Enabling Dynamic Sensor Configuration and Cooperation in Opportunistic Activity Recognition Systems

Marc Kurz, Gerold Hölzl, and Alois Ferscha

Institute for Pervasive Computing, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria

Correspondence should be addressed to Marc Kurz; kurz@pervasive.jku.at

Received 5 December 2012; Revised 18 April 2013; Accepted 21 April 2013

Academic Editor: Ming Liu

Copyright © 2013 Marc Kurz et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Opportunistic activity recognition as research discipline is characterized by the fact that human activities (and more generally the context) shall be recognized with sensors that are initially unknown to the system. In contrast to “traditional” applications—where sensors, their modalities, locations, and working characteristics have to be defined at design time—opportunistic systems do not rely on an initially defined and fixed sensing infrastructure. Sensors have to be utilized upon their spontaneous availability and activity recognition capabilities and dynamic sensor ensembles have to be configured at runtime with respect to maximized recognition accuracy and minimized energy consumption. This requirement contains two research challenges that this paper tackles: (i) estimating the accuracy of an ensemble without being able to compare the output in the form of recognized activity classes to a (labeled) ground truth and (ii) optimizing the accuracy/energy trade-off by applying exact and heuristic methods adapted for cooperative sensor ensembles.

1. Introduction

Opportunistic activity and context recognition is a rather new research field [1, 2]. It is motivated by the fact that sensor systems of different kinds are nowadays integrated into things of everyday life, which become more and more present (e.g., smart phones, wearable computers, gadgets, etc.). This makes the explicit deployment of sensors for specific activity recognition applications gratuitous. Opportunistic activity and context recognition aims at utilizing sensors upon their spontaneous availability and configuring sensor ensembles that are best suited to execute a certain recognition goal. This recognition goal can be defined at runtime by an application or user [3] and defines what the systems shall recognize but does not specify how [2]. Sensor ensembles combining available sensor devices to recognize certain activities of interest (as formulated in the goal definition) have to be dynamically configured upon the spontaneous availability of sensor devices. An easy approach would be to include all sensors that are capable of recognizing a specific activity (that is requested within the goal formulation). Nevertheless, this “brute-force” approach does not necessarily mean that the accuracy can be optimized/maximized; the only effect that can be guaranteed is the fact that the energy consumption of the sensor ensemble is increased. An ensemble consisting of heterogeneous sensor devices shall meet two important requirements: (i) maximize the accuracy, whilst (ii) minimizing the overall ensemble energy consumption (e.g., to ensure a long lifetime of the sensor devices if operated on battery mode).

These two requirements could be conflicting; therefore, a trade-off function has to be found to define sensor ensembles utilizing currently available sensor devices. An important aspect that has to be considered is the real-time practicability; the sensing infrastructure at any point in time cannot be assured to be static and fixed; instead it is of highly dynamic nature [1]. Another crucial and nontrivial aspect is the fact that the accuracy of an ensemble cannot be easily estimated. Usually, the accuracy for a (sensor) signal classification system is based on a labeled ground truth, where the predicted activity classes can be compared to the actual classes. Since an opportunistic system operates in real time and cannot initially fix the set of input sensor devices, a ground truth cannot be taken for granted. This paper tackles two
research challenges that emerge from opportunistic activity and context recognition.

(I) **Unsupervised Accuracy Estimation.** The expected accuracy of a (initially unknown) sensor ensemble has to be estimated. Possible input parameters could be the accuracy of single sensor devices, diversity measures between sensor combinations, and the technical properties (e.g., sampling rate, sensor signal modality, etc.).

(II) **Ensemble Optimization.** Based on the expected accuracy and the expected energy consumption of the ensemble (calculated as sum of the sensors’ individual energy requirements), an optimal trade-off (maximizing the accuracy whilst minimizing the energy demand) has to be found in order to configure ensembles at system runtime.

The remainder of the paper is structured as follows. Section 2 discusses related work in the broader fields of “traditional” activity recognition and utilizing sensors in a dynamic way. Recent achievements in the field of opportunistic sensing—which build the baseline for the research work within this paper—are also discussed in Section 2. The methodologies (i.e., how do we tackle the challenges of estimating the ensemble accuracy and optimizing the ensemble accordingly) are described in detail in Section 3. These methodologies are evaluated in Section 4 on a rich dataset that has been recorded including more than 70 sensors with 10 different modalities in a kitchen scenario (see [4, 5]). The paper closes with a conclusion and an outlook to future work in Section 5.

**2. Related Work and Recent Achievements**

Related work that matches the scope of this paper can be identified in two areas of research: (i) activity and context recognition and (ii) opportunistic sensing and opportunistic activity recognition. The following two Sections 2.1 and 2.2 contain overviews of related work (and recent achievements) within these two areas. Section 2.3 summarizes the findings and provides a discussion of the novelty of the targeted research challenges.

