Device-Free and Device-Bound Activity Recognition using Radio Signal Strength

Markus Scholz
scholz@teco.edu

Till Riedel
riedel@teco.edu

Mario Hock
hock@teco.edu

Michael Beigl
beigl@teco.edu

TECO, Karlsruhe Institute of Technology
76131 Karlsruhe, Germany

ABSTRACT

Background: We investigate direct use of 802.15.4 radio signal strength indication (RSSI) for human activity recognition when 1) a user carries a wireless node (device-bound) and when 2) a user moves in the wireless sensor net (WSN) without a WSN node (device-free). We investigate recognition feasibility in respect to network topology, subject and room geometry (door open, half, closed).
Methods: In a 2 person office room 8 wireless nodes are installed in a 3D topology. Two subjects are outfitted with a sensor node on the hip. Acceleration and RSSI are recorded while subject performs 6 different activities or room is empty. We apply machine learning for analysis and compare our results to acceleration data.
Results: 10-fold cross-validation with all nodes gives accuracies of 0.896 (device-bound), 0.894 (device-free) and 0.88 (accelerometer). Topology investigation reveals that similar accuracies may be reached with only 5 (device-bound) or 4 (device-free) selected nodes. Applying trained data from one subject to the other and vice-versa shows higher recognition difference on RSSI than on acceleration. Changing of door state has smaller effect on both systems than subject change; with least impact when door is closed.
Conclusion: 802.15.4 RSSI suited for activity recognition. 3D topology is helpful in respect to type of activities. Discrimination of subjects seems possible. Practical systems must adapt no only to long-term environmental dispersion but consider typical geometric changes. Adaptable, robust recognition models must be developed.

1. INTRODUCTION

Human activity recognition is a well researched domain in Pervasive Computing. Classically this domain involves attaching acceleration sensors on the user to deduce his activities. Depending on the type of activities multiple sensors and attachment locations may be necessary.

In this paper we explore how such activities are reflected in the link quality/packet signal strength (RSSI) of a low power wireless sensor network (WSN). We differentiate two cases: 1) Human carries transceiver (device-bound). Here we only consider wireless links between the WSN infrastructure and the body-worn transceiver. 2) Human carries no transceiver (device-free). Here we only consider wireless links in-between the WSN infrastructure.

We believe this to be relevant research to the Augmented Human community as it can:

- Simplify interaction by reducing user instrumentation while re-using existing infrastructure
- Enable recognition of activities/gestures hardly captured by single point sensors (think mobile phones)
- Enable novel contexts e.g. user id, user environment
- Improve recognition using fusion with typical sensors

Today, radio signal information is used in Pervasive Computing to infer a range of contexts. Whole communities are dedicated to researching RSSI for deriving a users' location for approaches[14], while other research investigates detection of humans[13, 15].

Recently, radio signal-based activity and gesture recognition is approached. The Humantenna project[3] can be considered a device-bound activity recognition (DBAR) system. Therein Electro-magnetic noise received by the human is evaluated to determine location or pose. Other DBAR research[11], investigates the impact on GSM signal strength for the activities "walking", "driving" and "no movement".

In contrast, device-free activity recognition (DFAR) using radio has only been investigated using Software Defined Radios (SDRs).[9, 10, 6, 7] While they show that some activities can be recognized using radio signal analysis the findings cannot be simply transferred to low power, low cost transceivers in the focus of the research presented here (see section 6). Thus, no statement can be made regarding feasibility of activity/gesture recognition using RSSI, the magnitude prevailing todays common radio hardware. Hence, the contributions of this paper are as follows:

- Feasibility and quality of 802.15.4 RSSI based Activity Recognition
- Recognition dependence on a 3D network topology
- Influence of subject on recognition
- Influence of room change on recognition
- Differences of device-free RSSI, device-bound RSSI, device-bound accelerometer

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
AH'13 March 07 - 08 2013, Stuttgart, Germany
Copyright 2013 ACM 978-1-4503-1904-1/13/03 ...$15.00.

The paper has the following structure: Sections 2-4 present the methods ranging from the sensor test bed, experiment design to the data preprocessing. Section 5 comprises our results, i.e. general feasibility of RSSI-based DFAR and DBAR and the parameter influence (topology, subject, door state). Section 6 covers the discussion of related work and results.

2. RADIO SENSOR TEST BED DESIGN

In order to design a potential test bed we examine related literature for requirements. Bouten et al.[2] showed that human physical activities do not exceed 20Hz, i.e. optimally requiring a sampling rate of at least 40Hz. Extending the research to device-free localization and DFAR literature we find that a higher number of nodes increases spatial resolution and possibly recognition accuracy.[8] Based on few device-free RSSI-based sensing research investigating radio frequency[8, 6, 13] we conclude that 2.4GHz offers a good chance of being influenced by human activities. Similar to WSN deployment strategies (infrastructure vs. ad-hoc) different designs of a radio sensor are imaginable[8]. In order to allow best possible control of the sensor we choose to design our test bed as an infrastructure-based sensor. Hence, all WSN nodes are predeployed, immobile and their locations are known. Additionally to the infrastructure, we require a single mobile node which is carried by the user in the experiments. This node should be outfitted with a 3-axis acceleration sensor, the transceiver and logging capabilities.

