

# Group Activity Recognition Using Belief Propagation for Wearable Devices

Dawud Gordon

TwoSense Labs

Karlsruhe, Germany

dawud@twosense-labs.com

Markus Scholz, Michael Beigl

Karlsruhe Institute of Technology

Karlsruhe, Germany

[firstname.lastname]@kit.edu

## ABSTRACT

Humans are social beings and spend most of their time in groups. Group behavior is emergent, generated by members' personal characteristics and their interactions. It is therefore difficult to recognize in peer-to-peer (P2P) systems where the emergent behavior itself cannot be directly observed. We introduce 2 novel algorithms for distributed probabilistic inference (DPI) of group activities using loopy belief propagation (LBP). We evaluate their performance using an experiment in which 10 individuals play 6 team sports and show that these activities are emergent in nature through natural processes. Centralized recognition performs very well, upwards of an F-score of 0.95 for large window sizes. The distributed methods iteratively converge to solutions which are comparable to centralized methods. DPI-LBP also reduces energy consumption by a factor of 7 to 40, where a centralized unit or infrastructure is not required.

## Author Keywords

Group activity recognition; computational social sciences; peer-to-peer; wearable computing; mobile computing;

## ACM Classification Keywords

I.2.11 Distributed Artificial Intelligence: Miscellaneous

## 1. INTRODUCTION

Human beings are social creatures, and as such we spend most of our time in groups [19]. Groups are better than individuals at accomplishing tasks, which is often why they are formed in the first place [7]. Understanding group behavior and context is then crucial for intelligent environments and assistive technologies through a process called group activity recognition (GAR) [8]. GAR has the potential to enable crowd management systems to be aware of and adapt to crowd behavior in real time, potentially saving lives in crowd emergencies [12]. However infrastructure is often the first casualty in emergency situations [4], motivating the research of P2P methods. These methods can also alleviate bandwidth bottlenecks due to the "curse of dimensionality," and improve privacy for social and smart environment applications by keeping behavioral data within the group. Mobile

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

ISWC'14, September 13-17 2014, Seattle, WA, USA

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-2969-9/14/09...\$15.00.

<http://dx.doi.org/10.1145/2634317.2634329>

devices such as Smart Phones present an attractive platform both for human activity recognition (HAR) and the recognition of group activities. The behavior of the group is emergent behavior, generated by the personal characteristics of the individual members and the group dynamic [16, 7]. Kurt Lewin, a pioneer of modern social psychology, uses the term "emergence" to signify that the properties of the behavior of the group are fundamentally different than the properties of the behavior of the individuals, or of the "sum" of those behaviors [16], a definition which we follow.

P2P GAR is therefore a challenging and interesting problem because the group behavior cannot be observed by any single device and can only be recognized by fusing distributed observations. In a parallel publication we address the problem of detecting group affiliation using wearable sensor data and P2P algorithms [10]. Here, once the members of the group have been identified, we now wish to recognize the behavior of that group. We present novel methods for GAR based on distributed probabilistic inference (DPI) combined with loopy belief propagation (LBP) [20], and compare the results with centralized approaches. For each group activity, the behavior is broken down into individual clusters using unsupervised methods. Each node estimates its belief over its local clusters for all group activities given current sensor observations, and communicates this information to its neighbors. All nodes iteratively update and re-communicate their beliefs based on the belief estimates received, and a model of individual-to-individual group dynamics. The network then iterates and converges towards a response estimation. We present two methods for LBP, one using linear regression over soft posterior probabilities of user behaviors (SLBP), and one using expectations of the most probable behavior (HLBP).

The novel algorithms are evaluated using an experiment in team sports with 10 subjects playing 6 different sports using Android phones. The experiment naturally creates emergent group behavior and the algorithms are evaluated in terms of recognizing that behavior (which sport is performed). Centralized inference of emergent group behavior when presented with the complete set of group sensor data is relatively straight-forward, approaching an F-score of 0.81 for a window length of 2 seconds and 0.96 for 10 seconds. However inference using solely the data of each subject independently is poor at around 0.55 for the same window. The DPI-SLBP approach begins at iteration 0 at the same value as with individual subject inference over 2s, but then rapidly improves with each iteration, surpassing the centralized naive Bayes approach after three iterations and converging to an F-score

of 0.84 after about 10 iterations. DPI-SLBP increases in the amount of local memory consumed and processing required, but reduces the amount of energy required overall for classification by a factor of almost 7. The simplified DPI-HLBP algorithm performs similarly but converges to a lower value of 0.81, just under the centralized approach. However DPI-HLBP requires less energy for classification than DPI-SLBP by a factor of more than 6, or 40 times less than that required by the centralized approach.

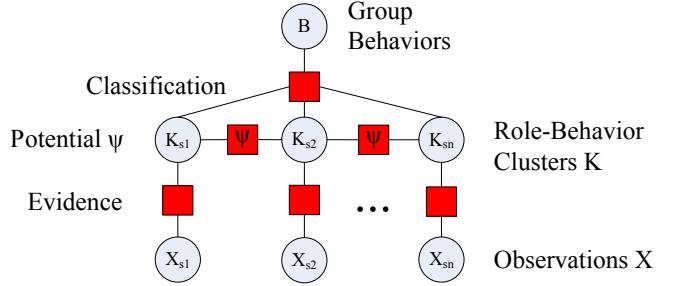
## 2. RELATED WORK

Inferring single-user activities from the distributed behavior of the body is also an emergent recognition problem [2]. However limbs don't change roles, and their interactions with each other are mechanical in nature, presenting a simpler problem than GAR. Multi-user activity recognition (MAR) is the process of recognizing the activities of multiple individuals in parallel [8]. Subjects may be interacting with each other or even members of the same group. Some work combines MAR and GAR where some labels are individual activities and other emergent group behavior [23, 13]. However these approaches do not address the problem from a decentralized standpoint. Distributed approaches to context recognition have been introduced [25] but cannot be applied to non-emergent phenomena. Distributed probabilistic inference has been effective for distributed sensor calibration [21] which is in itself a distributed phenomena, but loopy belief propagation does not converge, requiring a complex networking architecture for clique structuring and belief propagation.