2.1. **Activity Recognition.** This section provides a discussion on related work across different types of activities and different scales of activities’ complexity. The recognition of user’s activities and more generally the context [6] of persons by utilizing different body-worn and environmental sensor devices has been well evaluated in previous work (e.g., [7–9]). There, the authors attached a set of acceleration, inertial, and microphone sensors on the body of subjects, utilized different signal processing and machine learning algorithms for classification, to investigate in the quality of the detected activities in terms of recognition rate, accuracy, and other performance metrics [10]. Again, as we are interested in the recognition of activities of persons, the choice of sensors to achieve an accurate estimation of the activities together with the requirement to be as unobtrusive for the user as possible is a crucial task. In [11], Tapia et al. focused on recognizing activities with a set of simple and ubiquitous sensors (“tape on and forget” sensors). Logan et al. [12] investigated in the question which sensors and modalities in general are well suited for highly accurate activity recognition, whereas the findings are that infrared motion detection sensors outperformed others (e.g., RFID, built-in wired sensors). A third contribution that investigates accurate recognition of activities in user’s homes is presented in [13]. Again, the presented system relies on a set of specific wireless sensor nodes that have to be explicitly deployed in the home environment. Summarizing, the best sensor modality for on body activity recognition are accelerometers [14] due to their high accuracy and their unobtrusiveness. They are already integrated into things of everyday life like in smart phones or smart watches and can, therefore, be easily utilized.

2.2. **Opportunistic Sensing.** Opportunistic activity and context recognition as new research discipline draws from the fact that a plethora of sensor devices is already deployed in the world. Therefore, the application-specific sensor deployment is an effort that can be avoided. Important methods that have already been developed tackle the fact that an unknown and highly dynamic sensing infrastructure consists of heterogeneous devices, capable of delivering different functionality to an activity recognition system. Therefore, the concepts of sensor abstractions and sensor self-descriptions (see [1, 15, 16]) allow for the dynamic generation upon sensor availability of activity recognition chains (ARCs) dependent on the sensors capabilities as preserved in the self-descriptions. The OPPORTUNITY Framework [1] is a ready-to-use system combining different methods for enabling highly flexible, dynamic, and stable activity recognition. ExperienceItems as building blocks within the sensor self-descriptions memorize gathered experience in form of recognizable activities. The degree of fulfillment (DoF) value [1] is a metric (between 0 and 1) that is used to estimate the expected accuracy of a sensor/ensemble for a specific recognition goal. The DoF together with required machine learning and signal processing technologies (and accompanying parameters) is persistently stored in the ExperienceItems. These items can either be generated offline—in a traditional way by using a labeled ground truth—or (as shown in [17]) at systems runtime by applying the concept of transfer learning. There, a sensor (or sensor ensemble) is trained by another sensor/ensemble and the gathered experience is written into the corresponding self-description as new ExperienceItem. Still open in the research landscape of opportunistic activity recognition are two highly challenging issues: (i) estimating the accuracy for an initially unknown ensemble (by fusing single sensors using the MajorityVoting method [18]) and (ii) optimizing and shaping the ensemble according to the estimated accuracy and the expected costs (in form of energy consumption of the sensor devices); these items compose the core contribution of this paper.

2.3. **Related Work Summary.** Opportunistic sensing (and more specifically opportunistic activity recognition) is a rather
new research area. Particularly, the fact that no labeled and reliable ground truth is available to compute solid metrics to quantify an applications quality is a very challenging characteristic. The work presented in this paper tries to overcome this issue of not being able to reliably predict an ensembles accuracy and use this information subsequently to generate an ensemble that is very likely to be close to optimal in terms of expected accuracy and costs. To the best of our knowledge, these two challenges are novel in the area of sensor-based activity recognition and are identified as being important building blocks towards the vision of opportunistic activity recognition [1].

3. Methodology

3.1. Estimating the Accuracy of an Ensemble at Runtime. The definition of the quality criterion of having the “best” sensors configured in an ensemble (and not simply using all possible sensor candidates from the set \( S_C \), which is a subset of the set \( S \) containing all available sensors) is a nontrivial task. On the one hand, the accuracy of an ensemble for a certain set of activities that have to be recognized has to be maximized, and on the other hand, the energy consumption of the sensor(s) (ensemble) should be as low as possible. Particularly, the first criterion—maximizing the ensemble’s accuracy—is very difficult to predict. Having multiple ARCs from different sensors and the DoF of each, the accuracy of a fusion method that combines these sensors (e.g., Majority Voting fusion) cannot be predicted easily (see [18–21]). Nevertheless, an estimation of the overall ensemble DoF is needed to be able to maximize the ensemble’s performance in terms of accuracy and recognition rate for certain activities. Therefore, we present an approach how the diversity of multiple sensors can be used to approximate the accuracy of sensor combinations. The second criterion—minimizing the energy consumption—is important to ensure a long lifetime of a sensor network and the sensing, respectively infrastructure, which is seen as an equally important criterion. Therefore, we define the term “best ensemble” as being the sensor combination out of the set \( S_C \) that maximizes the estimated ensemble accuracy, whilst minimizing the energy consumption of the sensors in set \( S_C \) (the set of sensors that are included in the ensemble). Further criteria could also be included to provide a more complex definition of “the best ensemble” (e.g., a user-convenience factor or an initial-cost factor), but this is not in the scope of this paper.

Predicting the accuracy (or the degree of fulfillment (DoF)) of an ensemble, which consists of multiple sensors that are combined on the machine learning level by utilizing a decision fusion methodology, can be done in two different ways, respectively, processing stages of a system (see also Table 1):

(i) offline and manual: if a dataset is available that can deliver ground truth activity labels to compare the predicted activity classes from the fusion method with the actual classes and compute the overall accuracy of the ensemble. Obviously, this way of calculating the accuracy of an ensemble does not fit the opportunistic approach, since usually no ground truth and no predefined sensing infrastructure are available in a real-time, open-ended system;

(ii) online and autonomous: at runtime of the system. This fits the opportunistic approach, since neither a ground truth nor the available sensor infrastructure is available or known. Thus, the DoF of an ensemble has to be calculated (or rather estimated) at runtime of the system. This challenging task—constituting one major contribution of this paper—is discussed in the following.