Our requirements indicate a conflict: we require as much nodes as possible but we must also maintain a high sampling rate (i.e. sending and receiving) of 40Hz. For the radio sensor this means that each link in the network must be "used" 40 times per second. Thus, the more nodes are installed in the network the less time a node has for sending and receiving a packet. We choose the IEEE 802.15.4 NXP Jennic 5139 transceiver operating at 2.4GHz for which we seek to identify this boundary. To achieve this we tweak the Jennic wireless stack to switch off random send delays (up to 2ms) and set retries after collision to zero. Then, having a near deterministic sending behaviour we determine the minimum send time as close to 2ms. Adding 0.5ms safety we can theoretically achieve 40Hz with 10 nodes (2.5ms * 10 nodes * 40Hz = 1). We implement a time division multiple access (TDMA) algorithm on the Jennic directly on top of the 802.15.4 MAC layer. Therewith each node is assigned a time slot during which this node sends a broadcast packet. To compensate for the individual time drift of each node, we implement synchronization by adjusting all nodes to the clock of a predetermined master node which also participates in the TDMA. The accelerometer on the mobile node is a ADXL335 and also sampled at 40Hz. The RSSI of each received packet by each of the infrastructure nodes is relayed over USB to a PC where it is tagged by an experimenter with ground truth. Mobile node data is tagged based on timestamps after the tests.

3. EXPERIMENT

The experiment is conducted inside a small office room of 4.04 x 5.33m size, with gypsum walls on three sides and a window front. It has a bare concrete ceiling at the height of 2.70m. The room features typical pressboard furniture, two office chairs and two laptop computers. Among others Bao et al.[1] showed that considering human body halves separately increases recognition accuracy using accelerometers. Thus, we deploy 8 nodes in the experiment room in different heights. We choose 1.40m, approx. the height of the human torso and impacted by arm movement, for the nodes #1, #2, #3, #8 which are deployed in the rooms' corners. We choose 0.30m, approx. the height of the middle of an adult's shin, for the nodes #0, #4, #6, #7 located in the middle of the walls. Therewith node distance is between 2m to 6m a distance which has been shown to work for device-free RSSI-based detection of human motion[15]. Bao[1] also shows that an ideal location for recognizing a multitude of activities using a single sensor is a users' waist. Therefore, we attach the mobile node #5, which also has the accelerometer, on the hip of the subject performing the activities. Thus, in total we use 9 nodes in the experiment. A floor plan of the room with deployed nodes is depicted in fig. 1.

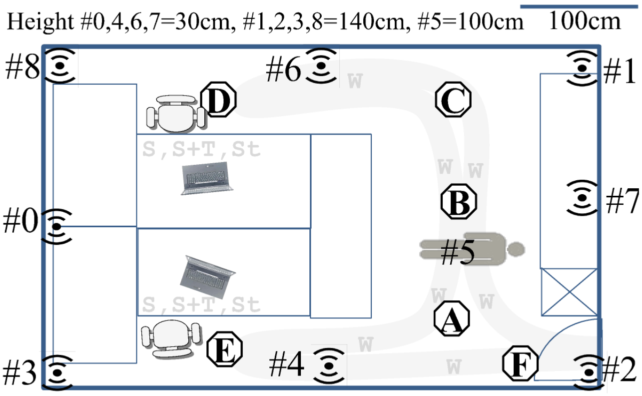

Figure 1: Room with furniture, node placements and symbolic locations

During the experiment the test subject performs a range of activities. We select activities of daily living, as well as activities that have been frequently investigated using accelerometers but also activities that seem hard to discriminate using accelerations. The selected activities are "walking" (W), "standing" (St), "sitting" (S), "sitting and typing" (S+T), "lie" (L), "lying and waving" (L+Wa) as well as being "outside the room" (O).

It is known, from existing research, that localization using RSSI is feasible. Thus, we need to ensure that really the activity and not the location of the person is recognized. Therefore activities are performed at five different locations (A-E) inside the room (see capital letters in fig.1). Location F marks the doors location. Crossing F indicates leaving or entering the room. The following table lists the four different scripted sequences that comprise the activities conducted by the subjects:

1	O, W(F-A-E), S(E), S+T(E), St(E), W(E-A-B-C-D), S(D), S+T(D), St(D), W(D-C-B-A), W(F), O
2	O, W(A), St(A), W (A,F), O, repeat sequence with B,C,D instead of A
3	O, W(A), L(A), L(A), L+Wa(A), St(A), L(A-F), O, repeat sequence with B,C,D instead of A

Each of these sequences was conducted by each of the two subjects while the door was either open, half open or closed. Thus, each sequence was repeated 6 times.

To clarify the experiment, an overlay in grey of the first sequence is shown in fig.1. The sequence starts with the subject entering the room (depending on sequence possibly interacting with the door at pos.F), walking from the door to pos.E and sitting at the chair there.

Activity sequences were conducted in three sessions on three consecutive working days. The 1-2 hour sessions started at 10pm, 3pm and 5pm. During the recordings the air pressure ranged from 992-1000 hPA, temperature varied from 21-23°C and rel. humidity from 28-29%. Over all three days we used IEEE 802.15.4 channel 25 (f=2.475GHz). We examined differing channel noise during all sessions due to office activity and Wi-Fi usage. The annotation was performed by a person located outside the room, hence only the person performing the activities was inside the room. Subjects conducting the activities were male, 175cm, 85kg (subject 0) and male, 176cm, 72kg (subject 1).

4. DATA PRE-PROCESSING

The data recorded during the experiments resulted in about 400,000 data samples 2:45h of data for the selected activity classes. Each sample contains the RSSI for all links in both directions (72 single way links for 9 nodes) and the 3-axis acceleration information occurring at node #5. According to recent studies link asymmetry is primarily an effect of transceiver power and receiver sensitivity[4], therefore we average link power on two way links to stabilize the signal strength information, creating 36 two-way links. Data with zero RSSI indicating lost packets were not removed from the data. Based on literature in classic Activity Recognition[1] and device-free RSSI-based localization[15] mean and variance over non-overlapping windows of 40 data instances (=1s) were selected as features. We chose non-overlapping windows to avoid including training information in test data when performing 10-fold cross validation or training/testing splits. Then, we adjusted the class prevalence i.e. number of instances per class to ensure that each class has the same number of representations as the class with the least representations. This leaves us with 40 instances per activity class per subject per door configuration. Random sampling the minimum instance count for every class/subject/door from the data set reduces the overall number of feature instances to 1680 instances (40 * 2 (subjects) * 3 (door stats) * 7 (activities) ≙ 28 minutes).

