile, 75% percentile.	
accelerations	mean, variance, median, skew, kurtosis, 25% percentile, 75% percentile.	spectral centroid, spectral spread, estimation of frequency peak, estimation of power of the frequency peak, signal power in different frequency bands.
magnetometer signals	mean, variance, median, skew, kurtosis, 25% percentile, 75% percentile, radius, two angles describing the vector of magnetic field.	spectral centroid, spectral spread, estimation of frequency peak, estimation of power of the frequency peak, signal power in different frequency bands, ratio between frequency bands 1-1.5 Hz and 0-5 Hz.
environmental light-intensity signal	mean, variance, median, skew, kurtosis, 25% percentile, 75% percentile.	spectral centroid, spectral spread, estimation of frequency peak, estimation of power of the frequency peak, signal power in different frequency bands, power of frequency band 80-100 Hz.
respiratory effort signal	mean, variance, median, skew, kurtosis, 25% percentile, 75% percentile, tidal volume, amplitude deviation, rate of ventilation.	spectral centroid, spectral spread, estimation of frequency peak, estimation of power of the frequency peak, signal power in different frequency bands, respiratory frequency, frequency deviation, spectral entropies.

Table 2. *Classifier results in (Pärkkä 2006, p. 125).*

	Custom Decision Tree	Automatic Decision Tree	Artificial Neural Network
Lie	**87**	83	74
Row	**69**	56	59
ExBike	79	**82**	75
Sit/ Stand	**96**	95	**96**
Run	**97**	**97**	22
Nordic walk	**90**	72	52
Walk	58	78	**79**
TOTAL	82	**86**	82

3.3 Palantir2003 data set

Palantir2003 was collected by VTT in 2003 during the Palantir project. The scenario and sensors used in the data collection system are listed in Table 3 and Table 4. The scenario reflected the real-world circumstances except that there was an annotator during the data collection.

Table 3. *Scenario for data collection (Pärkkä 2006, p. 121).*

Location	Task
Home	Sitting at home
	Lying
	Sitting & reading newspaper
	Putting clothes on, going out
Bus	Walking to a bus stop
	Waiting for bus
	Traveling in bus
Restaurant	Walking to restaurant
	Queuing
	Eating, drinking, talking
Library	Walking to library
	Sitting in library, reading
Shop	Walking to shop
	Walking in shop, shopping
Home	Walking back home
Outdoor activities	Nordic Walking
	Running
Indoor activities	Rowing (rowing machine)
	Walking
	Bicycling (exercise bike)
	Sitting, drinking

Table 4. *Signals and sensors for the data collection system (Pärkkä 2006).*

Signal	Sensor	Measurement site	Fs
Altitude	Air Pressure (Suunto X6HR)	Wrist	0.5
Audio	Microphone (AKG C417)	Chest, on rucksack strap	22000, mono, 16 bit
Body Position	Metal ball moves between resistors (ProTech Position)	Chest	200
Chest Accelerations	3D acceleration (2 x Analog Devices ADXL202)	Chest, on rucksack strap	200
Chest Compass	3D compass (Honeywell HMC-1023)	Chest, on rucksack strap	200
EKG	Voltage between EKG electrodes (Blue Sensor VL, Embla A10)	Below left armpit, on breastbone	200
Environmental Humidity	Humidity (Honeywell, HIH-3605-B)	Chest, on rucksack strap	200
Environmental Light Intensity	Light sensor with two output dynamics (Siemens SFH 203P)	Chest, on rucksack strap	200
Environmental Temperature	Temperature sensor (Analog Devices TMP36)	Chest, on rucksack strap	200
Event Button	Switch (Embla XN Oximeter)	Chest, on rucksack strap	-
Heart Rate	IR light absorption (Embla XN oximeter)	Finger	1
Heart Rate	IR light reflectance (Nonin XPOD)	Forehead	3
Heart Rate	Voltage between chest belt electrodes (Suunto X6HR)	Chest	0.5
Location	GPS satellite receiver (Garmin eTrex Venture)	Shoulder, on rucksack strap	Based on location
Pulse Plethysmogram	IR light reflectance (Nonin XPOD)	Forehead	75
Respiratory Effort	Piezo sensor (Pro-Tech Respiratory Effort)	Chest	200
SaO2	IR light absorption (Embla XN Oximeter)	Finger	1
SaO2	IR light reflectance (Nonin XPOD)	Forehead	3
Skin Resistance	Resistance between two metal leads (Custom-made)	Chest	200
Skin Temperature	Resistive temperature sensor (YSI 409B)	Upper back, below neck	200
Wrist Accelerations	3D acceleration (Analog Devices, ADXL 202E)	Wrist, dominant hand	40
Wrist Compass	2D compass (Honeywell HMC-1022)	Wrist, dominant hand	40

Palantir2003 consists of two parts, one video file showing how the data was collected, and a set of folders named as “PalantirXXX” storing the data from each volunteer. Each folder contained a text file and a named “Processed” folder. The text file stored the personal information of the volunteer, i.e., gender, age, weight and height. The “Processed” folder stored a cluster file and several other files in REC, PSV and MFLT formats.

The chest accelerations and chest compass were stored in REC format, and annotation information in PSV format. The REC format is an extension of the European Data Format (EDF). The EDF format was specified in (Kemp 1992) and commonly used in the field of biomedical engineering. The PSV format was developed in Palantir project (Pärkkä, 2004). It is based on the common CSV and TSV file formats. The acronym PSV stands for the prefixed separated values. The format is a textual format that is primarily used to represent tabular data as rows. Figure 2 gives an example that shows how the data is organized in a PSV file.

```
#HEADER|time|log=open|social=alone|location=VTT|weather=cloudy|activity=stand|activity=sit|
#FORMAT|T@s|N|N|N|N|N|N|N|N|N|N|N|N|N|N|N|N|N|N|N|-|N|N|N|-|N|N|N|N|-|-
|2003-11-03T11:43:49|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:43:50|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:44:13|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:44:14|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:44:17|1|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:44:19|1|1|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:46:58|1|1|1|1|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:50:08|1|1|1|1|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:52:05|1|1|1|1|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:54:13|1|1|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:55:29|1|1|1|1|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:56:50|1|0|1|1|0|0|0|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:58:12|1|0|0|1|0|0|0|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:58:16|1|1|0|1|0|0|0|1|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T11:59:01|1|1|0|1|0|0|0|1|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:00:46|1|1|0|1|0|0|0|1|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:06:35|1|1|0|1|0|1|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:09:35|1|1|0|1|1|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:09:56|1|1|0|1|0|0|0|0|0|0|0|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:12:32|1|1|0|1|0|0|0|1|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:14:41|1|1|0|1|0|0|0|0|0|0|0|1|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:17:01|1|1|0|1|0|0|0|1|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:17:06|1|1|0|1|0|0|0|0|0|0|0|1|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:24:31|1|1|0|1|0|0|0|1|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:24:39|1|1|0|1|0|0|0|1|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:26:26|1|1|0|1|0|0|0|0|0|1|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:29:11|1|1|0|1|1|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:29:28|1|1|0|1|0|0|0|0|0|1|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:32:46|1|1|0|1|1|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:33:50|1|1|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:34:47|1|1|1|1|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:36:54|1|1|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:38:43|1|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:38:51|1|1|0|1|0|0|0|0|0|0|1|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:45:05|1|1|0|1|1|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0
|2003-11-03T12:45:08|1|1|0|1|1|0|0|0|0|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0
```

Figure 2. *An example of PSV format.*

4. ACCELERATION DATA ANALYSIS

In this thesis work, one accelerometer attached on the chest was chosen for the human physical activity classification. This sensor was chosen as it provided the most important features for physical activity classification in the earlier study (Pärkkä 2006). The utilized tri-axial accelerometer consisted of 2 Analog Devices ADXL202, which are low power (<1.2 mA). Effects of different sampling frequencies and three training strategies were studied. The data analysis of chest accelerations was performed with the software MATLAB, which provides a multi-paradigm numerical computing environment. Finally, adaptive sampling rate logic was implemented to perform the adaptive sampling with 1 Hz for lying activity and 10 Hz for non-lying activity.

The author of this thesis implemented the whole data analysis process and the adaptive sampling rate logic independently.

4.1 Data pre-processing

In Palantir2003, the raw chest accelerations are continuous bio-signals and contain some noise. The stored annotation information was discrete, because it was the set of combinations of timestamps, location names, and activity names.

In the thesis work, the first step of the data pre-processing was to convert the discrete annotation information data sets into continuous class label sets. In details, each time went through each line of the PSV file; took the current timestamp as t0 and the timestamp of the closest following line as t1; set the class labels as the same one in the current line for the period between the two timestamps. There was a synchronization problem, because the annotations and measurement signals did not start and/or end at the same time. This problem could be solved by keeping the common parts of the annotations and measurement signals with the help of timestamp data.

The second step in the analysis was to discard 8-second data at the two ends of one activity for both the chest acceleration data set and its corresponding class label set. This step helped to remove or reduce annotation errors and transition periods between different activities. The choice of 8-second was empirical and might not be the optimal one.

The third step applied a simple Butterworth low-pass filter with 6 Hz being the 3 dB point, only to the chest acceleration data set to remove the high-frequency noise contained in the data. The filtered signals still kept almost all the needed information for physical activity recognition, because J. Pärkkä et al. (2006, p. 122) found that acceleration signals measured during activity running were 2.5-3 Hz oscillation signals.

The fourth step applied a 1-second-length window to segment the two data sets, because the classification algorithms could not work on the bio-signals or time-series data directly. With such a short-length window, the activity recognition was performed in “real-time”. For the class label set, the fourth step adopted the majority rule over each segment to get its class label.

In Palantir2003, the chest accelerations were originally sampled at 200 Hz. To analyse the effect of sampling frequency, the down-sampling was performed on the pre-processed data to get the new data sets for sampling frequencies, 20, 10, 8, 5 and 4 Hz separately. These new generated data sets were then used for physical activity recognition individually. In the down-sampling process, no anti-alias filter was used as it was supposed to simulate the situation where the sampling rate was dynamically tuned without modifying the analogue anti-alias filter. This might result in aliasing especially with lower sampling rate.

4.2 Feature extraction

Generally, features can be divided into two categories, i.e. time-domain features and frequency-domain features. In the thesis work, the extracted features are listed in Table 5 below.