As already described and analyzed in related work literature [19, 21], a correlation between the individual accuracies of the sensors together with their ARCs involved in an ensemble and the diversity between the particular sensors can be observed. This means that the performance of an ensemble is dependent on the accuracies (i.e., the degree of fulfillment (DoF)) and the diversities between the involved sensors. Meynet and Thiran propose in [21] an information-theoretic score (ITS) that expresses exactly this relation. The idea behind this approach is the fact that if two or more sensors have very low mutual information (MI) values (thus have a low diversity), they also agree on wrongly classified activity classes. Therefore, the converse argument comprises that if sensors in an ensemble that have high diversity values can avoid wrongly classified results. Chavarriga et al. follow up this approach in [19]. They use the accuracy of individual classifiers and the available pairwise MI values to calculate an ITS value of an individual sensor. Then, the authors show in simulations that the information-theoretic approach for ensemble creation and recreation outperforms a random method. They subsequently add the classifier to an ensemble that has the highest ITS value. Therefore, the information-theoretic score (ITS) is proven to be a candidate to estimate the accuracy and DoF of an ensemble. Generally, the higher the ITS value of an ensemble is, the higher is the accuracy. This finding is essential and will build the fundamental for further optimization methods within this paper. The MI value between two discrete variables \( X \) and \( Y \) according to Shannon and Weaver [22] and Kraskov et al. [23] is defined as (see also Meynet and Thiran [21]):

\[
\text{MI}(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right). \tag{1}
\]

| Table 1: Overview of possible ways of calculating the degree of fulfillment (DoF) for single sensors and ensembles. |
|---------------------------------|---------------------------------|
| (i) Offline/initial Ground truth available | (i) Offline/initial Ground truth available |
| Easy; calculation of accuracy by validation of initially trained classifier | Easy; calculation of accuracy by validation of initially trained fusion method |
| (ii) Online/runtime Ground truth unavailable | (ii) Online/runtime Ground truth unavailable |
| Difficult because no ground truth for validation is available | Difficult because no ground truth for validation is available |

\[\begin{array}{ll}
\text{Single sensor} & \text{Sensor ensemble} \\
\text{(i) Offline/initial Ground truth available} & \text{Easy; calculation of accuracy by validation of initially trained classifier} \\
\text{(ii) Online/runtime Ground truth unavailable} & \text{Difficult because no ground truth for validation is available} \\
\end{array} \]
According to Meynet and Thiran, the information-theoretic score (ITS) expressing an estimate of the accuracy can be calculated as [21]

\[
\text{ITS} = (1 + \text{ITA}) (1 + \text{ITA})^2 (1 + \text{ITD}).
\]  

(2)

The information-theoretic accuracy (ITA) value is the average accuracy of \(K\) sensors. There, the mutual information (MI) value of a sensor compared to the ground truth (i.e., \(\text{MI}(S_i; S_j)\)) is required, which can be approximated by the system in the same way as the pairwise diversity is calculated (see Figure 1), where an active ensemble with a high DoF is taken as reference for comparisons. The ITA value is calculated as [21]

\[
\text{ITA} = \frac{\sum_{i=1}^{K} \text{MI}(S_i; S_j)}{K}.
\]  

(3)

The second required parameter for computing the ITS value is the information-theoretic diversity (ITD), which describes the average diversity between sensors [21]:

\[
\text{ITD} = \frac{\left(\frac{\sum_{i=1}^{K} \sum_{j=i+1}^{K} \text{MI}(S_i; S_j)}{K(K-1)}\right)}{\sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \text{MI}(S_i; S_j)}.
\]  

(4)

In [19], Chavarriaga et al. had the accuracies for the individual classifiers and the pairwise mutual information values ready within their simulations. This is not the case in an opportunistic system and has to be achieved depending on the current sensing infrastructure and the recognition goal (thus, these MI values have to be calculated in the running system at real time), which is not an easy task and represents a major contribution of this paper. The mutual information (MI) value that is needed for calculation of the ITS value can be calculated as reciprocal of the diversity between two classifiers (sensors). Figure 1 illustrates how the mutual information (MI) value between two sensors can be calculated at runtime. This process is composed of four
The calculation of the mutual information (MI) value between two sensors, the preservation, and continuous update at runtime are key elements to estimate the accuracy of an ensemble at runtime and therefore optimize the ensemble in terms of energy consumption and expected recognition rate. The information-theoretic score (ITS) value of an ensemble can be calculated by utilizing the pairwise diversity and MI values of the sensors by applying the aforementioned equations. Thus, in order to maximize the accuracy, the ITS value of an ensemble has to be maximized. The following Sections 3.2.1 and 3.2.2 present two approaches to optimize the ensemble by utilizing this MI and ITS values that are calculated by the system at runtime. Evaluations of the method to dynamically calculate the diversity values at runtime are presented in Section 4.

3.2. Ensemble Shaping and Optimization. The following two sections present two concepts and methods on how the ensemble DoF can be estimated (upon usage of the Majority Voting fusion [18] method) and how the ensemble can be optimized and shaped with respect to the expected accuracy and the required energy for the sensor devices.