In the next step we split the data into three sets:

- Device-Bound ADXL (DBAR-ADXL) with 6 features: mean and variance for each axis of the accelerometer from the mobile node #5.
- Device-Bound RSSI (DBAR-RSSI) with 16 features: mean and variance for all 8 links to the mobile node
- Device-Free RSSI (DFAR-RSSI) with 56 features: mean and variance for all 28 links of the WSN infrastructure

Finally, the "outside room" class was excluded from the accelerometer data set.

5. RSSI-BASED ACTIVITY RECOGNITION

5.1 Brief Examination of Raw Data

Prior to applying machine learning for analysis on the recorded data we investigate a 5s raw data snapshot in fig.2. Therein we find an extract of raw data from a selected single link or axis from each of the sensors. Inspecting the figure most of the activities look significantly different. To our surprise we even find that the data from the waist-attached accelerometer also shows differences in the activities "sitting and typing" compared to "sitting", and "lie and waving" compared to "lie". When examining the experiments' video capture it becomes clear that this is due to a body posture change (e.g. "sitting still" performed leaned back; "sitting and typing" performed bowed forward). For the

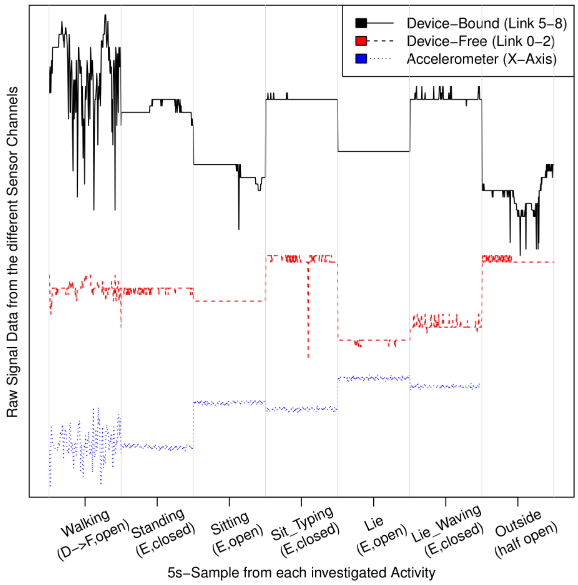

Figure 2: Examplary raw data of each sensor for each activity performed by the same subject.

RSSI sensors we find that "walking" seems to have a much stronger impact on the signal for the device-bound than for the device-free RSSI sensor. An expected effect, also observed e.g. by Woyach et al.[13], related to the strong signal changes the movement of the transceiver induces, another aspect leading to the observed volatility is probably related to arm motion of the subject in front of the waist-attached transceiver while walking. Comparing the specific performance of the activities by the subjects using video material we find that for the device-free sensor the level of magnitude change in the signal often correlates to size of an object being moved in the link (cf. "typing" vs. "waving" in fig.2), while for device-bound it may be an object being moved in the line of sight (LOS) or a strong motion of the transceiver itself. Both effects probably have their source in changed interaction between radio waves and objects (human, floor, ceiling, walls, table, etc.) causing different reflection and shadowing constellations.

Hence, for the device-free case the number of radio paths affected will change in dependence of the objects size (larger object = more propagation paths affected). In contrast, for the device-bound RSSI sensor, motion of the transceiver and motion of an object in its proximity will both change a large number of signal paths. This could be a source of ambiguities in the recognition.

Another reason for strong signal fluctuations are packet losses (cf. the activity "sitting and typing" for device-free in fig.2 at 3s). As these have been included down peaks shown in the figure indicate lost packets in one link direction. The data snapshot further suggests that both RSSI based sensors deliver a relatively stable signal for the static activity "lie". The device-free sensor also shows the same stable measurements for the static activity "sitting". The signal fluctuation in the device-bound can be currently only explained with other environmental noise. In the video no direct movement related to the change can be observed.

5.2 Applying Machine Learning on the Data

In order to avoid classifier dependent results three typical Activity Recognition classifiers are selected: k-Nearest Neighbours (kNN, k=10), naive Bayes (nB) and C4.5 decision tree. We evaluate these on the extracted data sets using the Orange data mining framework[1] in a 10-fold cross validation (10-CV). For kNN data is normalized before training/testing. Table 1 shows the results.

Table 1: 10-fold cross validation for device-based (DBAR) and -free (DFAR) activity recognition

	kNN	C4.5	nB
	Acc./F1	Acc./F1	Acc./F1
DBAR-ADXL	.856/.855	**.880/.880**	.825/.821
DBAR-RSSI	**.896/.896**	.791/.791	.689/.675
DFAR-RSSI	**.894/.892**	.836/.835	.874/.866

The best classification results for all three sensors are comparable at around 89%. kNN delivers best results for the radio sensors, while C4.5 does for the ADXL. Also, while DFAR-RSSI and DBAR-ADXL both have an average accuracy for all classifiers of around 85%, the DBAR-RSSI sensor shows a degraded performance comparing the kNN results to the C4.5 and the naive Bayes result. The naive Bayes result on the DBAR-RSSI data delivers the worst performance at only 70%.

