A multilabel classification approach for complex human activities using a combination of emerging patterns and fuzzy sets

Nehal A. Sakr¹, Mervat Abu-ElKheir², A. Atwan³, H. H. Soliman⁴

¹,³,⁴Department of Information Technology, Faculty of Computer and Information Sciences, Mansoura University, Egypt
²Department of Computer Science, Faculty of Media Engineering and Technology, The German University in Cairo, Egypt

ABSTRACT

In our daily lives, humans perform different Activities of Daily Living (ADL), such as cooking, and studying. According to the nature of humans, they perform these activities in a sequential/simple or an overlapping/complex scenario. Many research attempts addressed simple activity recognition, but complex activity recognition is still a challenging issue. Recognition of complex activities is a multilabel classification problem, such that a test instance is assigned to a multiple overlapping activities. Existing data-driven techniques for complex activity recognition can recognize a maximum number of two overlapping activities and require a training dataset of complex (i.e. multilabel) activities. In this paper, we propose a multilabel classification approach for complex activity recognition using a combination of Emerging Patterns and Fuzzy Sets. In our approach, we require a training dataset of only simple (i.e. single-label) activities. First, we use a pattern mining technique to extract discriminative features called Strong Jumping Emerging Patterns (SJEPs) that exclusively represent each activity. Then, our scoring function takes SJEPs and fuzzy membership values of incoming sensor data and outputs the activity label(s). We validate our approach using two different dataset. Experimental results demonstrate the efficiency and superiority of our approach against other approaches.

Keywords:
Complex activity recognition
Fuzzy sets
Multilabel classification
Pattern mining
Pervasive computing

INTRODUCTION

Human Activity Recognition (HAR) is a field in pervasive computing, that is concerned with recognizing the activities performed by a person using data collected from the monitoring devices. In the beginning of HAR, vision-based HAR used cameras to monitor humans within their surrounding environments [1]. The rapid developments of the monitoring devices from cameras to sensors lead to the release of sensor-based HAR [2]. In sensor-based HAR, humans wear sensors or install them in the surrounding environment to track their movements, and environmental conditions. Sensor-based HAR gains a lot of research interest due to the wide range of its application domains in medical, security, and military [3].

In this paper, we are interested in sensor-based HAR in which human activity is comprised of a set of sensor readings occurring within a specific time interval. Each of these readings forms an action that represents an atomic operation such as opening fridge. Within a specific time interval, a person can perform
multiple activities sequentially or in a complex scenario. Sequential (i.e. simple) activities are executed one after the other, while complex activities may be executed in a concurrent or an interleaved manner. In concurrent activities, actions of different activities are executed simultaneously. In interleaved activities, actions of different activities are executed interchangeably. Figure 1 illustrates different types of activity execution with examples.

Figure 1. Simple, interleaved, and concurrent activities with examples

Many research attempts have been introduced for simple activity recognition, but little work targets complex activity recognition and stills a challenging issue. Complex activity recognition is a multilabel classification problem, and a test instance is assigned multiple class labels. For instance, when user is drinking coffee while watching TV, the recognition model should identify these two activities. Existing data-driven approaches for complex HAR have some limitations. First, in order to recognize complex activities, they require a training dataset of complex activities. Consider we have $m$ activities, and then we have $m (m-1)$ possible combinations that form complex ones, which is difficult to obtain. Secondly, existing approaches for complex HAR limit the number of overlapping activities, and cannot recognize more than two overlapping (i.e. interleaved or concurrent) activities.

In this paper, we present a novel multilabel classification approach of complex human activities. In our approach, we can recognize multilabeled (i.e. complex) activities from single-labeled (i.e. simple) activities. This is achieved by making a combination of Emerging Patterns and Fuzzy Sets. First, we use a training dataset of simple activities and apply a pattern mining technique to extract discriminative features that are exclusively present in each activity called Strong Jumping Emerging Patterns (SJEPs). Then, our recognition model use a scoring function to identify activity label(s) using the extracted SJEPs with fuzzy membership values of incoming sensor events (actions).