Video systems present an advantage as they are able to view both individual and emergent group behavior simultaneously [5], as well as properties of individual roles [17]. However such approaches require infrastructure for communication and processing. Microphones also provides insight into the group activity [14] and allow extraction of certain types of role information [6]. While audio data is powerful for recognizing social context, it cannot directly sense physical behavior and scales with the number of users poorly. Centrally monitoring location has also been shown to give insight into emergent properties of larger groups or crowds [24]. Adding motion sensors also allows properties such as affiliation of users to each other and to groups [22].

Emergent behavior has also been addressed in the separate but related field of swarm intelligence in animals and insects [11]. Here the problems addressed usually have one of three different goals, either looking to simulate the emergent group behavior based on models of individuals (generation) [15], discover the rules governing individuals based on the emergent behavior produced (discovery) [18], or evaluate the correctness of assumptions about the relationship between local agents and emergent group behavior (evaluation) [15]. Our approach here differs from this field because we wish to **predict the emergent group behavior** based on observations of agents (humans) who are far too complex to model using expert knowledge.

## 3. CONCEPTS AND APPROACH



**Figure 1.** Factor Graph for DPI-LBP with Evidence  $p(K, X)$ , Potential  $\psi(K_{s_i}, K_{s_j})$  (for All  $i, j$ , Some Omitted), and Classification  $p(B, K)$

In this section we begin with the fundamental principles which govern group behavior from the fields of group dynamics and social psychology. Inspired by these abstract models and theories, we construct concrete methods for modeling and classifying group behavior in a probabilistic fashion.

### 3.1 From Field Theory to Probability Theory

Kurt Lewin's Field Theory states that an individual's behavior is dependent on their personal characteristics  $c \in C$  and the social environment of the group  $E$  [16]. The behavior of the individual is also a function of their role in the group [3], which we will denote as  $\rho \in R$ . For GAR we require probability  $p$  of all group behaviors  $B$  given the contributing factors, namely the personal characteristics and roles of the individuals:  $p(B) = p(B|C, R)$ .

GAR is focused on physical behavior, and we can make observations of those characteristics using sensors. We use the notation  $x_s^\tau$  to indicate a single observation, or observations over a window, for subject  $s$  at time  $\tau$ .  $X_s$  refers to all observations for subject  $s$ ,  $X^b$  refers to the evidence of all subjects for a single group activity, and  $X$  is the complete set of observations for all subjects and activities. We can now replace  $C$  with  $X$ , or  $p(B|X, R)$ .

Identifying the role of an individual requires explicit labeling of those roles by experts in that activity and social psychology. The evidence  $X$  is conditionally dependent on both the individuals characteristics, and the role of the individual in the group [7]. To avoid the need for specific role labeling, we cluster the evidence into clusters  $\kappa \in K$ , where  $\kappa_s^b$  is a cluster from subject  $s$  generated by group behavior  $b$  and their role  $\rho_s^b$  in that behavior:  $K = \forall_{b \in B} \forall_{s \in G} \text{clust}(X_s | b)$ . Each cluster now indicates a modality of behavior of an individual who has some unknown role in the group behavior. We refer to these as "role-behavior" clusters.

### 3.2 Modeling and Classifying Group Activities

The clustering approach is probabilistic using Expectation Maximization [20]. For each group activity and subject,  $X$  is separated into  $X_s^b$  and then clustered, yielding clusters  $K_s^b$ . The probability density function (PDF, or  $P$ ) of the clusters for a subject and group activity is given by a Gaussian mixture model (GMM) [7]:

$$P(X_s^b | K_s) = \sum_{\kappa_s^b \in K_s^b} \pi_{\kappa_s^b} \mathcal{N}(X_s | \mu_{\kappa_s^b}, \Sigma_{\kappa_s^b}) \quad (1)$$

Each node  $s$  has clusters  $K_s$  where each cluster  $\kappa_s$  is generated by a certain group behavior  $b$ , giving a subset of clusters for each group activity  $\kappa_s^b \in K_s^b$ . These clusters now build the evidence function for inference of group activities. The posterior probability distribution  $p(K|X)$  can be obtained using Bayesian inference, where each posterior is normalized using the following equation:

$$p(\kappa_s^b|x_s^\tau) = \underbrace{\text{Post}(\kappa_s^b|x_s^\tau)}_{\text{GMM posterior}} \frac{\text{like}(K_s^b|x_s^\tau)}{\sum_{b' \in B} \text{like}(K_s^{b'}|x_s^\tau)} \quad (2)$$