3.2.1. The Knapsack Metaphor for Ensemble Optimization. The knapsack problem is a known combinatorial optimization problem [26–28]. A set of items is given, each with a weight and a value, and the goal is to determine the combination of items that are packed in the knapsack, so that the total weight does not exceed a given maximum allowed weight, and that the sum of the values of the items in the knapsack are maximized. This problem is known to be in the complexity class NP-complete [29], which means that exact solutions can only be computed for very few items. Usually, if a solution for a bigger set of items shall be computed, this is done with mathematical approximations or heuristic algorithms, whereas it cannot be assured to find the optimal solution, but at least a rather good one (examples for such methods are simulated annealing [30], Branch and Bound [31], or Tabu Search methods [32]).

Nevertheless, the knapsack as metaphor for ensemble optimization is a promising approach, at least for a small set of available sensors, and if a maximum allowed total ensemble energy consumption is given. Thus, the degree of fulfillment (DoF) of the individual sensors can be seen as the value, the energy demand as the weight in the traditional definition of the knapsack problem. Then, by searching the complete solution space (i.e., all possible combinations of sensors, which is \(2^n\) for \(n\) sensors), it is possible to find the set of sensors that maximizes the accuracy (by estimating the DoF with the information-theoretic score (ITS) value, see Section 3.1), whilst not exceeding the allowed total energy consumption of the ensemble. If the solution space becomes too big (i.e., too many sensor candidates in the set \(S_C\)), a heuristic solution has to be utilized (e.g., the linear efficiency analysis (LEA) optimization method, see Section 3.2.2).

The following example of a particularly available sensing infrastructure (see Table 2) illustrates the working principle of the knapsack optimization method. In the example, 9
sensors (i.e., $2^9 = 512$ combinations in total in the solution space), the available ITS values (for demonstration purposes, the sensors feature specific ITS values; in an actual system, the pairwise mutual information (MI) values are available and the ITS value for sensor combinations is calculated as shown in the equations in Section 3.1), and the energy demands (as defined in the technical sensor self-description) for each single sensor are given.

The mathematical formulation of the exact knapsack ensemble optimization method is as follows [26–28]:

$$\max \sum_{i=1}^{n} x_i \cdot d_i$$

subject to

$$\sum_{i=1}^{n} x_i \cdot e_i \leq L$$

whereas the following variables/constants are used:

(i) $x_i$ . . . binary variable; 1 if sensor $i$ is in the ensemble, 0 otherwise;

(ii) $d_i$ . . . ITS value of sensor $i$;

(iii) $e_i$ . . . energy demand of sensor $i$;

(iv) $L$ . . . allowed energy limit of ensemble.

If, for example, the maximum allowed energy consumption of the resulting ensemble is set to 6.5 W (this could be defined in the goal description, where a parameter defining the maximum allowed energy value is provided), the method considers all possible combinations (in this case 512) and delivers the optimal solution: $\{S_4, S_5, S_7, S_8\}$, with a total information-theoretic score (ITS) value of 42 (again for demonstration purposes, the example is shown in Table 2 feature ITS values, and the method simply sums them up; in an actual system, the pairwise mutual information (MI) values are available and the ITS value for sensor combinations is calculated as shown in Section 3.1) and a total energy consumption of 6.4 W. Listing 1 displays the console-output for the exemplary sensor setting as defined in Table 2.

This method is feasible and applicable if two prerequisites are given:

(i) an upper limit for the total allowed energy consumption in the ensemble is given,

(ii) the number of ensemble candidates does not exceed approximately 10 sensors, since the ensemble definition and configuration has to be done in real-time.

In contrast to the knapsack optimization method, the linear efficiency analysis (LEA) method (see Section 3.2.2) needs neither a defined energy constraint, nor a number of sensors limitation.

### 3.2.2. The Linear Efficiency Analysis for Ensemble Optimization

The knapsack metaphor for ensemble shaping is an exact optimization method, which suffers from two major disadvantages: (i) the allowed total energy consumption has to be defined and provided to the system and (ii) the scalability (i.e., combinatorial solution) of the method is limited due to the real-time operability requirements by approximately 10 sensors. Therefore, heuristic solutions that do not expand the complete solution space (thus are not ensuring to find the optimal solution, but an acceptable one) have to be selected for ensemble optimization. The linear efficiency analysis (LEA) method operates on the efficiency of the sensor candidates that can be calculated by dividing the ITS value by the energy demand of the sensor. The approach operates in $O(n^2)$ (where $n$ is the number of sensor candidates), by simply starting from an initial ensemble that contains all sensors and subsequently removing the one with the worst efficiency. The occurring ensembles (that are constructed until only one sensor is left) are examined with respect to their efficiency. As a result, the ensemble with the highest efficiency is calculated and configured. The LEA method consists of three steps: (i) initialization, (ii) ensemble reduction, and (iii) resulting ensemble. Table 3 shows an example of a sensor infrastructure (again, for demonstration purposes).

| Sensor ID | ITS value | Energy (Watt) |
|-----------|-----------|---------------|
| S1        | 14        | 3.7 W         |
| S2        | 18        | 4.2 W         |
| S3        | 10        | 2.0 W         |
| S4        | 16        | 2.4 W         |
| S5        | 12        | 1.4 W         |
| S6        | 3         | 0.8 W         |
| S7        | 4         | 0.8 W         |
| S8        | 10        | 1.8 W         |
| S9        | 12        | 1.9 W         |
purposes, the sensors feature specific ITS values; in an actual system, the pairwise mutual information (MI) values are available and the ITS value for sensor combinations is calculated as shown in the equations in Section 3.1), and in the following Figure 2 and paragraph, the working principle of the LEA optimization is described (and demonstrated by application on the example shown in Table 3).

