A Hybrid Approach for Recognizing ADLs and Care Activities Using Inertial Sensors and RFID

Albert Hein and Thomas Kirste

University of Rostock, Institute of Computer Science, Rostock, Germany
{albert.hein,thomas.kirste}@uni-rostock.de

Abstract. In this paper we present a feasibility study regarding the recognition of high level daily living and care activities. We examine a hybrid discriminative and model based generative approach based on RFID and inertial sensor data. We show that the presented sensor configuration is able to deliver sensor readings and object sightings at a sufficient rate without forcing user compliance. We further evaluated the advantage of a model based approach over a static classifier, compared the individual contribution of each sensor type and could reach accuracy rates of 97% and 85%.

1 Introduction

The elderly currently are the fastest growing demographic group in the USA [1], a development which is similar to other for example european regions due to increased life expectancy and decrease of the birth rate. So the aging human is more and more getting into the focus of not only demographic research. Graceful aging is a common demand of the current and future older population.

The main motivation behind ambient assisted living is that 95% of the people wish to stay at home as long as possible. But the home care for elderly or sick people is a serious burden for family members, 30% of admissions to nursing homes are done not because of deterioration in the senior's conduit but because of so called "caregiver burnout". An alternative to stationary treatment is the professional ambulant elderly care at home. Especially this kind of service needs accurate documentation of care activities to allow correct accounting for the health insurances. The usual documentation process is to this day still done manually and takes up to 40% of the working time, is error-prone and mostly inaccurate because it does not happen in situ but afterwards.

Sensor based activity recognition is currently widely seen as an elemental technology for providing mobile assistance in AAL scenarios [2,3,4,5,6]. An ambient sensor infrastructure as in the iDorm [7] or the PlaceLab [8] is mostly unobstrusive, but (at least currently) too expensive and complex for home care settings where each apartment of every care patient would have to be equipped with environmental sensors. Wearable sensors as an alternative are already showing good results in human activity recognition, either through direct motion sampling [9,5] or through capturing object interactions [2]. In reality there are often ambiguities between different activities sharing the same motions or gestures like carrying a glass of water or carrying a pillbox and also ambiguities

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2009, LNCS 5615, pp. 178–188, 2009.

between different activities using the same objects, which is e.g. using an object or simply carrying it around. These two problems are not clearly distinguishable using only one of these methods.

This paper combines those former two approaches of direct motion measurement with inertial sensors and detection of object interaction with RFID for high level activity recognition in order to conduct a feasibility study. Our system uses hierarchical sensor fusion on different levels of abstraction for simultaneously integrating many channels of heterogenous sensor data. For inferencing high level activities we use a layered hybrid discriminative and model based generative approach. This will enable us to integrate prior knowledge into the decision process in the future to reduce the amount of training while keeping the probabilistic model simple. This approach was evaluated on two experimental settings with one Activity of Daily Living (ADL) breakfast scenario and one home care scenario where the combined approach reached accuracy rates of 97% and 85% respectively.

Our main objectives were to investigate whether such a combined sensor configuration is technically viable in terms of delivering reliable data without explicit compliance of the test subject, if a static classifier or a hybrid model is able to infer useable estimates on a continuous time trace and how much each sensor type actually contributes to the final classification results.

2 Problem Domain

There are several often cited publications reporting good results in low level activity recognition like [9,5] while high level activities are commonly seen as problematic and this kind of research is still at its beginnings [4]. These high level activities are mostly difficult to discriminate because of several general challenges :

- Interleaved or interrupted activities which are not executed sequentially
- Ambiguities between different activities sharing the same motions or gestures
- Ambiguities between different activities sharing common object handlings
- Variations in the activity performance between single or multiple subjects and distortions by uninvolved persons
- Different levels of complexity between elementary or compound activities
- Different levels of granularity between coarse motion and fine grained gestures
- Lack of representative training data containing examples even of special cases
- Highest possible unobtrusiveness by not requiring the user to explicitly handle or interact with the system

For this task usually different approaches are adopted ranging from simple model free pattern recognition methods to complex model based probabilistic approaches which are able to capture also causal and temporal dependencies (for a more general overview compare [10]). Depending on the chosen approach, the specific application and sensor data, furthermore methods for including prior knowledge and common sense (compare [11]) are needed for handling varying activity sequences without needing an exponentially growing amount of training data.