Here posteriors are generated over the Gaussian mixtures for each class  $K_s^b$  given an observation  $x_s^\tau$ , after which the posterior distribution is normalized by the likelihood of all activity cluster models for that subject. Both the likelihood of a GMM and the posterior of a cluster given an observation are obtained by applying Bayesian inference and the Law of Total Probability [20]. Due to the normalization in Eq. (2), the resulting probability distribution over all clusters for all activities for each subject ( $K_s$ ) sums to 1. This step is later necessary for nodes to learn relative probability distributions of neighboring nodes based on histories of these distributions generated by observations. Classification, or a behavioral estimator  $\hat{b}$ , of the current group activity at any point in time for a single subject is given by Eq. (3).

$$\begin{aligned} \hat{b}(x_{s_i}^\tau) &= \underset{b}{\operatorname{argmax}} p(b|K_{s_i}, x_{s_i}^\tau) \\ &= \underset{b}{\operatorname{argmax}} \sum_{\kappa_{s_i}^b \in K_{s_i}^b} \text{Post}(\kappa_{s_i}^b|x_{s_i}^\tau) \end{aligned} \quad (3)$$

Intuitively, each node classifies the group activity based on the most likely role-behavior cluster in the most likely activity. The classification approach of evaluating local posteriors using local evidence (Eq. (2)) can be used to evaluate the ability of a single node to infer the group activity based on local observations alone, which we call the **independent local inference (ILI)** method.

Two methods of inferring group behavior centrally are examined for comparison with the distributed approaches. The first is Bayesian inference using the complete posterior distributions over  $K$  given  $X$ . For this purpose,  $\bar{\xi}$  is constructed:

$$\bar{\xi}^\tau := \forall_{s \in G} \forall_{k_s \in K_s} \text{append}(\text{Post}(K_s|x_s^\tau)) \quad (4)$$

For each time-step  $\tau$ ,  $\bar{\xi}^\tau$  is then a vector of the complete normalized posters across  $K$ . Using this set as observations, a naive Bayesian classifier is constructed to model  $P(\bar{\xi}|B)$  and then to infer  $p(B|\bar{\xi}^\tau)$  for each time-step  $\tau$ . This method is referred to as **centralized cluster-based inference (CCI)**. The second centralized method is a more traditional Bayesian inference using the observations directly. Here  $P(X|B)$  is modeled as a GMM using the Expectation Maximization (EM) algorithm [7], and  $p(B|X)$  can then be inferred, referred to hereon as **centralized naive Bayes (CnB)**.

We approach the distributed problem using DPI with LBP. The missing information sampled by other nodes which is

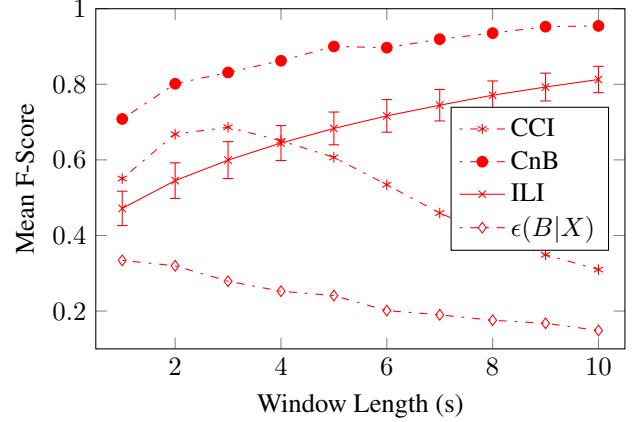


Figure 3. Performance for Centralized Algorithms, Distributed Independent Classification and the Resulting Degree of Emergence

necessary in order to infer the emergent behavior is propagated through the network in the form of beliefs. Non-loopy belief propagation has the advantage of being guaranteed to converge to a solution, but requires extensive networking support to avoid cyclical propagation and to factor the prior [21]. Loopy belief propagation (LBP) is simpler to implement, but whether or not it converges for a specific type of problem is unclear, and how many iterations are required is unknown. These questions are investigated here and as we will see in Sec. 5 the system does converge relatively quickly. The equation for DPI-LBP is given in Eq. (5).

$$p(K|X) = \prod_{s_i \in G} \underbrace{p(K_{s_i}|X_{s_i})}_{\text{local evidence}} \prod_{s_j \neq i \in G} \underbrace{\psi_{i,j}(K_{s_i}, K_{s_j})}_{\text{potential function}} \quad (5)$$

The potential function  $\psi$  can be any positive function which defines the relationship between the variables at subject  $s_i$  and  $s_j$ . Intuitively,  $\psi_{i,j}$  allows node  $i$  to tell node  $j$  what it believes about  $j$ 's behavior based on what  $i$  believes about its own behavior (Fig. 1). For this function we used linear regression [20] to model the relationship between the variables of each pair of subjects, or  $K_{s_i}$  and  $K_{s_j}$ . As stated before, the evidence function is trained using EM for unsupervised clustering of each subjects data for each group activity. Each potential function is trained using linear regression from the variables  $K_{s_j}$  of other subjects to each cluster  $\kappa_{s_i}$  separately. The resulting linear mapping takes the form:

$$\psi_{i,j} = \forall_{\kappa_{s_i} \in K_{s_i}} \forall_{s_j \neq i \in G} \alpha + [\beta_1, \beta_2, \dots, \beta_n] \times [p(K_{s_j})] \quad (6)$$

Where  $[p(K_{s_j})]$  is a column vector of all cluster posteriors  $\kappa_{s_j} \in K_{s_j}$ . This method we call **DPI with soft LBP (DPI-SLBP)** due to the “soft” posterior probability distributions which are mapped in  $\psi$ .