In the first phase the (i.e., initialization), the starting ensemble consisting of all available sensor candidates is calculated and stored in a list \( n \) times. Furthermore, the efficiencies for each single sensor candidate are calculated (see Table 3, column sensor efficiency). These values are important in the second phase (i.e., ensemble reduction), where in each step, the sensor with the lowest efficiency is subsequently removed from the ensemble. In this exemplary setting, the “worst” sensor is \( S_2 \), followed by \( S_1 \), and \( S_3 \). The \( n-1 \) worst sensors are removed step by step. For each resulting possible ensemble, the total efficiency is calculated (for that reason it is important to start with the initial ensemble that is stored in a data structure \( n \) times) and multiplied by the actual number of sensors and the ensemble with the highest efficiency is delivered as a result. The processing steps in the ensemble reduction phase in this example are listed in Table 4 (where \( E_9 \) stands for Ensemble_9).

Therefore, the resulting ensemble in this example would be \( \{S_1, S_2, S_3, S_4\} \) with an efficiency value of 554.60. This very simple approach can be further improved by considering a more sophisticated calculation of the ensemble ITS values (in this example, they have been simply added), and by adding further constraints (e.g., minimal number of required sensors in ensemble, etc.). Even a limit for the ensemble energy consumption can be provided (much like it is necessary in the knapsack solution method), then the expanded ensembles have to be examined to find the resulting sensor setting. Next Section 4 presents evaluations of the introduced methods in a real-time system operating in an opportunistic way.

4. Evaluation

The methods for estimating the ensembles accuracy based on Majority Voting fusion (i.e., the activity class where most of the classifiers agree on is selected as fusion result) [18] and the ensemble optimization and shaping are tested and evaluated on a real-world dataset. In this rich dataset, recorded in a kitchen scenario, 12 subjects performed activities for about 14 minutes per run (whereas each subject performed 6 runs) surrounded by 72 sensors with 10 different modalities. The dataset has been labeled to have a reliable ground truth available for evaluation purposes (i.e., the modes of locomotion have been added to the dataset). Details about the recordings and the dataset can be found in [4, 5]. The OPPORTUNITY framework enables the utilization of the sensors in the dataset as if the sensors were currently present by abstracting them in form of PlaybackSensors [1, 15]. This means that, the framework performs real-time activity recognition with the sensors that are simulated by replaying the prerecorded sensor data. As this paper aims at presenting and evaluating methods (i) for estimating the ensemble accuracy by utilizing the pairwise diversity between sensors and (ii) for optimizing an ensemble in terms of this estimated accuracy and the expected energy demand, a

---

**Table 3: Exemplary sensor infrastructure as input for the linear efficiency analysis optimization method.**

| Sensor ID | ITS value | Energy (Watt) | Sensor efficiency |
|-----------|-----------|---------------|------------------|
| \( S_1 \) | 21        | 0.37 W        | (21/0.37) = 56.76 |
| \( S_2 \) | 39        | 0.37 W        | (39/0.37) = 105.41 |
| \( S_3 \) | 78        | 0.54 W        | (78/0.54) = 144.44 |
| \( S_4 \) | 88        | 0.35 W        | (88/0.35) = 251.43 |
| \( S_5 \) | 15        | 3.20 W        | (15/3.2) = 4.69  |

---

**Table 4: Step-by-step working principle of the LEA optimization method.**

| Ensemble ID | Sensors | ITS | Energy (Watt) | Ensemble efficiency |
|-------------|---------|-----|---------------|---------------------|
| \( E_0 \)  | \( \{S_1, S_2, S_3, S_4\} \) | 241 | 4.83          | 249.48              |
| \( E_1 \)  | \( \{S_1, S_2, S_3\} \) | 226 | 1.63          | 554.60              |
| \( E_2 \)  | \( \{S_2, S_3, S_4\} \) | 205 | 1.26          | 488.10              |
| \( E_3 \)  | \( \{S_1, S_4\} \) | 166 | 0.89          | 373.03              |
| \( E_4 \)  | \( \{S_4\} \) | 88  | 0.35          | 251.43              |
subset of the available sensors is picked for recognizing the modes of locomotion (i.e., WALK, SIT, STAND, and LIE) as activity of interest. The sensors that are used for evaluation purposes are located on the body of the subject. In detail, the upper body is equipped with 5 sensors of type Xsens MTx, which provides 3D acceleration. The sensors are located on both arms (right/left lower/upper arm—RLA, RUA, LLA, and LUA), and on the upper back (BAC). The on-body sensor placement is displayed in Figure 3 (see also [17, 24]).

The first method evaluation tackles the dynamic generation of the diversity measurements. Therefore, 10 independent runs involving the five on-body sensor have been conducted, whereas every 300 samples (the number of samples per calculation iteration is kept intentionally low due to scalability reasons) the method described in Section 3.1 calculated the pairwise diversity value for all 10 possible combinations and stored this value (by utilizing the cumulative moving average) into the corresponding sensor self-description. Listing 2 shows a clipping of the self-description for one sensor (BAC) and its pairwise diversity value with another sensor (i.e., LUA).

The pairwise diversity value perfectly converges towards a value by utilizing the moving average value, which can be further used to define the information-theoretic score (ITS) value which can be used to estimate accuracies for the resulting ensembles, as described in Section 3.1. Figure 4 shows a graphical visualization for 280 iterations (the two top-level images, boxplot and plot chart), and 2500 iterations (lower-level images). The figure displays the pairwise diversity values for each of the 10 possible combinations (i.e., BAC-LLA, BAC-LUA, BAC-RLA, BAC-RUA, LLA-RLA, LUA-LLA, LUA-RLA, LUA-RUA, RUA-LLA, and RUA-RLA).

