Report

A methodology to use unknown new sensors for activity recognition by leveraging sporadic interactions with primitive sensors and behavioral assumptions

Author(s):
Calatrani, Alberto; Roggen, Daniel; Tröster, Gerhard

Publication Date:
2010

Permanent Link:
https://doi.org/10.3929/ethz-a-006286309

Rights / License:
In Copyright - Non-Commercial Use Permitted
A methodology to use unknown new sensors for activity recognition by leveraging sporadic interactions with primitive sensors and behavioral assumptions

Alberto Calatroni  
Wearable Computing Laboratory  
ETH Zurich, Zurich, Switzerland  
alberto.calatroni@ife.ee.ethz.ch

Daniel Roggen  
Wearable Computing Laboratory  
ETH Zurich, Zurich, Switzerland  
droggen@ife.ee.ethz.ch

Gerhard Tröster  
Wearable Computing Laboratory  
ETH Zurich, Zurich, Switzerland  
troester@ife.ee.ethz.ch

ABSTRACT
In the perspective of activity recognition systems operating for long periods of time in environments susceptible of upgrades, one question that arises is how to exploit a-priori unknown newly discovered sensors for activity recognition. We present a methodology to exploit these unknown new sensors. This methodology uses sporadic interactions with primitive sensors together with behavioral assumptions to confer activity recognition capabilities to a newly discovered sensor. The behavioral assumptions are used to infer additional information from the primitive sensors (e.g. simple reed switches) in ways that go beyond their initially foreseen function (e.g. detecting walking). We explain the methodology on the example of learning how to use an unknown new on-body sensor to detect modes of locomotion, when the user sporadically interacts with instrumented furniture.

Author Keywords
Activity recognition, Transfer learning, Behavioral assumptions

ACM Classification Keywords
I.2.6 Artificial Intelligence: Learning; I.5.2 Pattern Recognition: Design Methodology

General Terms
Algorithms, Experimentation, Performance

PROBLEM STATEMENT
Ambient intelligence environments (AmI) are operating for long periods of time (months to years or more). During this time they are likely going to undergo upgrades, such as being enhanced by new sensors. In the same way, as time goes by, a user may buy new smart items (e.g. mobile phones, shoes) or garments which may contain new sensors.

In both cases, one common research question is how to exploit this a-priori unknown newly discovered sensors for activity recognition. This is a key problem when designing an opportunistic activity recognition system, where the system ought to make best use of any kind of available resources.

An additional challenge is that in the general case, the functionality (e.g. modality, location) of the sensor is not predictable. This unpredictability comes from the fact that different players may provide context-aware applications and new sensors (e.g. building automation company or shopping spree of the user v.s. context-aware application designer). Therefore, mechanisms are required to be able to exploit this new information, in these challenging open ended environments.

In this paper we present a new methodology to exploit newly discovered sensors for activity recognition. This methodology assumes sporadic user interactions with primitive pre-existing sensors (e.g. simple reed switches). The information collected from these primitive sensors is enhanced by behavioral assumptions to infer activities other than the ones foreseen by the system designer (e.g. a reed switch is only designed to detect opening and closing of windows or doors in a home automation system). By behavioral assumptions we mean hypotheses about likely human behavior, for example somebody is going to be standing while opening a window, while she will be walking towards it and away from it shortly before and after her interaction with the window.

This combination of sporadic interactions with pre-existing sensors, extended by the behavioral assumptions, can be used to train an unknown, newly discovered on-body sensor. Eventually, the on-body sensor will then gain the ability to recognize the activities learned from these interactions, even when the user moves to a totally “AmI-free” environment.

RELATED WORK
The traditional approach to use a new sensor for activity recognition consists in collecting a training dataset, and to
train a new classifier offline [1].

Another approach consists in sporadically collecting labels from the user to reduce the initial training overhead - so-called experience sampling [10].

Further reducing user involvement, transfer learning attempts to exploit knowledge in an old domain (e.g. pre-existing sensors with their features and classifier models) and apply it to a new domain where only a few labelled data exist [4].

One common limitation of these approaches is the need for at least some labels of activities when the sensors is discovered. This is cumbersome for the user. Also, transfer learning approaches need to have identical feature spaces, or need to learn a mapping between old and new feature spaces [4]. This is in many cases not feasible if the sensor domains are too different or if the feature spaces are not accessible.