Each iteration consists of a local inference step followed by several update and classification steps. In the inference step, each node  $s_i$  generates a posterior distribution over its clusters using its local evidence function from Eq. (5), creating an



Figure 2. Device Placement (Left) for the Team Sports Group Activity Experiment (Middle) with Relative Subject Locations on the Field (Right)

initial estimate of the group activity based only on local behavior estimates. In the first update step, this information is propagated to all neighboring nodes  $s_j$ , i.e. all nodes within range of one-hop communication. These nodes then convert this estimation of the posterior probability distribution over  $K_{s_i}$  to a belief over  $K_{s_j}$  using  $\psi$ . These beliefs are then combined with the current beliefs of node  $s_j$  over  $K_{s_j}$  and the resulting classification of the group behavior is reevaluated using Eq. (3) in the classification step. The update and classification steps are then repeated until the network is satisfied that convergence has been reached, where we will empirically evaluate how many update steps are required in Sec. 5.

We also present a simplified version of the aforementioned DPI with LBP approach. SLBP requires each node to broadcast its posterior  $Post(K_s|X_s)$  to all neighboring nodes. Probabilistic classification works on the assumption that the most likely model given specific evidence is the correct model for that instance. Based on this assumption, the most valuable information  $p(K_s|X_s)$  is the most likely cluster in the most likely activity, namely  $\text{argmax}_{k_{s_j}^b} p(k_{s_j}^b)$ . We present a simplified method where beliefs are calculated using only this information, instead of the full cluster posteriors  $p(K_s|X_s)$ . This simplified method takes the same form as Eq. (5) with a modified potential function presented in Eq. (7).

$$\psi_{ij}^{\text{simp.}} = p(K_{s_i} | \text{argmax}_{k_{s_j}^b} p(k_{s_j}^b)) \quad (7)$$

We refer to this method as **DPI with hard LBP (DPI-HLBP)** due to the hard role-behavior classification in the potential function.

For Lewin, emergent behavior is only observable given the whole, and not when observing the individuals. We define a metric for activity recognition which expresses this. The “**degree of emergence**”  $\epsilon$  is the proportional increase of activity recognition scores given all data (CnB), to the mean of activity recognition of all nodes using their local observations (IL), quantified using the F-score.

$$\epsilon(B|X) = \frac{\text{F-score}(\hat{b}(B|X)) - \frac{\sum_{s \in G} \text{F-score}(\hat{b}(B_s|X_s))}{|G|}}{\text{F-score}(\hat{b}(B|X))} \quad (8)$$

This measure is dependent on and specific to the models used, the subjective observations (labels), and only for the behavior recognition problem, and does not necessarily be general-

ized over these parameters, other definitions of emergence, or other recognition problems.

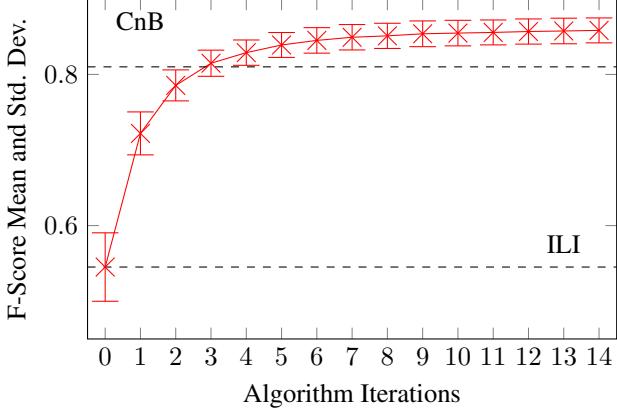
#### 4. EXPERIMENT AND PROCEDURE

To evaluate the approach detailed in the preceding section we constructed an experiment with emergent group activities. The activities performed were team sports, where the emergent behavior is the sport being played itself, based on the observations of the physical behavior of the individuals. LG Nexus 4 Android devices sampled the accelerometer, gyroscope and magnetometer at 50 Hz, 50 Hz and 20 Hz respectively, as well as GPS location for simulating the communication range only.

The devices were attached at the right hip (see Fig. 2), shown to be effective for activity sensing [2]. 6 different team sports were performed by all subjects: **volleyball, badminton, football (soccer), ultimate Frisbee, touch rugby, and flunk-y-ball** for 10 minutes each. The experiment was conducted outdoors in a 15m by 20m field, and a video recording was made from an elevated standpoint of the experiment. The subjects were made up of 7 males and 3 females, with experience levels averaging 4.5 with a variance of 3.5 on a scale of from (no experience in all sports) to 10 (very experienced in all sports)

The data recorded was synchronized, hold-resampled to 50 Hz, and input into an offline sensor replay mechanism in MATLAB. This sensor data was windowed from 1 to 10 seconds, where the window is advanced by 0.5 seconds each iteration over which features were calculated. For each window length, all models are retrained and reevaluated using the features generated over the windows. 50% of the data is used to train the algorithms, and the other 50% (2-fold cross-validation) for evaluating algorithmic performance after randomization. This creates a more difficult recognition problem than e.g. 10-fold cross validation. The goal is to gain an understanding of the relation between distributed and centralized approaches, where a more difficult recognition problem is advantageous. The features used were the mean and variance of the total acceleration signal, the mean and variance of the azimuth orientation with respect to the subject’s body, and the mean and variance of the rotation around the X and Z axes (see Fig. 2 for orientation). These features calculated for subject  $s$  then represent the observations  $X_s$  of the subject, where  $\tau$  is the last timestamp of a sensor data window.