For the evaluation of the optimization methods, the ongoing dynamic updates of the diversity measures have been stopped after approximately 2500 iterations to have a reliable (and comparable) sensor self-description base for comparing the two optimization methods. To be able to compare the optimization techniques and to highlight the benefits in terms of optimizing towards reduced energy consumption and maximized accuracy, one session (lasting approx. 14 min, involving the 5 on-body sensors described above) was conducted by combining all available sensors into an ensemble. The results for this session are shown as confusion matrices in Figure 5 (top left illustration), in Figure 6 (illustrating the accuracies and the efficiencies) and in Table 5 (method ALL).

One major benefit of using a prerecorded dataset is the fact that a labeled ground truth is available which makes the
calculation of the actual accuracy possible. The accuracy (i.e.,
telling how many of the predicted activity classes are correctly
classified, manifested as DoF value in the self-description) is the most important metric that quantifies the quality of a classifier or an activity recognition application. The accuracies for the four ensembles that have been generated by utilizing the optimization methodologies can be calculated out of the confusion matrices that display the number of correctly classified classes (the main diagonal of the matrix) and the false positives/negatives. The two optimization methods as described in Sections 3.2.1 (knapsack optimization) and 3.2.2 (linear efficiency analysis (LEA)) have been utilized in equal playback sessions within the dataset [4, 5] used by the system as PlaybackSensors (simulating sensors by using prerecorded data). The knapsack method was subject to evaluation twice, since the method needs a specific maximum allowed energy boundary parameter (therefore, this value was set to 5 W in the first run and 7.5 W in the second run). The results for the four sessions are displayed in the form of confusion matrices in Figure 5, whereas the top left matrix shows the results for the ensemble consisting of all five available sensor (i.e., BAC, LUA, LLA, RUA, RLA); the top right matrix shows the classification result for the LEA optimization methods, where the method configured an ensemble consisting of only one sensor (i.e., RLA); the two lower images show results for the knapsack optimization methods with different maximum allowed energy boundaries (lower left has an energy boundary of 7.5 W and an ensemble consisting of the BAC and RLA sensors; lower right has an energy boundary of 5 W and an ensemble consisting of the LLA sensor). The results in form of the accuracies and efficiencies are summarized in Table 5, where the accuracies are calculated out of the confusion matrices for all four evaluation ensembles. This summary highlights the benefits of optimizing ensembles instead of using all available and suitable sensors.

Table 5: Overview of the evaluation result for the three different methods (the knapsack optimization method was started twice: one run with 5 W as boundary for the energy consumptions and one run with 7.5 W).  

| Method | Ensemble | Energy (Watt) | Accuracy | Efficiency |
|--------|---------|---------------|----------|------------|
| ALL    | BAC, LUA, LLA, RUA, RLA | 13.5 W | 0.759 | 0.056 |
| LEA    | RLA     | 2.7 W | 0.641 | 0.237 |
| KS5    | LLA     | 2.7 W | 0.561 | 0.208 |
| KS7.5  | BAC, RLA | 5.4 W | 0.680 | 0.126 |

An additional evaluation of the two ensemble optimization methods is done without the dataset. A set of ExperienceItems is artificially generated, including a random value for the DoF of the activity WALK (between 0.0 and 1.0), a random value for the pairwise MI values (between 0.0 and 4.9, which is a realistic approximation), and a randomly generated energy consumption (between 0.5 W and 4.9 W) per sensor. Both methods are applied within the framework based on the set of constructed ExperienceItems, where the focus is on the aggregated energy consumption of the sensors selected for the ensemble and the ITS value per ensemble. As already discussed, the “best” ensemble minimizes the energy consumption whilst maximizing the accuracy. The information-theoretic score (ITS) value is used for estimating the expected accuracy; the higher the ITS value of an ensemble is, the higher is the expected accuracy. Both methods are compared against a random ensemble generation method and the selection of all candidate sensors (i.e., “select all” method). The results are illustrated in Figure 7.

The upper two illustrations in Figure 7 visualize the knapsack ensemble optimization method. On the left-hand side, the energy consumption for the method compared to a random and “select all” method is illustrated. The knapsack method needs an upper bound for the maximum allowed energy consumption (which is set to 15 W in this run); therefore, the right-hand side illustration (showing the ensemble ITS value in relation to the number of sensors in the candidate set) is more meaningful. In most of the cases, the knapsack method outperforms the random and “select all” methods in terms of the reachable ITS value. The two lower images show the results for the LEA ensemble optimization method for up to 40 sensors in the candidate set. As shown on the left-hand side, the energy consumption level of LEA is constantly kept low. The most impressive and convincing visualization on the lower right-hand side shows the calculated ITS values for the ensembles in relation to the

Figure 6 visualizes the calculated accuracies (gray bars) and the efficiency (black bars) of the four methods. As it can be seen, utilizing all sensors in the ensemble results in the highest accuracy (resp., degree of fulfillment (DoF) value), but has the lowest efficiency value. This is obvious, since the ensemble consisting of all sensors has the highest energy consumption. The linear efficiency analysis (LEA) method performs best in this experiment session in terms of efficiency. An interesting issue is that the knapsack method (configured with 5 W upper boundary) selected the LLA sensor (resulting in an accuracy of 0.561), and the LEA method selected the RLA sensor (resulting in an accuracy of 0.641). The reason for this discrepancy is the fact that both sensors have the same initial DoF value in their corresponding self-description, so they are equal in terms of energy consumption, expected accuracy, and thus also efficiency. Coincidentally, one method selected the RLA sensor, the other method the LLA sensor. The different values for the accuracy are caused by the fact that these accuracies are the actual ones, calculated by comparing the predicted classes to the actual classes from the available ground truth within the dataset.

