Towards Cooperative Self-adapting Activity Recognition

Andreas Jahn\(^1\), Sven Tomforde\(^2\), Michel Morold\(^1\), Klaus David\(^1\) and Bernhard Sick\(^2\)

\(^1\)Communication Technology Group, University of Kassel, Kassel, Germany
\(^2\)Intelligent Embedded Systems Group, University of Kassel, Kassel, Germany

Keywords: Activity Recognition, Self-adapting, Organic Computing, Cooperation, Active Learning.

Abstract: Activity Recognition (AR) aims at deriving high-level knowledge about human activities and the situation in the human’s environment. Although being a well-established research field, several basic issues are still insufficiently solved, including extensibility of an AR system at runtime, adaption of classification models to a very specific behaviour of a user, or utilising of all information available, including other AR systems within range. To overcome these limitations, the cooperation of AR systems including sporadic interaction with humans and consideration of other information sources is proposed in this article as a basic new way to lead to a new generation of “smart” AR systems. Cooperation of AR systems will take place at several stages of an AR chain: at the level of recognised motion primitives (e.g. arm movement), at the level of detected low-level activities (e.g. writing), and/or at the level of identified high-level activities (e.g. participating in a meeting). This article outlines a possible architectural concept, describes the resulting challenges, and proposes a research roadmap towards cooperative AR systems.

1 MOTIVATION

The well-established research field of activity and context recognition (AR) aims at deriving high-level knowledge about human activities and the situation in the human’s environment from simple sensors such as acceleration sensors, microphones, or gyroscopes (Lara and Labrador, 2013; Shoaib et al., 2014). Often, mobile devices such as smartphones or smart watches are used for this purpose (Shoaib et al., 2015). Examples for application fields range from monitoring maintenance tasks (Roy et al., 2013) through traffic applications (Liao et al., 2006) or monitoring sports activities (Ermes et al., 2008) to medical applications (Maurer et al., 2006).

Today’s AR systems have a number of limitations, including the following: a) they typically rely on a fixed configuration of sensors that are available in a given device, b) their pre-trained classification models are often not customised to the specific user, and c) the set of activities that have to be recognised in an AR system is fixed at the design-time of this system.

Based on the observation that we face an ever-increasing number of (smart) sensors and devices in our daily environment (already able to host AR systems), we suggest to overcome these limitations by means of a fundamentally different way of developing and deploying AR systems. We claim that a cooperation of AR systems including sporadic interaction with humans and consideration of other information sources whenever possible will lead to a new generation of “smart” AR systems with: i) the capability to self-adapt to the activities and contexts of a specific user at runtime (including semi-autonomous extension to new kinds of activities or contexts), and ii) an increased recognition accuracy and reduced energy consumption.

In this article, we outline our vision of a cooperative self-adapting AR System. Cooperation of AR systems will take place at several levels of an AR chain: at the level of recognised motion primitives (e.g. arm movement), over the level of detected low-level activities (e.g. writing), to the level of identified high-level activities (e.g. participating in a meeting or activities of daily living).

To illustrate, consider a managerial meeting as use-case to highlight the possible benefits of such a collaborative approach: An AR system running on a smartphone recognises that its user is sitting. In its vicinity are several other smartphones (also running AR systems), a smart pen is activated and used, and room-based sensors signal the utilisation of the room (e.g. movement detectors are activated or a projector is used). The recognition can be supported by considering even other information sources such as sitting detectors in the chairs. Once a multitude of AR sys-
tems recognise the sit down activity, they might share this information and conclude cooperatively that the meeting is going to start. Important energy savings might be achieved, if the whole recognition cycle is not required for all participants, but only for a few. For all other participants, the activity can be concluded purely by e.g. proximity. Another important improvement, other than increasing the recognition accuracy, will be the ability to include new, unforeseen activities such as sitting on a table instead of a chair, and still having this meeting. As a consequence of “meeting ongoing”, incoming calls may be muted and the calendar shows “not available”. In the course of the meeting, it can be cooperatively concluded that the (high-level activity) meeting is still ongoing – even though a participant might have left already – based on (basic) arm movements, i.e., the (low-level) activity of writing of several participants.