Time-domain features were extracted from each segment, but frequency-domain features from four segments, i.e. three right former segments and the current one, to which the features referred. The reason for changing to four segments was to improve the frequency resolution, because the spectrograms of each activity data showed that some activities were separable with frequency resolution 0.25 Hz, but not with frequency resolution 1 Hz. The spectrograms of each activity data also gave the choices of targeted frequencies and frequency bands for physical activity recognition.

Table 5. *Extracted candidate features for physical activity recognition*

Time-domain	Maximum, mean, median, variance, covariance, range, interquartile, magnitude, third-order central moments, skewness, kurtosis, total energy, and peak energy.
Frequency-domain	Peak frequency, spectral entropy, power of targeted frequencies, energy of frequency bands, and energy ratio of frequency bands.

4.3 Custom decision tree

In the thesis work, the binary CDT was used to implement the physical activity classification, the same method used by Pärkkä (2006, p. 123). The CDT (Figure 3) was built by using domain knowledge and visual inspection of the chest accelerations. Normally, binary decision tree algorithms perform a binary partition in the feature space. For a continuous variable and binary tree, the splitting rule may simply be determining a threshold that will be used to partition the dataset at each testing node. This simple thresholding mechanism has been successfully used in (Pärkkä 2006; Ermes 2008b). The decision tree can employ more complex node questions involving a linear or non-linear combination of variables, which in turns results in partitions based on more complex decision surfaces. In this case, a richer hypothesis set could be generated. But this can cause overfitting problem in the absence of a sufficiently large training sample, and will increase the complexity of the algorithm and the computational workload. So the thresholding method on a single variable is preferred and adopted in this thesis work. Because the extracted features in this thesis work were continuous, threshold values needed selecting for each testing node. This thesis work adopted the knowledge from the information theory in threshold values selection. In details, the entropy gain worked as the selecting criterion and this method was explained in (Mitchell 1997). It selected the threshold that resulted in the largest entropy gain for each testing node.

The built tree has 6 leaf nodes and 5 binary testing nodes. At each testing node, a sample data goes to the right branch, if the value of its some specific feature is larger than the threshold value, otherwise it goes to the left branch. The leaf nodes store the class labels, and each one only has a distinct class label. When a sample data finally reaches at a leaf node, it has its class label as the one stored in that leaf node.

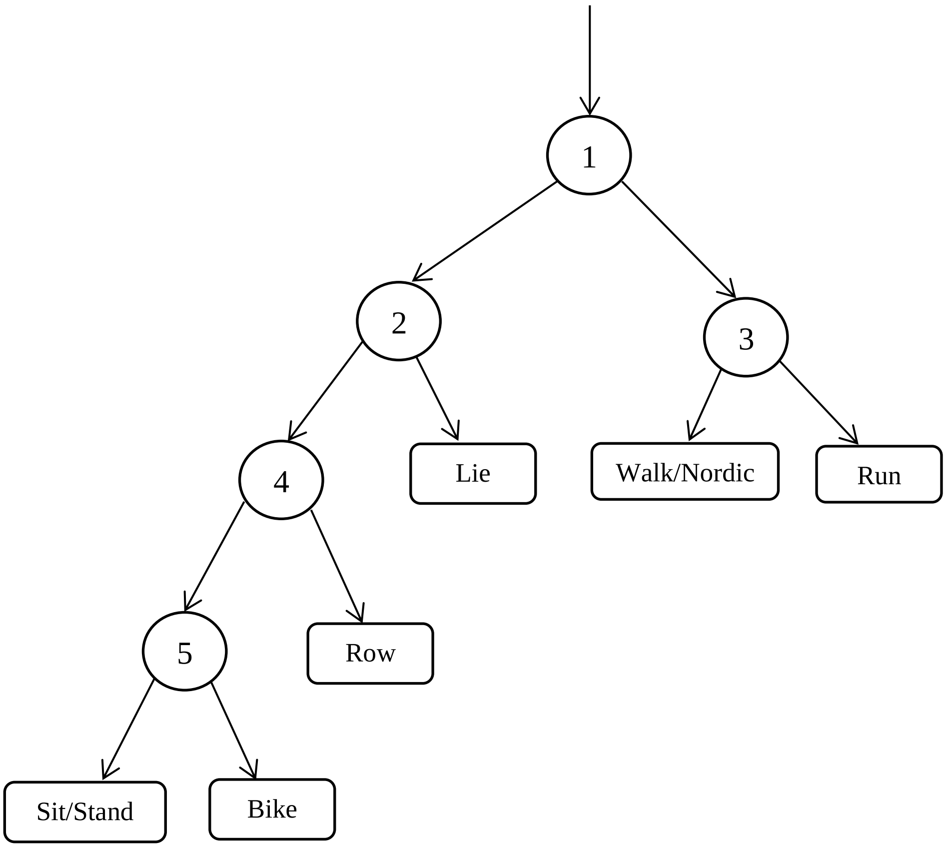

Figure 3. *Custom decision tree.*

The CDT used in the thesis work is slightly different from the one (Figure 4) used in (Pärkkä, 2006). Specifically, the two leaf nodes for “walking” and “Nordic walking” were merged into one leaf node labeled as “Walk/Nordic”, and the corresponding testing node was removed. The reason for this change was that J. Pärkkä et al. (2006, p.122) used the sum of variances of three-dimensional wrist accelerations to distinguish “walking” from “Nordic walking”, and in the thesis work, it was found that the classification accuracies for “Nordic walking” and “walking” were not good when only the chest accelerations were employed. More specifically, the obtained classification accuracies of “Nordic walking” and “walking” were about 73% and 55% respectively, while J. Pärkkä et al. (2006, p. 125) obtained 90% and 58% correspondingly. When these two activities were combined into one activity as “walking”, the classification accuracy obtained in the thesis work for this combined activity was 76.93%. Based on the confusion matrix showed as Table 6 (Pärkkä 2006, p. 124), the classification accuracy for the combined activity “walking” was 78.89%. Comparing these two results, it was decided to use the combined activity “walking” instead of the separate two activities “walking” and “Nordic walking” in this study. In addition, to make it more clear to show how the sample data flows in the custom decision tree, a modification was made on the placement of the sub-branches of each testing node. In details, the right branch indicates that some specific feature value is larger than the threshold value at that testing node, and vice versa.

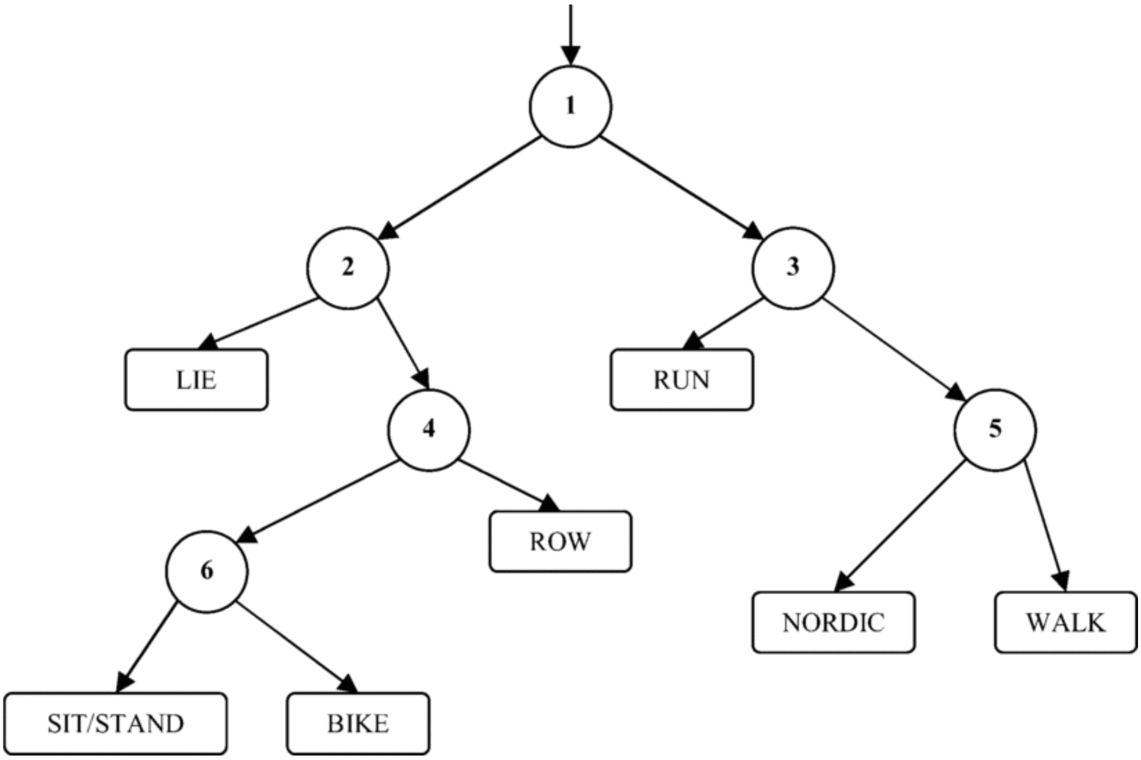

Figure 4. *Custom decision tree used in (Pärkkä 2006, p. 123).*

Table 6. *Confusion matrix for custom decision tree in (Pärkkä 2006, p. 124).*

Annotation	Recognized Activity						
	Lie	Row	Ex-Bike	Sit/ Stand	Run	Nordic walk	Walk
Lie	1417	0	0	205	0	0	0
Row	0	1646	0	717	0	0	23
ExBike	0	0	2461	612	0	0	29
Sit/ Stand	121	40	53	34083	4	340	962
Run	0	0	0	44	2284	21	5
Nordic walk	0	1	0	256	39	4507	194
Walk	0	16	4	5412	15	3964	12797

4.4 Feature selection

In the thesis work, many different time-domain and frequency-domain features were extracted as shown in Table 5. Some of the extracted features were not useful or performed very poorly in physical activity recognition.

The objective of the feature selection is to find powerful and useful features. As discussed in the section 4.3, the simple thresholding method on a single variable at each testing node was adopted in the construction of the binary CDT. As a consequence, the feature selection problem is constrained to find one feature for each testing node. Here, the study adopted distribution bar graphs or boxplots analysis in feature selection. This method starts with making distribution bar graphs or boxplots of each feature signal during different activities. Then it makes a comparison and selects the features according to the following rule. The less the distributions overlaps, the better the feature is for the discrimination of activity. Thus, the selected feature at each testing node is the one that most separates the two groups.

