Mobile Netw Appl
DOI 10.1007/s11036-012-0415-x

Towards Collaborative Group Activity Recognition Using Mobile Devices

Dawud Gordon · Jan-Hendrik Hanne · Martin Berchtold · Ali Asghar Nazari Shirehjini · Michael Beigl

Abstract In this paper, we present a novel approach for distributed recognition of collaborative group activities using only mobile devices and their sensors. Information must be exchanged between nodes for effective group activity recognition (GAR). Here we investigated the effects of exchanging that information at different data abstraction levels with respect to recognition rates, power consumption, and wireless communication volumes. The goal is to identify the tradeoff between energy consumption and recognition accuracy for GAR problems. For the given set of activities, using locally extracted features for global, group activity recognition is advantageous as energy consumption was reduced by 10 % without experiencing any significant loss in recognition rates. Using locally classified single-user activities, however, caused a 47 % loss in recognition capabilities, making this approach unattractive. Local clustering proved to be effective for recognizing group activities, by greatly reducing power consumption while incurring a loss of only 2.8 % in recognition accuracy.

Keywords Group activity recognition · Context recognition · Distributed systems · Multi-user · Wearable

D. Gordon (✉) · A. A. N. Shirehjini · M. Beigl
Karlsruhe Institute of Technology (KIT),
Vincenz-Prießnitz-Strasse 1, 76137 Karlsruhe, Germany
e-mail: Dawud.Gordon@kit.edu

A. A. N. Shirehjini
e-mail: ali.nazari@kit.edu

M. Beigl
e-mail: michael.beigl@kit.edu

J.-H. Hanne
Technische Universität Braunschweig, Pockelsstrasse 12,
38106 Braunschweig, Germany
e-mail: jan@janhanne.de

M. Berchtold
AGT Group GmbH, Hilperstrasse 20a,
64295 Darmstadt, Germany
e-mail: mberchtold@agtgermany.com

1 Introduction

Context and activity recognition provide intelligent devices in the environment with the ability to act proactively in the interest of users [23]. Many of us now carry around one or more intelligent devices constantly, and the number of intelligent systems in our environment such as entertainment systems, vending machines and informational displays is steadily increasing [2, 29]. Implicit pro-active interaction based on situational awareness is increasingly more important in order to prevent us from entering a state of permanent distraction and informational overload [12]. This state is a result of constantly having to administrate and respond to the myriad of intelligent devices in our immediate environment [3]. One vision within pervasive and ubiquitous computing sees environments progressing from single-user, private devices to include multi-user devices running private applications for those users who are present. A challenge then becomes not only recognizing the context of the single user interacting with the device as with mobile phones [3], but now attempting to recognize the activity of a group of individuals interacting with the system or the environment [13].

In this work we define multi-user activity recognition (MAR) as the recognition of distinct activities of

Published online: 19 October 2012

multiple users over a period of time. Group activity recognition (GAR) is the recognition of a single activity for an entity consisting of multiple individuals (see Section 2). The group activity is not necessarily the same as the sum of the activities of the individuals in it [16]. The activity or context of a group is a function of the activity or context of all individuals in the group.

Wearable technology has been proven to be effective for human activity recognition (HAR) [1, 3, 16] and is ever more prevalent, and is therefore an attractive platform for MAR and GAR as it is already ubiquitous. Using a distributed wearable platform for both the sensing and processing aspects of activity recognition is advantageous in that it allows the system to operate independently of existing infrastructure and therefore widens the field of applications [15]. Furthermore, in times of emergency, when GAR may be needed most, the conditions of infrastructure are at their worst [7]. In order to both combat the scalability challenges, and to be robust to infrastructure collapses or partial breakdowns, methods for recognizing group activities using the devices of the individuals within those groups are advantageous. These devices are intrinsically wearable, therefore motivating the field of GAR using mobile devices with wearable sensing modalities.

When using wearable technology such as badges [29], mobile phones [3], coffee cups [2, 13], etc. for group activity or context recognition, it is inherently a hierarchical problem, where data from wearable sensors on multiple users must be aggregated in order to infer the group context [16]. This poses a problem for such systems, as energy storage is a very limiting factor and reducing energy is a main priority [26]. Activity recognition approaches must therefore also be acutely aware of this issue and make every effort to reduce their energy consumption footprint on the system as a whole [12]. Preprocessing data locally reduces its volume and therewith the energy required for transmitting that data, but at the same time this process discards information which may be vital for classification [22, 25]. Transmitting unprocessed, raw data guarantees that the maximum amount of information is available for GAR, but the cost of communication is high. The low-power requirements of mobile devices must now be reconciled with the hierarchical nature of GAR, where again a tradeoff between recognition rates and energy consumption is evident.

The main contributions of this work are as follows. A system for recognizing group activities using only a distributed network of sensor nodes and mobile phones is introduced, a sketch of which was already presented [14]. A mobile phone is used as a central node for GAR, and wireless sensor nodes are attached to coffee mugs (Smart Mugs) to monitor the activities of the individual subjects. The Smart Mugs can process measured sensor data locally to different abstraction levels before forwarding that data to the mobile phone for GAR. The capability of the system to handle real-world MAR and GAR problems is, however, not evaluated here. To the best of our knowledge, there is no other work that describes an approach for GAR using wearable devices for sensing and recognition.

An experiment was designed to create a simple collaborative GAR problem as it poses issues where nodes must exchange information in order to infer group activities (see Section 2). The experiment is used to evaluate different levels of abstraction at which the information exchange occurs in terms of its effects on the distributed sensing (energy consumption), information exchange (communication volumes) and recognition (recognition rates) systems. The goal is to identify which abstraction level is optimal for collaborative GAR in terms of the energy savings and loss of recognition values.

Different levels of data processing result in different levels of abstraction [25], from low-level raw sensor data to high-level single-user activity information processed using single-user activity recognition (SAR) techniques. The later approach introduces the problem of having to doubly-label training data in terms of single-user and group activities in order to train both local SAR classifiers on the Smart Mugs and global GAR classifiers on the mobile phone. The term local is used to refer to processes which occur at a single node, while global refers to processes which occur on the mobile phone which has a global view of the network, sensor data, and activities. Two methods for avoiding the doubly-labeling problem are presented and evaluated here: separate training sessions for local and global activities, and using unsupervised clustering techniques. These different modes of operation are evaluated in terms of distributed energy consumption and GAR rates in experiments with multiple subjects.

Here we present the results of that experiment in detail, extending an initial publication [13]. Furthermore, we present a more advantageous unsupervised learning approach, which solves one of the main problems of GAR using single-user activity information: the doubly-labeling issue. Finally, the data set gathered during the course of these experiments is published (see Section 6) to enhance reproducibility and to make this information available for future work within the community.

We analyze the difference between multi-user and group activity recognition in Section 2 and introduce a definition for these terms, and a categorization of GAR and MAR approaches into cooperative and collaborative recognition problems. We then provide explicit, formal definitions of cooperative and collaborative single-user, multi-user and group activity recognition problems. A brief survey of related work based on these definitions and categorizations is then presented in Section 3.

The rest of this paper is organized as follows: Section 2 provides a formal definition of SAR, MAR and GAR and describes their characteristics. Related work is then described in Section 3. Section 4 describes the GAR approach and the system proposed in this paper, followed by a description of the experiment in Section 5. The data set which represents a portion of the contribution of this work is described in Section 6. Section 7 presents results that are analyzed and discussed in Section 8. In Section 9, we conclude the paper and present plans for future work.