The remainder of this article is organised as follows: Section 2 presents the research statement, Section 3 briefly summarises relevant work from the state of the art, Section 4 presents an architecture concept for collaborative AR systems, and Section 5 describes the resulting research roadmap. Finally, Section 6 concludes the article.

2 RESEARCH STATEMENT

To address the vision above, we have to develop and investigate the foundations for a new generation of cooperating, self-improving AR systems based on the confluence of ideas from two well-established research areas: 1) Human activity and context recognition in pervasive/ubiquitous systems. 2) Organic Computing (OC) techniques for runtime self-organisation and self-adaptation in technical systems. From a scientific point of view, the above vision implies several research directions including communication in ad-hoc networks, sensor self-description, service discovery, robustness, security, and leveraging new sources of information. In this article, we focus on the latter as the other aspects are well covered by current research.

The vision is briefly sketched in the motivation essentially amounts to a cooperating and self-improving AR system. In this context, “cooperating” refers to an AR system that is able to perform purpose-oriented interaction with other AR systems. As cooperation partners for the AR system, we not only consider other AR systems being available in its communication range, but also its human user who may provide answers - assuring that the user is only asked very sporadically. Thus, we aim at transforming an AR system from a (traditionally) static system into an evolving system that adapts to the time-dependent and changing behaviour of its user. In particular, cooperation has to be established with the following objectives:

Objective 1. Cooperation based recognition accuracy improvement of ongoing human activities at several levels of abstraction (i.e., from motion primitives such as arm movements to high-level activities such as participating in a meeting).

Objective 2. Reduction of energy consumption in mobile devices hosting an AR system without deterioration of the AR system’s accuracy by cooperation.

Objective 3. Customisation of pre-trained AR systems (i.e., the AR system will adapt to the unique activities of a specific user) supported by cooperation.

Objective 4. Detection of hitherto unknown kinds of activities (i.e., “novel” activities a specific AR system was not pre-trained for) and self-extension of the activity repertoire by cooperation.

The objectives above are not only relevant for the field of AR, but also for the area of OC (Tomforde et al., 2017; Müller-Schloer and Tomforde, 2017). In particular, OC focuses on enabling autonomous technical devices with capabilities to self-organise, continuously self-assess the success of their behaviour, and consequently self-adapt and self-improve at runtime. As a result, traditional design-time decisions are transferred to runtime and into the responsibility of the systems themselves. This transfer of design-time decisions to runtime is necessary, since not all required features of a system can be anticipated offline, i.e., at design-time.

Consequently, we can summarise that the main objectives under the common theme is to develop and investigate novel techniques based on the principles of OC to develop a new generation of “smart” AR systems. In our approach, we focus on the specific questions from the field of AR systems, but we assume that the main insights will also be transferable to many other application domains.

Within this article, we refer to the object that performs the AR as “entity”. An entity might be a smartphone or another smart device. For AR the entity accesses internal and external information sources. Internal sources might be sensors that are available within the entity, for instance, the built-in accelerometer or gyroscope of the smartphone. External sources are not part of the entity such as infrastructure sensors or other entities. Further, the term “activity” summarises activities at all level of abstraction, i.e. mo-
tion primitives as well as low- and high-level activities.

3 STATE OF THE ART

The following paragraphs summarise contributions from the state of the art that are closely related to this article in several domains.