For the RSSI sensors the kNN F-Measure indicates a higher precision and recall than for the ADXL. On the other hand, when looking at the kNN classifier the difference between the RSSI sensors themselves is minimal although the device-bound sensor only has a fraction of the features (16 vs 56) compared to the device-free system to gain information from.

One aspect which could explain this superiority of kNN may lie in its distance calculation. As kNN treats all features equally a feature which has a strong information gain or probability in the training data will not be treated differently when testing. In contrast, C4.5 might use a feature providing good discrimination in the training set as top level decision feature failing on the test data.

Tables 2, 3, 4 show the confusion matrices for discrimination of activities for the device-bound ADXL (DBAR-ADXL), device-bound RSSI (DBAR-RSSI) and device-free RSSI (DFAR-RSSI) data sets for the best classifiers, respectively.

Table 2: Confusion Matrix for DBAR-ADXL C4.5 decision tree.

True\Pred.:	W	St	S	S+T	L	L+Wa
W	**221**	17	0	0	0	2
St	10	**230**	0	0	0	0
S	3	0	**232**	5	0	0
S+T	0	0	10	**230**	0	0
L	0	0	0	0	**172**	68
L+Wa	0	0	0	0	93	**147**

From the confusion matrices we make the following observations:

- ”Walking” (W) is best discriminated using the ADXL data with DBAR-RSSI showing similar results. This makes sense as both sensors are largely influenced by the movement of the object the sensor is attached to.

[1]http://orange.biolab.si/

Table 3: Confusion Matrix for DBAR-RSSI kNN.

True\Pred.:	W	St	S	S+T	L	L+Wa	O
W	**217**	9	2	0	0	1	11
St	19	**200**	2	1	5	5	8
S	6	1	**230**	3	0	0	0
S+T	3	0	5	**230**	0	0	2
L	0	0	0	0	**229**	11	0
L+Wa	0	1	0	0	23	**216**	0
O	21	8	0	2	12	13	**184**

Table 4: Confusion Matrix for DFAR-RSSI kNN.

True\Pred.:	W	St	S	S+T	L	L+Wa	O
W	**170**	18	7	1	0	9	35
St	10	**229**	0	0	0	0	1
S	0	0	**240**	0	0	0	0
S+T	1	0	8	**230**	0	1	0
L	0	0	0	0	**221**	17	2
L+Wa	0	0	0	0	38	**200**	2
O	2	5	5	0	11	5	**212**

- ”Standing” (St) is discriminated well by DBAR-ADXL and DFAR-RSSI. DBAR-RSSI has some confusion with ”walking” probably due to slight rotational body movements during the activity. While in the ADXL data this will only be reflected on a single axis for DBAR-RSSI all links will be affected.

- ”Sitting” (S) is discriminated equally well using any sensor.

- ”Lie” (L) and ”Lying and Waving” (L+Wa) are worst discriminated using the ADXL data, while both RSSI sensors detect these activities with similar accuracy.

- The ”outside room” activity (O) is classified best by DFAR-RSSI.

Looking at the confusions we particularly see that activities ”lying and waving” and ”lying” are more often confused for the ADXL data set then for the RSSI based data. RSSI-based systems confuse these activities to a lesser extend with each other, but rather with other seemingly unrelated activities (mostly ”outside room”). The activities ”sitting” and ”sitting and typing” are more often confused using the ADXL data. In contrast, the RSSI-based systems have problems discriminating ”walking” and ”standing” and ”outside room”. In fact, most classification errors using the RSSI-based data sets occur by confusing the ”outside room” activity with any of the other activities.

5.3 Influence of Network topology

For the evaluation regarding network topology all k-combinations of node connections were simulated based on the complete RSSI data sets described in sec.4. For any number of node pairs a 10-fold cross-validation (10-CV) was performed.

5.3.1 DBAR-RSSI

Due to the nature of the device-bound sensor the network topology always constitutes a star topology in which all nodes of the infrastructure communicate with the body attached node #5. The evaluation results for the device-bound RSSI sensor are visualized in the box plot given in fig.3. Thereby each box represents a selection of 2, 3, etc. nodes from the set of available nodes in the infrastructure. However, in each selection node #5 is included. For instance, for the first box the data set is reduced to only a single link (two nodes), i.e. we investigate the link #5-#1 in a single 10-CV investigation, then #5-#2 in a single investigation, then #5-#3 and so on. For the second box we reduce the data set to contain only two links for each evaluation. I.e. we look at the recognition rate when #5-#1 and #5-#2 are in the data; then #5-#1 and #5-#3 and so on. Thus, the number of links varies as the number of nodes to include increases. The last box yields the same 10-CV-result given in the result table 1 for each classifier as all nodes are included in the data.

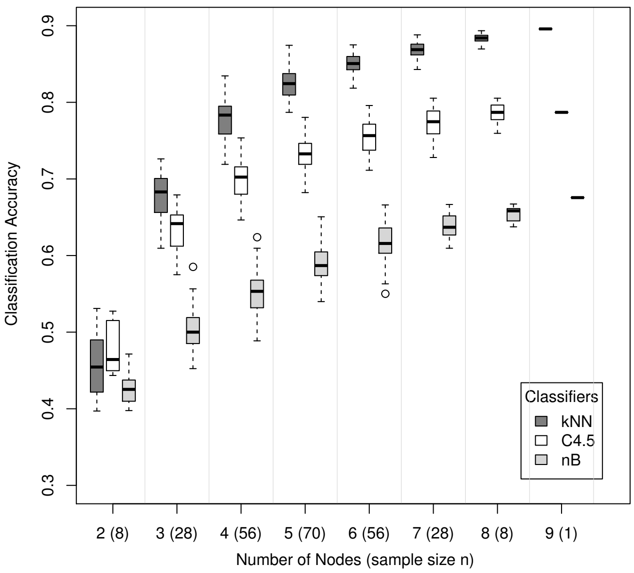

Figure 3: DBAR-RSSI accuracy for changing node numbers.