We then simulated performance for a communication range  $\phi$  of 5m, 10m, 15m, 20m and  $\infty$  sequentially, compared to the



**Figure 4.** DPI-SLBP for a Window of 2 Seconds and Full Inter-Connectivity ( $\phi = \infty$ )

diagonal of the field of 25m which is also a good approximation of the radius of the group. We used the relative Euclidean distance between two subjects  $\text{dist}(s_i, s_j)$  based on their GPS coordinates, and judged them to be able to communicate if  $\text{dist}(s_i, s_j) \leq \phi$ . The timestamp used to evaluate  $\text{dist}(s_i, s_j)$  is the final timestamp of the window  $\tau$ , as this is the point where the network is able to evaluate the distributed evidence functions and communicate beliefs. No multi-hop communication is implemented, simulated or required for the methods presented here.

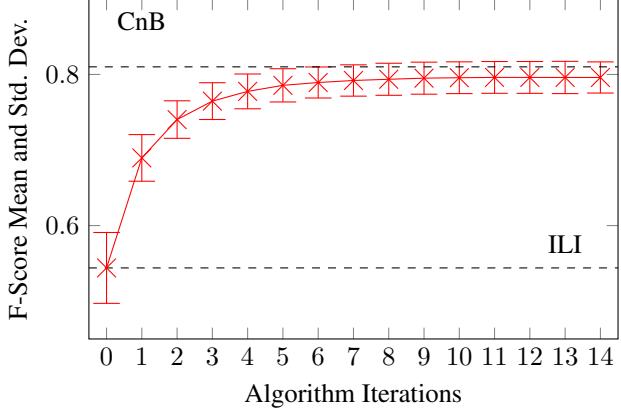
## 5. EVALUATION

### 5.1 Centralized Recognition

To analyze performance of centralized inference using the complete picture of sampled sensor data, we looked the performance of CnB, CCI, ILI and the degree of emergence  $\epsilon$  of this specific problem. In Fig. 3 the performance is shown over varying lengths of the feature analysis window. The CnB algorithm performs the best, with F-scores of 0.71 for a window of 1 second, increasing up to a recognition rate of 0.96 for a window of 10 seconds. For the given scenario and set of conditions, the emergent group behavior can be recognized using relatively straight forward methods, *if* observations of all members of the group are present.

The CCI approach yields an F-score of 0.52 for an observation window length of 1 second, with an optimum of 0.70 for a window length of 3 seconds, after which it subsides towards random classifications with an F-score of 0.31 at 10 seconds. This would appear to indicate that posteriors over role-behavior clusters do not contain the pertinent information required to infer group behavior. However, as we will see later, this is not the case. The implication is therefore only that naive Bayesian inference is not the correct method for inference using these posteriors. Bayesian inference using GMMs separates the data probabilistically using EM for clustering, but the posteriors themselves do not separate well into such clusters.

For a window size of 1 second, the mean F-score of all nodes across all experiments (ILI) was 0.48, with a variance of 0.05. For 10 seconds, the mean increases to 0.82 and variance drops



**Figure 5.** DPI-HLBP for a Window of 2 Seconds and Full Inter-Connectivity ( $\phi = \infty$ )

slightly to 0.03. The longer the time-line of data used to classify the group activity, the better the group activity can be recognized, both for the centralized as distributed evidence functions. Also the quantified emergence of the group activity shrinks with the size of the window from 0.34 for 1 second to 0.16 for 10 seconds.

The reason for the reduction in  $\epsilon$  over time is that sports activities in general are very dynamic in nature, where players change roles rapidly. Over time, a classifier observes the majority of role-behaviors from a single subject, improving classification of the emergent behavior. This effect cannot be generalized to other forms of group activities such as social gatherings or meetings and is specific to the experiment conducted here. For the remaining evaluation of the distributed algorithms, a window of 2 seconds has been selected.

### 5.2 DPI with LBP

The results of DPI with SLBP for a window size of 2 seconds and a communication range of  $\phi = \infty$  are displayed in Fig. 4. The shape of the curve presented demonstrates clearly that the distributed algorithm does indeed converge to a solution. This solution is reached after 15 iterations at an F-score of 0.86 (precision 0.86, recall 0.85). At iteration 0, the lower bound is given by the evaluation of the local evidence functions of each node separately, and corresponds to the value for a window size of 2s in Fig. 3. This value even exceeds the centralized approach at 0.81 (precision 0.83, recall 0.82) after 3 iterations where 95% of convergence, a value of 0.84 is already reached after 6 iterations. It must be noted here that the indication is not distributed inference performs better, but that the potential performance using posteriors over  $K$  is higher than the performance of a nB classifier over  $X$ . The standard deviation across nodes is 0.045 for iteration 0, but drops to 0.027 after one iteration and converges to 0.021.