AR aims at deriving knowledge about human activities and the situation in the human’s environment. In the recognition process, commonly, sensor data are processed and matched to activities. That means, the continuously incoming data stream is segmented, for instance by applying a Sliding Window or Sliding Window And Bottom-up (SWAB) method (Keogh et al., 2001). For each segment application-specific characteristics, i.e. features, are extracted. Common features are time domain based features such as mean or variance. These features are often used as they are easy to calculate and simultaneously provide comparatively well insights into the data characteristics. Further, frequency domain based features have been investigated, which can extract unseen patterns and trends in the data (Chaovalit et al., 2011). For example, a Fourier transform can be used to uncover data characteristics that support the recognition of a user’s fall (Delahoz and Labrador, 2014). The extracted feature values are passed to a machine learning algorithm which generates an AR model. A variety of machine learning algorithms are available such as clustering algorithms, Support Vector Machines (SVMs), or Bayesian classifiers. The generated model identifies a user’s activity based on the incoming feature values. To improve the AR performance, a well investigated approach is the integration of additional information sources, often referred to as sensor fusion. Commonly, a combination of sensor sources is done in terms of raw sensor data. The data are provided by additional (internal and external) sensors. A multitude of sensors have been combined such as accelerometer and gyroscope (Shoaib et al., 2014), accelerometer and pressure sensor and microphone (Khan et al., 2014), as well as accelerometer and various combinations of infrastructure sensors such as radio-frequency readers, object tags, or video cameras (Roy et al., 2013).

Along integrating additional sensors, a (passive) collaborative approach was investigated. The project “Collaborative Context Recognition” (CoCoRec) considers other entities and information at different abstraction levels, i.e. raw data, low-level contexts, and high-level contexts, were data are exclusive at these levels. The researchers investigate the activity monitoring (Kampis and Lukowicz, 2014), or the efficient information distribution (Kampis et al., 2015). Though, the project results are mainly in information monitoring and distribution. In this article, we assume the monitoring and distribution of data as being solved and focus on cooperative knowledge gathering and processing in AR. We investigate cooperative entities that dynamically send and receive data needed for AR at multiple levels of abstraction, and consider these in the corresponding stages of the AR process. In comparison to our vision, active and purpose-oriented cooperation with other sensors and human users is not considered in AR systems today.

Organic Computing (OC) is a recent paradigm of designing and developing self-adapting and self-organising technical systems acting in the real world. OC systems are designed to process so-called self-x properties that allow them to be self-adaptive and self-organising at runtime. In this article, we aim at cooperative and self-adaptive AR systems, which require a system design allowing for internal adaptation. OC and related initiatives (such as Autonomous Computing (Kephart and Chess, 2003)) have proposed a variety of architectural blueprints. Examples include the generalised observer/controller (O/C) framework (Tomforde et al., 2011) and the Monitor-Analyse-Plan-Execute(-Knowledge) cycle, called MAPE(-k) (Kephart and Chess, 2003). For both concepts (i.e., O/C and MAPE-k), multi-layered extensions have been proposed as well as system-of-systems concepts.

The utilisation of machine learning techniques is a key factor for self-organised and self-adaptive technical systems. These systems have to adapt themselves in response to the ever-changing environmental conditions and at the same time have to guarantee the compliance to restrictions preventing faulty behaviour (Prothmann et al., 2009). Especially Autonomous Learning is considered to be a key feature in OC systems: At design-time, only incomplete information is available as basis for learning processes (e.g., for training purposes).

Collaborative Learning (CL) is a topic related to this article as we aim to enable individual AR systems to interact with each other to further improve their recognition process. Distributed intelligent systems that work in a collaborative manner recently became an active research issue (Panait and Luke, 2005). In the majority of these approaches, information is locally acquired and pre-processed by the intelligent systems and then sent to special processing units (which can be either centralised or distributed). Less common, but closer related to our approach, is work dealing with collaborating agents that learn
from each other by exchanging locally inferred rules such as in (Jakob et al., 2008). In (Tan, 1993), the authors investigate the exchange of different kinds of knowledge (i.e., observations, observation-action-reward vectors, and learned state transitions) between agents that are equipped with reinforcement learning techniques. Furthermore, agents equipped with SVM exchange newly learned support vectors in (Jändel, 2009). Correctly classified samples yield a reward which is used by the agents to adapt their SVM to changes in the environment. A periodic exchange of knowledge between agents with different learning paradigms (i.e., table based Q-learning and neural networks trained with backpropagation) is presented in (Gifford and Agah, 2009). This approach, however, is based on an application- and learner-specific intermediate knowledge representation that must be defined in advance. Additionally, the choice of representation greatly influences the performance of the overall agent system.