The ensemble consisting of all sensors has an energy demand of 13.5 W (2.7 W for each single sensor) and reaches an accuracy for the modes of locomotion of 0.759. Compared to that, the ensemble generated by using the LEA method consists of only one sensor (i.e., RLA); therefore, it needs only 2.7 W but reaches an impressive accuracy of 0.641. The same holds for the knapsack optimization ensemble, where the accuracies are also very close to the original ensemble consisting of all sensors, but demand much less energy.
number of sensors in the candidate set (the random method selects approx. 50% of the available candidates; the "select all" method combines all available candidates into the ensemble). It can be safely said that the LEA method outperforms both naïve approaches in terms of energy consumption and expected accuracy.

5. Conclusion

The work presented in this paper tackles two challenging features in an opportunistic activity and context recognition system: (i) the estimation of the accuracy of a sensor ensemble without relying on a labeled ground truth and (ii) optimizing the ensemble (resp., the involved sensor systems) in terms of accuracy and expected energy consumption. Regarding the first research challenge, the method proposed in this paper utilizes the mutual information (i.e., the diversity) between two sensors to calculate an information-theoretic score. The basic idea is that if two sensors have a high diversity value, it is very unlikely that they do not agree on wrongly classified activities, thus resulting in a higher accuracy if such sensors are utilized together in an ensemble. This correlation can be broadened to the available sensors resulting in an estimate of the ensemble accuracy for a specific recognition goal in form of the degree of fulfillment (DoF). One major contribution presented in this paper is the ability of calculating this diversity value at system’s runtime which results in a stable value for sensor combinations and can be persistently stored in the corresponding self-descriptions. Thus, the sensor (and the entire system) learns at runtime and is able to enhance its capabilities.

Optimizing and shaping sensor ensembles is the second major contribution of this paper. Particularly in large-scale
systems, utilizing all available sensors for activity recognition might not necessarily improve the recognition capabilities. However, the lifetime of the system is minimized since sensors can run out of energy. Therefore, optimizing the ensemble in terms of estimated accuracy and expected energy consumption is being identified as important aspect in an opportunistic system. Two methods are introduced: (i) the knapsack as metaphor for ensemble optimization and (ii) the linear efficiency analysis (LEA) method. The two solutions differ on the one hand in the required input parameters (i.e., the knapsack needs a maximum allowed energy consumption as boundary) and on the other hand in the algorithmic complexity. The knapsack expands the complete solution space (i.e., $O(2^n)$), which results in an optimal solution. The LEA method runs in $O(n^2)$ time and does not need a boundary for the allowed energy consumption. Thus, the solution is not necessarily an optimal solution (but good enough, as the results show).

These two challenging features are identified as being important building blocks that still have mission in the vision of opportunistic activity recognition. The results of the evaluation of the presented methods are promising,
Figure 7: Illustration of the analysis of the knapsack and LEA ensemble optimization method compared to a random and “select all” method.

which means that the methods will be further expanded and investigated.

Acknowledgments

The project OPPORTUNITY acknowledges the financial support of the Future and Emerging Technologies (FET) Programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open Grant no. 225938.

References

[1] M. Kurz, G. Hölzl, A. Ferscha et al., “The opportunity framework and data processing ecosystem for opportunistic activity and context recognition,” International Journal of Sensors, Wireless Communications and Control, vol. 1, no. 2, pp. 102–125, 2011.

[2] D. Roggen, K. Förster, A. Calatroni et al., “Opportunity: towards opportunistic activity and context recognition systems,” in Proceedings of the 3rd IEEE WoWMoM Workshop on Autonomic and Opportunistic Communications (AOC ’09), IEEE CS Press, Kos, Greece, June 2009.

[3] G. Hölzl, M. Kurz, and A. Ferscha, “Goal oriented opportunistic recognition of high-level composed activities using dynamically configured hidden markov models,” in Proceedings of the 3rd International Conference on Ambient Systems, Networks and Technologies (ANT ’12), Niagara Falls, Ontario, Canada, August 2012.

[4] D. Roggen, A. Calatroni, M. Rossi et al., “Collecting complex activity data sets in highly rich networked sensor environments,” in Proceedings of the 7th International Conference on Networked Sensing Systems (INSS ’10), IEEE Computer Society Press, Kassel, Germany, June 2010.
[5] D. Roggen, A. Calatroni, M. Rossi et al., *Walk-Through the Opportunity Dataset For Activity Recognition in Sensor Rich Environments*, 2010.

[6] D. Salber, A. K. Dey, and G. D. Abowd, “The context toolkit: aiding the development of context-enabled applications,” in *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit (CHI ’99)*, pp. 434–441, ACM, New York, NY, USA, 1999.

[7] L. Bao and S. Intille, “Activity recognition from user-annotated acceleration data,” in *Pervasive Computing*, A. Ferscha and F. Mattern, Eds., vol. 3001, pp. 1–17, Springer, Berlin, Germany, 2004, Lecture Notes in Computer Science.