Based on this selection method, features selected for the thesis work are (listed according to the order of their targeted testing nodes): 1) energy of frequency band 1.75-3 Hz of up-down chest acceleration; 2) energy of left-right chest acceleration; 3) energy of frequency band 2.5-3 Hz of up-down chest acceleration; 4) energy of frequency band 0.25-1.5 Hz of back-forth chest acceleration; 5) energy of frequency band 1-1.5 Hz left-right chest acceleration. These chosen features were totally different from the ones selected in (Pärkkä 2006, p. 123). The difference might result from the fact that J. Pärkkä et al. performed the feature selection with a prior knowledge and the distribution bar graphs analysis (Pärkkä 2006).

4.5 Three training models

The thesis work studied and compared three training models in physical activity recognition, i.e. impersonal models, personal models and hybrid models. The reason why these models were studied in this work, was that G. M. Weiss (2012, p. 98) claimed they found the personal model could obtain much better classification accuracies than the impersonal model. Thus, it is more likely that a lower sampling rate could be found with the personal model, when considering the tradeoff between the sampling rate and classification accuracies.

These three models were first formally introduced in (Weiss 2012). Some of them were called with different names in other studies. For example, the leave-one-subject-out training method (Bao 2004, p. 11) and the subject-independent analysis (Tapia 2007) both correspond to the impersonal model, and the user-specific training (Bao 2004, p. 11) and the subject-dependent analysis (Tapia 2007)to the personal model. When the whole data is used with K-fold cross-validation (Siirtola 2009), it is actually in the hybrid model.

Essentially, they were based on the way of partition of the data into training and test data sets. In impersonal models, training data sets and test data sets are from different users. There is no overlapping between them. Impersonal models have the advantage that the classifier can be built once for all users, so they do not require the users to provide their own data for training or building the classifier. For personal models, both the labeled training data and test data are from only the user from whom the classifier is intended. Personal models have the advantage that they consider the individual difference and may match the idiosyncrasies of the intended user, but they require each user to provide training data. Obviously, the hybrid model is a mixture of above two models. In the thesis work, the leave-one-subject-out cross-validation applied in impersonal models, and 10-fold cross-validation both in personal models and hybrid models.

The impersonal model has the advantage that the classifier can be built once for almost all users and has the information from many users. When people buy the product, they can start using it directly without providing their own data for training. But in some application(s), the interpersonal or individual differences play an important role. To overcome this kind of problems, the personal model can be employed. When using the personal model, there is no built or trained classifier before people using the product. Actually, people are required to provide the training data individually. With the personal model, it is possible used to build a private or personal classifier for each user to match the idiosyncrasies of the intended user. However, it cannot use the information from the other users in building the classifier.

4.6 Adaptive sampling rate logic

Normally, people spend about one third of their time in sleeping per day. During the sleeping, most of the activities may be viewed as lying. When performing the feature selection in section 4.4, it was noticed that the lying activity could be well distinguished from the other activities with a low sampling frequency, e.g., 1 Hz. This fact is appealing because a lot of energy will be saved without much loss of the classification accuracies for all the targeted activities, if using a low sampling frequency for activity lying, and a higher sampling frequency for other activities. Adaptive sampling rate logic is built for this purpose.

The flow chart for the designed sampling logic is demonstrated in Figure 5. There are only two choices of sampling frequencies, i.e., 1 and 10 Hz. This logic aims to sample the signals from activity lying with 1 Hz, and changes the sampling frequency to 10 Hz for the other activities. The “threshold4SL” is used to evaluate whether the energy of 1-second-length left-right chest acceleration is larger than it or not when the data is sampled at 1 Hz. It is a scalar type and must be assigned a fixed value before using the adaptive sampling rate logic, which will keep constant once the logic is used. The “lieCount” counts the number of classification result as lying, but it is only used when the sampling frequency is 10 Hz and the right former 1-second-length data is also sampled at 10 Hz. To be more specific, it records the successive lying activity on a second-by-second basis. The possible values of “lieCount” vary from 0 to 4. With 0, it means there is no successive lying activity. With 4, it indicates there is 4-second successive lying activity. The “shiftFlag” is a Boolean variable. It indicates whether the former one-second data is sampled at 1 Hz. The value 1 refers to yes, and 0 to no. If its value is 1, the following data is sampled with 10 Hz for 4 seconds, otherwise for 1 second. The “fs” stores the sampling frequency value, either 1 or 10. Initially, the “lieCount”, “shiftFlag” and “fs” are set as 0, 0 and 10, respectively. The “oldData” is a 3-second-length data, sampled at 10 Hz. It is the right former 3 seconds’ data of the current segment. The “testData” is a 4-second-length data, consisting of the “oldData” and the current 1-second-length segment. It is also sampled at 10 Hz. The reason for using it is that frequency-domain features are extracted from 4-second-length data for the sake of frequency resolution. The “E(ChestAcc, y)” represents the energy of 1-second-length left-right chest acceleration when the sampling frequency is 1 Hz. It is a feature in time-domain. With 1 Hz sampling frequency, it is actually a single value.

The adaptive sampling rate logic is initialized with a built classifier, a fixed “threshold4SL”, and its parameters, the “lieCount”, “shiftFlag” and “fs”, being set with default values. The logic starts working with sampling the signals at 10 Hz for 3 seconds, and initializing the “oldData” with the sampled data. Then it checks whether the current choice of sampling frequency is 10 Hz. If yes, it then checks whether the “shiftFlag” is true. If yes, it first sets “shiftFlag” as false, and then samples the signals at 10 Hz for 4 seconds. This sampled 4-second data is stored as “testData”. The “oldData” is set as the later 3-second data of “testData”. Features are extracted from “testData”, and later used in classification. Because the classification result is only for one second data, an additional step assigns the same classification result to each of the former four segments. If the value of “shiftFlag” is false, it just samples the signals at 10 Hz for only 1 second. This new sampled data is then jointed with the “oldData” to get “testData”. After this step, the “oldData” is updated as the later 3-second-length data of “testData”. This combined “testData” is used to extract features for classification. If the classification result is lying, the value “lieCount” increases by 1, otherwise “lieCount” is set as its default value 0. For nonzero “lieCount”, a comparison is made with 3. If “lieCount” is larger than 3, the sampling frequency is changed to 1 Hz and itself is set as 0.

At the beginning, a check is made on “fs”. If the decision is that its value is 1 Hz, the adaptive sampling rate logic performs the sampling of the signals at 1 Hz for 1 second. Then the feature “E(ChestAcc, y)” is calculated based on the sampled data. This is used in comparison with “threshold4SL”. If it is larger than “threshold4SL”, the class label for the sampled 1-second data is set as lying, otherwise the sampling frequency is changed to 10 Hz by setting “fs” as 10, the “shiftFlag” is set as 1.

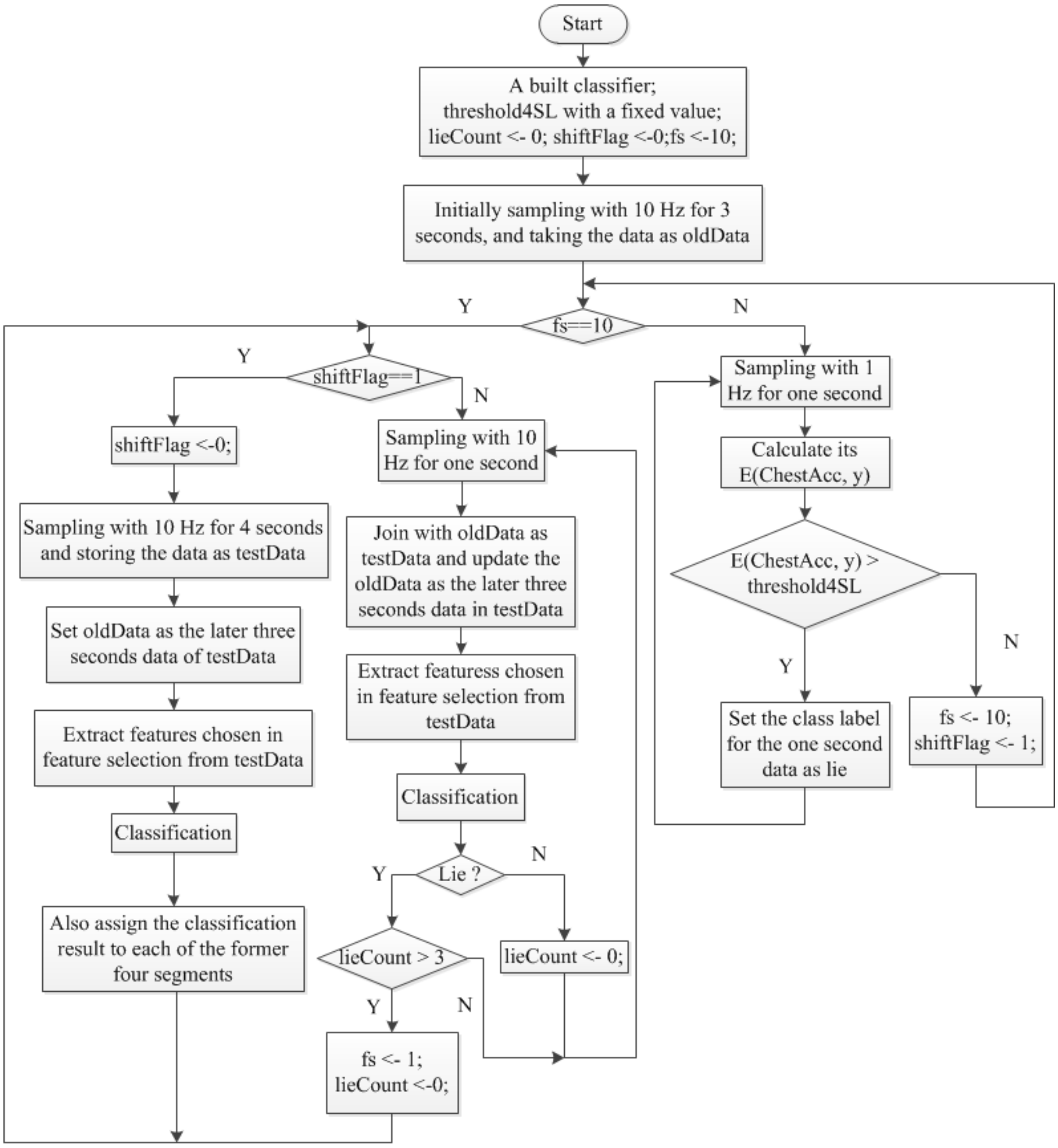

Figure 5. *Flow chart for the adaptive sampling rate logic.*

5. RESULTS AND DISCUSSION

There are three sections in this chapter. Section 5.1 compares the classification results in (Pärkkä 2006) with the ones in this thesis work to show that the processing methods used in data pre-processing step are suitable. In section 5.2, different classification results for different sampling frequencies and training models are demonstrated, compared and analyzed. Section 5.3 shows the results obtained by the adaptive sampling rate logic. Compared with the results showed in section 5.2, the ones in section 5.3 reflect both the effectiveness and efficiency of the built sampling logic.