Active Learning (AL) provides powerful approaches to create flexible systems which are able to adapt themselves to a changing environment (Settles, 2009). These methods interact with their target system to investigate which information might optimise their model, and they actively acquire this information. In classification (also in regression) problems, AL algorithms actively request the target value of an instance (feature vector) (Aggarwal et al., 2014). Three basic AL approaches exist: 1) query synthesis (the query instance is generated), 2) pool-based AL (the query is an instance from a pool of unlabelled instances), and 3) stream-based AL (instances successively appear and the AL algorithm decides if the label should be acquired) (Aggarwal et al., 2014). One of the main challenges is to balance the exploration of new regions in the feature space and the exploitation of the existing knowledge to refine the trained model (Settles, 2009). The most popular method is uncertainty sampling, although it solely exploits the model by acquiring labels from instances near the classifier’s decision boundary (Settles, 2009). More sophisticated methods extend this approach by adding exploratory components, density information, or class priors (Reitmaier and Sick, 2013).

As a conclusion from this discussion of the state of the art, we can state that traditional research in AR does not consider cooperation sufficiently. Thus, possibly available knowledge to improve the efficiency and quality of solutions is not taken into account. To address this issue, especially techniques and insights from the domains of OC, CL, and AL are promising and have to be extended accordingly.

In this article, we claim that the state of the art needs to be improved by a new approach for AR which is taking advantage of cooperation by enabling AR systems to be flexible to environmental changes by means of cooperation. This specifically includes:

- the flexibility to consider the knowledge of other AR systems about recognised activities at various stages of a recognition chain,
- the capability to adapt to the behavioural patterns and activities of a specific user at runtime (including semi-autonomous extension to new kinds of activities), and
- improving activity recognition in terms of its key performance indicators such as accuracy and reduced energy consumption.

4 AN ARCHITECTURAL CONCEPT FOR COOPERATIVE SELF-ADAPTING AR SYSTEMS

We present an architectural blueprint for cooperating AR systems that is based on design concepts from OC (Tomforde et al., 2011). For the design of self-organising and self-adapting systems, the OC community has proposed a generalised design concept that distinguishes between a “System under Observation and Control” (SuOC) and an “Observer/Controller” (O/C) tandem that is responsible for adapting the behaviour of the SuOC to changing conditions (Tomforde et al., 2011). The O/C tandem may be realised in hierarchies of layers with increasing abstraction (Tomforde and Müller-Schloer, 2014).

Fig. 1 shows a customised variant of the generic O/C architecture for cooperating, self-adaptive AR systems. Here, the concept of the SuOC is instantiated by the human user (or several humans) in a sensor enhanced environment. The architecture consists of two layers: The Reaction Layer is responsible for reactions to observed behaviour, i.e., it realises the tasks of augmenting a traditional AR system with mechanisms for cooperative reaction according to the Objectives 1 and 2. The Adaptation Layer is responsible for long-term improvements by adapting the Reaction Layer at runtime, i.e., the upper layer monitors and modifies the behaviour of the bottom layer according to Objectives 3 and 4.

The Reaction Layer is organised in three main components: an observer component containing a four-stage recognition chain, a controller component triggering actuators, and a component for cooperative reaction.