2 Single-user, multi-user and group activities

In this work we differentiate between single-user activity recognition (SAR), multi-user activity recognition (MAR) and group activity recognition (GAR), which are all forms of human activity recognition (HAR). The research here is focused on wearable systems, but results should be generalizable to other approaches such as video or audio recognition systems. These terms, specifically MAR and GAR, have been multiply defined across related work, sometimes synonymously [28] while other times defined in a contradictory manner by different works (compare [16] and [22]). Here a new definition of MAR and GAR is presented with a further classification into collaborative or cooperative recognition.

For all types of HAR, labels from the label space are usually assigned to areas of the activity space in such a way as to make the activities for all labels mutually exclusive. Although this does not have to be the case, overlapping activities create a different kind of recognition problem which must be addressed.

The spaces for activities and for labels are infinite, and the mapping between them is subjective [23]. For example, when observing Fig. 1a, one person may consider it "chopping" and another considers it "cooking", and two people may describe the activity "making coffee" very differently. Therefore defining single-user, multi-user and group activities by trying to make distinctions using labels or activity names is not a valid approach. We can, however, differentiate between these concepts by examining what is necessary in order to infer labels based on the physical characteristics of the behavior. In other words, there is no fundamental difference between single-user, multi-user and group activities per se, but rather the difference arises only when attempting to distinguish activities from each other in the process of recognition (SAR, MAR and GAR).

2.1 SAR, MAR and GAR problems

Single-user activity recognition (**SAR**) is the problem of recognizing what a user is doing based on sensor measurements taken from that user's body, possessions or environment [16]. This can be seen on the left side of Fig. 1a, where the activities of the single user (subject 1 is "chopping" vegetables) are being monitored. Here, SAR is only concerned with monitoring environmental parameters directly influenced by that subject, e.g. body-worn sensors or utensils which they are using.

Multi-user activity recognition (**MAR**) is the recognition of separate activities of multiple users in parallel, where two or more users are involved [16]. This is demonstrated in Fig. 1b, where the system recognizes several activities, one for each subject.

Group activity recognition (**GAR**) is the process of recognizing activities of multiple users, where a single group activity is a (complex) function of the behavior of each and every user within the group [13, 15]. The activity of the group (crowd) can be observed as spontan-

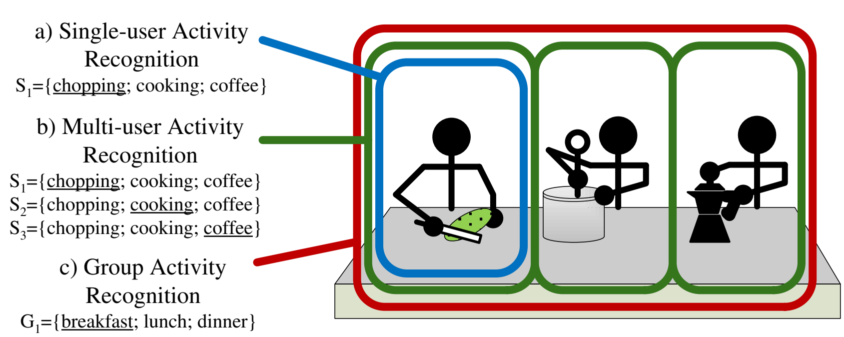

Fig. 1 An example of single, group and multi-user activity recognition for a group G_1 consisting of three subjects $S_1, S_2, S_3 \in G_1$. Individual activities are chopping, cooking, and making coffee, while group activities are which meal is being prepared

ious emergent behavior, generated from the activities and interactions of the individuals within it [6, 27, 30]. Figure 1c shows this where without knowledge of all of the users, it is improbable that the system will infer the correct activity, as the activities of each user are ambiguous with respect to the group activities (e.g. "chopping" could be preparation of any meal). Only the 3-equation-problem given by observing all subjects provides enough evidence for accurate inference.

In the same way that multiple sensors placed on the human body, each sensing only "arm", "leg" or "hip" parameters can be used to infer the activity of the entire person (SAR) [1, 3], we propose that sparse sensing of the actors within the group can be used to infer the activity of the group as a whole [13, 15]. This is analogous to treating the group of individuals as an organism in and of itself, rather than the sum of its parts. Although group activities are to the individuals in the group as single-user activities to the limbs of the user, the same methods can not necessarily be applied. Interactions between humans are far more complex than those between e.g. "knee" and "hip," and therefore their relationship to the behavior generated is far more complex as well. Further research is required to understand what can and can not be used from SAR for GAR.

2.2 Cooperative and collaborative recognition

For MAR and GAR problems we define two distinct classes: cooperative recognition problems and collaborative recognition problems. Cooperative and collaborative here do not refer to the type of human interactions occurring between the the users, but rather the type of interaction required between the subject activity models in order to recognize behavior.

Cooperative recognition problems are those in which only sensors local to a single individual are required to infer the activity. Based on these sensors, a model can be constructed in order to recognize behavior based on readings. This can be conducted independently of other individuals in the environment. The activity models, however, may be imperfect which can cause errors in recognition. Using information about neighboring activity (e.g. if one subject is "chopping" then other subjects may be more likely to be "cooking") can help to correct these errors, therefore the cooperative nature of the recognition approach.

Collaborative recognition problems are those in which the activity being performed is fundamentally dependent on information from multiple subjects. This class of problems requires that data from multiple individuals be fused in order to infer their activity, while not restricting whether or not it is one activity per user (MAR) or one activity per group (GAR). Since recognition is not possible without fusion, meaning activities cannot be modeled without observing multiple users, this class is referred as "collaborative."

An example Imagine an indoor track and field area in which we are monitoring activities. Now imagine a single subject using that track whose activity we are monitoring, and let's say he or she can only perform one of three things: "run," "walk" and take a "break". This is a SAR problem since we are monitoring a single subject (see Fig. 1a) and cooperative problem, since sensor information from that user is used to infer their activity. If they were to be wearing multiple sensing devices, where each device is not capable of inferring activity on its own, we could then refer to this as being collaborative SAR problem.

Now imagine the same situation but one more subject enters the track, where the two do not know each other, are of different skill levels, etc., so that they are not in any way interacting. This is now a MAR problem (see Fig. 1b) and a cooperative problem since we are recognizing individual activities for multiple subjects, where the dependence of each recognition problem is only on the sensors of that subject.

Going one step further, imagine the same situation, except where the two are now acquaintances and take a "break" together at some point, during which we don't know what they do, but only that they do it together. Now the system must recognize "run," "walk" and "break" for both subjects (MAR), but because we cannot be sure if they are taking a break we must now observe both subjects in order to find out, making this a collaborative MAR problem. Notice how two of the activities "run" and "walk" do not change, but depending on what you are trying to distinguish them from, it is either a collaborative or a cooperative problem.

Now imagine the same situation with more users (no effect with respect to MAR/GAR, cooperative/collaborative), where they are all members of a team, meaning all individuals perform one of the activities in unison. Combining all activities together where when every one is walking the single group activity is "walk" converts a cooperative MAR problem in a GAR problem due to the fact that we are recognizing a single activity for the group (see Fig. 1c). Similarly, the group activity could be obtained by observing only one subject, since what he or she is doing is also what the group is doing, therefore it is a cooperative GAR problem.

Finally, observing the last example where a team performs the activities "run," "walk" and "break," and changing it such that the team takes breaks together but each individual has a varying skill level such that the activity "walk and run" is possible, this fundamentally changes the nature of the recognition problem. Now, all members of the group must be observed since it cannot be assumed that if one subject is walking that the group is walking and so on. By adding one activity, the problem becomes one of collaborative GAR. Once again notice that the difference between cooperative and collaborative problems is not an intrinsic property of the activities, but rather of the recognition problem.