Based on the different fields of application, three different complex application domains can be pointed out, which can but don't need to build on each other:

1. *Activity Profiling*
   The activity profiling task usually tries to answer the question how often and how long which activity is performed over a given period. So Activity Profiling is an acquisition of activity data in a lexical manner which allows simple qualitative and quantitative analyses of activities on an absolute scale. Any causal or temporal dependencies remain out of consideration. This application domain is already well explored for base level activities usually using statistical classifiers (e.g. [9,12]).
2. *Behaviour Assessment*
   Applications out of this field ask the question whether there are long-term variations in the daily routine or any temporal abnormal behavior. This can be seen as a syntactical analysis of changes in specific behaviour patterns and allows reasoning about mental and physical decay or hazardous situations. All those measurements are done on a relative scale, where temporal but no causal dependencies have to be considered during the recognition process. This task is pretty well explored for simple quantitative measurings of low level activities. On a qualitative level for activities of higher abstraction still only a few mainly connectionist [7] or model based approaches can be found.
3. *Proactive Assistance*
   For proactive assistance systems it is important to not only infer the plain context but also the current intention of the user. In analogy to the other two application domains this implies a semantic analysis of the activity trajectory which allows autonomous assistance and the prediction of future events. Such a system must consider both, causal and temporal dependencies of activities. Today in this field there are basically model based intention recognition approaches built upon environmental infrastructural sensors and location data [13].

In this work we are mainly concerned with the first application domain of activity profiling. In addition to prior work we focus on recognizing *high level* activities by additionally including temporal dependencies.

3 Activity Recognition System Overview

Due to the very heterogeneous nature of the sensor data the underlying activity recognition system must be able to cope simultaneously with its very different characteristics to avoid any unintentional loss of any potentially important information during the recognition process. Therefor it must be able to handle nominal and numerical data with different time bases (discrete and event driven) of a varying number of sensors. Because of the good experience with statistical classifiers for activity recognition and their good performance but the need of temporal modelling and a comprehensable decision making process we have chosen a hybrid and layered approach (Fig. 1). While each processing layer makes use of sensor data on different levels of abstraction the sensor fusion process is separated hierarchically into several steps.

The individual raw sensor data channels are being synchronized and then processed in the feature extraction module. This module calculates 562 different features from half overlapping windows of 1.28s consisting of frequency domain, statistical, curve,

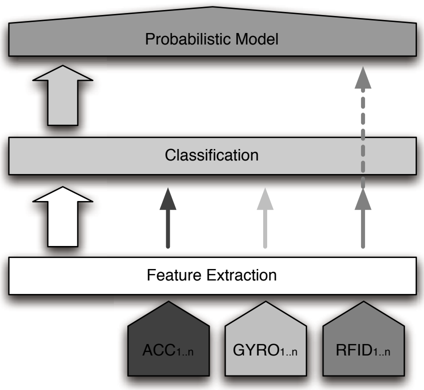

Fig. 1. Inference Process and Sensor Data Fusion

physical, correlation and step detection features (for a detailed explanation see [12]) and handled object and object class information. In this step for many features already several sensors and sensor axes are combined to more abstract representations (e.g. physical features or step detection, the broad white arrow) while in others single sensor channels are still contained seperately (the thin arrows).

All generated features are processed by an embedded Weka C4.5 decision tree [14] which has often shown superior accuracy over comparable algorithms in context of human activity recognition in the past [9,6,12] and is also capable of handling both nominal and numerical data. In this step all sensor channels are merged and mapped onto the different class probabilities for each data window (broad grey arrow). The RFID features are mirrored and passed through into the probabilistic model (dotted line).

For the model layer we have chosen a Hidden Markov Model (HMM) with one hidden state for each activity class, which is also a common choice for sensor/RFID based activity recognition [2,5,11]. They are robust to sensor noise, able to reflect complex temporal properties, are human readable and adjustable and have well known algorithms for inference, which also allow prediction. We use the HMM in conjunction with a particle filter. An example HMM for the simplest experimental setting is shown in Fig. 2.

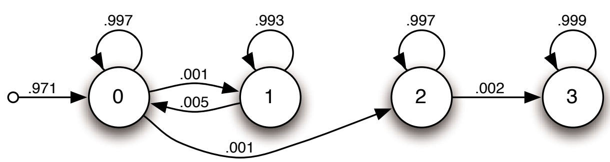

Fig. 2. Simplified HMM state diagram for a model automatically generated by the recognition system (Setting 1, high level context). The four states represent breakfast (0), make coffee (1), talk (2) and clear table (3). Only transition probabilities greater than 0.001 are shown as arrows.