The four stages of the recognition chain in the observer component are:
The Adaptation Layer – which addresses Objectives 3 and 4 – has to be developed and investigated but contains the O/C tandem as well. The observer component is responsible for evaluating the behaviour of the Reaction Layer. In particular, this means to assess the classification success of the AR system and the appropriateness of consecutive actions. As a result, the AR system becomes self-aware concerning its own performance. The controller component improves the behaviour of the Reaction Layer over time by combining the two modules learning and adaptation. Conceptually, this establishes a control loop on-top of the control loop of the Reaction Layer that increases the ability of the system to positively react to new situations arising at runtime and to adapt the AR system to a specific user or a new kind of activity. Similar to the Reaction Layer, the Adaptation Layer contains a module for cooperative adaptation that (passively or actively) collects activity information from other AR systems or additional sources (e.g., external sensors) and humans who can occasionally be asked to provide information about a current situation or an event that occurred shortly before.

In order to finally establish a cooperative AR system based on the presented blueprint, we have to address several research challenges in the first place. These are briefly outlined in the following section.

5 RESEARCH ROADMAP

The vision of cooperative, self-adaptive AR systems leads to the following research roadmap.

Challenge 1 – Assessment and Selection of Information Available Via Cooperation at the Reaction Layer.

The first step towards cooperation at the Reaction Layer is to assess the knowledge of other entities in order to identify the most beneficial information for AR. Therefore, all information arriving via broadcasting needs to be managed, organised, and structured. To identify information that are beneficial for AR a promising approach is to estimate the "recognisability". As the information are possibly available at different abstraction levels, the "recognisability" estimation algorithm needs to be able to respect the different informational content of the received data. Further, information are provided from dynamically changing information sources. Techniques are necessary that flexibly manage the (un-)available sources of information.
Challenge 2 – Improving the Classification Accuracy Through Cooperation.

To improve the AR classification accuracy, the information beneficial for AR needs to be integrated into the AR process. As it is possible that previously unknown information sources appear, these sources are analysed and introduced into the AR system. Depending on the level of abstraction the information are considered at the corresponding recognition stages of the recognition chain, i.e., pre-processing, motion-primitives, low-level, or high-level activities (see Section 4). The motion-primitives might be recognised by applying feature extraction and machine learning algorithms such as kNN and SVM. At the further stages, an HMM might be applied to process the higher-level information. The applied classifiers need to be extended to be able to handle dynamically changing information sources.

Challenge 3 – Energy Efficiency.

As an important advantage of cooperation, single devices may (partly) switch off their AR systems to save energy. This is possible whenever information obtained from other, "similar" entities can be transferred to maintain their own functionality without local AR. Thus, "similar" entities have to be identified. To quantify the similarity between entities, the alignment algorithm might be applied (Sigg et al., 2010). As the alignment algorithm calculates the similarity for one information abstraction level, techniques are needed that are able to handle all levels. Based on these results, "footprints" of each entity might be used to identify "similar" entities. Once a similarity is agreed, only one entity processes the AR. The AR is conducted as discussed in Challenge 2 and, hence, includes all stages. The results are provided to all other entities.

Challenge 4 – Short-term Adaptation of the Reaction Layer.

As shown in various publications, the classification success largely depends on the selected window size (when using the sliding window method). Thus, the AR accuracy might be improved when the window sizes are dynamically adapted to the current situation at runtime. To "spot" activity changes in the continuous sensor data, a promising approach is analysing the low- and high-frequency components of the acceleration data. The components can be derived by applying a Butterworth low-pass filter (Suarez et al., 2015). Secondly, activity changes might be seen by monitoring the activity frequency characteristics within the sensor data. Applying Dual-Tree Complex Wavelet Transformation to emphasise areas of frequency changes in sensor data as suggested in (Weickert et al., 2009) might give insights into the activity frequency characteristics to adjust the window size.

Challenge 5 – Runtime Customisation of the AR System at the Reaction Layer.