5.1 Classification results comparison

In the thesis work, the chest accelerations in Palantir2003 served as the data source. Because J. Pärkkä et al. had performed one study on Palantir2003, it was advisable to get similar classification results as Pärkkä's with their specifications before performing the analysis for other sampling frequencies. But it is very hard to get the exact same results because of possible different methods used in the data preprocessing.

As introduced in the section 4.1, in this work the data-preprocessing involved four steps. After the data pre-processing, there were 72 961 one-second-length segments left, compared with 72 272 segments obtained by J. Pärkkä et al. (2006, p. 124). The Figure 6 demonstrates almost the same portions of different activity data as the corresponding ones showed in the Figure 7. The differences are that the percentage of biking increases by about 2%, and the combined percentage of sitting and standing reduces by about 2%.

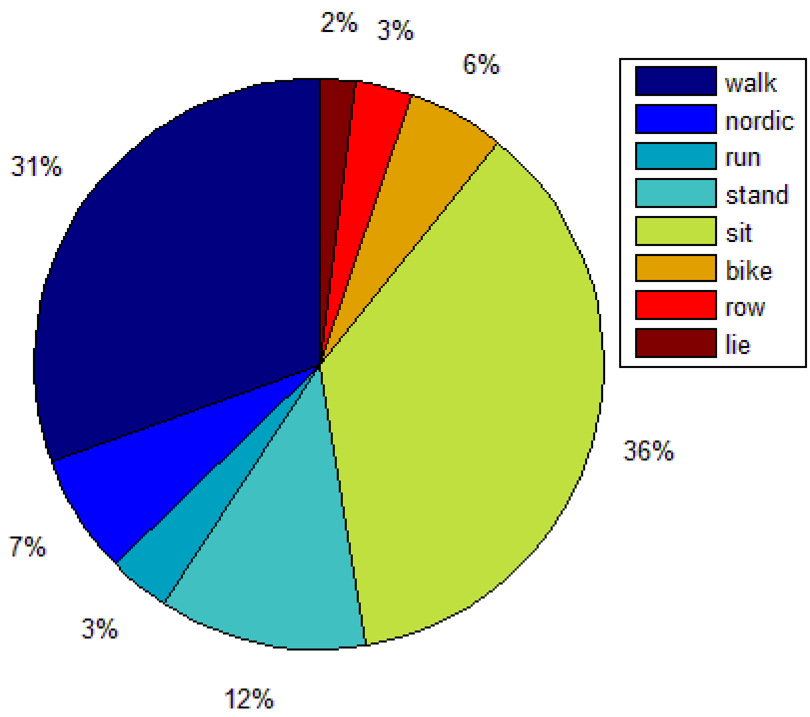

Figure 6. *Portions of activities in annotation. Activities clockwise from 12 o'clock: lying, rowing, sitting, standing, running, Nordic walking, walking.*

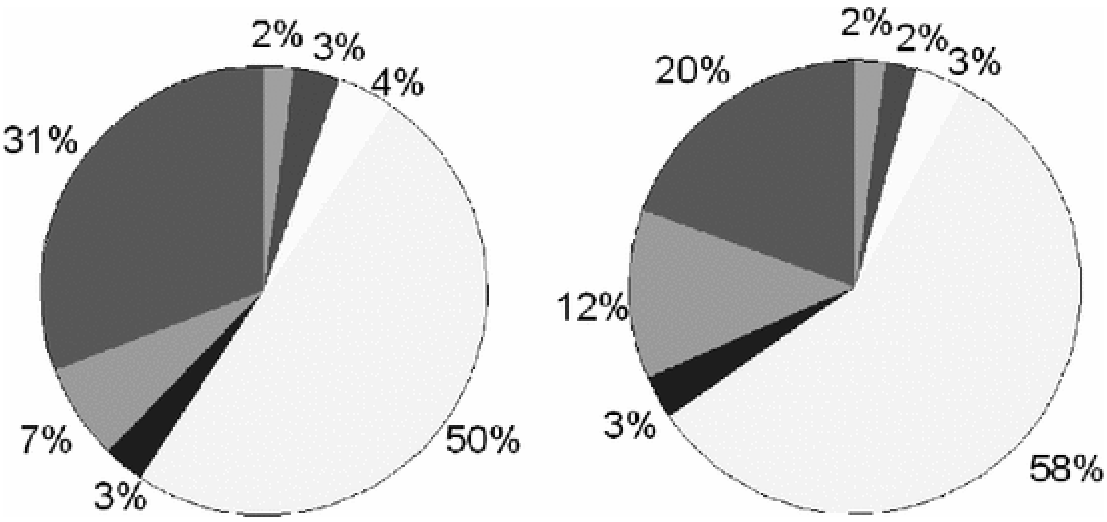

Figure 7. *Portions of activities in annotation (left) and results of CDT (right). Activities clockwise from 12 o'clock: lying, rowing, cycling, sitting/standing, running, Nordic walking, walking (Pärkkä 2006, p. 126).*

The Table 7 demonstrates two sets of classification results with the CDT as the classification algorithm and 200 Hz as the sampling frequency. In the (Pärkkä 2006), the activity Nordic walking and activity walking were not combined. The classification accuracy of combined activity Nordic walking and walking in first column was calculated based on the confusion matrix for CDT in (Pärkkä 2006, p. 124). The absolute and relative differences between each pair of the two classification accuracies for each activity show that except the classification accuracies for activity rowing, the left are similar. For rowing, the relative classification accuracy increases by 16.84%. Table 8 and Table 9 demonstrate the corresponding confusion matrices. Comparing these two confusion matrices, it is noticed that the misclassifications of the two combined activities, sitting and standing, and Nordic walking and walking, increase, but the left are similar. These two comparisons indicate that the methods used in data pre-processing and selected features are suitable in this study

Table 7. *Classification accuracies with CDT.*

Activity	**Classification accuracies (%) with CDT in the impersonal model**		**Absolute differences (%)**	**Relative differences (%)**
	in (Pärkkä 2006)	**in this thesis work**		
lying	87.36	88.90	+1.54	+1.76
rowing	68.99	80.61	+11.62	+16.84
biking	79.34	78.44	-0.90	-1.13
sitting and standing	95.73	93.45	-2.28	-2.38
running	97.03	96.92	-0.11	-0.11
Nordic walking and walking	78.89	76.93	-1.96	-2.49
Unequally weighted overall	87.66	86.05	-1.61	-1.84
Equally weighted overall	84.56	85.88	+1.32	+1.56

Table 8. *Confusion matrix obtained via CDT in the impersonal model by J. Pärkkä et al. (Pärkkä 2006, p. 124).*

		Predicted activities						Total
		bike	**lie**	**row**	**run**	**sit/stand**	**Nordic/walk**	
Annotated activities	**bike**	2461	0	0	0	612	29	3102
	lie	0	1417	0	0	205	0	1622
	row	0	0	1646	0	717	23	2386
	run	0	0	0	2284	44	26	2354
	sit/stand	53	121	40	4	34083	1302	35603
	Nordic/walk	4	0	17	54	5668	21462	27205
Total		2518	1538	1703	2342	41329	22842	

Table 9. *Confusion matrix obtained via CDT in the impersonal model in this thesis work.*

		Predicted activities						Total
		bike	**lie**	**row**	**run**	**sit/stand**	**Nordic/walk**	
Annotated activities	**bike**	3177	0	62	0	515	296	4050
	lie	8	1337	15	0	134	10	1504
	row	15	0	1900	0	201	241	2357
	run	5	0	3	2427	14	55	2504
	sit/stand	744	120	617	3	32959	826	35269
	Nordic/walk	816	2	275	49	5152	20983	27277
Total		4765	1459	2872	2479	38975	22411	72961

5.2 Effects of sampling frequencies and training models

Aliasing may have affected some results and the results might differ if anti-alias filter was used before down-sampling. But it is unlikely that they would affect the results significantly.

Figure 8-13 show the classification results for each activity with different training strategies and varying sampling frequencies. For activity “biking”, “rowing”, and “lying”, the personal model achieves clearly better classification accuracies than the impersonal model and hybrid model do. For activity “running”, the joint activity “sitting and standing”, and the combined activity “Nordic walking and walking”, these three models obtain almost the same results. According to these comparisons, the personal model is preferred in terms of the classification accuracy.

In Figure 8 and Figure 13, classification accuracies for activity “biking”, and “Nordic walking and walking”, increase at 4 Hz and 5 Hz. In Figure 11, the classification accuracy of activity “running” suddenly drops at 5 Hz. These may result from aliasing. Figure 9 shows an interesting finding that the classification accuracy of activity “lying” almost keeps stable when the sampling frequency changes. This also gives the support for building the adaptive sampling rate logic.