The box plot shows the expected trend: a larger number of nodes lead to an improved recognition rate. We also find the previously observed accuracy differences across classifiers maintained: kNN performs best (with exception of node count 2) while C4.5 performs second best and naive Bayes provides least performance. All boxes describe a saturation tendency with a limit at around 90% for kNN, 80% for C4.5 and ca. 70% for the nB. Looking at the ascent of the curve kNN and C4.5 accuracy improves significantly from 2 to 3 nodes (kNN median increase by 25%, C4.5 median increase by 20%) and from 3 to 4 (kNN median increase by 10%, C4.5 by 5%). For further investigation of this effect we give the links yielding the best performance for kNN for node counts 2 to 5:

Node Count	Best Links	Acc.
2	#5-#8	53%
3	#5-#8, #5-#4	71%
4	#5-#8, #5-#4, #5-#2	83%
5	#5-#8, #5-#4, #5-#2, #5-#6	86%

We find that especially #4 and #8 seem helpful for the discrimination of activities. Looking at the floor plan in fig.1 we see that #8 is located in the corner near location D at 140cm height, while #4 is next to location D at 30cm. Node #2 on the other hand is again at 140cm in the corner next to the door, while #6 is next to D in 30cm height.

The varying height of the top selected nodes may be an indication, that similar to classical accelerometer-based activity recognition, splitting up the radio sensor WSN in two different planes for upper and lower body parts supports the recognition.

To evaluate this assumption we compare the kNN 10-CV confusion matrices of the single link data (#5-#8) with the dual link data (#5-#8,#5-#4). We find the following increase in correctly classified instances for (#5-#8,#5-#4): S(+77), L(+66), S+T(+54), O(+46), St(+36), L+Wa(+27), W(+9). While it seems surprising that "lying and waving" is only advanced by 27 additional correct classifications, we have to keep in mind that we are looking at the device-bound RSSI sensor data here. In this sensor movement or location of the sensor itself have typically a stronger impact than motion performed between the LOS or its multiple paths to the on-body node. The activities S(itting) and L(ying) whose recognition improvements are significant, obviously profit from the position of node #4 (height: 30cm), which is on a vertical axis closer to #5 during these activities (S: waist at ca. 55cm, L: waist at ca. 15cm) than node #8, same is true for the "sitting and typing" ("typing" at ca. 75cm). Better discrimination of these activities reduces confusion with the outside activity. Lastly, activities St, L+Wa, W probably profit simply due to additional signal information. Thus, we find our assumption partially proven as indeed these activities affect the nodes installed at a low position.

5.3.2 DFAR-RSSI

The network topology of the device-free RSSI WSN is a non-regular shape consisting of two planes which are rotated by 45 degrees of which all plane vertices are connected to each other. Node #5 is not part of this evaluation.

Similar to the previous investigation we present the box plot for 10-CV for all three classifiers for changing node configuration for DFAR-RSSI in fig.4. Here each box describes the number of nodes for the evaluation and the sample size is the total number of links evaluated for this box. The evaluation is similar to DBAR-RSSI but all interconnections between all nodes are evaluated.

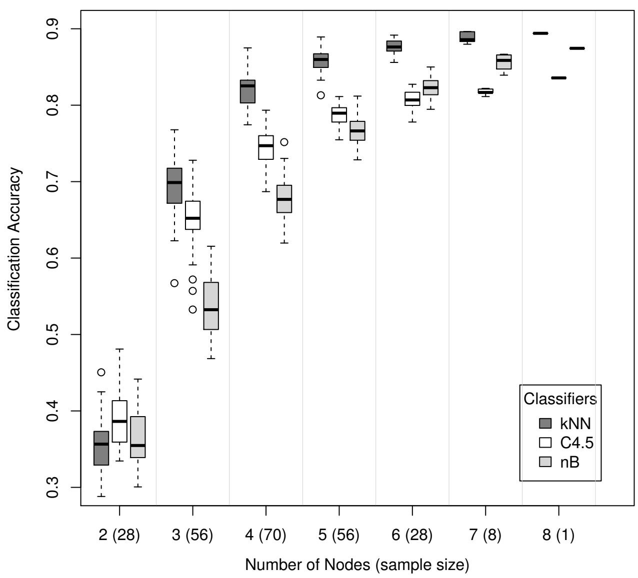

Figure 4: DFAR-RSSI accuracy for changing node numbers.

As for DBAR-RSSI more links provide better discrimination with kNN showing a superior performance. Interestingly, for two nodes C4.5 outperforms kNN. The difference between classifiers regarding performance is similar to the

DBAR investigation until node count 5. Then all classifiers converge to a performance between 80% to 90%. Starting from 6 nodes naive Bayes outperforms C4.5 showing a greater spread of accuracies among the link configurations.