2.3 Formalization

When observing the definition of MAR vs. GAR and cooperative vs. collaborative from a Bayesian probabilistic viewpoint, we can now formalize it using the prior and belief networks. Figure 2a shows the factor graph for SAR, where the hidden variables (activity states) for subject 1 $x_1 \in X$ are governed by a prior $Pr\{X\}$ and are connected to the observable variables (sensors) $y_1 \in Y$ by the belief $Be\{Y|X\}$. Similarly, Fig. 2b shows the factor graph for the prior and belief function for collaborative MAR, showing the conditional independence of both belief functions and priors between hidden and observable variables. Figure 2c shows the belief and priors for collaborative MAR, where subject activities are dependent on information of other subjects.

Figure 2d shows the factor graph for cooperative GAR. The grayed relationships indicate a "one-of-each" relationship between the hidden variable and the observable variables, meaning that the posterior $p(x|y)$ can be evaluated given any y. Collaborative GAR is shown in Fig. 2e, where the hidden variables are dependent on all observable variables.

2.4 Discussion

This begs the question, why are these distinctions necessary? Usually, they are not of importance to the system used to recognize these activities, as posteriors for hidden variables are inferred using priors, belief functions and observed states of all observable variables. These dependencies can also be modeled explicitly in order to improve system performance [28].

In this work we address approaching this problem from a distributed point of view, where the distributed wireless sensing network is also the platform conducting recognition. Each dependency requires communication between nodes, as they would otherwise not have access to the states of remote variables for inference. Interdependency between users for inference must be explicitly modeled and accounted for, and a distinction between which types of problems require this communication and which types do not must be made.

In summary, MAR is used where the goal is to obtain the activities of multiple individuals in parallel, where GAR is used to recognize the activity of a group as an organism, often where the activity of the group is not directly evident when observing the activity of the individuals independently: emergent behavior. Both MAR and GAR can either be cooperative or collaborative, depending on the activities recognized. For cooperative problems, knowledge of other subjects can help to reduce error caused by model errors or simplicity. Collaborative issues, however, require exchange of information about other group members in order to

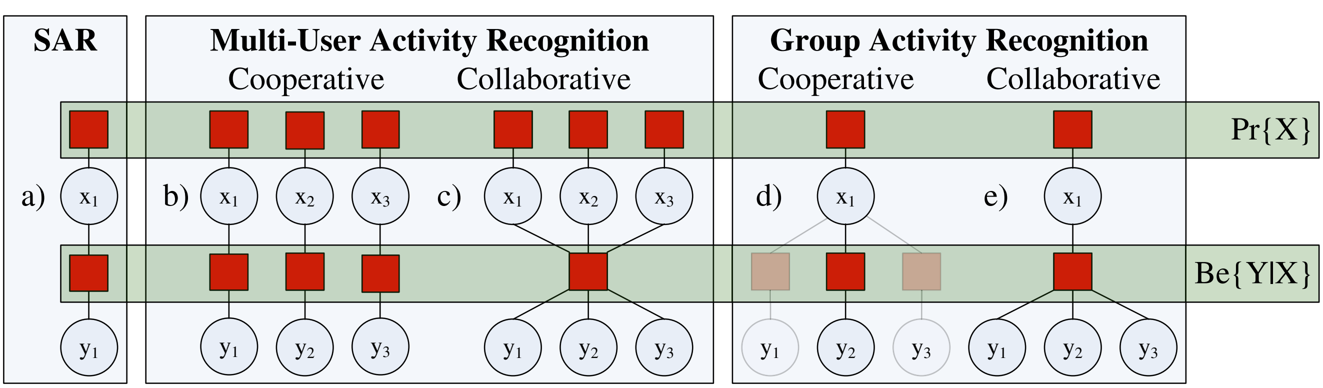

Fig. 2 Factor graphs for belief $Be\{Y|X\}$ and priors $Pr\{X\}$ over hidden variables (activities) $x_1, x_2, x_3 \in X$ and observable variables (sensors) $y_1, y_2, y_3 \in Y$ for **a** Single-user Activity Recognition (SAR), **b** cooperative Multi-user Activity Recognition (MAR), **c** collaborative MAR, **d** cooperative Group Activity Recognition (GAR) and **e** collaborative GAR

model the activities. In the remainder of this work we focus on collaborative recognition problems which pose issues for distributed, in-network recognition which are novel. Specifically, the level of abstraction at which the information exchange occurs is evaluated in terms of its effects on the distributed sensing (energy consumption) and recognition (accuracy) systems, as well as the information exchange (communication volumes) between them.

3 Related work

The majority of all context and activity recognition work is focused on human subjects and concentrates on single-user activity and context recognition. Traditionally, this is conducted using body-worn acceleration sensors [1, 12] which forward sampled data to a central server for classification. Other approaches range from embedded recognition [12, 26], where emphasis is on the tradeoff between energy expenditure and recognition quality, to server based approaches which optimize classification results using crowd-sourcing [3].

First simple attempts at recognizing the activity of a group as a whole were pioneered with the Active Badge [29] and MediaCup [2] projects, where the status of a user (including meetings or gatherings) was updated based on their location and the location of others. These approaches where not learning-based, but rather static code which recognized activities mostly based on location, proximity, and some sensor measurements.

Other approaches use audio classification to recognize multi-user group activities, such as concurrent chatting activities [18], or for classifying roles of individuals in conversations [9]. These methods have proven effective, but rely heavily on infrastructure for recognition. Theoretically, embedded GAR approaches using audio sensors would be possible [26], but the authors are unaware of research in this direction.

Camera-based systems are well suited to collecting information about multiple individuals within the field of vision. This advantage has been put to use for the purpose of group activity recognition, for example for monitoring activities of groups of individuals in a prison yard [8] or cargo and logistics activities [11, 21]. Another great example of uniquely group-related activities, is recognition of American Football plays based on TV feeds [20]. There, Li et al. track individual trajectories and activities of single users, and then use this information to recognize which play is being orchestrated. The large drawbacks of video-based systems is

Table 1 Analysis and comparison of existing multi-user and group activity approaches

Reference	Application domain	Activity type	Dependency type	Architecture	Sensor tech.	Issues
Chang et al. [8]	Prisoner activity recognition	GAR	Collaborative	Centralized	Video	Dependent on infrastructure, video requires instrumentation
Want et al. [29]	Office activities	MAR/ GAR	Collaborative	Centralized	Multiple	Dependent on infrastructure, static logic, not capable of learning, domain specific
Beigl et al. [2]	Office activities	MAR/ GAR	Collaborative	Centralized	Wearable	Dependent on infrastructure, static logic, not capable of learning, domain specific
Hsu et al. [18]	Human conversation	GAR	Collaborative	Centralized	Audio	Dependent on infrastructure, domain specific
Wirz et al. [30], Roggen et al. [22]	Pedestrian flocking	MAR/ GAR	Collaborative	Centralized	Wearable	Domain specific, focused on group affiliation
Hwang et al. [19]	Behavioral singularities	MAR	Collaborative	Centralized	Wearable	Application-specific outlier detection
Gu et al. [16], Wang et al. [28]	ADLs in home	MAR	Collaborative	Centralized	Wearable	Dependent on infrastructure, applicability for GAR unclear
Li et al. [20]	American football plays	GAR	Collaborative	Centralized	Video	Dependent on infrastructure, video requires instrumentation
Gong et al. [11], Loy et al. [21]	Logistics and public places	MAR	Cooperative	Centralized	Video	Dependent on infrastructure, video requires instrumentation, highly domain specific
Dong et al. [9]	Conversational roles	MAR	Cooperative	Centralized	Audio	Dependent on infrastructure, application for HAR unclear
Present work [13, 14]	Generic (office) activities	GAR	Collaborative	Distributed	Wearable	

that they require pre-instrumentation of recognition environments, and commonly require infrastructure to connect sensors and processing architectures.