The rest of paper is organized as follows. In Section 2, we review the previous work proposed for complex HAR. Then, we formulate our problem and introduce the required preliminaries in Section 3. Section 4 introduces the proposed approach for complex HAR. In Section 5, we present our empirical results. Finally, we conclude our work and discuss our plans for future work in Section 6.

2. RELATED WORK

The existing approaches for HAR can be divided into two prominent approaches data-driven and knowledge-driven approaches [4]. In this paper, we are interested in data-driven approaches that require a predefined dataset to build the classification model using machine-learning techniques, and then use the built model to identify unlabeled activities. As a multilabel classification problem, the problem of complex HAR can be handled with three main approaches: data transformation, method adaptation, and ensemble of classifiers [5].

The data transformation approach converts the multilabel classification problem into a number of binary classification problems (e.g. binary relevance, classifier chain), multiclass classification problems (e.g. label power set), or use ensemble methods. For example, in [6] the authors applied a dynamic segmentation approach to split complex activities into multiple simple activities, and then apply the Random Forest technique with a voting mechanism to identify activity labels.
On the other hand, method adaptation approach extends the traditional methods for binary classification to support multilabel classification. The existing attempts are probabilistic graphical models, pattern mining-based models, or are based on time series classification. In [7, 8] the authors extend the traditional Conditional Random Fields (CRFs) to handle complex activities. They proposed a Factorial CRF (FCRF), and a Skip chain CRF (SCCRF) respectively. In [9], the authors presented a modified Hidden Markov Model (HMM) to deal with interleaved activities called Interleaved HMM (IHMM). Also [10] presented a modified Naïve Bayes (NB) called Interval Temporal Bayesian Network (ITBN). Another attempt presented a pattern mining approach to recognize simple, interleaved, and concurrent activities [11], and time-series based classification in [12].

3. PRELIMINARIES

Before going in depth to the proposed model, we formulate the task of sensor-based HAR as follows. Consider we have a set $S$ of sensors attached to human or installed in the environment. Using these sensors to monitor human behavior, we obtain a dataset $D = \{t_1, t_2, ..., t_n\}$ that consists of a set of activity traces $t$. Each trace is as a sequence of triggered sensors (i.e. events), whereas each event is a tuple of $\langle ts, sn, sv, l \rangle$ where, $ts$ represents time stamp, $sn$: sensor name, $sv$: sensor value, $l$: the activity label from the set $A$ of $n$ activity labels. Our main goal is to build a mapping function (MF) that assign new activity traces for simple or complex activities with the correct label(s). Below we list a set of definition required in our proposed approach [13].

Itemset, in our task an item $i$ corresponds to a pair of sensor name and value, such that a set of items forms an itemset $I = \{i_1, i_2, ...\}$.

The support, for an itemset $X$ in dataset $D$, referred as $supp(X, D) = count(X, D)/count(D)$. The Growth rate, for an itemset $X$ from dataset $D_1$ to dataset $D_2$ referred as $GR(X, D_1, D_2)$ is computed as:

$$GR(X, D_1, D_2) = \begin{cases} 0 & \text{if } supp(X, D_1) = 0, \text{and } supp(X, D_2) = 0 \\ \infty & \text{if } supp(X, D_1) = 0 \text{ and } supp(X, D_2) > 0 \\ \frac{supp(X, D_2)}{supp(X, D_1)} & \text{otherwise} \end{cases}$$

Emerging Pattern (EP), a pattern $X$ is EP, $EP(X) = \{X: supp(X, D_2) > \mu, \text{and } GR(X, D_1, D_2) \geq \rho\}$, where $\mu$ is the threshold of minimum support, and $\rho$ the threshold of minimum growth rate. With regard to the task of HAR, EPs represent the discriminative features between activity classes. Therefore, EPs make absolute separation. In addition, the number of the extracted EPs will be extremely high. In order to reduce the number of the extracted EPs, we introduce the MinEPs.