For the node counts until 3 the performance increments (median) are larger than for DBAR-RSSI: for kNN from 2 to 3 (31%) and from 3 to 4 (11%). For 4 nodes kNN is already around 87% close to the maximum with 8 nodes. This is not surprising, as the number of links for DFAR-RSSI (and therewith the spatial coverage) increases with (N*(N-1)) while for DBAR-RSSI the number of links increases with N-1, when N is the number of nodes. Thus, for N=4 we have 12 links with DFAR-RSSI but only 3 with DBAR-RSSI. For DFAR-RSSI we investigate the top selected nodes for node counts 2 to 4:

Node Count	Nodes Giving Best Accuracy	Acc.
2	#0,#2	45%
3	#0,#2,#3	76%
4	#3,#4,#6,#7	87%

In respect to WSN topology we find that for the first three top link configurations selected nodes alter between high and low nodes. The selected high nodes are both attached to a corner on the same wall. This changes with 4 nodes, this configuration particularly contains nodes #4, #6, #7, which are all low nodes. Interestingly, in this configuration only a single high node (#3) is included. Nodes #2, #6 and #4 were also chosen for the DBAR-RSSI as top contributors. Nodes #7 and #3 seem to provide specific benefit for the DFAR-RSSI recognition in this setup (room, locations, and activities). Node #8 which provides 53% performance for DBAR-RSSI is not selected, although it is on the same height as #3. This could be explained by the position of node #5 (mobile node) which is always attached to the right side of the waist. Due to this, #8 probably had a near perfect LOS connection to #5, e.g. when the subject sat at D and E. In contrast, the subjects' body would be in-between LOS of #3 and #5 at the same locations.

5.4 Influence of Subject

In this evaluation we investigate the impact of subject on recognition. We split our data in test and training sets based on the two different subjects and use either to test classification on the other set. In tab.5 we give the average accuracy for generalization from one subjects' data to the other.

Table 5: Results for 50/50 training/test split on subject. SX→SY indicates classifier trained on subject X and tested on Y.

	kNN	C4.5	nB
	Acc./F1	Acc./F1	Acc./F1
DBAR-ADXL			
S0→S1	.499/.339	.521/.504	**.706/.640**
S1→S0	.678/.614	.550/.467	**.675/.680**
DBAR-RSSI			
S0→S1	.579/.583	.356/.348	**.583/.566**
S1→S0	.530/.521	.501/.498	**.593/.593**
DFAR-RSSI			
S0→S1	.442/.438	.418/.421	**.451/.418**
S1→S0	**.538/.537**	.510/.519	.512/.471

Performance of the classifiers on the ADXL data is also given as it serves as a known benchmark of what to expect and to factor out differences in how an activity was performed. We are aware that the achieved results do not challenge latest research in Activity Recognition but believe they can still be used to grasp the trends for certain activities. Thus, using our un-optimized standard features on the ADXL data we only achieve around 68% average accuracy. For the RSSI sensors the difference to the 10-CV baseline (cf. sec.5.2) is more drastic: DBAR-RSSI achieves a maximum of 60%, while DFAR-RSSI achieves 45%/54%. A drop of 20 percentage points compared to the ADXL. In contrast, to the previous investigation naive Bayes gives best performance.

From the confusion matrices (omitted here) we find for ADXL the confusion of "lying" and "lying and waving", "sitting" and "sitting and typing" cause the major reduction in accuracy. In the case (S0→S1) only a single correct instance (true positive) of "lying and waving" from 115 was identified. For the RSSI sensors misclassifications cannot be pinned to a single activity. Best recognized activities for DBAR-RSSI are "walking" (true positives: 113/115) and "lying" (105/115). For DFAR-RSSI best recognized were "walking" (115/115), "lying" (101/115) and "sitting and typing" (74/115).

A challenge in traditional Activity Recognition has been the transfer of recognition models in-between subjects. Major reasons are differences in activity execution and sensor attachment (slight rotation, not perfectly fixed, etc). While DBAR-RSSI can certainly be affected by sensor fixation differences, the DFAR-RSSI is robust to such changes. Both RSSI sensors however, are certainly affected by activity execution differences as the radio signals propagate in all directions and are also affected by the movement of body parts not detected by the single point accelerometer. This might partially explain the larger accuracy drop compared to the ADXL. However, this may not be the only reason. Instead, we believe that also the individual performing the activity makes a difference.

In fact, the subjects' height, weight, volume and clothing must affect RSSI. As an extreme example imagine a person in heavy firefighter apparel with a metallic oxygen flask vs. a barely dressed victim. For simplicity, the reader may consider the different shadows of those individuals.

Assuming there is this individual radio distortion pattern, we can explain why DBAR-RSSI is not reduced to the same extend as DFAR. As DBAR-RSSI is strongly affected by the motion of the body attached node a strong activity signal can be probably taken from this information which is similar to the ADXL information. In contrast, the DFAR system is solely based on the link strengths in-between the infrastructure. It cannot "rely" on being attached to a certain place on the body of the users undergoing typical movements. Thus, a DFAR system is also more susceptible to an individuals' "noise signature". As our sample size is very small, we plan further investigations on the impact of different subjects.

5.5 Change of Door state

In this evaluation we investigate the influence of a small geometrical change in the room (door open, closed, half open). We split the data in train and test sets, whereby the test set contains all activities of all users with the door state under investigation. The remainder constitutes the training data. In table 6 we give the average classification results for all sensor systems. Again we include the classification accuracy on ADXL data as baseline to separate effects induced by the door vs. effects induced by differences in activity execution.

As the ADXL results indicate, activity performance does

Table 6: Results for all 2/1 training/test splits for different door states

	kNN	C4.5	nB
	Acc./F1	Acc./F1	Acc./F1
DBAR-ADXL			
open	.779/.779	.776/.768	**.826/.820**
closed	.753/.753	.830/.829	**.852/.852**
half	.795/.790	.775/.773	**.809/.805**
DBAR-RSSI			
open	.407/.392	.388/.381	**.477/.428**
closed	.594/.598	.530/.536	**.699/.710**
half	.556/.557	.369/.331	**.688/.688**
DFAR-RSSI			
open	**.593/.589**	.537/.536	.521/.500
closed	.591/.587	**.619/.634**	.523/.497
half	.577/.572	.49/.504	**.595/.579**

not vary greatly: accuracies and F-measures are similar throughout all investigations. Thus, we don't expect a strong impact in the other sensors' values by activity differences which are detected by the accelerometer.