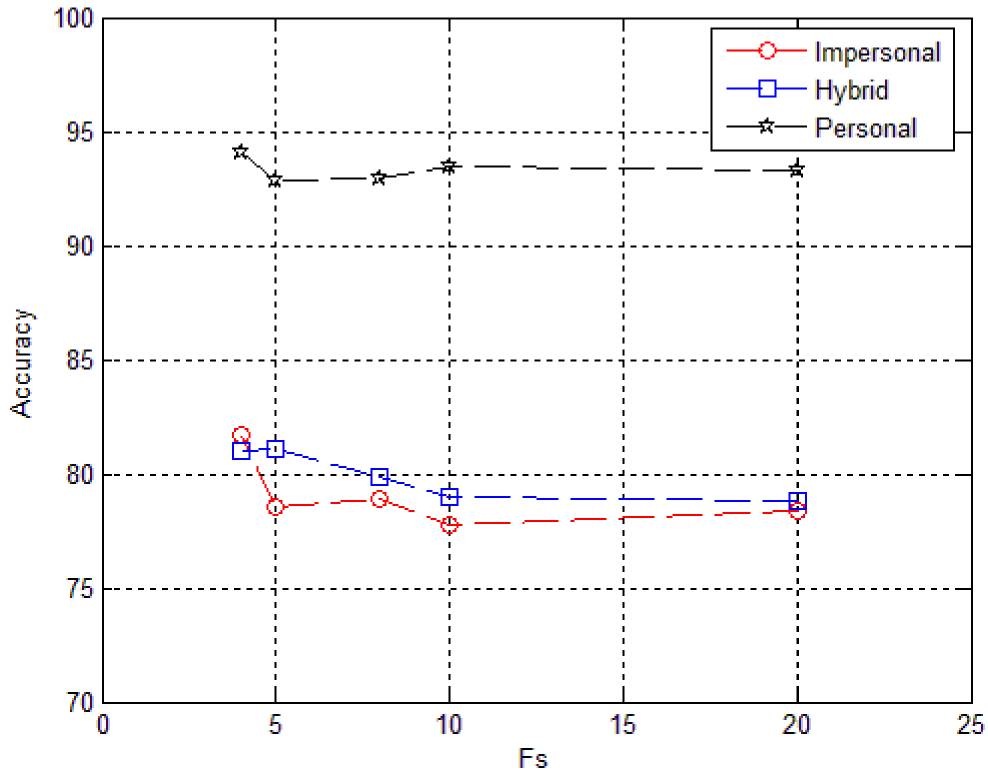

Figure 8. *Classification results for activity biking.*

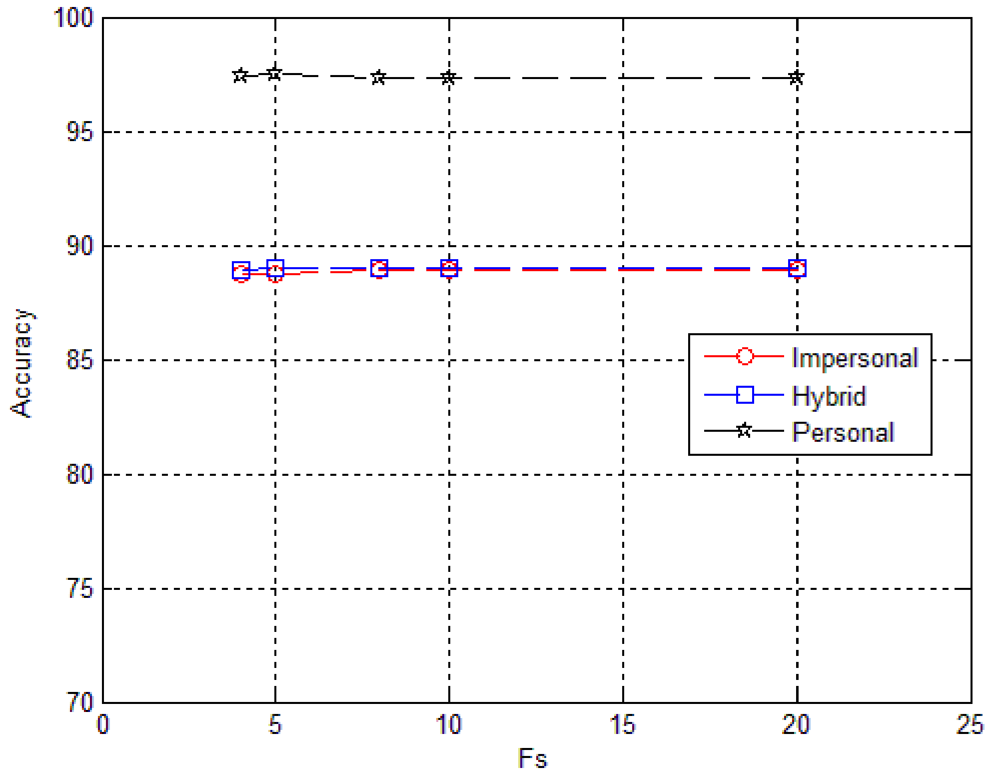

Figure 9. *Classification results for activity lying.*

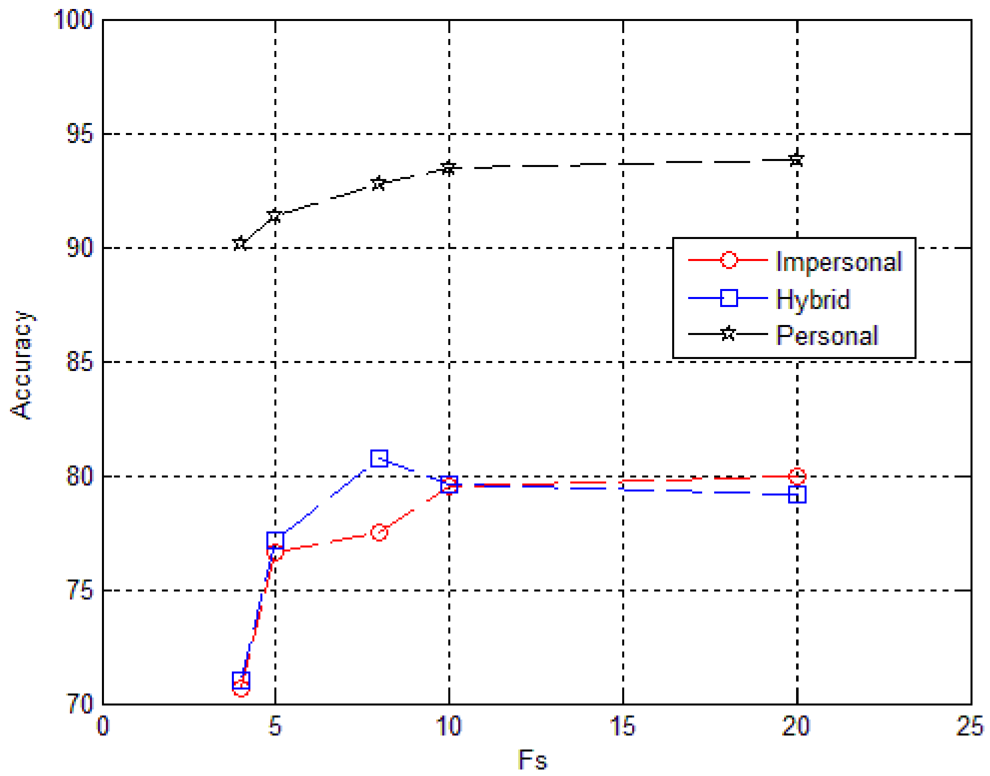

Figure 10. *Classification results for activity rowing.*

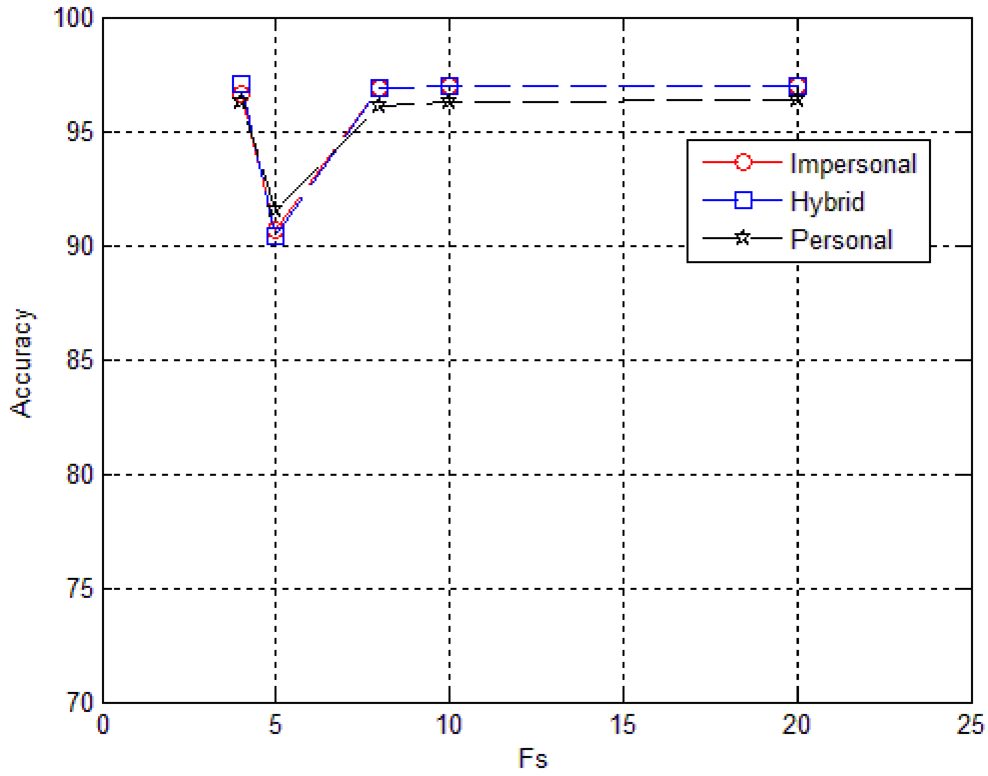

Figure 11. *Classification results for activity running.*

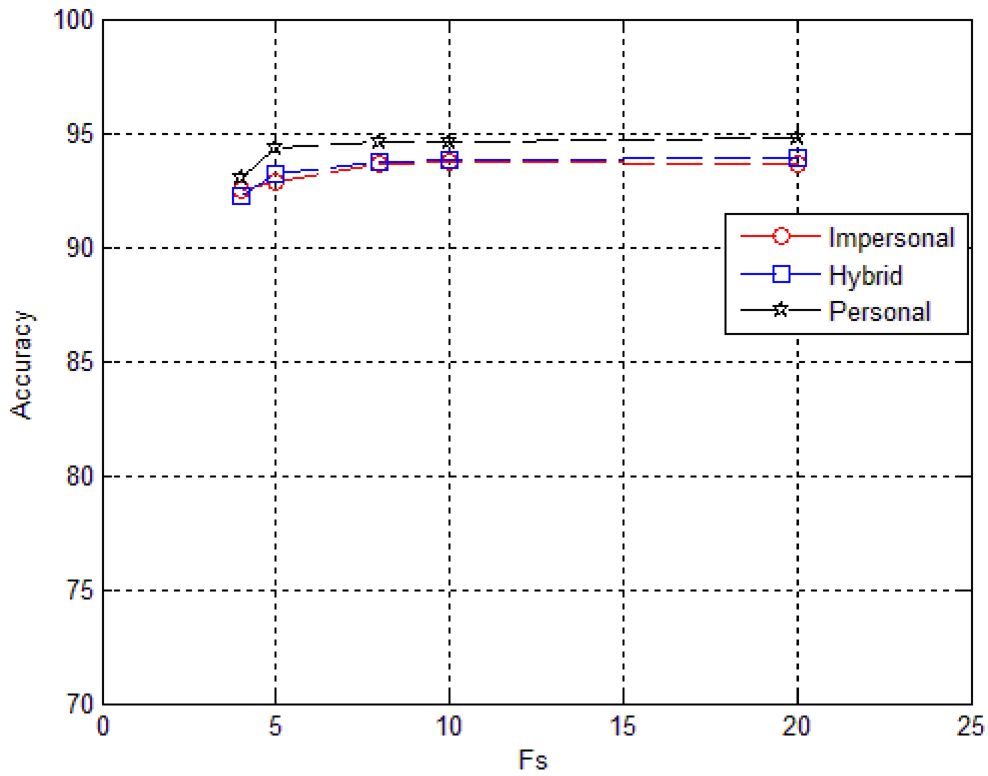

Figure 12. *Classification results for combined activity sitting and standing.*

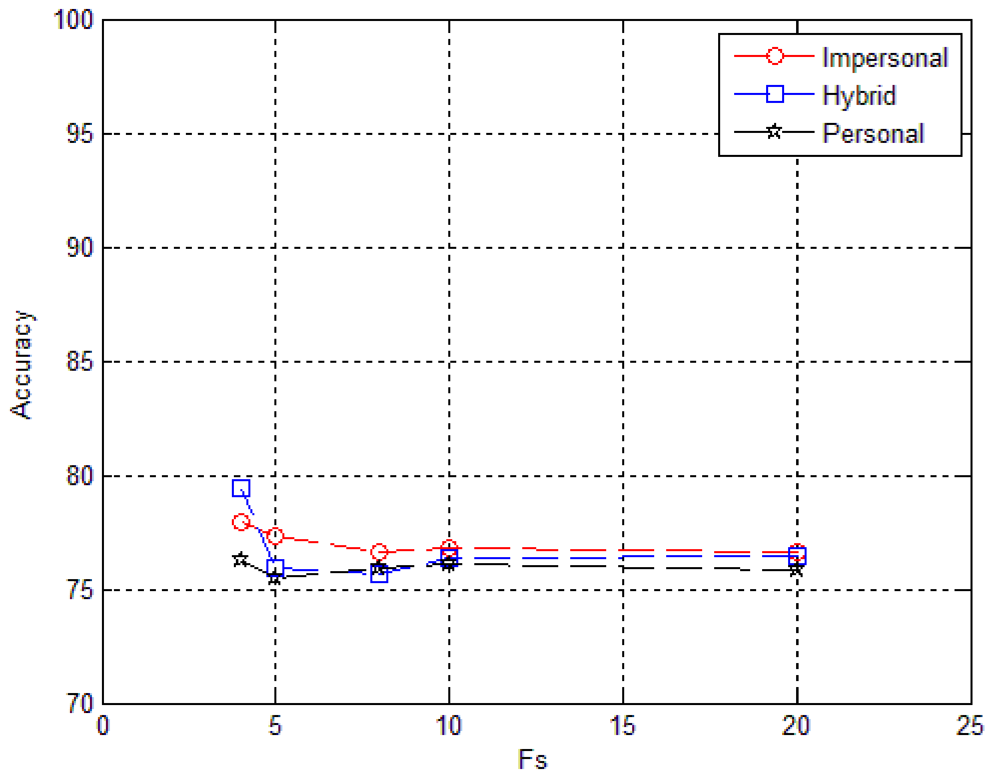

Figure 13. *Classification results for combined activity Nordic walking and walking.*

Figure 14-16 demonstrate the classification results for all activities for each training model. Comparing the Figure 14and Figure 15, it is found that the impersonal model achieves similar results as the hybrid model does. Specifically, they both obtain very good classification accuracies for activity “running”, “lying”, and joint activity “sitting and standing”, and good accuracies for activity “rowing”, “biking”, and