Minimal Emerging Pattern (MinEP), a pattern $X$ is MinEP from $D_1$ to $D_2$. $MinEP(X, D_1, D_2) = \{X: GR(X, D_1, D_2) \geq \rho, \text{and } \exists Y \subseteq X \text{ and } GR(Y, D_1, D_2) \geq \rho\}$.

Jumping Emerging Pattern (JEP), a pattern $X$ is JEP from $D_1$ to $D_2$. $JEP(X, D_1, D_2) = \{X: GR(X, D_1, D_2) = \infty\}$.

Strong Jumping Emerging Pattern (SJEP), a pattern $X$ is SJEP from $D_1$ to $D_2$. $SJEP(X, D_1, D_2) = \{X: GR(X, D_1, D_2) = \infty, \text{and } \exists Y \subseteq X \text{ and } GR(Y, D_1, D_2) = \infty\}$. Our main goal is the recognition of complex human activities using a training dataset of simple activities, so, we need discriminative features that are exclusively present in each activity class. This is achieved with the specific type of EPs called SJEPs that combine the discriminative characteristics of JEPs and MinEPs. A detailed discussion about different types of EPs with example will be found in [14].

4. THE PROPOSED APPROACH

In this paper, we present a novel approach for the recognition of simple and complex human activities using a combination of Emerging Patterns and Fuzzy Sets. This is accomplished in two main phases training and testing as shown in Figure 2. In the first phase, we use a training dataset of simple activities and produce; the discriminative SJEPs for each activity class, besides sensor/activity correlation matrix. Then, in the second phase, we segment the incoming sensor data, and then each segment is assigned with the corresponding label(s) using our constructed scoring function. Below, we discuss these two phases in details.
4.1. The training phase

The main goals of this phase are the mining of the SJEPs, computation of the discriminative score of SJEPs, and computation of the sensor/activity correlation matrix required at the testing phase.

4.1.1. Mining the strong jumping emerging patterns

From the characteristics of SJEPs introduced before, they make one-versus-all absolute separation (i.e. separate one activity class from the remaining classes). Given a pre-processed dataset of trace sequences for simple activities (i.e. represented as a pairs of sensor name and value), Algorithm 1 outputs the corresponding SJEPs.

Algorithm 1. Mining of Strong Jumping Emerging Patterns

**Input:** Preprocessed dataset of observation sequences for the simple activities

**Output:** The set of SJEPs.

**Steps:**
1. Begin
2. For each activity class \( a_i, i = 1, \ldots, k \)
3. Consider dataset \( D_i \) contains observation sequences from class \( a_i \), and \( D'_i \) contains observation sequences from \( k - 1 \) classes, \( D'_i = \bigcup_{j=1}^{k} D_j, i \neq j \)
4. Compute ItemSet_List of all itemsets in dataset \( D_i \) using the FP-growth algorithm [15]
5. For each candidate itemset in ItemSet_List,
6. Compute supp(candidate, \( D_i \))
7. Compute supp(candidate, \( D'_i \)).
8. Compute GR(candidate, \( D_i, D'_i \))
9. If \( GR(candidate, D_i, D'_i) = \infty \)
10. Add candidate to the set of SJEPs for class \( a_i \)
11. End If
12. End For
13. End For
14. End

4.1.2. Computation of correlation matrix

The other feature computed in the training phase is the sensor/activity correlation matrix. For each sensor \( s \) from the set of sensors \( S = \{ s_1, s_2, \ldots \} \) used for monitoring activity \( a \) from the set of activities \( A = \{ a_1, a_2, \ldots a_n \} \), we compute a correlation matrix that measures the interdependency between each sensor and activity using the support \( supp(s, D_a) \) defined in Section 3.

4.2. The testing phase

The input to this phase is a stream of sensor readings for the performed human activities to be recognized. In order to identify these activities, the input stream should be segmented beforehand. Regarding trace segmentation, there exist two main approaches, fixed-size segmentation, and dynamic segmentation. In this paper, we are interested in fixed-size segmentation that uses a fixed-length sliding window to split the given trace into equal-sized time segments, which add the following challenge. The size used for segmentation is considered a challenge, because activities occur with different durations. As a result, activity
can be distributed among a number of segments. To resolve this issue, while computing score at specific segment, we take into account the score in previous segment with parameter α that is the \( \text{cum} \_\text{score} \) refer (2).