Research into MAR and GAR using wearable sensors has only recently been introduced to the scientific community, an overview of which can be found in Table 1. Gu et al. [16] and Wang et al. [28] combine patterns of individual activities to recognize concurrent multi-user activities using probabilistic methods. Here the activities which are recognized range from single-user activities as well as concurrent and conflicting multi-user activities, making this approach collaborative in nature.

Wirz et al. approach recognition of cluster formations and flow patterns in groups of pedestrians [30]. The work presented here expands on that done by Roggen et al. [22], where the concept of "behavioral primitives" are introduced as single-user activities. Here, group membership for each subject is monitored (MAR), but also crowd behavior is addressed (GAR), both of which can only be evaluated with knowledge of other group members (collaborative). Similarly Hwang et al. track behavioral singularities in children on field trips, where the behavioral singularity is tracked for each child (MAR), but can only be calculated in comparison with other subjects (collaborative).

Sigg et al. [25] researched the optimal context abstraction level for prediction of future contexts. This was also addressed for a different application, namely sensor control for embedded SAR using prediction [12]. Since GAR using wearable sensors is inherently a hierarchical problem, these same issues are also present here as well, but with focus on GAR instead of context prediction. A case study on GAR to evaluate the optimal context abstraction level for GAR using sensors from wearable devices was presented in a preliminary poster abstract [14]. A requirements analysis for distributed recognition in peer-to-peer networks of mobile devices was also presented [15]. Preliminary results provided insight into the power-accuracy trade-off for GAR, and uncovered several novel research questions [13].

4 System design

The system used here was made up of a wireless sensor network and a mobile phone. Wireless sensor nodes equipped with 3D acceleration sensors were attached to coffee mugs in a university/office setting. The nodes sampled activity and context data at the mugs, processed this data to a specified local abstraction level, and then forwarded it to the smart phone for further classification to the group activity as shown in Fig. 3.

The smart phone was tasked with recognizing group activities based on the data sampled by the wireless nodes on the coffee mugs. These nodes forwarded either raw sensor measurements, extracted sensor signal features, local clustering information, or locally recognized single-user activities to the smart phone. The different modes were evaluated in terms of power consumption and recognition accuracy. The classifiers used in this paper are the k-Nearest-Neighbors (kNN) (k = 10, Euclidean distance, no feature weighting), Decision Tree (DT) (C4.5), and Naive Bayes (nB) (no kernel estimation, single Gaussian, no covariance modelling) algorithms, selected for their simplicity for embedded purposes.

A hard K-Means clustering algorithm was used which outputs a single cluster candidate (top-1) for each vector, and uses subtractive clustering to identify the number of clusters present [13]. Each node outputs the index of the cluster which is identified given the k-Means clustering algorithm. These values are then fused by the mobile phone into a single group activity using trained classifiers. Since clustering algorithms do not require labels for training, local labels

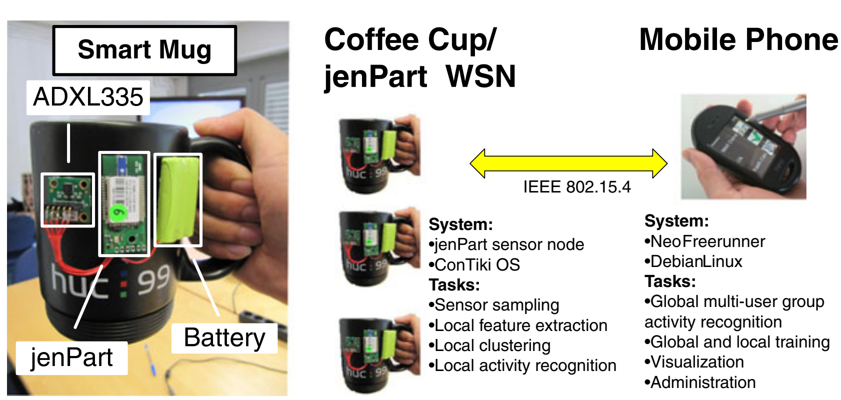

Fig. 3 *Left:* the smart mug with an acceleration sensor, jenPart node and battery. *Right:* network topology of the group activity recognition system and abstraction level experiment

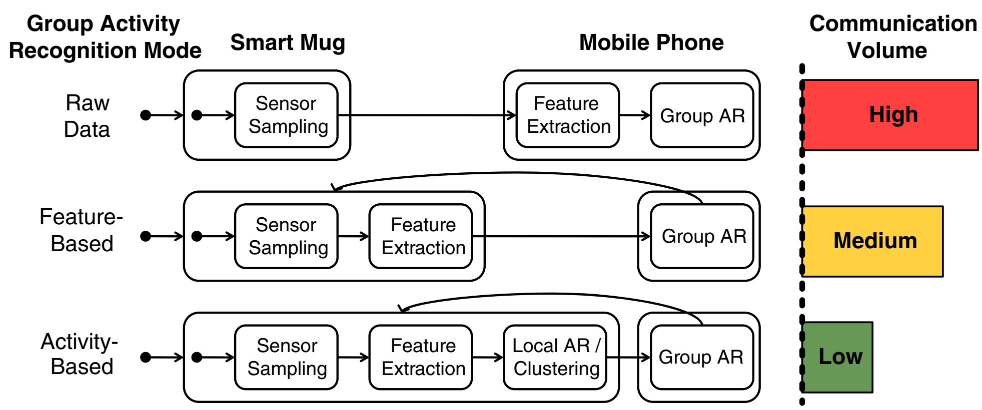

Fig. 4 State charts for the three different system modes for GAR with associated approximate communication volumes

are not required for GAR, making these approaches advantageous.

Here a second method for evaluating the potential of unsupervised clustering as a method of skirting the doubly-labeling issue was investigated. Expectation maximization (EM) for Gaussian mixture models (GMM) [4] was used to cluster the data and a soft clustering approach was used as it has been shown to be advantageous for other approaches [5]. The Gaussian mixture is given by:

$$p(x) = \sum_{k=1}^{K} \pi_k P(x|\mu_k, \Sigma_k), \quad \text{where} \quad \sum_{k=1}^{K} \pi_k = 1 \qquad (1)$$

For each Gaussian component of the GMM, the probability that the vector was generated by that component is calculated. This probability is then normalized and output as the feature indexed by that component. For example, if training of one node yields a 3-component GMM, the output vector is then of length 3, where the kth feature is the posterior for the kth component of the GMM given the feature vector x, or $p(\mu_k, \Sigma_k|x)$.

The wireless sensor nodes used were jenParts [24] from the open-source Jennisense Project.[1] The nodes are based on the JENNIC JN5139 wireless microprocessor, the ConTiki operating system [10], a battery and an analog 3D acceleration sensor.[2] The nodes sample the sensors at a rate of 33 Hz and segment the sample data into windows (1 window = 16 samples ≈ 250 ms with 50 % overlap). Based on the operational mode, the windows are then processed and forwarded to the Neo: either the raw sensor data is forwarded (**Raw Data** mode, low-level data [25]), or the sensor signal features mean and variance are forwarded (**Feature** and **Training** mode), or single-user activity information from a classifier or clusterer is forwarded (**Classification** mode, high-level data [25]).

A Neo Freerunner[3] was connected to a jenPart bridge in USB host mode for communication with the Smart Mugs. The Neo serves as a mobile platform for classifying the group activity based on the data aggregated from all nodes in the WSN. This involves a training mode and a classification mode for the global classifier. At training time, a vector consisting of data from the local nodes (either raw, features, clusters or classes) and a global group activity label is input into the global classifier. In classification mode, an unlabeled data vector consisting of the local data from the distributed nodes is input into the classifier, which then outputs the classification, or group activity estimation for that vector.