combined activity “Nordic walking and walking”. In the Figure 16, it demonstrates that the personal model clearly improves the classification accuracies for activity “lying”, “rowing”, and “biking”. Additionally, all activities except the “Nordic walking and walking” were very well recognised with the personal model. These three figures all indicate that there is indeed no uniform optimal sampling frequency for all activities in terms of classification accuracy only.

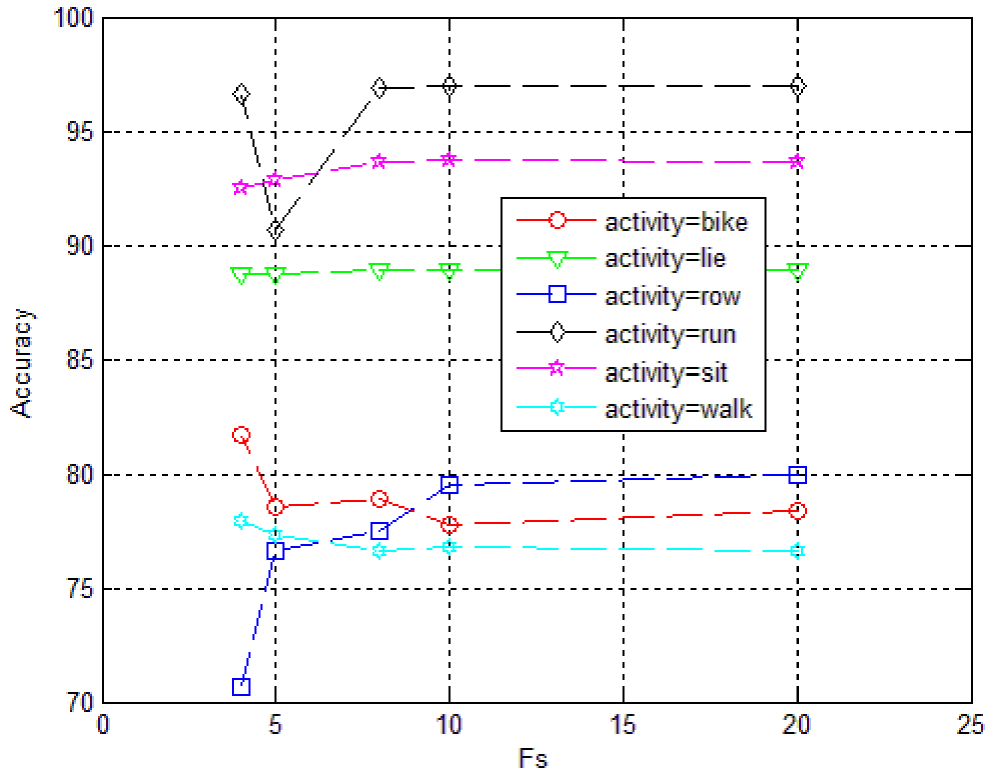

Figure 14. *Classification accuracies for different activities with the impersonal model.*

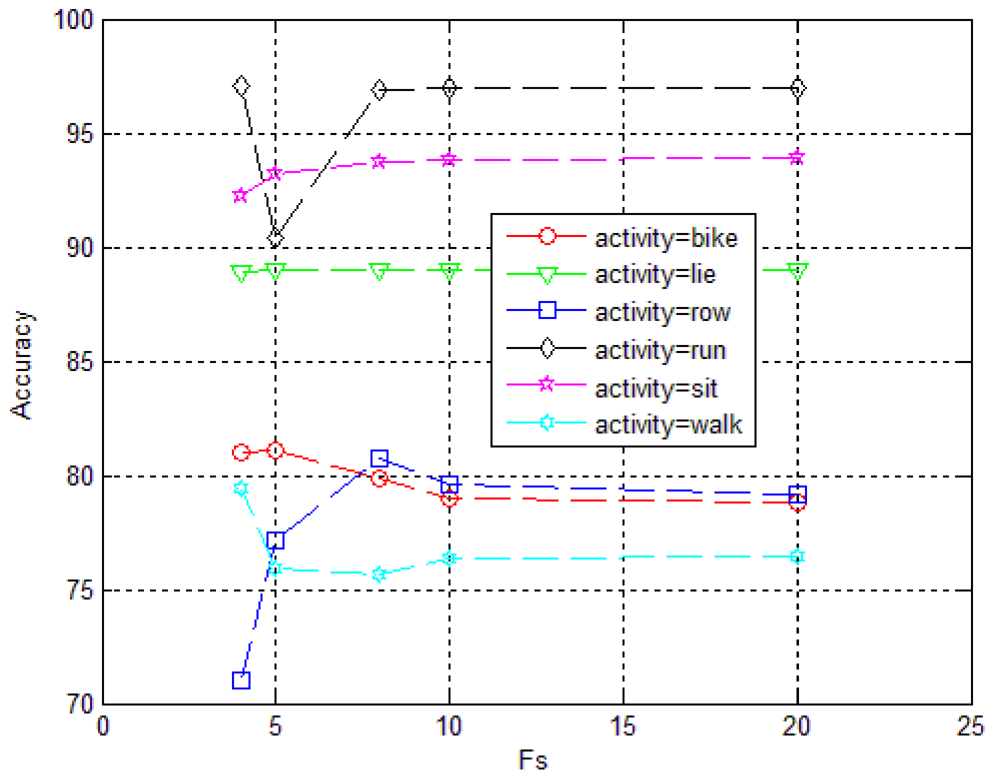

Figure 15. *Classification accuracies for different activities with the hybrid model.*

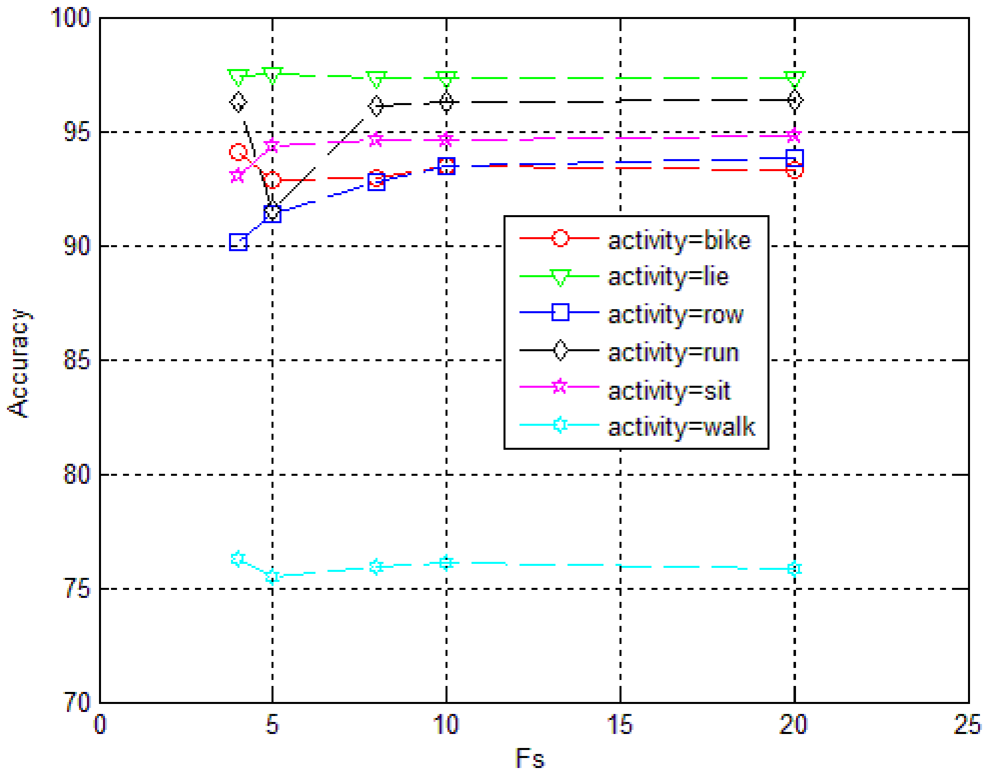

Figure 16. *Classification accuracies for different activities with the personal model.*

Table 10 demonstrates the suitable choices of sampling frequencies for each activity over the three user models. In addition, the choices are made based on the classification accuracies and energy consumption.