For both RSSI sensors the closed door allows for the optimal detection. A pragmatic explanation of this observation is, that when closed the door provides an additional reflector and thus allows for a better room coverage. Another more thorough attempted explanation, is to remind one that the classifiers have been trained on the data for "open" and "half". If in those cases signals are, for instance, only reflected at a wall outside the room the distance of a path between nodes compared to a closed door increases. In turn, this will result in a weaker impact of an activity on the signal (cf. [5] who showed that impact of motion on RSSI decreases with distance). Thus, signal changes or regions which have previously been identified in the data are now emphasized due to the stronger impact of activities, possibly making the discrimination of activities simpler for classifier. For DFAR-RSSI both of the investigations show a similar performance.

For DBAR-RSSI the open door reduces accuracy by 20 percentage points. This may be due to the fact that the open door covers node #2, which was shown as important for DBAR-RSSI recognition (cf. topology investigation in sec.5.3.1). Activities affected significantly by this geometry change are "sitting" (true positives: 2 of 115) and "sitting and typing" (10/115). These are confused to a large extend with "standing" and "empty". Thus, node #2 must be especially important for DBAR-RSSI for these activities.

Not surprisingly (cf. topology investigation, sec.5.3.2) DFAR-RSSI is only slightly affected by the occlusion of #2 by the door. With the "door open" the system has problems discriminating "lying and waving" from activities "lying" and "standing". Activity "standing" is only rarely recognized and confused with nearly all other activities. Opposed to DBAR-RSSI, the best detected activities are "walking", "sitting" and "empty".

6. RELATED WORK AND DISCUSSION

Prior to this work others have used radio reception distortions to derive activity or gesture related contexts in a device-bound but also in a device-free fashion.

The humantenna[3] constitutes a radio receiver installed on a subjects' skin, treating him/her as a morphable, noisy, wide-band antenna. The authors show that analyzing the received radio noise a users' location and pose can be derived. As the received noise depends on emissions from the environment, the system is very location dependend. Sohn et al.[11] present another device-bound work. Using link quality between mobile phones and GSM base stations they discriminate: "walking", "driving" and "still" on one month data from 3 subjects. In all of this research sensor parameters such as sampling frequency, emitter location, signal quality and/or carrier frequency (Humantenna) are uncertain i.e. not controlled by the investigator. While this aspect makes the systems very interesting, the development challenge for practical models seems very hard as e.g. signal quality can be heavily affected by external factors. We may also experience such influences (e.g. Wi-Fi on same channel) but we know what to expect from our system. Ultimately we also plan to reuse existing infrastructure but due to the noise of the measured quantity we believe that the development of robust models in a more controlled environment first is a valuable strategy.

Only few research exists which investigating device-free radio-based activity recognition (DFAR) and all rely on software defined radios (SDRs). In [7], two SDRs are setup parallel to an office room door. One SDR transmits at 900MHz, the second SDR samples the received signal at 320KHz. The online system recognizes the activities: "Walking"/"Not Walking", "Phone call"/"No Phone Call" and "Door open"/ "closed" online. Recognition uses three distinct threshold-based classifiers which calculate the average magnitude of the received signal either every 50ms (detection of "walking", detection of "phone call") or every second (detection of "door"). Thresholds are set during a training period of 10s. Accuracies are up to 95%, 81%, 90% for the "door", "walking"/"still" and "phone" context, respectively. Confusions occur when the subject walks through the door or stands close to the door (door is recognized as closed). Interestingly, despite the simple setup but more sophisticated radio hardware, recognition of the activity "walking" is also reduced when the door is open compared to closed door.

In [6], Sigg et al. apply machine learning algorithms (kNN, naive Bayes, C4.5, Orange decision tree) on data from two or three SDRs placed in the corners of a small room (15sqm). One SDR transmits continuously at 900MHz/2.4GHz, the other SDRs sample the signal at 320KHz. Features calculated over windows of 16000 I/Q values (50ms) are: root mean square power, signal to noise ratio and average magnitude squared. Training and evaluation are performed offline. For each activity 600 feature instances are recorded: 100/500 samples are used as training/test data. Activities are "sitting", "walking" and "standing". Best accuracy was 64% (2.4 GHz, three SDRs, C4.5) compared to 62% (900 MHz, three SDRs, C4.5). Reducing the number of SDRs reduces accuracy to 17% (2.4 GHz, Orange Tree/Rule Learner) and 61% (900MHz, C4.5). While accuracies are much lower than those reported in this investigation, our accuracies have been calculated using 10-CV which can be seen as upper recognition boundary. The low accuracy may also be explained with the insufficient spatial covering of the three or two links, respectively.

Shi et al.[9, 10] recognize activities ("empty", "lying", "running", "standing", "crawling", "walking") using a single SDR receiving FM radio at 82.5MHz. In both investigations the SDR is placed next to the rooms' door. Activities are conducted in a very restricted 1x2m rectangle around the SDR. They propose a two-stage classifier[10] to discriminate static (e.g. "standing") and dynamic activities (e.g. "walking") and thereafter identify the actual activity. They achieve a slight improvement from 87% vs 84% over a single-stage classifier in 5-fold cross validation for all activities. The confusion matrices show that especially the activities "crawling" and "walking" profit from the dynamic/static separation. Activities which are always (independent of the approach) classified correctly are: "lying", "empty", "standing" (around 90% accuracy). In contrast, all three dynamic activities seem hard to discriminate (90%-55%). As in the reviewed device-bound publications Shi et al. have no control of the sending entities (e.g. power, location, signal) making it hard to estimate the repeatability of this work. Also the activities are conducted only in a small restricted area which: 1) makes the system less practical and 2) inflicts hardly predictable near field effects in the receivers antenna.