The Neo also serves as a classifier training platform for the mugs in the WSN. Following the approach presented by Berchtold et al. [3], after being set in training mode by the Neo, each mug gathers data and forwards it to the Neo along with a local annotation indicated by segmenting activities using the button on the jenParts. Once this process is complete, the Neo trains the selected classifier, segments the trained classifier into packet-sized chunks, and sends these chunks sequentially to the nodes in a JSON format. The Mugs are equipped with a JSON interpreter which then reconstructs the classifiers locally and places them in memory so that they can be executed as a module.

For all wireless communication tasks, the data is vital at the receiver. Wireless communication must either be designed in a reliable fashion, or measures must be taken to reconstitute missing data, although these questions are outside the scope of this work. For this experiment the system was designed such that packet

[1] The Jennisense Project: http://github.com/teco-kit/Jennisense/wiki.

[2] ADXL335 3-Dimensional Acceleration Sensor: http://www.analog.com.

[3] http://www.openmoko.org/.

loss was not an issue under the controlled conditions of usage, but in a real deployment this must be addressed (see Figs. 3 and 4).

5 Experiment

This experiment was designed to evaluate different levels of data abstraction carried out by the mugs in terms of energy consumption, communication volumes and GAR rates. To this end the experiment represents a collaborative GAR problem where single-user activities map to group activities in such a way as to be ambiguous for individual subjects, but explicit when observing all subjects (see Section 2). Processing data to the activity abstraction level [25] poses the problem of having to doubly-label the training data in terms of local, single-user activity labels and global, group activity labels. This must either be done using video recordings and offline annotation (time consuming) or multiple annotators in real time, both of which are too elaborate to allow easy deployment in new scenarios.

To counteract this, two methods of skirting the doubly-labeling issue are employed and evaluated. First, local classifiers and global classifiers are trained in two sessions where each session must only be labeled with local or global activities respectively. Second, local activity classifiers are replaced with a hard, top-1, unsupervised k-means clustering, and soft, probabilistic clustering [5], which does not require local activity labels, and can therefore be trained on the same data basis as the group activity classifier. Although the system was implemented on the distributed heterogeneous platform, the classification results were generated offline using the WEKA toolkit [17] for analytical purposes but were cross-checked with online results.

5.1 Activity recognition experiment

During the course of this experiment, three subjects performed seven different activities, three of which were group activities and four of which were single-user activities involving the Smart Mugs. In total, over 45 min of data were collected, making over 22,700 sample windows, although some data was discarded at random to ensure that experimental data was independently and identically distributed (i.i.d.). The experiments were conducted in a meeting room in a university setting over the course of a single day. In the first phase, local classifiers were trained and evaluated, followed by the global classifiers in the second.

5.1.1 Phase 1: local single-user classifiers

In the first phase of the evaluation, each user performed a set of activities, each one for a duration from approximately 2–15 min with the mug in training mode, meaning features and labels were extracted locally and uploaded to the Neo. The activities were local to the mugs, and were not performed as part of group activities, as doubly labeling local and group activities in real time is impractical. The local single-user activities were as follows: the subject has placed the mug on the **table** (or other surface), the subject is holding the mug in their **hand**, the subject is **drinking** from the mug, and the subject is **gesticulating**.

After each activity was performed for the specified period of time, a button press on the node updated the label on the feature vector sent to the Neo and the next activity was performed. The first half of the data generated in this phase was used to train the local classifiers, and the second half was used to evaluate their performance. After all local activities were performed, the local classifiers were trained and communicated to the Smart Mug using JSON packets. The procedure of the process conducted in phase 1 is displayed in the upper portion of the sequence diagram in Fig. 5.

5.1.2 Phase 2: global group training and evaluation

The evaluation of the global classifier was conducted offline using the data generated in this phase, where again half of the data was used for training and the other for performance evaluation. The subjects conducted the following activities together for 4–5 min each using the same mugs they trained in the previous phase: **Meeting**, **Presentation** (users 1, 2 and 3) and **Coffee break**. The mappings of group to single-user activities are as follows: meeting consists of all subjects either setting their mugs on the table, holding them in their hand or drinking. In a presentation one subject will be gesticulating or holding their mug while others are either holding, drinking from, or have set the mugs down, and in a coffee break all are either holding, gesticulating with, or drinking from their mugs.

During this period, the nodes transmitted the full locally extracted feature vector, as well as the local classifications of the single-user activities listed previously. The raw sensor data was ignored for reasons which will be explained later. The process flow for phase two is shown in the lower portion of Fig. 5 where feature vectors and local activity classifications are transmitted simultaneously to train global classifiers for each data type respectively.

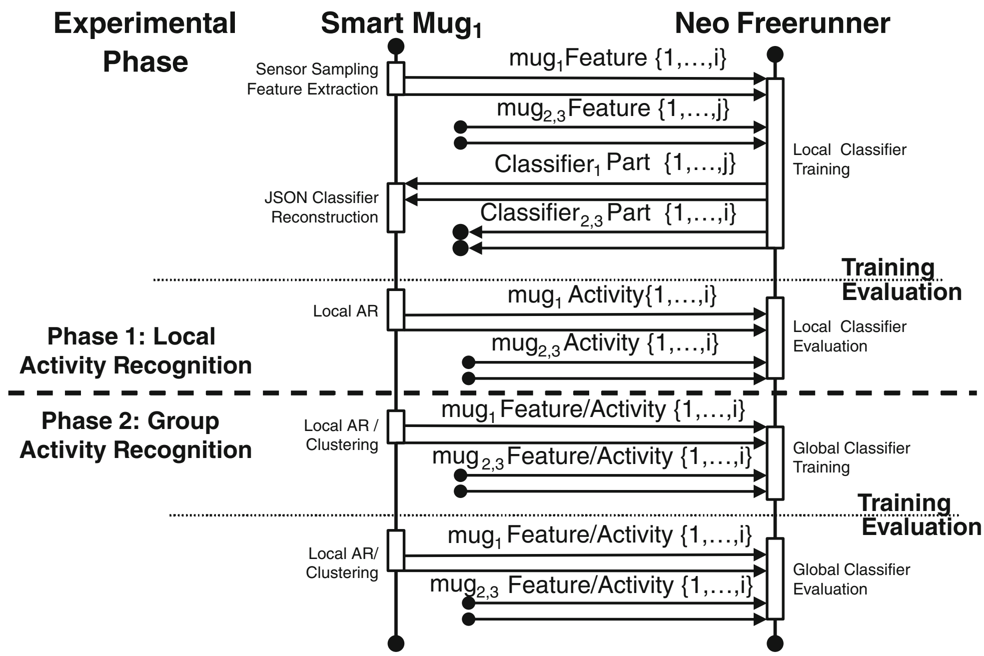

Fig. 5 Sequence diagram for the two-phase group activity recognition experiment

5.2 Power measurements

The power consumption of each device was measured by placing the node in serial with a low error tolerance resistor and measuring the drop in voltage across the resistor. For each of the modes (raw sensor data, extracted feature data and classifier/cluster data) the average rate of consumption and the amount of energy consumed was calculated. The amount of energy consumed over the period of time beginning at t_0 and ending at t_1 is then given by $\int_{t_0}^{t_1} V_{\text{supply}} \times I_{\text{supply}}\, dt = \int_{t_0}^{t_1} V_{\text{supply}} \times \frac{V_{\text{meas}}}{R_{\text{meas}}}\, dt$ where V_{supply} is the supply voltage, I_{supply} is the current drawn by the node, which is given by the voltage drop (V_{meas}) over the measurement resistor with resistance R_{meas}.