The results of DPI with HLBP with the same parameters ( $\phi = \infty$ ,  $ws = 2$ ) are shown in Fig. 5. Iteration 0 also begins at the same lower bound as in Fig. 4. A similar convergence is also clearly visible, but convergence occurs at 0.80 (precision 0.82, recall 0.79), as compared to a value of 0.86 for the full potential method. The standard deviation also drops dramat-

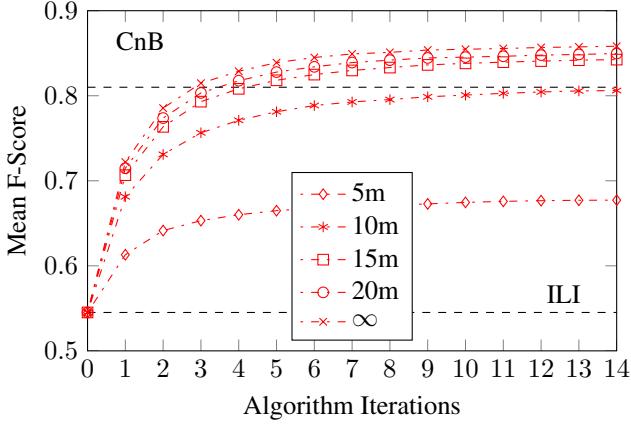


Figure 6. Convergence Curves for DPI-SLBP for Varying Ranges  $\phi$

ically after one iteration from 0.045 to 0.037, and then iteratively converges to 0.031. This value is however greater than the standard deviation of 0.021 for the regression-based potential function. Here again, 95% of convergence is reached fairly quickly after 5 iterations. The effects of the simplified potential function HLBP are clear. Convergence occurs slightly faster (1 iteration less for 95%), but converges to an optimum 7% less than when using SLBP, and the standard deviation across nodes also increases by 68%. However, the reduced F-score and increased standard deviation come with reductions in resource consumption, which can be advantageous for certain applications.

### 5.3 Effects of P2P Communication

The two novel distributed methods were also simulated for various communication ranges. The range  $\phi$  was simulated for 5m, 10m, 15m, 20m, and  $\infty$ , or full connectivity. The mean F-score results for SLBP are displayed in Fig. 6. There the value for  $\phi = \infty$  corresponds to the data in Fig. 4. Mean values for 20m and 15m perform similarly to full connectivity, converging to a value of 0.85 and 0.84 respectively, compared to 0.86 for full connectivity. Reducing communication to 10m and 5m converges to 0.80 and 0.68 respectively.

Similar behavior was also observed for performance using HLBP for the same simulated communication distances in Fig. 7. Communication ranges of 20m and 15m iteratively incur a loss of less than one F-score point, although 95% of convergence requires 6 iterations. At 10m, convergence occurred at an F-score of 0.76 with 95% reached after 8 iterations. Reducing communication further to 5m also required 8 iterations and converged to an F-score of 0.65.

A survey of convergence values for both algorithms after 5 iterations can be seen in Tab. 1, where the coverage is simply the ratio of the of  $\phi$  to the diameter of the group, assumed to be the diagonal of the field 25m. From full connectivity to 15m range there is little effect on the convergence times, although the using the simplified potential function incurred a greater reduction of 4.9 percentage points (pp) as apposed to 2.4 pp for regression-based potentials. This effect is due to the speed of belief propagation for the two algorithms. For

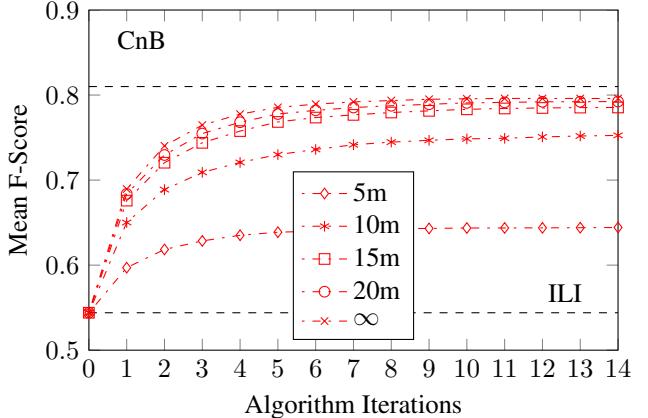


Figure 7. Convergence Curves for DPI-HLBP for Varying Ranges  $\phi$

Table 1. Coverage and Convergence in % After 5 Iterations

| Range    | Coverage (%) | Convergence SLBP (%) | Convergence HLBP (%) |
|----------|--------------|----------------------|----------------------|
| $\infty$ | 100          | 91.2                 | 94.6                 |
| 20m      | 80           | 89.6                 | 91.5                 |
| 15m      | 60           | 88.8                 | 89.7                 |
| 10m      | 40           | 86.5                 | 86.5                 |
| 5m       | 20           | 91.1                 | 91.7                 |

$\psi^{\text{simpl.}}$  the propagation takes more “effort” as a node must receive enough belief contrary to its current state before its internal belief about its most probable cluster changes. For the regression-based approach, this occurs more quickly as beliefs are integrated and propagated in a continuous manner. For these communication ranges, the large majority of nodes are in the same network with occasional disconnection of individuals as they leave the group, e.g. to collect the ball. Hence, only the small changes in recognition rates over these ranges as belief propagates over intermediary nodes throughout the network.

For a communication distance of 10m, both algorithms propagate information quickly, but the network breaks apart into disjoint sub-networks as groups of nodes and individuals are out of range of each other. This is the cause of the reduced recognition rates in Figs. 6 and 7 for a range of 10m, where necessary information cannot propagate to all nodes due to the lack of a link between nodes in different sub-networks. For a range of 5m the network becomes disjunct, and nodes only have one or two other nodes in the same sub-network. The results can be seen clearly in the low convergence rates in Figs. 6 and 7. However, convergence occurs quickly, as beliefs are only propagated to small subgroups of  $G$ .