The emission probabilities $P(O_t|S_t)$ for the observations O_t and the state S_t for each timestep t are calculated as the product of the probability of the sensor sightings given the current state $P(acc_t|S_t)$ (maximum likelihood estimation calculated from the Weka class probabilities at the current timestep) and the probabilities of the RFID object feature sightings (objects and object classes) for the current state $P(rfid_i|S_t)$ (Eq. 1).

$$P(O_t|S_t) = P(acc_t|S_t) \prod_{i=0}^{N-1} P(rfid_i|S_t) \tag{1}$$

A priori and transition probabilities of the model are entirely calculated from experimental data ground truth. This way no additional parameter learning with the expectation maximization algorithm is needed, which can be very time consuming. To avoid overfitting, biases for the model accuracy, the RFID reliability and the weka class probabilities (trust in sensor data) were set, which also allows the inclusion of expert's domain knowledge. While the first two were adjusted manually, the latter was determined by model averaging for the weka model weighted with its cross validated classification accuracy and a set of rudimental models always predicting one class and their particular accuracy [15].

4 Sensors

We used two types of wireless sensors for our recording which usually are treated separately in literature: Inertial Motion Sensors (IMU = Inertial Measurement Unit) and RFID. There have been two approaches in the recent past combining accelerometers for the direct measurement of human bodies' motions with a wearable RFID reader for identifying handled objects – the work of Stikic et al. [16] who attached the Intel iBracelet and a custom-made sensor board to the dominant wrist, and Wang et al. [11] also using the iBracelet and a comparable mobile sensor platform and a similar setup.

For our initial tests we were recording raw data with a probably higher number of sensors than actually required for the final appliance, so that it is possible to evaluate single sensor channels or combinations of subsets afterwards on the original sensor data later.

For the inertial motion recordings we used three SparkFun IMU 6-DOF v3 sensor boards. These are equipped with a 3-axis Freescale MMA7260Q accelerometer, 3-axis InvenSende IDG300 gyroscopes and a 2-axis Honeywell HMC1043 magnetometer. The LPC2138 ARM7 microcontroller is also capable of preprocessing the raw data onboard. We used the IMU to sample relative motion and rotation at a rate of 50Hz at a range of 6g to fully capture normal human motion as described in [17]. The raw data was instantaneously transmitted via a class 1 bluetooth link with a max. operating distance of 30 to 100m. Because of the compact size (51x41x23mm) the board could be attached at unobtrusive positioned at the dominant wrist for recording gestures and object motion without the need of attaching a sensor directly to the handled objects, at the chest/upper back and at the hip. These sensor positions have been shown to operate well in the literature and in own prior work.

RFID is a popular technology for contactless identification of objects. A basic system consists of a reader module with an antenna and several active or passive tags in the

Fig. 3. RFID Wrist Antenna and Inertial Sensor Board

form of small boxes, stickers or even implants. Especially passive RFID stickers are a cheap, battery free solution for reliable object detection. RFID practically doesn't return any false positive object sightings. Because currently no wearable RFID modules are available commercially, we used a Texas Instruments S4100 multi function reader with a custom-made wrist antenna (see Fig. 3). Depending on object geometry and material it has a reading range between 10 and 30cm. For tagging the objects we used ISO 15693 Standard HF stickers in the two sizes 43x43mm and 18x36mm. The reader module was plugged to an external class 1 serial/bluetooth adaptor.

All data streams were wirelessly transmitted to a laptop computer where they were immediately formatted, synchronized and saved to disk.

5 Experimental Setup

The experiments were conducted to evaluate the feasibility of our combined approach for recognizing sufficiently realistic daily living (ADL) and health care activities. The chosen repertory only consists of compound activities of a high level of abstraction by trying to consider all of the challenges specified in Section 2 like ambiguities and interleaved activities.

The general experimental setup is related to the household ADLs of Stikic et al., also with a comparable sensor equipment. Attempts like the PlaceLab [8] or iDorm [7] represent basically opposing approaches, because no environmental sensor infrastructure was used in the experiment. As we were interested in a general proof of concept, we were not doing tests out of the lab with authentic subjects inside a nursing home at this early stage. It is best practice to carry out initial experiments in a controllable environment under optimal observability. The test runs have additionally been accompanied by video and audio surveillance to facilitate later manual annotation.