For the task of activity recognition, within a specific test segment the existing activities executed sequentially or in parallel may be fully contained or partially contained. From this perspective, we make use of fuzzy set theory; that is an activity exists with a specific degree of membership value. Using a combination of the SJEPs extracted in the training phase and the fuzzy set theory, we propose a SJEP-based algorithm for the recognition of simple and complex human activities that consist of three main phases as presented in Algorithm 2. First, we compute the matching-score that identifies the candidate activities existing in the current segment refer (1). Then, we compute the likelihood score (like_score) for each candidate activity using a scoring function refer (3). Finally, using the computed likelihoods and a predefined threshold, we make a decision on the winning activities from these candidates using the following rule.

“If \( \text{cum} \_\text{score}(a_i, s_t) \geq \lambda \times \text{mean(like} \_\text{score)} \) Then output \( \text{a}_t \)”, where \( \text{a}_t \) refers to the candidate activity, \( s_t \) current trace segment, and \( \lambda \) a user-defined threshold used for tuning the threshold.

**Algorithm 2.** The Proposed Recognition Algorithm.

**Input:** A set of preprocessed trace segments: \( S = \{s_1, s_2, \ldots, s_t\} \). The extracted SJEPs, \( \text{max} \) correlation matrix

**Output:** Winning Activity label(s) for each segment: \( \text{Winning} \)

**Steps**

1. Begin
2. For each segment \( s_t \in S \) preprocessed as a sequence of \( l \) items \( s_t = \langle e_1, e_2, \ldots, e_l \rangle \)
3. For each SJEP in the sets SJEPs
4. Compute \( \text{match} \_\text{score}(S_t, \text{SJEP}) \)
5. If \( \text{match} \_\text{score} > 0 \)
   6. Label = getActivityLabel(SJEP)
   7. Add Label to candidates
8. End If
9. End For
10. For each candidate activity \( a_i \) in candidates with cardinality \( k \)
11. For each item \( e \) in the current segment \( S_t \) with cardinality \( l \)
   12. Compute fuzzy_membership(\( e, a_i \))
13. End For
14. Compute like_score \( (a_i, s_t) \), and \( \text{cum} \_\text{score}(a_i, s_t) \)
15. End For
16. mean_score = mean(like_score)
17. For each activity in candidates
18. If \( \text{cum} \_\text{score}(a_i, s_t) \geq \lambda \times \text{mean} \_\text{score} \)
19. Add \( a_i \) to \( \text{winning} \( S_t \) \)
20. End If
21. End For
22. End For
23. End

**Definition 4.1:**

The matching score (match_score) of the set \( S_t \) of SJEPs with the set \( S_2 \) of sensor events within a specific test is defined as follows

\[
\text{match} \_\text{score}(S_t, S_2) = \begin{cases} 
0 & \text{if } S_t \cap S_2 = \emptyset \\
1 & \text{if } S_t \cap S_2 = S_2 \\
\frac{|S_t \cap S_2|}{|S_2|} & \text{otherwise} 
\end{cases} \quad (1)
\]

\[
\text{cum} \_\text{score}(a_i, s_t) = \alpha \times \text{like} \_\text{score} (a_i, s_{t-1}) + (1 - \alpha) \times \text{like} \_\text{score} (a_i, s_t) \quad (2)
\]

\[
\text{like} \_\text{score} (a_i, s_t) = \sum_{\nu=1}^{l} \text{fuzzy} \_\text{membership}(e_\nu, a_i) \quad (3)
\]

\[
\text{fuzzy} \_\text{membership}(e, a_i) = \frac{\text{supp}(e, a_i)}{\sum_{j=1}^{k} \text{supp}(e, a_j)} \cdot \sum_{j=1}^{k} \text{fuzzy} \_\text{membership}(e, a_j) = 1 \quad (4)
\]
5. EXPERIMENTAL RESULTS AND EVALUATION