**APPROACH**

The approach we propose to address the limitations of the state of the art consists of:

- **Sporadic sources of labels**: we assume that the system is receiving occasionally labels as a result of the user’s interactions.

- **Enhancement of label richness by behavioral assumptions**: in many cases, the information provided by the sporadic sources of labels is limited, but can be extended by taking into account human behaviors that are likely to occur.

- **Derived sensor data**: many sensors can provide information that they were initially not designed to provide and that can be retrieved after simple processing. Essentially, this also provides a source of (sporadic) labels.

- **Transfer learning based on sporadic labels**: incremental learning algorithms can be used to learn the mapping between the signals measured on the newly discovered sensors, and the activity classes indicated by the labels.

**METHODOLOGY**

In this section we propose a methodology to explore the design aspects to use unknown new sensors, following the approach outlined above.

**Sporadic interactions with systems capable of delivering labels**

A first aspect consists in investigating the availability of sensors that can become label sources. This should be done taking into account the future large-scale availability of ambient intelligence environments. Essentially, the outcome consists of an exhaustive list of these resources along with the labels that they can provide. These information sources may be classified in:

- **Primitive ambient sensors**: pre-existing set of sensors, typically of modalities such as magnetic switches or RFID tags, are deployed as part of a the initial ambient intelligence environment. As the user interacts with the environment, these sensors can provide information about her actions.

- **Software sensors**: knowledge about the user’s activities available from software sources such as calendars can also provide labels.

**Enhancement of label richness by behavioral assumptions**

The second step consists in investigating for each listed source of labels, if there are typical patterns of human behavior that are recurring whenever a label is provided. The result may be also be presented in the form of a table.

For example, when a reed switch in a drawer is activated, we can infer that the user is performing the actions ‘open drawer’ or ‘close drawer’. The behavioral assumption here could be that the user is with high probability standing while operating that drawer. Furthermore, the user will likely be walking before and after interacting with the drawer (see fig 1).

We note that some assumptions may be “trained”, following a statistical approach, or “taken for granted” through common knowledge. In any case, the goal is to make these assumptions as robust as possible, or to be able to indicate the confidence with which labels are provided, which could be exploited afterwards in the transfer learning.

**Derived sensor data characteristics**

Activities can also be inferred by a simple processing of some sensor data. For example, from a sequence of frames grabbed by a vision system (after the user sporadically passes in front of a camera), a tracking algorithm can deduce the speed at which a person is moving, thereby detecting if the person is walking, running or having a static posture. The same kind of information can be obtained by building motion vectors from successive locations read from a positioning system like Ubisense. This system is not one with which...
**Table 1. Illustration of a possible survey of the source of sporadic labels in an ambient intelligence environment, their foreseen usage, and the possible alternate usage in an opportunistic scenario.**

<table>
<thead>
<tr>
<th>Activity kind</th>
<th>What</th>
<th>Typical usage</th>
<th>Common knowledge / derived properties</th>
<th>Opportunistic exploitation</th>
<th>Act. kind</th>
<th>How</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reed or normal switches</td>
<td>Detection of opening and closing of items (e.g. door, drawer) or activation (e.g. light switch)</td>
<td>User performing a opening or closing gesture or switch toggling</td>
<td>2-class problem: on-body sensors can learn to detect opening or closing</td>
<td>G</td>
<td>BA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Opening or closing implies a reach gesture before and a release gesture afterwards</td>
<td>1-vs-null-class problem: on-body sensors can learn to spot reaching for items</td>
<td>G</td>
<td>BA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lock/unlock shortly before the open, but after the reach (reach - unlock - open - release) or (reach - close - lock - release)</td>
<td>On-body sensors can learn to detect standing v.s. not-standing (typically other modes of locomotion), or even standing v.s. walking v.s. other activities</td>
<td>MOL</td>
<td>BA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>While interacting, the user is very likely standing. Before and after interaction with item, the user is likely to walked or (typically other modes of locomotion), or even standing v.s. walking v.s. other activities.</td>
<td></td>
<td>MOL</td>
<td>BA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Localization system</td>
<td>Position of the person in the room. Infer prior probabilities of activities (e.g. near sink - higher prior on washing dishes), or directly activities (if very localized)</td>
<td>With additional assumption on the speed of displacement, walking and running can be differentiated by their speed</td>
<td>MOL</td>
<td>DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rate of variation of position can be mapped to walking v.s. static posture</td>
<td>On-body sensors can learn to recognize walking</td>
<td>MOL</td>
<td>DC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proximity infra-red</td>
<td>Movement detection likely indicates a walking person.</td>
<td>On-body sensor can learn to detect walking</td>
<td>MOL</td>
<td>DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Camera</td>
<td>Person identification tracking, speed of displacement, walking, or running</td>
<td>On-body sensors can learn to recognize walking, running vs static posture.</td>
<td>MOL</td>
<td>DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instrumented objects with e.g. RFID, acceleration.</td>
<td>Object use [7]</td>
<td>On-body sensor can learn to detect the same activities</td>
<td>G</td>
<td>DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The user is moving or carrying or using the object in the hand</td>
<td>Training hand gesture recognition system to recognize carrying/moving objects</td>
<td>G</td>
<td>BA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Shortly before the object is moved, the user must perform a reach gesture. Shortly after the object is left static, the user must perform a release gesture</td>
<td>On-body sensors can learn to spot reaching for objects</td>
<td>G</td>
<td>BA</td>
</tr>
</tbody>
</table>