To avoid biasing, the test subjects were not involved in planning and setting up the experiment or analyzing the data afterwards in any way. Each subject was instructed to behave as natural as possible, to try to ignore the attached sensores and especially not to pay attention to the rfid labels on the objects. This strongly distinguished these experiments from others found in prior publications where the subjects were explicitly instructed to wait until the rfid reader has scanned the current object [2], which significantly increases the number of object sightings but is in conflict with our requirement not to assume specific user interaction. Therefor here also the tag placement was done by a person not involved any further with the experiment which, as a matter of fact, resulted in a high number of tag placements due to the lack of knowledge of the most relevant objects.

We utilized two main experimental settings as follows:

5.1 Setting 1: Breakfast

The breakfast ADL setting follows the example of Patterson et al. [2], who used RFID gloves for detecting routine morning activities. In our experiments also one sensor equipped subject and four other participants were involved. No particular actions or action sequences were initially scheduled to be performed during the test. The test utilized real equipment and real food and was conducted distributed over two rooms – an office and a kitchen. 118 RFID tags were placed on 55 different objects at several different positions. The course was observed by a fisheye camera, which was able to oversee the whole setting.

Activities of two levels of granularity were recorded: the coarse high level abstract context with ”breakfast”, ”make coffee”, ”talk” and ”clear table” and more detailed actions consisting of 27 activity classes like ”make bread”, ”drink”, ”hand over”, ”carry”, ”stir” or ”collect dishes” which come closer to the work of [2]. Altogether 1h and 8min (590mb) of raw data have been sampled.

5.2 Setting 2: Home Care

The home care setting is part of ”MArika”, a subproject of the current state research project ”Mobile Assistance” [18]. For this setting we roughly rebuilt the floor plan of an apartment consisting of a bedroom, a bathroom, a living room and a kichenette in our SmartLab. The test runs were performed by professional care personnel (a geriatric nurse, a student helped out as a patient). We put 43 tags on 33 objects again partially at different positions for increasing the probability of detection. The scenery was observed by a fisheye and a ceiling mounted dome camera. This time a general preselection of care activities was given, as this is common for a care plan. The test agenda and the scenario has been developed in close cooperation with a nursing service, which also provided authentic equipment for the tests. We have sampled two runs of an authentic sequence of morning care activities taken from a real person. The activities were directly taken from the service accounting catalogue of the health insurances: ”general service” (greeting, fetching newspaper, ...), ”big morning toilet” (including washing whole body, brushing teeth), ”micturition and defecation”, ”administration of medications”, ”injections”, ”bandaging”, ”preparation of food” and ”documentation”. We collected 14min (317mb) and 12 minutes (289mb) of raw data.

6 Results

Our first objective was to determine the distribution of RFID object sightings with the given sensor configuration and without explicit user compliance. In the first setting 316 Tags were read during the experiment at an average rate of $m = 12,9$s and a standard

Table 1. Comparision of the overall accuracies for the different experimental settings

		RFID	*IMU*	*both*
Breakfast coarse	C4.5	50.2%	87.3%	86.9%
	HMM	67.4%	97.7%	97.8%
Breakfast fine	C4.5	62.5%	65.5%	67.4%
	HMM	65.8%	73.3%	76.7%
Care	C4.5	51.4%	52.2%	57.1%
	HMM	84.9%	80.3%	85.1%

deviation of $\sigma = 33,1$s. The second setting 60 object sightings were detected at an average rate of $m = 21,5$s and a standard deviation of $\sigma = 43,7$s. While there were several bursts of frequent sensor readings there were also gaps up to several minutes. The three IMUs delivered an overall good performance without any data loss.

The static classification accuracy was evaluated using a 10-fold stratified cross validation. The output class probabilities were used as part of the observations for the HMM as described in Sec. 3. The decoding of the HMM was done by sequential monte carlo filtering using a particle filter with 100000 particles. The overall accuracies for the different experimental settings are itemized in Table 1 while a continuous time trace is shown in Fig. 4 for the breakfast example.

The decision tree already showed surprisingly good classification results on inertial sensor data, especially for the coarse high level breakfast activities, while, as expected, it was completely unusable on RFID data only (the apparently high recognition rate is misleading, it only predicted the activity class with the highest prior probability). The combination of both sensor types did not bring any significant advantages. In all cases the static classifier produced many temporal glitches.

