# LSTM with Uniqueness Attention for Human Activity Recognition

Zengwei Zheng[1], Lifei Shi[1,2], Chi Wang[1,2], Lin Sun[1], Gang Pan [2]
1 Hangzhou Key Laboratory for IoT Technology and Application, Zhejiang University City College, Hangzhou, China
2 College of Computer Science and Technology, Zhejiang University, Hangzhou, China
Contact Author: (shilf@zju.edu.cn)

ICANN19
28th International Conference on Artificial Neural Networks

## Introduction

The research of human activity recognition(HAR) aim to identify human activity from the data acquired by sensors on human. In earlier studies, researchers used traditional machine learning methods to deal with this problem. Later, RNN became very popular as it is suitable for processing time-series data. But there are also **problems** with these approaches:

- Traditional machine learning methods' ability to handle raw data and the performance to recognize activity was limited.
- In HAR, only a few parts of time-series data are relevant to human activities. This makes it difficult to capture important parts of human activity.

To solve these problems, we combined attention mechanism with the LSTM, a type of RNNs, so that it can focus on the more important parts. Every human activity consists of many atom motions, such as running and walking. Some of these atom motions are common to both running and walking, and thus not appropriate for distinguishing human activity. There is also motion that only happens when you are running, which is more suitable for classification. The degree to which an atom motion uniquely belong to one single activity is called uniqueness of this motion in this paper. Notice that our definition of atom motion is very broad. Every movement for a short period can be considered as atom motion.

**Objectives**: To design a model that can calculate the uniqueness attention of atom motion so that we can use this attention to focus on the parts of human activities that are more important for classification

## Method

The model we design that can calculate attention is shown in Fig. 1. $x_t$ denotes a sensor reading at time t. We use LSTM to encode these $x_t$ , and for each $x_t$ it output $h_t$. The *genAtt* module will calculate attention weight $att_t$ for each $h_t$ and combine these into $H = \sum_{t=1}^{T} att_t \cdot h_t$. The right elliptic region in Fig. 1 represents the internal structure diagram of its left *genAtt* module.

This generic model don't specify how to calculate attention weights. In order to compare with the attention model proposed below, we specify two way to calculate the attention weight, and thus form two benchmark models: **LSTM with Last Attention**, and **LSTM with Mean Attention**. LSTM with Last Attention only sets the attention weight of the last vector to 1 and the rest to 0. LSTM with Mean Attention gives the same attention weight to all of these vectors.

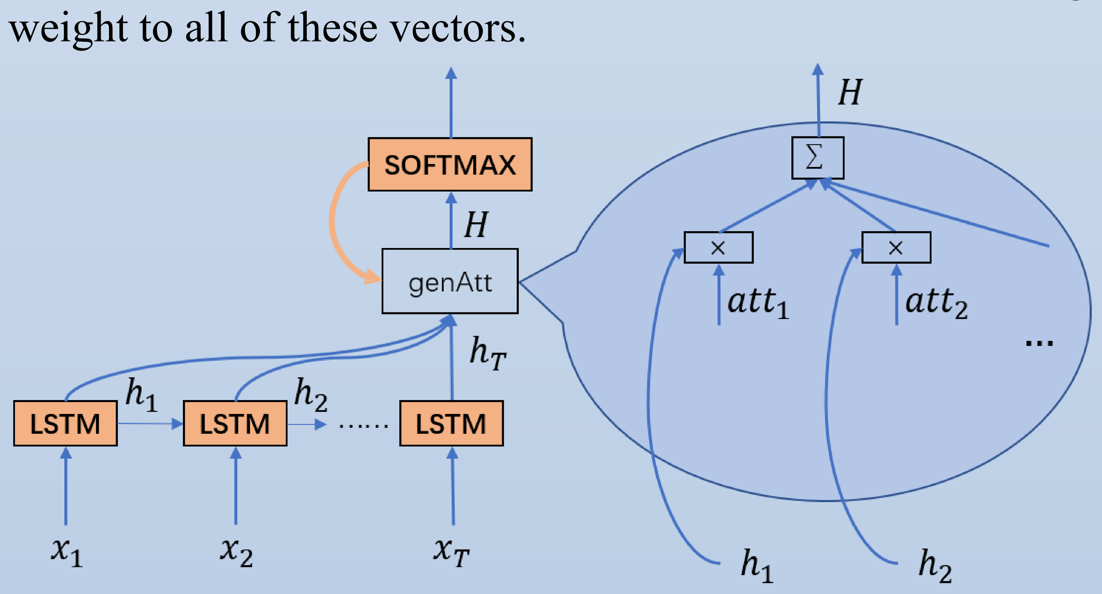


**Fig. 1.** Structure of LSTM with Generic Attention

Fig. 2 show structure of **LSTM with Uniqueness Attention**. The right elliptic region in Fig. 2 represents the internal structure diagram of its left *uniqAtt* module. These operations in this diagram can be defined by the following formula:

- $prob_t = softmax(h_t)$
- $score_t = max_1(prob_t) - max_2(prob_t)$
- $att_t = uniq(prob_t) = \frac{score_t}{\sum_{t=1}^{T} score_t}$
- $H = \sum_{t=1}^{T} att_t \cdot h_t$

In Fig. 2, each $h_t$'s attention weight is only determined by itself and *softmax* module. $h_t$ is first processed by *softmax* module, and then turned into $prob_t$ . If the current atom motion is unique to an activity, the difference between the largest and second largest value in vector of $prob_t$ should be quite large. So *uniq* module use this difference to calculate the uniqueness attention of $h_t$. Then, we use the calculation defined in the last formula to get H, which will finally result in the activity recognition result through the *softmax* module.