All currently published DFAR research relies on SDRs which differ significantly from end-user radio hardware. Fundamentally different parameters to the presented testbed system include: channel sampling rate differences on the order of multiple magnitudes (SDR: 256kHz vs. testbed: 40Hz), channel bandwidth (256kHz vs. 2MHz), available signal information (I/Q values vs. RSSI), signal resolution (float vs. single byte RSSI), receiver sensitivity, etc. Thus, findings based on SDR hardware are not simply transferable to measurements conducted with common radio hardware.

To the best of our knowledge this paper is the first of its kind reporting on purely RSSI-based activity recognition. Using most basic features and unoptimized algorithms we were able to show the impact of different parameters on activity discrimination. Depending on the sensor type (DFAR-RSSI/DBAR-RSSI) certain topology requirements are sufficient to allow for discrimination. We further found that subject as well as minor geometrical changes have a noticeable impact on classification performance.

However, we are aware that the sample size for the subject dependency evaluation is small and the influence of the door and WSN topology must depend on the layout of the room and the space where activities are conducted.

Additionally, applying cross validation for classifier evaluation is prone to overfitting and thus, the reported results may be overconfident. Still dependencies on investigated parameters must be seen as valid and, most importantly, parameter alterations are indeed reflected in the RSSI. Thus, while this work does not present a polished recognition algorithm it raises challenging research questions and demands for further studies. Besides investigating new contexts such as subject discrimination/identification, gestures or maybe even subject health state, the most important challenge is the creation of an adaptable, robust sensor algorithm and a practical and predictable radio sensor. The device-free localization community has developed models with similar goals but for localization, e.g.[5, 12]. These may provide a basis for developing recognition algorithms and provide insights into further sensor parameters.

7. CONCLUSION

In this paper we showed the general feasibility of activity recognition using RSSI on simple transceiver hardware. We further showed the influence of a number of relevant sensor parameters. Surprisingly, even activities with relatively small motion (e.g. "typing") affect RSSI and are detectable with few WSN nodes, showing potential for gesture recognition. Another potential context may be the identification of subjects. Understanding the influence of such parameters and incorporating these in adaptable models, algorithms and radio sensors will make RSSI-based activity recognition feasible in practice.

8. REFERENCES

[1] L. Bao and S. Intille. Activity recognition from user-annotated acceleration data. In A. Ferscha and F. Mattern, editors, *Pervasive Computing*, volume 3001 of *Lecture Notes in Computer Science*, pages 1–17. Springer Berlin / Heidelberg, 2004.

[2] C. Bouten, K. Koekkoek, M. Verduin, R. Kodde, and J. Janssen. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Transactions on Biomedical Engineering*, 44(3):136 –147, Mar. 1997.

[3] G. Cohn, D. Morris, S. Patel, and D. Tan. Humantenna: using the body as an antenna for real-time whole-body interaction. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, pages 1901–1910, New York, NY, USA, 2012. ACM.

[4] P. Misra, N. Ahmed, D. Ostry, and S. Jha. Characterization of asymmetry in low-power wireless links: an empirical study. In *Proceedings of the 12th ICDCN'11*, Berlin, Heidelberg, 2011. Springer-Verlag.

[5] N. Patwari and J. Wilson. Spatial models for human motion-induced signal strength variance on static links. *Information Forensics and Security, IEEE Transactions on*, 6(3):791 –802, sept. 2011.

[6] M. Reschke, J. Starosta, S. Schwarzl, and S. Sigg. Situation awareness based on channel measurements. In *Proceedings of the fourth Conference on Context Awareness for Proactive Systems (CAPS)*, 2011.

[7] M. Scholz, S. Sigg, G. Bagschik, T. Guenther, G. von Zengen, D. Shishkova, Y. Ji, and M. Beigl. Sensewaves: Radiowaves for context recognition, video submission. In *Pervasive*, 2011.

[8] M. Scholz, S. Sigg, H. R. Schmidtke, and M. Beigl. Challenges for device-free radio-based activity recognition. In *Workshop on Context Systems, Design, Evaluation and Optimisation*, 2011.

[9] S. Shi, S. Sigg, and Y. Ji. Activity recognition from radio frequency data: Multi-stage recognition and features. In *VTC2012 Fall*, 2012.

[10] S. Shi, S. Sigg, and Y. Ji. Passive detection of situations from ambient fm-radio signals. In *SegaWare Workshop in Conjunction UbiComp '12.*, 2012.

[11] T. Sohn, A. Varshavsky, A. Lamarca, M. Y. Chen, T. Choudhury, I. Smith, S. Consolvo, J. Hightower, W. G. Griswold, and E. D. Lara. Mobility detection using everyday gsm traces. In *Ubicomp 2006*, pages 212–224. Springer, 2006.

[12] J. Wilson and N. Patwari. A fade-level skew-laplace signal strength model for device-free localization with wireless networks. *Mobile Computing, IEEE Transactions on*, 11(6):947 –958, june 2012.

[13] K. Woyach, D. Puccinelli, and M. Haenggi. Sensorless sensing in wireless networks: implementation and measurements. In *2nd International Workshop on Wireless Network Measurement (WiNMee)*, 2006.

[14] M. Youssef, M. Mah, and A. Agrawala. Challenges: device-free passive localization for wireless environments. In *Proceedings of the 13th MobiCom*, pages 222–229, New York, NY, USA, 2007. ACM.

[15] D. Zhang and L. Ni. Dynamic clustering for tracking multiple transceiver-free objects. In *PerCom*, march 2009.