In general the results of the HMM were temporally much more smoothed, but also prone to a short delay on transitions between activity states which, in the given application domain of activity profiling, is not a drawback at all. Altough the RFID object sightings were very irregular, the model did relatively well on inferring the high level activities. Using only IMU data, the HMMs behaved primarily as a temporal smoother for the static classification probabilities, which resulted in a nearly perfect recognition rate (breakfast coarse) or at least significant increase (care setting). Due to its ”lethargic” behaviour, the model had problems estimating the short term activities in the second breakfast setting.

The combination of both sensor types inside the HMM generally resulted in the highest recognition accuracy, but in detail the combined result is not significantly higher than the best single sensor type performance.

General problems came up while disambiguating activities with both, shared motions and objects. In the care setting e.g. ”micturition and defecation” was mistaken for ”morning toilet”. No additional representative object sightings were detected during the test runs, although characteristic tags were present (”toilet chair”, ”toilet brush”, etc.). By contrast the coarse abstract activities in the first breakfast setting implied very distinct motion patterns, which lead to a nearly perfect classification performance.

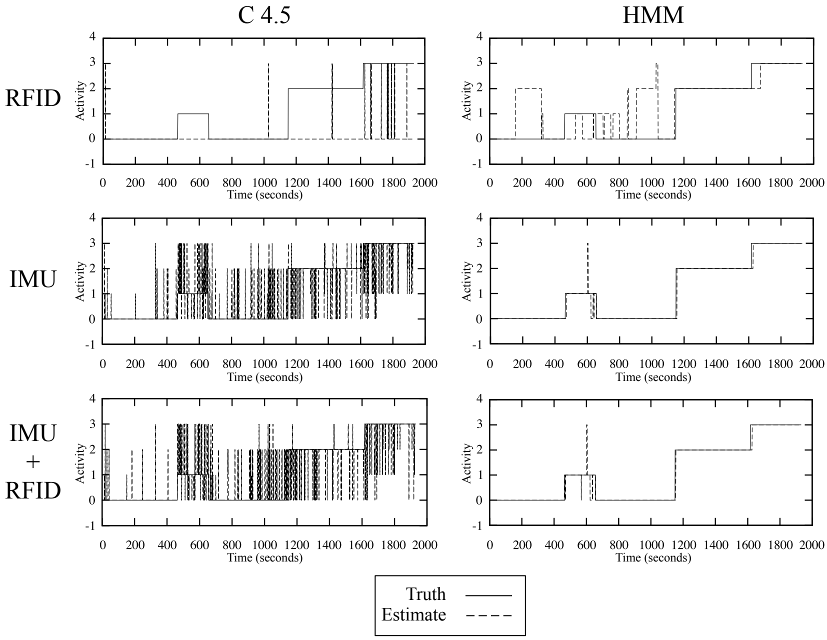

Fig. 4. Accuracy of Inference for Setting 1, high level context. The four activities represent breakfast (0), make coffee (1), talk (2) and clear table (3). Results of RFID only, IMU only and the combination of both are compared against each other for the C4.5 decision tree and the HMM.

7 Conclusions and Future Work

The experiments were conducted as a feasibility study. Regarding our first main objective it could be shown that the presented combined sensor configuration can deliver sensor readings and object sightings at a sufficient rate without requiring explicit user compliance or interaction. Probably many of the sporadic gaps in the RFID data and missed objects can be avoided by not only instrumenting the dominant wrist, as many items were utilized by the other hand, too. For future experiments a second antenna / reader module has already been built. Apart from that different subsets of sensors will be evaluated in order to increase the wearing comfort and the unobtrusiveness.

Our second finding is, that the hybrid discriminative and model based approach is basically able to infer high level daily living or care activities. In addition, the model layer in general significantly outperforms simple static classification using the C4.5 decision tree. Our approach was able to handle the main general challenges regarding the classification of abstract high level activities. A solution for ambiguous classes sharing motion patterns *and* objects could be breaking down compound activities into smaller atomic actions, which can then be used as building blocks in a multilevel model.

The third objective of this work was to find out how much each sensor type can contribute to the recognition process. This question can not clearly be answered, although it seems that in case of reliable IMU based recognition the additional knowledge of object handling does not bring noticable advantages. As this is a primary problem of weighting and biasing in respect of sensor fusion, further investigation is needed.