In this section, we present the experimental evaluation of the proposed approach. First, we introduce the datasets used, followed with our settings for the evaluation, and then we present and discuss our obtained results. Regarding the datasets used, we evaluate our approach using two datasets with different monitoring approaches; CASAS benchmarking dataset [16], and dataset of simple, interleaved and complex activities refereed as SICA from [11]. CASAS dataset was collected from a smart home equipped with different types of environmental sensors (e.g. motion, door, temperature), in which 20 participants collected a set of simple and complex observation sequences for eight different ADL activities presented in Table 1. On the other hand, SICA contains observation sequences of simple and complex execution of 26 different activities collected by four volunteers using two sets of wearable sensors as listed in Table 2.

Table 1. Activities Collected in CASAS dataset [16]

<table>
<thead>
<tr>
<th>No.</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Take Medicine</td>
</tr>
<tr>
<td>2</td>
<td>Watch DVD</td>
</tr>
<tr>
<td>3</td>
<td>Water Plants</td>
</tr>
<tr>
<td>4</td>
<td>Answer phone</td>
</tr>
<tr>
<td>5</td>
<td>Write Birthday Card</td>
</tr>
<tr>
<td>6</td>
<td>Prepare Meal</td>
</tr>
<tr>
<td>7</td>
<td>Clean</td>
</tr>
<tr>
<td>8</td>
<td>Select an outfit</td>
</tr>
</tbody>
</table>

Table 2. Activities collected in SICA dataset [11]

<table>
<thead>
<tr>
<th>No.</th>
<th>Activity</th>
<th>No.</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>making coffee</td>
<td>10</td>
<td>brushing teeth</td>
</tr>
<tr>
<td>2</td>
<td>making tea</td>
<td>11</td>
<td>washing hands</td>
</tr>
<tr>
<td>3</td>
<td>making oatmeal</td>
<td>12</td>
<td>washing face</td>
</tr>
<tr>
<td>4</td>
<td>frying eggs</td>
<td>13</td>
<td>washing clothes</td>
</tr>
<tr>
<td>5</td>
<td>making a drink</td>
<td>14</td>
<td>ironing</td>
</tr>
<tr>
<td>6</td>
<td>applying makeup</td>
<td>15</td>
<td>eating meal</td>
</tr>
<tr>
<td>7</td>
<td>brushing hair</td>
<td>16</td>
<td>Drinking</td>
</tr>
<tr>
<td>8</td>
<td>Shaving</td>
<td>17</td>
<td>taking medication</td>
</tr>
<tr>
<td>9</td>
<td>Toileting</td>
<td>18</td>
<td>cleaning a dining table</td>
</tr>
</tbody>
</table>

To evaluate our work, we use 3-fold cross validation with CASAS and 10-fold cross validation with SICA dataset, besides the three standard evaluation metrics: Precision, Recall, and F-Measure.

\[
Precision = \frac{\text{number of correctly identified activities}}{\text{total number of detected activities}} \tag{5}
\]

\[
Recall = \frac{\text{number of correctly identified activities}}{\text{total number of expected activities}} \tag{6}
\]

\[
F - Measure = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{7}
\]

Using the two different datasets, and fixed-time segmentation approach with different sliding windows moving around the average window length (L_{avg}), we perform two experiments. In each experiment, first we compare the performance of our approach for simple activity recognition with three popular classifiers, Support vector Machine (SVM), Hidden Markov Model (HMM), and K-Nearest Neighbor (K-NN) as shown in Figure 3. For the recognition of complex activities, we compute the precision, Recall, and F-measure for our approach and compare it against methods in [6] and [11] as depicted in Figure 4 and Figure 5 respectively, where the X-axis refers to the window size and the Y-axis refers to the evaluation metric. Regarding the parameters \( \alpha, \lambda \), we reach the best results with \( \alpha = 0.3, \lambda = 0.1 \) in CASAS dataset, and \( \alpha = 0.3, \lambda = 0.8 \) in SICA dataset.