5.3 Raw data issues

Since the features calculated by the mobile phone and the distributed nodes are identical, the recognition rates for both modes would be identical as well. Theoretically, the Neo is capable of calculating far more complex and extensive feature sets than the sensor nodes, meaning that recognition rates for the raw data phase could be higher than for locally extracted features. That certain features provide better or worse recognition values is however a known fact [12], and the field of feature selection is a different area of research, making this comparison out-of-scope in the context of this work. For this reason, the raw data phase was only used to evaluate data volumes and energy consumption rates, and not to compare classification values.

6 The data set

One of the most difficult and time consuming steps in HAR research is the collection of a data set for evaluating hypotheses and algorithms. Publishing these experimental data sets is a great step towards increasing reproducibility within the research field of activity recognition. Without this, progress in the field is slowed as scientist must redundantly record private data sets independent of each other, and the effort required for reproducing published results is prohibitive. For these reasons, part of the contribution of this publication is the data set used for evaluating this work.

The data set has been made available online.[4] All data has been published in the `.ARFF` file format for compatibility with the WEKA toolkit [17]. The files contain features generated over the sensor data streams for the two phases of the experiment. The raw sensor data was not transmitted in order to reduce the volume of communication on the wireless channel, avoiding collisions (causing possible loss of data) and freeing up bandwidth. The root folder contains two folders `/local/` and `/global/`, which correspond to Phase 1 and Phase 2 respectively.

[4] http://www.teco.edu/gordon/GAR/data_set.zip.

6.1 Data set from phase 1

In the folder `/local/`, there are two files for each user, one containing training data and one containing testing data. Each file contains a set of vectors of length 7, each containing the average and variance for each axis of the accelerometer, and an activity label. The activities in these files are the local single-user activity labels from Phase 1. The identity of each feature in the vector is given by the header of each file, for example:

```
@relation subject_1_train
@attribute mean_x numeric
@attribute mean_y numeric
@attribute mean_z numeric
@attribute var_x numeric
@attribute var_y numeric
@attribute var_z numeric
@attribute label{table,wave,drink,hold}
@data
```

6.2 Data set from phase 2

Folder `/global/` contains the data from Phase 2 of the experiment. In the folder `/global/single vectors/`, the files `features_train.arff` and `*_test.arff` contain the combined data for all three subjects in one single vector. These vectors have a length 19, with 18 features (six from each subject) and activity label. The labels here are Phase 2 activities, i.e. group activities. Again, the identity of each dimension of the feature vector is given by the header.

Local single-user classification vectors from from Phase 2 are stored in six files in the folder `/multiple vectors/`. The file names begin with the classifier name (`j48, knn, nb`) followed by `_train.arff` or `_test.arff`. Each dimension in the vector represents the activity of the subject indexed by that dimension. These are generated by classifying each vector of the global feature data with the classifiers trained using the data from Phase 1. The value of each position is a Phase 1 single-user activity, where the labels in this file are group activities. Since no doubly-labeling was done, information about the exact correctness of these classifications is not known.

The data set presented here has been normalized using a min-max normalization. The minimum and maximum cannot be taken from the entire set. The testing data is used to simulate and evaluate online operation of the system. This data would therefore not be available at the point where the min and max values have to be fixed for the system. In order to correctly model this, the minimum and maximum values for normalization where acquired from the training data only, and used to normalize the entire data set.

7 Results

The results are two fold, first the classification rates for local and group activities are presented, followed by the evaluation of the communication load and power consumption of the nodes. The implications of these results and the insights they provide into the field of collaborative GAR will be discussed in Section 8.

7.1 Classification results

The classification results will be presented in two parts. First the local single-user classification rates achieved by the mugs themselves of their local activities are presented, followed by the recognition rates of the global classifier for GAR based on local features and local activities will be presented.

7.1.1 Phase 1: local classification

In phase 1 the mugs were trained using the following four classes: subject has set the mug down, subject is holding the mug, subject is drinking and subject is gesticulating. Table 2 displays the results of the evaluation of the local classifiers trained in phase 1 of the experiment. The accuracy, precision, recall and F-measure of each mug, as well as the average over the three mugs is displayed. All classifiers for local, single-user activities performed at around 95 %, where minimal variance across mugs, activities and classifiers was observed.

Table 2 Classification rates for local single-user activity recognition

Data basis	C4.5 Decision tree				k-Nearest-neighbors				Naive–Bayes			
	Acc.	Prec.	Rec.	F-meas.	Acc.	Prec.	Rec.	F-meas.	Acc.	Prec.	Rec.	F-meas.
Node 1	0.976	0.976	0.976	0.976	0.971	0.972	0.970	0.971	0.985	0.985	0.985	0.985
Node 2	0.948	0.947	0.947	0.947	0.936	0.941	0.935	0.938	0.906	0.910	0.905	0.908
Node 3	0.951	0.951	0.951	0.951	0.955	0.955	0.954	0.955	0.932	0.940	0.932	0.936
Average	0.958	0.958	0.958	0.958	0.954	0.956	0.953	0.955	0.941	0.945	0.941	0.943

Table 3 Classification rates for global group activity recognition

Data basis	C4.5 Decision tree				k-Nearest-neighbors				Naive–Bayes			
	Acc.	Prec.	Rec.	F-Sc.	Acc.	Prec.	Rec.	F-meas.	Acc.	Prec.	Rec.	F-meas.
Features	0.962	0.962	0.962	0.962	0.894	0.900	0.896	0.898	0.565	0.612	0.575	0.593
Soft clust.	0.935	0.919	0.931	0.925	0.914	0.904	0.892	0.898	0.569	0.646	0.753	0.695
Hard clust.	0.767	0.772	0.656	0.709	0.767	0.764	0.662	0.710	0.716	0.624	0.561	0.591
Activities	0.507	0.531	0.516	0.524	0.424	0.557	0.429	0.485	0.491	0.514	0.495	0.505

7.1.2 Phase 2: global classification

Similar to phase 1, the global GAR classifier used half of the data generated in phase 2 for training and the other half for classifier evaluation. Table 3 displays the results of the evaluation of the global GAR classifiers from phase 2. Each row of the table represents a different data abstraction level of the mugs: either feature transmission, transmission of local activities (the local classifier algorithm is always the same as the global one, e.g. the first column is local single-user DT, with a global GAR DT), or transmission of local clustering results, either soft or hard. In total 12 global GAR classifiers were trained and tested, three classifiers (DT, kNN, nB) for each type of local data abstraction.

Table 3 indicates that local single-user classification provided poor results with a accuracies of 51 % (DT), 49 % (nB) and 42 % (kNN). Local hard clustering provided better GAR results, with accuracies of 77 % (DT, kNN) and 71 % (nB). Local soft clustering resulted in a variance across different classifiers, achieving recognition rates of 94 % and 91 % for the DT and kNN classifiers respectively, but only 57 % for the nB classifier (see Section 8 for a details). The best results were achieved using local features and a DT classifier (96 %), where the kNN algorithm achieved relatively high recognition rates (89 %), while the nB classifier was only able to achieve GAR with an accuracy of 56 % (compare with 33 % at random).

7.2 Data transmission and energy consumption

In order to analyze the requirements of the three different system modes in terms of resource consumption the nodes were monitored over different modes of operation. The effects of each mode was analyzed in terms of communication time and volume as well as energy consumption. Table 4 displays the amount of time required for communication per second (T_{tx}) and the amount of data communicated per second for each node. The results indicate a drop in data volume of 73.5 % between transmitting raw data and features, 88.5 % between features and classes/hard clusters, and a 96.9 % drop in the amount of data communicated from raw data mode to local context classification mode. Values for soft clustering are approximate as they vary across nodes depending on the number of Gaussian components in the GMM. The values shown in Table 4 are achieved using linear approximation based on the average data volume per node.