Table 10. *Suitable sampling frequencies.*

	Suitable choices of sampling frequencies (Hz) in terms of classification accuracies and energy consumption		
	Impersonal model	**Hybrid Model**	**Personal model**
Bike	4, 5, 8, 10	4, 5, 8, 10	4, 5, 8, 10
Lie	4, 5, 8, 10	4, 5, 8, 10	4, 5, 8, 10
Row	5, 8, 10	5, 8, 10	4, 5, 8, 10
Run	4, 8, 10	4, 8, 10	4, 8, 10
Sit/stand	4, 5, 8, 10	4, 5, 8, 10	4, 5, 8, 10
Nordic/walk	4, 5, 8, 10	4, 5, 8, 10	4, 5, 8, 10

5.3 Results for the adaptive sampling rate logic

In the thesis work, adaptive sampling rate logic, described in section 4.6, was implemented both in the impersonal model and the personal model. Its objective is to reduce energy consumption by using a low sampling frequency for activity lying, and a higher sampling frequency for other activities.

In the impersonal model, those same 12 subjects used in (Pärkkä 2006) were studied. After data pre-processing, there were 72 996 one-second-length segments left in total. In the personal model, 6 folders of the 12 ones were analysed, because each of the left out folders only contained data from several activities out of these 6 chosen activities. There were 37 068 segments left after data pre-processing. The logic obtained an overall 86.00% classification accuracy for all activities with the impersonal model, and 87.82% with the personal model. When taking 10 Hz as the uniform sampling frequency for all activities, the corresponding classification accuracies were 86.05% and 88.00%. This resulted in little loss in both cases.

The Table 11 and Table 12 demonstrate the confusion matrices obtained by the adaptive sampling rate logic in the impersonal model and personal model. In Table 11, most of misclassifications of activity biking are sitting and standing, and Nordic walking and walking. This is because the activity biking can be roughly viewed as the combined activity of sitting and footsteps. The activity lying is mainly misclassified as sitting and standing. It may results from the case that both the activities lying and sitting are static and do not involve any footsteps and arm-movements. For rowing, it involves sitting and movements. Those are also two main sources of misclassifications of rowing. The activity running is almost perfectly classified. For sitting and standing, one more source of misclassifications is that some data from activity working is marked down as standing, and vice versa. In Table 12, the misclassifications show the same trend as the one demonstrated by Table 11.

Table 11. *Confusion matrix obtained by the adaptive sampling rate logic in the impersonal model.*

		Predicted activities						Total
		bike	**lie**	**row**	**run**	**sit/stand**	**Nordic/walk**	
Annotated activities	**bike**	3186	1	55	0	558	250	4050
	lie	1	1339	22	0	134	9	1505
	row	16	1	1877	0	233	231	2358
	run	5	0	2	2427	16	56	2506
	sit/stand	716	125	491	4	33179	790	35305
	Nordic/walk	858	1	206	51	5384	20772	27272
Total		4782	1467	2653	2482	39504	22108	72996

Table 12. *Confusion matrix obtained by the adaptive sampling rate logic in the personal model.*

		Predicted activities						Total
		bike	**lie**	**row**	**run**	**sit/stand**	**Nordic/walk**	
Annotated activities	**bike**	1669	0	18	0	74	34	1795
	lie	0	787	5	0	15	3	810
	row	19	0	1285	0	63	7	1374
	run	2	0	2	1996	17	60	2077
	sit/stand	302	87	149	3	16740	392	17673
	Nordic/walk	192	0	62	39	2973	10073	13339
Total		2184	874	1521	2038	19882	10569	37068

Table 13 shows the classification accuracies obtained by the adaptive sampling rate logic in the impersonal and personal model. The classification accuracies achieved with the impersonal model, are used to make a comparison with the ones demonstrated in Table 7. Based on the comparison, it indicates that the two sets of results are almost the same. So the adaptive sampling rate logic works well with the impersonal model. For the personal model, activities biking, lying, rowing, running, and sitting and standing are almost perfectly classified, except Nordic walking and walking with decent classification accuracy. In Table 13, the absolute differences and relative differences give the information about how better the personal model performs over the impersonal model. The plus symbol indicates for the corresponding activity the personal model performs better than the impersonal model, and the minus symbol reflects the opposite case. The personal model achieves much better classification accuracies than the impersonal model for lying, rowing, and biking. But these two models obtain similar classification results for running, the combined activity sitting and standing, running and the jointed activity Nordic walking and walking. The overall classification accuracies for the impersonal and personal models mentioned above, are actually weighted means of all activities' classification accuracies, by the amounts of each activity data. Their values are 86.00% and 87.82% separately, which shows small difference around 2% absolute and relative differences. But when leaving out the effect of different amounts of each activity data, the equal weighted overall classification accuracies for the two users now become 85.71% and 91.68%, which yield about 6% absolute difference and about 7% relative difference.

Table 13. *Classification results obtained by the adaptive sampling rate logic in impersonal and personal model respectively.*

	Classification accuracies (%)		**Absolute differences (%)**	**Relative differences (%)**
	Impersonal model	**Personal model**		
Lying	88.97	97.16	+8.19	+9.21
Rowing	79.60	93.52	+13.92	+17.49
Biking	78.67	92.98	+14.31	+18.19
Sitting and standing	93.98	94.72	+0.74	+0.79
Running	96.85	96.19	-0.66	-0.68
Nordic walking and walking	76.17	75.52	-0.65	-0.85
Unequally Weighted overall	86.00	87.82	+1.82	+2.11

Equally weighted overall	85.71	91.68	+5.97	+6.97

To measure the efficiency of the implemented adaptive sampling rate logic, four percentages were calculated as the indirect indicators (as listed in Table 14): P1 for the proportion of the number of data sampled at 1 Hz to the total number of data annotated as lying; P2 for the proportion of the number of data sampled at 1 Hz to the total number of data classified as lying; P3 the proportion of the number of data (classified as lying) sampled at 1 Hz to the total number of data annotated as lying; P4 for the proportion of the number of data (classified as lying) sampled at 1 Hz to the total number of data annotated as lying. Table 13 and Table 14 showed that the logic worked very well in this task.

Table 14. *Percentages of data sampled at 1 Hz in impersonal and personal model respectively.*

	Percentages of data sampled at 1 Hz (%)	
	Impersonal model	**Personal model**
P1	92.16	103.58
P2	94.55	96.00
P3	91.10	102.59
P4	93.46	95.08

6. CONCLUSIONS

The thesis work aimed to find suitable sampling frequencies to reduce the energy consumption in the sampling process and later data analysis process. Additionally, simple methods were adopted in data pre-processing and classification to reduce the computational workload.

This study was based on the one (Pärkkä 2006), the first experimental study performed on the Palantir2003 data set. So it was advisable to evaluate the methods used in data pre-processing and the selected features before analyzing the effect of different sampling frequencies. The comparison indicated that the data pre-processing methods and the selected features were suitable, because similar results were obtained as J. Pärkkä et al. did in (Pärkkä 2006) with the CDT when the sampling frequency was 200 Hz. After this step, the pre-processed signals were down-sampled to get the data sets for 20, 10, 8, 5 and 4 Hz sampling frequencies respectively. These data sets were analyzed to study the effect of sampling frequencies on classification accuracies for each activity in different training models.

Figure 8-13 demonstrate the classification results over each training model on an activity-by-activity basis. These figures indicate that the personal model is preferred in this study. To use the personal model, users are required to provide their own activity data for training the classifier. In practice, user-annotated data collection tends to cause more annotation errors compared with the one performed by a separate annotator or supervisor. One obvious reason is that more activity transition data will be wrongly annotated. This may vary a lot among different users. Thus it may be hard to decide how much transition data to be discarded during the continuous data collection for all the targeted activities. One possible method to solve this problem is discretization by recording each activity for a fixed time-period instead of recording all activities continuously. In this case, users can select which activity to perform and the selection is taken as the annotation. After the data collection of each activity, only the middle part of its recorded data is kept to reduce the effect of activity transition. If the CDT is employed as the classification algorithm, one more problem may occur when the user does not perform all the targeted activities, because the structure of the CDT is fixed. Possible solutions to this problem involve adopting the automatic decision tree as the classification algorithm or providing several options of CDTs.

As discussed in the section 5.2, it is sufficient to adopt 10 Hz as the sampling frequency for all activities. But 10 Hz is not the optimal choice for some activity, e.g., biking. There are two strategies for choosing the sampling frequency. One is adaptively using the optimal sampling frequency for each activity. If there are many sampling frequency candidates, the smallest one is used. The other is adopting a uniform sampling frequency for all activities with some loss of accuracies. The adaptive sampling rate logic implemented in this thesis work can be viewed as an application of the first strategy but with a little modification by using a lower sampling frequency only for activity lying. Or, it can be treated as an application of the combination of the two strategies, with the adaptive sampling implemented only for lying activity and non-lying activity. The results obtained by the sampling logic were demonstrated and described in the section 5.3, which indicated the method of using a low sampling frequency for activity lying and a higher sampling frequency for other activities, performed very well in this study.

For the future work, it would be interesting to implement the adaptive sampling for all activities and to compare the energy consumptions in that situation.

REFERENCES

Aminian, K., & Najafi, B. (2004). Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications. Computer Animation and Virtual Worlds, 15(2), 79-94.