From the results shown in Figure 3, there is a close competition between our approach, SVM and K-NN except HMM with the lowest value. However, our approach achieves the highest performance for the recognition of simple activities in both datasets. Regarding Precision, Recall, and F-Measure for complex activities, the results shown in Figure 4 and Figure 5 indicate the superiority of our approach against the others in [6, 11] at large windows containing multiple overlapping activities. In Figure 4, using CASAS
A multilabel classification approach for complex human activities using a combination of... (Nehal A. Sakr)

dataset, our approach achieves the highest F-measure rate at window size $1.5 \times L_{avg}$ which is 92.3%, while the other in [6] reaches 81.1%. The same in SICA dataset, we reach 92% for F-Measure, that is almost 10% more than the other in [11].

Figure 3. The comparison Results for Simple Activity Recognition, (a) CASAS Dataset, (b) SICA Dataset

Figure 4. Comparison results of our SJEP-based approach and [6] using CASAS dataset

Figure 5. Comparison results of our SJEP-based approach and [11] using SICA dataset

A multilabel classification approach for complex human activities using a combination of... (Nehal A. Sakr)
6. CONCLUSION

The main objective of this paper is to design a unified framework that recognize both simple and complex activities efficiently and overcome the following challenges: the requirement of a training dataset of complex activities and the ability to recognize a maximum of two overlapping activities. This was achieved using a combination of discriminative features called Strong Jumping Emerging Patterns and Fuzzy sets. The proposed model consists of two main phases training and testing. In the training phase, we extract SJEPs the represent each activity class besides another discriminating feature. While in the testing phase, we use the extracted SJEPs and feature as inputs to a scoring function that outputs the activity labels for the existing activities.

Using two datasets with different monitoring approaches and size for the evaluation, the experimental results indicate the efficiency and superiority of the proposed approach against others. Our proposed approach assumed that all activities are linearly separated and it is designed for single residents. Therefore, in the future, our plans are to extend the proposed model to deal with nonlinear separated activities and multi-residents.

REFERENCES


BIOGRAPHIES OF AUTHORS

Nehal A. Sakr received her B.Sc. in Computer Sciences from Faculty of Computer and Information Sciences, Mansoura University of Egypt in 2007. She received her M.Sc. degree in Information Technology from Faculty of Computer and Information Sciences, Mansoura University of Egypt in 2013. She is currently a PhD student and teaching assistant in Information Technology Department at Faculty of Computer and Information Sciences, Mansoura University.
A multilabel classification approach for complex human activities using a combination of ...

(Nehal A. Sakr)

Mervat Abu-ElKheir received her B.Sc. degree in Information Systems from Mansoura University in 2000. She received her M.Sc. degree in Information Systems from Ain Shams University in 2007. Through an Egyptian-funded joint scholarship, she received her PhD degree from Mansoura University with collaboration from Queen’s University, Canada in 2012. She was an Assistant Professor of Computer and Information Sciences at Mansoura University, Egypt. She has served as a workshop co-Chair for the IEEE Wireless Local Networks and the IEEE Convergent IoT workshop. She is an IEEE member since 2014. She is currently an Associate Professor of Media Engineering and Technology at The German University in Cairo.

A. Atwan received his B.Sc. in Electronics and Communication from Faculty of Engineering, Mansoura University, Egypt in 1988. He received his M.Sc. degree in Electrical Communication from Faculty of Engineering, Mansoura University, Egypt in 1998. He received his PhD degree in Communications from Mansoura University in 2004. He was a Vice Dean of Student Affairs, Faculty of Computer and Information Sciences, Mansoura University. He is currently a visiting Professor at Faculty of Computing and Information Technology, Northern Border University, Saudi Arabia.

H. H. Soliman received his B.Sc. and M.Sc. degrees in 1983 and 1987 respectively from Faculty of Engineering, Mansoura University, Egypt. Through an Egyptian-funded joint scholarship, he received his PhD degree from Mansoura University with collaboration from a University in Germany in 1993. He is currently the Dean of Faculty of Computer and Information Sciences, Mansoura University.