During the course of these experiments, the energy consumption rates of the different devices were also monitored. Table 4 displays the results for the energy measurements for both the mug hardware and the Neo mobile phone as they carried out the necessary operations. The results indicate a decrease in average energy consumption (Avg(P)) at the mugs of 1.4 % from raw data to feature modes, a decrease of 4.5 % from feature mode to classification mode, and a total drop of 5.8 % from raw data to classification mode. For the Neo, a drop of 2.7 % in average energy consumption was registered from raw data to features, a drop of 1.33 % from features to classes, and a total drop of 4.0 % from raw data to classification mode.

Due to the difference in the the ratio of operational to transmission power consumption between the 2 device types, the energy consumption due to transmission could only be directly measured accurately at the mugs, but not at the Neo. The right-most column in Table 4

Table 4 Communication volumes and power consumption results

Mode	Data volume	Neo Freerunner	Wireless node (Mug)	
	(B/s)	Avg(P) (W)	Avg(P) (mW)	E_{Tx} (mJ)
Raw data	404.25	1.771	24.574	1.012
Features	107.25	1.723	24.233	0.909
Soft clusters	Variable (avg. 24.75)	Variable (≈ 1.703)	Variable (≈ 23.296)	Variable (≈ 0.648)
Hard clusters	12.375	1.700	23.140	0.605
Local activities	12.375	1.700	23.140	0.605

indicates the amount of energy consumed by a node for the purpose of wireless communication each second (E_{Tx}). This indicates a 10.17 % drop in energy consumed when transmitting features compared to raw data, and a decrease of 33.44 % from features to classes, with a total decrease of 40.22 % from raw data to classes.

8 Analysis and discussion

8.1 Using single-user activities for GAR

One of the most important issues is selecting local activities relevant to discrimination between the global group activities. Here the experiment was designed to avoid this problem by engineering a collaborative GAR problem which can be directly mapped onto the single-user activities in order to evaluate the effects of the abstraction level, rather than the scenario. For real scenarios, either intuitive or experimental knowledge of the relationship between group and individual activities is required for activity selection, otherwise global recognition rates will deteriorate.

In this experiment, global classifiers were trained using the output of the local classifiers in the local classification mode, meaning that local classifier error was present in the training data for global classifiers. Alternatively, doubly-labeling activities would have allowed for training local and global classifiers on the ground truth labels simultaneously. The effects on global rates is unknown; using local labels could allow for the global classifier to account for and correct local errors, though it may also worsen results by distorting global classifier mappings. Furthermore, in this experiment a great deal of the GAR error when using locally classified activities was due to the fact that the data generated in Phase 1 of the experiment differed greatly from the data generated in Phase 2. Although subjects were instructed to conduct local activities as they would in a meeting, they were incapable of reproducing their own behavior under the group activity conditions. This becomes apparent when comparing the averaged maximum feature values for signal average (812 local vs. 1,324 global) and variance (6,621 local vs. 148,271 global) of the two datasets. Eliminating this discrepancy would involve labeling local activities during group activities which would greatly increase labeling effort.

Table 4 indicates that the energy consumed by the nodes for the purpose of transmission dropped by 33 % when the nodes only transmit a locally classified situation instead of locally generated features. When compared with Table 3, it becomes clear that these values come at a high price in terms of the recognition rates for global classification.

In the previous section, the nB classifier performed badly as a global classifier. Both the nB and DT classifiers performed comparably locally, but there is a disparity of up to almost 50 % for global group activities based on local features. This indicates that GAR presents problems which are not present for single-user AR, and that not every classifier algorithm used for single-user HAR is appropriate for multi-user GAR. Data analysis indicates that often times group activities create multiple clusters in the multi-dimensional feature (18 dimensions) and activity (three dimensions) space, for instance group activity ”Presentation” consists of three clusters, one for the “flavor” of the activity when each different user presents. The nB classifier used here uses a single Gaussian to model each activity without kernel estimation. For GAR, the poor results imply that a probabilistic approach must be combined with clustering and covariance modeling in order to model multiple clusters and dependencies, as the naive Bayes assumption can be detrimental.

8.2 The energy-recognition tradeoff

The ratio of how much of the total energy consumption is used for communication can be seen in Table 4, and is very much system and implementation dependent, where the volume of data falls by 75 %, meaning that a large portion of the energy consumed for communication is in overhead.

Also currently, Table 4 indicates that the energy used for transmitting data is only around 4 % of the total amount of energy consumed by the node, which is due to this fact. The short sample window length (500 ms) means each communication contains only half of one second’s worth of data. Increasing the length of this window would increase the amount of data per packet and reduce the packet overhead ratio. These values are heavily system and scenario dependent, where factors such as number of sensors and features, as well as window length and sample rate play a large role.

Changing these parameters could tip the energy and GAR rate trade-off and would require a new evaluation. In this system, only two features are calculated, whereas in the literature, activity recognition systems often implement multiple features in the time and frequency domains [1, 3, 12, 25, 26, 30]. Increasing the number of features calculated would further tip this ratio in favor of local classification and clustering, also increasing the overall effect of the energy consumed for communication on the total energy consumed by the system.

8.3 Using clustering for GAR

Although the results of GAR using local hard clustering were significantly lower than using local features (77 % as opposed to 96 %, 20 % drop), clustering is quite promising. Clustering does not require a separate phase for local training as local labels are not required (unsupervised learning), and reduces the energy consumption due to transmission by 33 %. The 20 % drop in GAR rates is prohibitive for most applications.

Soft probabilistic clustering, which showed significant promise for other applications [5], proved to be an effective tool here. The GAR rates in Table 3 indicates an accuracy of 94 % for soft clustering, compared to the maximum of 96 % when using features. This indicates a loss of recognition accuracy of 2.8 % from GAR using features to GAR using soft clustering, while maintaining energy savings at approximately 29 %. The implications of these results are two-fold. Depending on the number of clusters identified, a parameter which can be controlled by the system designer, the resulting impact on energy reserves can be varied as well. The amount of data communicated is proportionate to the number of Gaussian components, therefore less clusters means lower consumptions.

Conversely, it stands to reason that increasing the number of clusters increases the quality of the information transferred, thereby increasing recognition rates. In this specific instance, between 1 and 3 clusters were detected per node, with a total of 6 clusters. This configuration generated power consumption values only slightly greater than hard clustering (double the data volume), but already produced high recognition values. Soft clustering allows the application designer to tune the tradeoff between energy consumption and recognition by increasing or decreasing the number of clusters. The range for tuning is given by the recognition rate using raw data (features, in this case 96 %) which is the maximum, with the minimum being the values for hard clustering (here 77 % accuracy).

Using local activities reduces cost, but also reduces GAR accuracy by an unacceptable 47 %. Hard clustering maintained the cost reductions but with an accuracy loss of 20 %, the advantage being that unsupervised learning does not require single-user labels. Soft clustering enables high recognition rates without requiring local labels, representing a real and viable solution to the doubly labeling problem. In this instance a loss of only 3 % of recongition could be achieved without requireing local labels.

8.4 Generalization of results

As a strong caveat, the absolute values for GAR rates and energy consumptions cannot be assumed for other collaborative GAR problems. The activities here were designed to present a collaborative GAR problem for experimentation, and to be fairly straight-forward to recognize. Results for different scenarios, as is also the case with SAR and MAR, would depend on a multitude of factors such as sensing modalities, type and number of activities, amount of training data, etc.. What can be generalized, however, are the relative rates for energy consumption, communication volumes and recognition rates with respect to the abstraction levels under similar conditions.