### 5.4 Resource Consumption Analysis

The resource consumption is only for the classification phase of GAR. The values presented in Tab. 2 are approximations, calculated from the bitrate and power consumption of differ-

ent communication technologies [1]<sup>1</sup>, based on computation time and communication volumes from the simulation. The model assumes an Android Nexus 4 Device with processing on a single core with a consumption of 0.5W. For the DPI algorithms, 10 iterations are assumed which is well over the amount required for 95% convergence presented in Tab. 1. DPI-SLBP reduced power consumption due to communication by 84% compared to CnB, and DPI-HLBP presents a reduction of 97.5%. DPI-SLBP increases response time by a factor of 2.5, although server-side calculations for CnB are not taken into account [8]. DPI-HLBP however reduces the reaction time of the system by 51% with respect to CnB, which is around 5.5 times less than the reaction time of DPI-SLBP. CnB only requires the amount of memory to store 1 window of sensory data. For DPI-SLBP, around 30 times more storage is required or almost 100 kB. DPI-HLBP only requires around 5 times more memory than CnB, representing a reduction of over 83% compared to DPI-SLBP due to the reduced size of the expectation look-up table compared to linear regression mappings. It is important to note that the necessity to communicate with a server or centralized instance is removed for DPI-LBP algorithms.

## 6. DISCUSSION

The large reductions in resource consumption and low convergence time make DPI-HLBP an attractive approach. However, the effect of reducing communication range was more pronounced than for DPI-SLBP. For both algorithms, convergence time increases as the group grows proportional to the communication range (see Tab. 1), but it grows slower for DPI-SLBP than for DPI-HLBP. For applications where the surface area of the group is large proportional to the communication range, e.g. groups or crowds in public areas, propagation rates for DPI-SLBP could be greatly affected. For such applications the indications are that DPI-SLBP is the best approach to take, although performance and scalability to large groups was not evaluated here. Each node is only dependent on neighboring nodes, meaning the approach is very scalable, limited by the time needed for beliefs to propagate.

For small groups such as the one analyzed here, this time is negligible. However if the required response time drops relative to iteration time, the algorithms may not converge. For both algorithms however, it is important that the communication range be proportional to the surface area of the group such that the vast majority of group members are connected to at least one other member by one link, and to all members by at least one multi-hop path so that belief may propagate. In the case of sport activities, this is a range of around 12.5m-15m, or 50% of the surface area of the group. However there are other aspects of GAR which are not addressed here [9]. Group members can come and go over time, leading to changing group sizes and changes in individual and group behavior characteristics. These aspects are outside the scope of this work and are the subject of continuing research [13], for GAR in general and using DPI-LBP.

<sup>1</sup>[http://www.csr.com/sites/default/files/white-papers/comparisons\\_between\\_low\\_power\\_wireless\\_technologies.pdf](http://www.csr.com/sites/default/files/white-papers/comparisons_between_low_power_wireless_technologies.pdf)

## 7. CONCLUSION

Group activities are emergent from the individual characteristics of group members, their roles in the group, and the group dynamic [7]. The group behavior therefore has properties which are different from the properties of the behavior of the individuals, as well as the “sum” of those individual properties [16]. We have shown that the emergent behavior of the group can be inferred using centralized inference methods with F-scores upwards of 95% possible for this scenario. Using clustering to address the problem of inference without explicitly requiring role, we presented two methods of inferring emergent behavior in a distributed fashion, using local estimations (DPI) and exchange of belief estimates (LPB). The first (DPI-SLBP) propagates beliefs based on linear potentials over posteriors from subject to subject. The second (DPI-HLBP) propagates beliefs as expectation of the most likely behavior of an individual.

DPI-SLBP and DPI-HLBP converged to F-scores of 0.84 and 0.80 respectively compared to a centralized inference of 0.81 for the same parameters. Reducing the the communication range to 50% of the diameter of the group only marginally affected the value which the distributed algorithms converged to, as long as the range did not create disjunct networks out of the single group. However it did affect convergence time, where the effect on DPI-HLBP was greater, increasing the number of iterations needed. For larger groups or crowds where local communication range is small in proportion to the surface area of the group, DPI-SLBP is then preferable. However, DPI-HLBP greatly reduces local resource consumption compared to DPI-SLBP, making it attractive for small group applications. HLBP and SLBP increase locally memory consumption though remaining under 100 kB, where the former reduces response time and the latter increases it slightly compared to a centralized system. The conclusion is the following: 1) DPI-LBP over role-behaviors does converge for GAR, 2) the convergence accuracy is dependent on how many nodes are connected to the P2P network, 3) convergence speed is dependent on the degree of connectivity of that network. In total, the distributed approaches allow inference of emergent group behavior using only wearable devices, with implications for crowd emergency management, intelligent environments, and privacy-aware social applications.

## 8. REFERENCES

1. Balasubramanian, N., Balasubramanian, A., and Venkataramani, A. Energy consumption in mobile phones: a measurement study and implications for network applications. In *Proceedings of the Internet measurement conference*, ACM (New York, NY, USA, 2009), 280–293.
2. Bao, L., and Intille, S. S. Activity recognition from user-annotated acceleration data. In *Pervasive* (2004).
3. Benne, K. D., and Sheats, P. Functional roles of group members. *Journal of Social Issues* 4, 2 (1948), 41–49.
4. Boin, A., and McConnell, A. Preparing for Critical Infrastructure Breakdowns: The Limits of Crisis Management and the Need for Resilience. *Contingencies and Crisis Management* 15 (2007), 50–59.