An AR system continuously analyses sensor data to estimate the current user activities. Once deployed, the detection mechanism runs continuously and provides a classification. However, the correctness of this classification depends highly on the pre-training at design-time by means of example data. Assuming that all detectable activities have been part of the training data, we still face the problem that pre-training is not necessarily done with the particular user, and classifications based on sensor data will occasionally be wrong (depending on variances in the user’s behaviour). A framework of measures based on (Fisch et al., 2016) and techniques, e.g., for transductive learning will allow for an adaptation of the classification system to changing conditions and a customisation to the specific user.


The knowledge of the Reaction Layer comprises all activities that have been part of the training process. However, this seldom covers all distinct activities the respective user will experience when using the AR system: Novel activities may appear and others may become obsolete due to changes of the user’s behaviour. Consequently, techniques are needed that foster the set of known activities that are considered by the Reaction Layer at all stages of the recognition system. A promising approach may be found in the combination of two aspects: (i) determine if the considered classes of known activities are sufficient, e.g., by means of developing techniques for anomaly detection, and (ii) introduce appropriate novel classes of behaviour and update the recognition system accordingly (e.g., based on exchanging class information among systems).

Challenge 7 – Cooperation with Human Users for Long-term Improvement.

The human user is the instance in the entire system with the best knowledge about activities that supports long-term improvement of the AR system. The challenge is to actively collaborate with the user by considering efficiency, acceptability, and comfort issues. Therefore, the status and the preferences of a user
have to be analysed and assessed continuously. As an approach to solve this challenge, techniques for updating the user model that are based on the ideas of “Active Learning” (Settles, 2009) have to be developed. Active Learning allows for covering the task of efficiently acquiring knowledge from the user and improving the underlying knowledge models accordingly.

Challenge 8 – Cooperation with Other AR Systems for Long-term Improvement.

Due to mobility, the user’s smartphone running the AR system is surrounded by other users that may also carry AR systems or other “smart” devices – which may have experienced different user behaviour and varying sequences of activities. This existing knowledge may be beneficial for the AR system under consideration: as basis for evaluating the success of its Reaction Layer, for identifying novel classes of activities, or as additional source to improve the certainty of a classification decision. The challenge in this context maps the previous human-related challenge to find efficient and beneficial ways to query and incorporate user knowledge to technical devices. An approach to solve this challenge will be based on techniques for modelling the knowledge, the expertise, and the mutual experiences made with other technical devices.

Challenge 9 – Estimating the Success of Cooperation.

The vision postulated in this article is that such a cooperative AR system is more efficient, more robust, and more successful than traditional approaches. The final challenge is to proof these assumptions. Therefore, scenarios are necessary that cover the different aspects of the previous challenges. In order to determine the success of the system two setups need to be compared: 1) No cooperation allowed: This represents the recognition performance achievable already today and, thus, provides the initial benchmark; 2) Assuming a perfect cooperation: This evaluates whether applying cooperation improves the recognition performance. It is guaranteed that all knowledge is immediately accessible to all entities. The (cooperatively) achieved recognition results are quantified by appropriate evaluation metrics and the calculation of statistically relevant results. To gain deeper insights, the time span of adaptation (for long-term success analysis) might be monitored.

6 CONCLUSION

In this article, we claimed that the next promising step in activity recognition (AR) research is to focus on cooperative solutions. Therefore, we outlined an architectural concept and the resulting challenges, followed by deriving a research roadmap towards cooperative and self-adapting AR systems. Cooperative self-adapting AR is a basic new way to lead to a new generation of “smart” AR systems that address several basic issues that are still insufficiently solved in the research field of AR, including extensibility of an AR system at runtime, adaption of classification models to a very specific behaviour of a user, or utilisation of all information available, including other AR systems within communication range. Cooperation of AR systems will take place at all stages of the AR chain: at the level of recognised motion primitives (e.g. arm movement), the level of detected low-level activities (e.g. writing), and/or at the level of identified high-level activities (e.g. participating in a meeting).

REFERENCES