9 Conclusion

This paper introduced a system for group activity recognition using only wearable and mobile devices for both sensing and recognition purposes. The problems of multi-user (MAR) and group activity recognition (GAR) where defined and further classified into cooperative and collaborative problems, where information exchange between nodes was analyzed. An experiment was designed to investigate the effects of the abstraction level for information exchange on energy consumption and recognition rates.

The experiment was conducted in an office scenario where nodes attached to mugs were used to monitor user's activities and perform collaborative group activity recognition on a mobile phone. Different levels of context preprocessing at the mugs were examined and evaluated in terms of power consumption and activity recognition rates. Specifically, using raw data, signal features, locally classified single-user activities and local clustering were examined as the basis for GAR and evaluated in terms of the cost of transmission incurred as well as GAR rates. The dataset was presented as part of the scientific contribution of this work.

Results indicate that for the given set of activities, the optimal recognition was achieved using locally extracted features, with GAR accuracy of 96 % and a 10 % drop in the amount of energy consumed for the purpose of wireless communication. Locally classifying activities and using these to classify the global group activity reduced power consumption by a further 33 % to 40 % total, but incurred a 47 % drop in global multi-user GAR rates due to subjects' inability to recreate their own behavior under different conditions. Using local hard clustering showed potential by maintaining the reduced power consumption at 40 %, but still incured a recognition drop of 20 %.

The investigations presented here into soft clustering for GAR showed two major insights. First, probabilistic soft clustering using Gaussian mixtures can be used to tweak the tradeoff between accuracy and power consumption of the GAR application. Second, by reducing power consumption (here only twice the communication volume compared to local activities, saving approximately 36 %) and maintaining high GAR accuracy (only a loss of 2.8 %), probabilistic soft clustering represents a method for tackling the doubly-labeling issue which is intrinsic in GAR.

References

1. Bao L, Intille SS (2004) Activity recognition from user-annotated acceleration data. In: Pervasive, pp 1–17
2. Beigl M, Gellersen H-W, Schmidt A (2001) Mediacups: experience with design and use of computer-augmented everyday. Comput Netw 35:401–409
3. Berchtold M, Budde M, Gordon D, Schmidtke H, Beigl M (2010) ActiServ: activity recognition service for mobile phones. In: ISWC'10: Proceedings of the 14th international symposium on wearable computers. IEEE Computer Society, Seoul, pp 83–90
4. Bishop CM (2006) Pattern ecognition and machine learning. Springer, 1st edn. Corr. 2nd printing edn, October 2006
5. Blanke U, Schiele B (2009) Daily routine recognition through activity spotting. In: Proceedings of the 4th international symposium on location and context awareness, LoCA '09. Springer, Berlin, pp 192–206
6. Blumer H (1951) Collective behavior. Irvington Publishers
7. Boin A, McConnell A (2007) Preparing for critical infrastructure breakdowns: the limits of crisis management and the need for resilience. JCCM 15(1):50–59
8. Chang M-C, Krahnstoever N, Lim S, Yu T (2010) Group level activity recognition in crowded environments across multiple cameras. In: IEEE conference on advanced video and signal based surveillance. IEEE Computer Society, Los Alamitos
9. Dong W, Lepri B, Cappelletti A, Pentland AS, Pianesi F, Zancanaro M (2007) Using the influence model to recognize functional roles in meetings. In: Proceedings of the 9th international conference on multimodal interfaces, ICMI '07. ACM. New York, pp 271–278
10. Dunkels A, Grönvall B, Voigt T (2004) Contiki—a lightweight and flexible operating system for tiny networked sensors. In: Proceedings of the 1st IEEE workshop on embedded networked sensors (Emnets-I). Tampa, Florida, USA
11. Gong S, Xiang T (2003) Recognition of group activities using dynamic probabilistic networks. In: Proceedings of the 9th IEEE international conference on computer vision, ICCV '03, vol 2. IEEE Computer Society, Washington, DC, p 742
12. Gordon D, Czerny J, Miyaki T, Beigl M (2012) Energy-efficient activity recognition using prediction. In: 16th international symposium on wearable computers (ISWC), pp 29–36
13. Gordon D, Hanne J-H, Berchtold M, Miyaki T, Beigl M (2011) Recognizing group activities using wearable sensors. In: 8th annual international mobile and ubiquitous systems: networking services, MobiQuitous. MobiQuitous '11
14. Gordon D, Hanne J-H, Berchtold M, Miyaki T, Beigl M (2011) An experiment in hierarchical recognition of group activities using wearable sensors. In: Proceedings of the 7th conference on modeling and using context (CONTEXT). Springer, pp 108–114
15. Gordon D, Scholz M, Ding Y, Beigl M (2011) Global peer-to-peer classification in mobile ad-hoc networks: a requirements analysis. In: Proceedings of the 7th conference on Modeling and using context (CONTEXT). Springer, pp 108–114
16. Gu T, Wu Z, Wang L, Tao X, Lu J (2009) Mining emerging patterns for recognizing activities of multiple users in pervasive computing. In: The 6th annual international conference on mobile and ubiquitous systems, pp 1–10
17. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The weka data mining software: an update. SIGKDD Explor Newsl 11:10–18
18. Hsu JY-J, Lian C-C, Jih W-R (2011) Probabilistic models for concurrent chatting activity recognition. ACM Trans Intell Syst Technol 2:4:1–4:20
19. Hwang I, Jang H, Park T, Choi A, Lee Y, Hwang C, Choi Y, Nachman L, Song J (2012) Leveraging childrens behavioral distribution and singularities in new interactive environments: study in kindergarten field trips. In: 10th international conference on pervasive computing
20. Li R, Chellappa R, Zhou SK (2009) Learning multi-modal densities on discriminative temporal interaction manifold for group activity recognition. In: IEEE conference on computer vision and pattern recognition, CVPR 2009, pp 2450–2457
21. Loy CC, Xiang T, Gong S (2010) Time-delayed correlation analysis for multi-camera activity understanding. Int J Comput Vision 90(1):106–129
22. Roggen D, Wirz M, Tröster G, Helbing D (2011) Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods. Networks 6(3):521–524
23. Schmidt A (2003) Ubiquitous computing—computing in context. PhD thesis, Lancaster University
24. Scholl PM, Van Laerhoven K, Gordon D, Scholz M, Berning M (2012) Jnode: a sensor network platform that supports distributed inertial kinematic monitoring. In: 9th international conference on networked sensing systems (INSS)
25. Sigg S, Gordon D, von Zengen G, Beigl M, Haseloff S, David K (2011) Investigation of context prediction accuracy for different context abstraction levels. IEEE Trans Mob Comput 11(6):1047–1059
26. Stäger M, Lukowicz P, Tröster G (2007) Power and accuracy trade-offs in sound-based context recognition systems. Pervasive Mob Comput 3:300–327
27. Sumpter DJT (2006) The principles of collective animal behaviour. Phil Trans R Soc B 361(1465):5–22
28. Wang L, Gu T, Tao X, Chen H, Lu J (2011) Recognizing multi-user activities using wearable sensors in a smart home. Pervasive Mob Comput 7(3):287–298
29. Want R, Hopper A (1992) Active badges and personal interactive computing objects. IEEE Trans Consumer Electron 38(1):10–20
30. Wirz M, Roggen D, Tröster G (2009) Decentralized detection of group formations from wearable acceleration sensors. In: IEEE on proceedings of the international conference on computational science and engineering, pp 952–959