Analog Devices. Low-cost, low-power, complete dual-axis iMEMS® accelerometer. Available at: http://www.analog.com/media/en/technical-documentation/obsolete-data-sheets/ADXL202_210.pdf

Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2012). Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In Ambient assisted living and home care (pp. 216-223). Springer Berlin Heidelberg.

Avci, A., Bosch, S., Marin-Perianu, M., Marin-Perianu, R., & Havinga, P. (2010, February). Activity recognition using inertial sensing for healthcare, welling and sports applications: A survey. In Architecture of computing systems (ARCS), 2010 23rd international conference on (pp. 1-10). VDE.

Bao, L., & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. In Pervasive computing (pp. 1-17). Springer Berlin Heidelberg.

Brand, M., Oliver, N., & Pentland, A. (1997, June). Coupled hidden Markov models for complex action recognition. In Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on (pp. 994-999). IEEE.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public health reports, 100(2), 126.

Culhane, K. M., O'Connor, M., Lyons, D., & Lyons, G. M. (2005). Accelerometers in rehabilitation medicine for older adults. Age and ageing, 34(6), 556-560.

De Vries, S. I., Garre, F. G., Engbers, L. H., Hildebrandt, V. H., & Van Buuren, S. (2011). Evaluation of neural networks to identify types of activity using accelerometers. Med Sci Sports Exerc, 43(1), 101-7.

Duong, T. V., Bui, H. H., Phung, D. Q., & Venkatesh, S. (2005, June). Activity recognition and abnormality detection with the switching hidden semi-markov model. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 838-845). IEEE.

Ermes, M., Pärkkä, J., & Cluitmans, L. (2008a, August). Advancing from offline to online activity recognition with wearable sensors. In Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE (pp. 4451-4454). IEEE.

Ermes, M., Pärkkä, J., Mantyjavi, J., Korhonen, I. (2008b). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. Information Technology in Biomedicine, IEEE Transactions On, 12(1), 20-26.

Foerster, F., Smeja, M., & Fahrenberg, J. (1999). Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring. Computers in Human Behavior, 15(5), 571-583.

Godfrey, A., Bourke, A. K., Olaighin, G. M., Van De Ven, P., & Nelson, J. (2011). Activity classification using a single chest mounted tri-axial accelerometer. Medical engineering & physics, 33(9), 1127-1135.

Godfrey, A., Conway, R., Meagher, D., & OLaighin, G. (2008). Direct measurement of human movement by accelerometry. Medical engineering & physics, 30(10), 1364-1386.

He, J., Li, H., & Tan, J. (2007, August). Real-time daily activity classification with wireless sensor networks using hidden Markov model. In Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International conference of the IEEE (pp. 3192-3195). IEEE.

He, Z., & Jin, L. (2009, October). Activity recognition from acceleration data based on discrete consine transform and SVM. In Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on (pp. 5041-5044). IEEE.

Junker, H., Lukowicz, P., & Troster, G. (2004, October). Sampling frequency, signal resolution and the accuracy of wearable context recognition systems. In Wearable Computers, 2004. ISWC 2004. Eighth International Symposium on (Vol. 1, pp. 176-177). IEEE.

Kemp, B., Värri, A., Rosa, A. C., Nielsen, K. D., & Gade, J. (1992). A simple format for exchange of digitized polygraphic recordings. Electroencephalography and clinical Neurophysiology, 82(5), 391-393.

Kim, E., Helal, S., & Cook, D. (2010). Human activity recognition and pattern discovery. Pervasive Computing, IEEE, 9(1), 48-53.

Krause, A., Ihmig, M., Ranking, E., Leong, D., Gupta, S., Siewiorek, D., … &Sengupta, U. (2005, October). Trading off prediction accuracy and power consumption for context-aware wearable computing. In Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on (pp. 20-26). IEEE.

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. ACM SigKDD Explorations Newsletter, 12(2), 74-82.

Lee, M. W., Khan, A. M., & Kim, T. S. (2011). A single tri-axial accelerometer-based real-time personal life log system capable of human activity recognition and exercise information generation. Personal and Ubiquitous Computing, 15(8), 887-898.

Long, X., Yin, B., & Aarts, R. M. (2009, September). Single-accelerometer-based daily physical activity classification. In Engineering in Medicine and Biology Society, 2009. EMBS 2009. Annual International Conference of the IEEE (pp. 6107-6110). IEEE.

Luhr, S., Bui, H. H., Venkatesh, S., & West, G. A. (2003, March). Recognition of human activity through hierarchical stochastic learning. In null (p. 416). IEEE.

Mannini, A., Intille, S. S., Rosenberger, M., Sabatini, A. M., & Haskell, W. (2013). Activity recognition using a single accelerometer placed at the wrist or ankle. Medicine and science in sports and exercise, 45(11), 2193-2203.

Mathie, M. J., Celler, B. G., Lovell, N. H., & Coster, A. C. F. (2004a). Classification of basic daily movements using a triaxial accelerometer. Medical and Biological Engineering and Computing, 42(5), 679-687.

Mathie, M. J., Coster, A. C., Lovell, N. H., & Celler, B. G. (2004b). Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. Physiological measurement, 25(2), R1.

Maurer, U., Smailagic, A., Siewiorek, D. P., & Deisher, M. (2006, April). Activity recognition and monitoring using multiple sensors on different body positions. In Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on. IEEE.

Mitchell, T. M. (1997). Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 45.

Najafi, B., Aminian, K., Paraschiv-lonescu, A., Loew, F., Bula, C. J., & Robert, P. (2003). Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. Biomedical Engineering, IEEE Transaction on, 50(6), 711-723.

Olguin, D. O., & Pentland, A. S. (2006, October). Human activity recognition: Accuracy across common locations for wearable sensors. In Proceedings of 2006 10th IEEE International Symposium on Wearable Computers, Montreux, Switzerland (pp. 11-14).

Pärkkä, J., Cluitmans, L., & Korpipää, P. (2004). Palantir context data collection, annotation and PSV file format. In Proceedings of the Pervasive, Workshops (pp. 9-16).

Pärkkä, J., Ermes, M., Antila, K., van Gils, M., Manttari, A., & Nieminen, H. (2007, August). Estimating intensity of physical activity: a comparison of wearable accelerometer and gyro sensors and 3 sensor locations. In Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE (pp. 1511-1514). IEEE.

Pärkkä, J., Ermes, M., Korpipää, P., Mäntyjärvi, J., Peltola, J., & Korhonen, I. (2006). Activity classification using realistic data from wearable sensors. Information Technology in Biomedicine, IEEE Transactions on, 10(1), 199-128.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., … & Wilmore, J. H. (1995). Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. Jama, 273(5), 402-407.

Ravi, N., Dandekar, N., Mysore, P., & Littman, M. L. (2005, July). Activity recognition from accelerometer data. In AAAI (Vol. 5, pp. 1541-1546).

Reiss, A., & Stricker, D. (2011, May). Towards global aerobic activity monitoring. In Proceedings of the 4th International Conference on Pervasive Technologies Related to Assistive Environments (p. 12). ACM.

Reiss, A., & Stricker, D. (2014). Aerobic activity monitoring: towards a long-term approach. Universal Access in the Information Society, 13(1), 101-114.

Siirtola, P., Laurinen, P., Haapalainen, E., Röning, J., & Kinnunen, H. (2009, March). Clustering-based activity classification with a wrist-worn accelerometer using basic features. In Computational Intelligence and Data Mining, 2009. CIDM'09. IEEE Symposium on (pp. 95-100). IEEE.

Staudenmayer, J., Pober, D., Crouter, S., Bassett, D., & Freedson, P. (2009). An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. Journal of Applied Physiology, 107(4), 1300-1307.

Sun, L., Zhang, D., Li, B., Guo, B., & Li, S. (2010). Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. In Ubiquitous intelligence and computing (pp. 548-562). Springer Berlin Heidelberg.

Tapia, E. M., Intille, S. S., Haskell, W., Larson, K., Wright, J., King, A., & Friedman, R. (2007, October). Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In Wearable Computers, 2007 11th IEEE International Symposium on (pp. 37-40). IEEE.

Veltink, P. H., Bussmann, H. B., De Vries, W., Martens, W. L., & Van Lummel, R. C. (1996). Detection of static and dynamic activities using uniaxial accelerometers. Rehabilitation Engineering, IEEE Transaction on, 4(4), 375-385.

Wang, Y., Lin, J., Annavaram, M., Jacobson, Q. A., Hong, J., Krishnamachari, B., & Sadeh, N. (2009, June). A framework of energy efficient mobile sensing for automatic user state recognition. In Proceedings of the 7th International conference on Mobile systems, applications, and service (pp. 179-192). ACM.

Webb, A. R. (2011). Statistical pattern recognition. John Wiley & Sons.

Weiss, G. M., & Lockhart, J. W. (2012, July). The impact of personalization on smartphone-based activity recognition. In AAAI Workshop on Activity Context Representation: Techniques and Languages.

World Health Organization. (2009). Global health risks: mortality and burden of disease attributable to selected major risks. World Health Organization.

World Health Organization. (2015, January). Physical activity. Available at: http://www.who.int/mediacentre/factsheets/fs385/en/ [Accessed 10.07.2015].

Yamato, J., Ohya, J., & Ishii, K. (1992, June). Recognition human action in time-sequential images using hidden markov model. In Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on (pp. 397-385). IEEE.

Yan, Z., Subbaraju, V., Chakraborty, D., Misra, A., & Aberer, K. (2012, June). Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach. In Wearable Computers (ISWC), 2012 16th International Symposium on (pp. 17-24). IEEE.

Yang, C. C., & Hsu, Y. L. (2010). A review of accelerometry-based wearable motion detectors for physical activity monitoring. Sensors, 10(8), 7772-7788.

Yang, J. (2009, October). Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. In Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics (pp. 1-10). ACM.

Yang, J. Y., Wang, J. S., & Chen, Y. P. (2008). Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers. Pattern recognition letters 29. 16 (2008): 2213-2220.

Yurur, O., Liu, C. H., Liu, X., & Moreno, W. (2013, October). Adaptive Sampling and Duty Cycling for Smartphone Accelerometer. In Mobile Ad-Hoc and Sensor Systems (MASS), 2013 IEEE 10th International Conference on (pp. 511-518). IEEE.

Zhang, S., Murray, P., Zillmer, R., Eston, R. G., Catt, M., & Rowlands, A. V. (2012). Activity classification using the GENEA: optimum sampling frequency and number of axes. Medicine and science in sports and exercise, 44(11), 2228-2234.