**Table 2. Resource Consumption Analysis for All Algorithms for 1 Classification after 10 Iterations**

| Approach      | Memory Used (kB) | Comm. Per Class. (B) | Comm. Time (ms) | Comm. En. (mJ) | Proc. Time (ms) | Proc. En. (mJ) | Total Time (ms) | Total En. (mJ) |
|---------------|------------------|----------------------|-----------------|----------------|-----------------|----------------|-----------------|----------------|
| CnB (3G)      | 3.6              | 3600                 | 71.53           | 32.47          | 2.13            | 0              | 73.65           | 32.47          |
| CCI (3G)      | 14.48            | 51.5                 | 1.02            | 0.46           | 7.16            | 0              | 8.18            | 0.47           |
| ILI           | 14.48            | 0                    | 0               | 0              | 2.56            | 0              | 2.56            | 0              |
| DPI-SLBP (BT) | 99.35            | 5150                 | 119.5           | 4.57           | 76.31           | 0.38           | 195.81          | 4.96           |
| DPI-HLBP (BT) | 16.6             | 800                  | 18.56           | 0.71           | 17.7            | 0.09           | 36.26           | 0.8            |

5. Chang, M.-C., Krahnstoever, N., Lim, S., and Yu, T. Group level activity recognition in crowded environments across multiple cameras. In *Advanced Video and Signal Based Surveillance, IEEE Conference on*, IEEE (Los Alamitos, CA, USA, 2010).
6. Dong, W., Lepri, B., Cappelletti, A., Pentland, A. S., Pianesi, F., and Zancanaro, M. 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, NY, USA, 2007), 271–278.
7. Forsyth, D. *Group Dynamics*, 4th ed. International student edition. Thomson/Wadsworth, 2006.
8. Gordon, D., Hanne, J.-H., Berchtold, M., Shirehjini, A., and Beigl, M. Towards collaborative group activity recognition using mobile devices. *Mobile Networks and Applications* (2012), 1–15.
9. Gordon, D., Scholz, M., Ding, Y., and Beigl, M. 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 (2011), 108–114.
10. Gordon, D., Wirz, M., Roggen, D., Beigl, M., and Tröster, G. Group affiliation detection using model divergence for wearable devices. In *Proceedings of the 2014 International Symposium on Wearable Computers*, ISWC '14, ACM (New York, NY, USA, 2014).
11. Grünbaum, D., Viscido, S., and Parrish, J. Extracting interactive control algorithms from group dynamics of schooling fish. In *Cooperative Control*, LNCIS. Springer Verlag, 2005, 447–450.
12. Helbing, D., Brockmann, D., Chadeaux, T., Donnay, K., Blanke, U., Woolley-Meza, O., Moussaïd, M., Johansson, A., Krause, J., Schutte, S., et al. How to save human lives with complexity science. Available at SSRN 2390049 (2014).
13. Hirano, T., and Maekawa, T. A hybrid unsupervised/supervised model for group activity recognition. In *Proceedings of the 2013 International Symposium on Wearable Computers*, ISWC '13, ACM (New York, NY, USA, 2013), 21–24.
14. Hsu, J. Y.-J., Lian, C.-C., and Jih, W.-R. Probabilistic models for concurrent chatting activity recognition. *ACM Trans. Intell. Syst. Technol.* 2 (2011), 4:1–4:20.
15. Huth, A., and Wissel, C. The simulation of fish schools in comparison with experimental data. *Ecological Modelling* 75-76, 0 (1994), 135 – 146. State-of-the-Art in Ecological Modelling proceedings of ISEM's 8th International Conference.
16. Lewin, K. *Field theory in social science: selected theoretical papers*. Social science paperbacks. Harper, New York, 1951.
17. Li, R., Chellappa, R., and Zhou, S. Learning multi-modal densities on discriminative temporal interaction manifold for group activity recognition. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (june 2009), 2450 –2457.
18. Lukeman, R., Li, Y.-X., and Edelstein-Keshet, L. Inferring individual rules from collective behavior. *Proceedings of the National Academy of Sciences* 107, 28 (2010), 12576–12580.
19. Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., and Theraulaz, G. The Walking Behaviour of Pedestrian Social Groups and Its Impact on Crowd Dynamics. *PLoS ONE* 5, 4 (2010), 7.
20. Murphy, K. P. *Machine learning: a probabilistic perspective*. MIT Press, Cambridge, MA, 2012.
21. Paskin, M. A., and Guestrin, C. E. Robust probabilistic inference in distributed systems. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, UAI '04, AUAI Press (Arlington, Virginia, 2004).
22. Roggen, D., Wirz, M., Tröster, G., and Helbing, D. Recognition of Crowd Behavior from Mobile Sensors with Pattern Analysis and Graph Clustering Methods. *Networks* 00, 0 (2011), 1–24.
23. Wang, L., Gu, T., Tao, X., Chen, H., and Lu, J. Recognizing multi-user activities using wearable sensors in a smart home. *Pervasive Mob. Comput.* 7, 3 (June 2011), 287–298.
24. Wirz, M., Franke, T., Mitleton-Kelly, E., Roggen, D., Lukowicz, P., and Tröster, G. Coenosense: A framework for real-time detection and visualization of collective behaviors in human crowds by tracking mobile devices. In *Proceedings of European Conference on Complex Systems*, Springer (2012).
25. Wittenburg, G., Dziengel, N., Wartenburger, C., and Schiller, J. A system for distributed event detection in wireless sensor networks. In *IPSN '10: Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks* (2010).