**PRELIMINARY RESULTS**

We consider using a newly discovered sensor on-body to evaluate reach and locomotion. A large dataset contains also numerous instances of interactions between the user and doors and drawers, which were all instrumented with sensors. Such an approach has so far been validated on a large activity dataset containing natural interactions of a large number of subjects with their environment, with an embedded sensor for bread, move to bread cutter, operate bread cutter. The dataset consists of temporally unfolding situations and in each situation the user is operating in a naturalistic environment. Suitable datasets may include e.g. [7] or [8].

**Evaluation in rich scenario**

We consider using a newly discovered sensor on-body to track the user's interactions with the environment. The procedure works as follows: The sensor learns the activity recognition sensor network and its time processing activity recognition sensor network and its time label along with time stamps each time that a label is recognized. For example, each time that a door is being opened, the corresponding magnetic sensor sends a label with start and end times of the opening sensor. Upon receipt of interest label, the body-worn sensor processes the instance of interest in its signal buffer, and subsequently incorporates the instance of interest into account confidence values. As an example, weighted KNN Classifier.

**Transfer learning based on sporadic labels**

The labels (possibly sporadic) obtained in the previous steps may include e.g. [7] or [8]. Transfer learning can be used to initiate a transfer learning process. We can use the transfer learning algorithm to achieve similar kind of information.

Again, these may be listed in table form.
recognize modes of locomotion. The user is equipped with Ubisense tags and interacts with three drawers, a dishwasher and a fridge. Those are instrumented with reed switches that indicate when they are opened/closed. The behavioral assumption is that when the user opens/closes the item, she stands, and that she walks shortly before and after (see Figure 2). This corresponds to line 4 in table 1.

We characterize the suitability of the combination of sporadic labels and behavioral assumptions - according to the parameters of the assumptions - at correctly providing the relevant labels. We use the following measures for characterization:

- **Degree of “purity”** of the labels generated by the system. This is the proportion of time for which the generated label correctly corresponds to the ground truth. Higher values indicate that the training data include only the correct activity. Referring to Figure 2, the “walk” label would have 100% purity (see regions labeled a and e), whereas the “stand” label would have a lower purity, calculated as \( \frac{c}{\delta_S} \).

- **“Accuracy”** of the labels generated by the system. Proportion of time in which there is an agreement between system generated labels and ground truth. In Figure 2 it would be \( \frac{a + c + c}{a + b + c + d + e} \).

- **Visual inspection** of the label correspondence, to gain a better understanding of the system behavior.

We run simulations using different parameter sets to test the validity and the limits of the approach. In Figure 3 we see the purity of the “stand” label as we sweep through the possible lengths of the generated label (\( \delta_S \)). Indeed, for short lengths we can obtain very pure labels, which is very useful for the successive training that is needed for the on-body sensor to learn the activity mapping. The same evaluation on the accuracy gives much lower values (around 30%), due to the inevitably numerous areas where “standing” is not detected just because the user is not interacting with the environment.

When we also include the “walk” label, the overall purity decreases, because the assumption about walking cannot be as precise as the one about standing. Nevertheless, we see in Figure 4 that a good parameter choice can still provide a reasonable combined purity (around 65% in this example).