[8] H. Junker, O. Amft, P. Lukowicz, and G. Troster, “Gesture spotting with bodyworn inertial sensors to detect user activities,” *Pattern Recognition*, vol. 41, no. 6, pp. 2010–2024, 2008.

[9] P. Lukowicz, J. Ward, H. Junker et al., “Recognizing workshop activity using body worn microphones and accelerometers,” in *Pervasive Computing*, A. Ferscha and F. Mattern, Eds., vol. 3001, pp. 18–32, Springer, Berlin, Germany, 2004, Lecture Notes in Computer Science.

[10] J. A. Ward, P. Lukowicz, and H. W. Gellersen, “Performance metrics for activity recognition,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 1, pp. 1–6, 2011.

[11] E. Tapia, S. Intille, and K. Larson, “Activity recognition in the home using simple and ubiquitous sensors,” in *Pervasive Computing*, vol. 3001, pp. 158–175, Springer, Berlin, Germany, 2004, Lecture Notes in Computer Science.

[12] B. Logan, J. Healey, M. Philiope, E. M. Tapia, and S. Intille, “A long-term evaluation of sensing modalities for activity recognition,” in *Proceedings of the 9th international conference on Ubiquitous computing (UbiComp ’07)*, pp. 483–500, Springer, Berlin, Germany, 2007.

[13] T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, “Accurate activity recognition in a home setting,” in *Proceedings of the 10th international conference on Ubiquitous computing (UbiComp ’08)*, pp. 1–9, ACM, New York, NY, USA, 2008.

[14] J. Mantyjarvi, J. Himberg, and T. Seppanen, “Recognizing human motion with multiple acceleration sensors,” in *IEEE International Conference on Systems, Man, and Cybernetics (SMC ’01)*, vol. 22, pp. 747–752, 2001.

[15] M. Kurz and A. Ferscha, “Sensor abstractions for opportunistic activity and context recognition systems,” in *Proceedings of the 5th European Conference on Smart Sensing and Context (EuroSSC ’10)*, K. G. Lukowicz and P. Kunze, Eds., Springer, Berlin, Germany, November 2010.

[16] M. Kurz, G. Höflzl, A. Ferscha, H. Sagha, J. R. Millan, and R. Chavarriaga, “Dynamic quantification of activity recognition capabilities in opportunistic systems,” in *Proceedings of the 4th Conference on Context Awareness for Proactive Systems (CAPS ’11)*, pp. 15–16, Budapest, Hungary, May 2011.

[17] M. Kurz, G. Höflzl, A. Ferscha, A. Calatroni, D. Roggen, and G. Troester, “Realtime transfer and evaluation of activity recognition capabilities in an opportunistic system,” in *Proceeding of the 3rd International Conference on Adaptive and Self-Adaptive Systems and Applications (ADAPTIVE ’11)*, pp. 73–78, Rome, Italy, September 2011.

[18] J. Kittler, M. Hatef, and R. Duin, “On combining classifiers,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, 1998.

[19] R. Chavarriaga, H. Sagha, and J. d. R. Millan, “Ensemble creation and reconfiguration for activity recognition: an information theoretic approach,” in *IEEE International Conference on Systems, Man and Cybernetics (IEEE SMC ’11)*, 2011.

[20] L. I. Kancheva and C. J. Whitaker, “Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy,” *Machine Learning*, vol. 51, pp. 181–207, 2003.

[21] J. Meynet and J. P. Thiran, “Information theoretic combination of pattern classifiers,” *Pattern Recognition*, vol. 43, no. 10, pp. 3412–3421, 2010.

[22] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, Ill, USA, 1949.

[23] A. Kraskov, H. Stögbauer, R. G. Andrzejak, and P. Grassberger, “Hierarchical clustering using mutual information,” *EPL*, vol. 70, no. 2, p. 278, 2005.

[24] M. Kurz, G. Höflzl, and A. Ferscha, “Dynamic adaptation of opportunistic sensor configurations for continuous and accurate activity recognition,” in *Proceedings of the 4th International Conference on Adaptive and Self-Adaptive Systems and Applications (ADAPTIVE ’12)*, pp. 22–27, Nice, France, July 2012.

[25] Y. I. Chou, *Statistical Analysis, with Business and Economic Applications*, Holt, Rinehart and Winston, New York, NY, USA, 2nd edition, 1975.

[26] S. Martello, *Knapsack Problems: Algorithms and Computer Implementations*, John Wiley & Sons, New York, NY, USA, 1990.

[27] J. Puchinger, G. Raidl, U. Pferschy, and editors, “The core concept for the multidimensional knapsack problem,” in *Evolutionary Computation in Combinatorial Optimization*, J. Gottlieb and G. Raidl, Eds., vol. 3906, pp. 195–208, Springer, Berlin, Germany, 2006, Lecture Notes in Computer Science.

[28] G. Zápeř, R. Braune, and M. Bögl, *Metaheuristic Search Concepts: A Tutorial with Applications to Production and Logistics*, Springer, Berlin, Germany, 2010.

[29] M. R. Garey and D. S. Johnson, *Computers and Intractability, A Guide To the Theory of NP-Completeness*, W. H. Freeman, New York, NY, USA, 1990.

[30] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[31] E. L. Lawler and D. E. Wood, “Branch-and-bound methods: a survey,” *Operations Research*, vol. 14, no. 4, pp. 699–719, 1966.

[32] F. Glover and M. Laguna, *Tabu Search*, Kluwer Academic Publishers, Norwell, Mass, USA, 1997.