As the presented classification results are based on very few experimental training data, they have to be treated carefully in respect of generalization. Future experiments will have to follow under real world conditions to allow a reliable evaluation of a comprehensive set of care activities and multiple test subjects. This is expected to provide more realistic results allowing an outlook on every day use.

As generally very few training data is available for a high number of complex activities, additional modelling and the inclusion of expert's domain knowledge is inevitable. So a more direct way to involve RFID events, model the average length of activities or integrate uncertainty is desireable. Probably other models than HMMs will allow more control. An automatic base model generation from simple task models, as known from software engineering, is currently under development.

Acknowledgements

Marika [18] is funded by the state of Mecklenburg-Vorpommern, Germany within the scope of the LFS-MA project. The care experiments were supported by Informatik-Forum Rostock e.V..

References

1. Roush, R.E.: Smart home technology for aging in place longer and better. Technical report, Roy M. and Phyllis Gough Huffington Center on Aging (June 2004)
2. Patterson, D.J., Fox, D., Kautz, H., Philipose, M.: Fine-grained activity recognition by aggregating abstract object usage. ISWC, 44–51 (2005)
3. Smith, J.R., Fishkin, K.P., Jiang, B., Mamishev, A., Philipose, M., Rea, A.D., Roy, S., Sundara-Rajan, K.: Rfid-based techniques for human-activity detection. Commun. ACM 48(9), 39–44 (2005)
4. Huỳnh, T., Blanke, U., Schiele, B.: Scalable recognition of daily activities with wearable sensors. In: Hightower, J., Schiele, B., Strang, T. (eds.) LoCA 2007. LNCS, vol. 4718, pp. 50–67. Springer, Heidelberg (2007)
5. Lester, J., Choudhury, T., Kern, N., Borriello, G., Hannaford, B.: A hybrid discriminative/generative approach for modeling human activities. In: Kaelbling, L.P., Saffiotti, A. (eds.) IJCAI, pp. 766–772. Professional Book Center (2005)
6. Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., Korhonen, I.: Activity classification using realistic data from wearable sensors. IEEE Transactions on Information Technology in Biomedicine 10(1), 119–128 (2006)
7. Rivera-illingworth, F., Callaghan, V., Hagras, H.: Detection Of Normal and Novel Behaviour. In: Ubiquitous Domestic Environments. The Computer Journal (2007) bxm078
8. Hightower, J., LaMarca, A., Smith, I.E.: Practical lessons from place lab. IEEE Pervasive Computing 5(3), 32–39 (2006)
9. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
10. Hein, A., Kirste, T.: Activity recognition for ambient assisted living: Potential and challenges. In: Ambient Assisted Living Ambient Assisted Living Ambient Assisted Living, pp. 263–268. VDE Verlag (January 2008)

11. Wang, S., Pentney, W., Popescu, A.M., Choudhury, T., Philipose, M.: Common sense based joint training of human activity recognizers. In: Veloso, M.M. (ed.) IJCAI, pp. 2237–2242 (2007)
12. Hein, A.: Echtzeitfähige merkmalsgewinnung von beschleunigungswerten und klassifikation von zyklischen bewegungen. Master's thesis, University of Rostock (November 2007)
13. Giersich, M., Forbrig, P., Fuchs, G., Kirste, T., Reichart, D., Schumann, H.: Towards an Integrated Approach for Task Modeling and Human Behavior Recognition. In: Jacko, J.A. (ed.) HCI 2007. LNCS, vol. 4550, pp. 1109–1118. Springer, Heidelberg (2007)
14. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques, 2nd edn. Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann Publishers Inc, San Francisco (2005)
15. Hastie, T., Tibshirani, R., Friedman, J.H.: The Elements of Statistical Learning. Springer, Heidelberg (2001)
16. Stikic, M., Huynh, T., Van Laerhoven, K., Schiele, B.: Adl recognition based on the combination of rfid and accelerometer sensing. In: 2nd International Conference on Pervasive Computing Technologies for Healthcare 2008 (2008)
17. Bouten, C., Koekkoek, K., Verduin, M., Kodde, R., Janssen, J.: A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. IEEE Transactions on Biomedical Engineering 44(3), 136–147 (1997)
18. Landesforschungsschwerpunkt (April 2009), http://marika.lfs-ma.de/