**Analysis**

In the previous section we showed the purity values that are achievable. A closer look to the underlying mechanisms is needed to gain a deeper understanding about how the assumptions can be used, when they work and how they could be improved. In Figure 5 we see a portion of generated labels, where the ground truth is compared to what is generated by behavioral assumptions from the reed switches (second line) and from derived sensor characteristics (processing of the Ubisense data). It appears clear how the assumption that the user is standing (grey) while the reed switches fire is very reasonable. This is true both in case of a totally isolated interaction (e.g. user opening a door and then going away) and in case of multiple, repeated interactions with items (e.g. user opening and closing many drawers one after the other). The assumption about the user walking (white) before and after the interaction is much more critical. While it can be reasonably correct in case of an isolated interaction, it is only well approximating reality at the beginning and at the end of a series of short interactions with different items, but not at the boundaries of each short interaction (the
user is always standing, while the system predicts instances of walking as well). On the other hand, using information from the Ubisense (derived characteristics) is very useful to detect walking. Therefore, the most promising approach will be to fuse in a clever way information coming from different sources. The Ubisense is taken here as an example to explain that it is crucial to extract the best piece of information from each modality, although it is not a sensor with which the user has a sporadic interaction - it is rather a continuous monitoring. Nevertheless, a similar result would be obtained if the user would sporadically transit in front of a videocamera, which could, like the Ubisense, provide the motion vectors.

**CHALLENGES AND HOW TO OVERCOME THEM**

The degree of label purity has a key influence on the training of classifiers operating on the new sensors. A high purity allows to make better activity models while a lower purity means that activities of other kinds corrupt the class models. A key challenge is to find the right tradeoffs to obtain high purity, while at the same time maintaining a high enough number of label instances and label length, to ensure sufficient training data.

In the example of the detection of modes of locomotion, the inferred labels include “standing”. However, when the user is not standing the behavioral assumptions indicates an “other” activity which in this case we assumed to be always “walking”, although the user may as well be sitting, standing or lying. Therefore, the labels may be not directly useable to train an activity classifier. Instead, clustering of the data when the assumption indicates an activity other than standing may be required. For instance, k-means may be used to cluster the features derived from the new sensors. Afterwards, whenever an inferred label has a high confidence, the entire cluster may be associated to that label and used to train a classifier. This means that the classifiers can also wait until enough information is collected, before performing the model update.

Parameters underlying the assumptions may be user-specific. One of the key challenges is to find assumptions that are robust against inter-user differences in behavior. This might include finding a relevant set of parameters underlying an assumption, or assumptions that are robust per se. The assumptions may also be trained, to some extend.

There may be conflicting labels originating from different sources. This must be taken into account in the training of the classifiers operating on the newly introduced sensors. Approaches may include fusion at the level of the labels, according to confidence in their accuracy.

Fusion of multiple sources of sporadic labels is likely a key to reach higher performance. Finding appropriate assumptions (extending Table 1) derived from multiple interactions remain to be explored.

Another challenge to address is to detect when the information content of the newly introduced sensor with respect to the activity is actually relevant to mandate a classifier training. As an example, a newly discovered sensor may be placed on a location (e.g., shoe) which is not suitable to detect hand interactions with objects, and thus this sensor ought to be discarded for this recognition task and its training would be useless. An alternative is to measure online the performance of the trained classifiers on the new sensors upon reception the sporadically inferred labels.

Eventually in a real-world multi-user deployment of such a system, identifying the user interacting with the environment will have to be addressed, so that only the relevant sensors are taken into account.

**CONCLUSION**

We have proposed a methodology to exploit newly discovered, a-priori unknown sensors, for activity recognition. Unlike other related work, this approach does not require explicit user intervention. It allows a fully autonomous training of the new sensors to recognize activities. The source of labels is derived from pre-existing sensors. A key novelty of the methodology is that the information derived from these pre-existing sensors is enhanced by the introduction of assumptions about human behavior. Thus, information for whose provision the pre-existing sensors were not conceived, can be uncovered and exploited.

One of the next key steps is now to investigate up to which extend such an approach can be used to learn how to use new sensors. We speculate that the simplest activities that may be recognized in this way may be the modes of locomotion, while manipulative gestures are likely more challenging.

Among the most important research issues we outline: exhaustively listing sources of sporadic labels and behavioral assumptions, selecting and developing features and classifiers which are robust to label noise, and validating the complete approach (including the transfer learning) on a realistic and complex enough scenario. Furthermore, more investigation has to be carried out on the interconnection of oppor-
tunistic ubiquitous ecologies in a coherent framework [6].

ACKNOWLEDGEMENTS
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 number: 225938.

REFERENCES