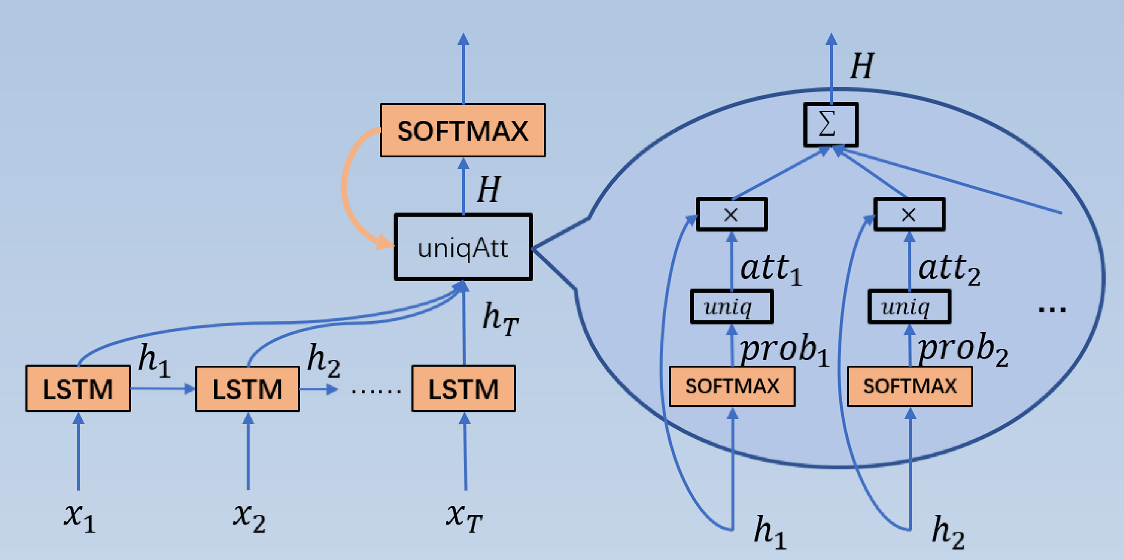


**Fig. 2.** Structure of LSTM with Uniqueness Attention

## Result

| Models | $F_m$ |
|---|---|
| deepConvLSTM | 0.7480 |
| Temporal Attention | 0.8052 |
| Continuous + Temporal Attention | 0.8629 |
| LSTM with Mean Attention | 0.7922 |
| LSTM with Last Attention | 0.7568 |
| **LSTM with Uniqueness Attention** | **0.8796** |

**Table 1.** $F_m$ results of the PAMAP2 dataset

| Models | Locomotion | Gesture |
|---|---|---|
| LDA | 0.590 | 0.690 |
| QDA | 0.680 | 0.530 |
| NCC | 0.540 | 0.510 |
| kNN | 0.850 | 0.850 |
| Cstar | 0.630 | 0.870 |
| SStar | 0.640 | 0.840 |
| CNN[1] | 0.878 | 0.883 |
| LSTM with Mean Attention | 0.873 | 0.894 |
| LSTM with Last Attention | 0.875 | 0.889 |
| **LSTM with Uniqueness Attention** | **0.892** | **0.904** |

**Table 2.** $F_w$ results of the Opportunity dataset

**Results on the PAMAP2 dataset**
In the PAMAP2 (Physical Activity Monitoring for Aging People ) dataset, we compare our method with other deep learning methods, such as deepConvLSTM and Temporal Attention. Furthermore, to demonstrate that our uniqueness attention is indeed effective, we also compare our method with two LSTM benchmark models: LSTM with Mean and Last Attention.

**Results on the Opportunity dataset**
Chavarriaga et al. introduced the Opportunity challenge, and provided four well-known classification methods: Linear Discriminant Analysis(LDA), Quadratic Discriminant Analysis(QDA), k-Nearest Neighbours (k-NN), and Nearest Centroid Classifier(NCC), together with nine contributions submitted to this challenge. We use methods of two best contributions, CStar and SStar, and four classification methods mentioned before as baseline to verify our method. We also use a CNN model for comparison.

**Visualization results of Uniqueness Attention**
The visualized results are shown in Fig. 3. There are two noteworthy characteristics in this figure:

- Uniqueness attention changes periodically.
- The cycle time of uniqueness attention is same as that of walking (or running).
- Uniqueness attention is almost zero at the beginning.

Walking and running are ideal periodical activity. So if there's an atom motion that's only occur in walking (or running), it should occur periodically as the subjects walks (or runs). The final characteristic is also reasonable because we shouldn't pay too much attention to the time before the subjects walking or running

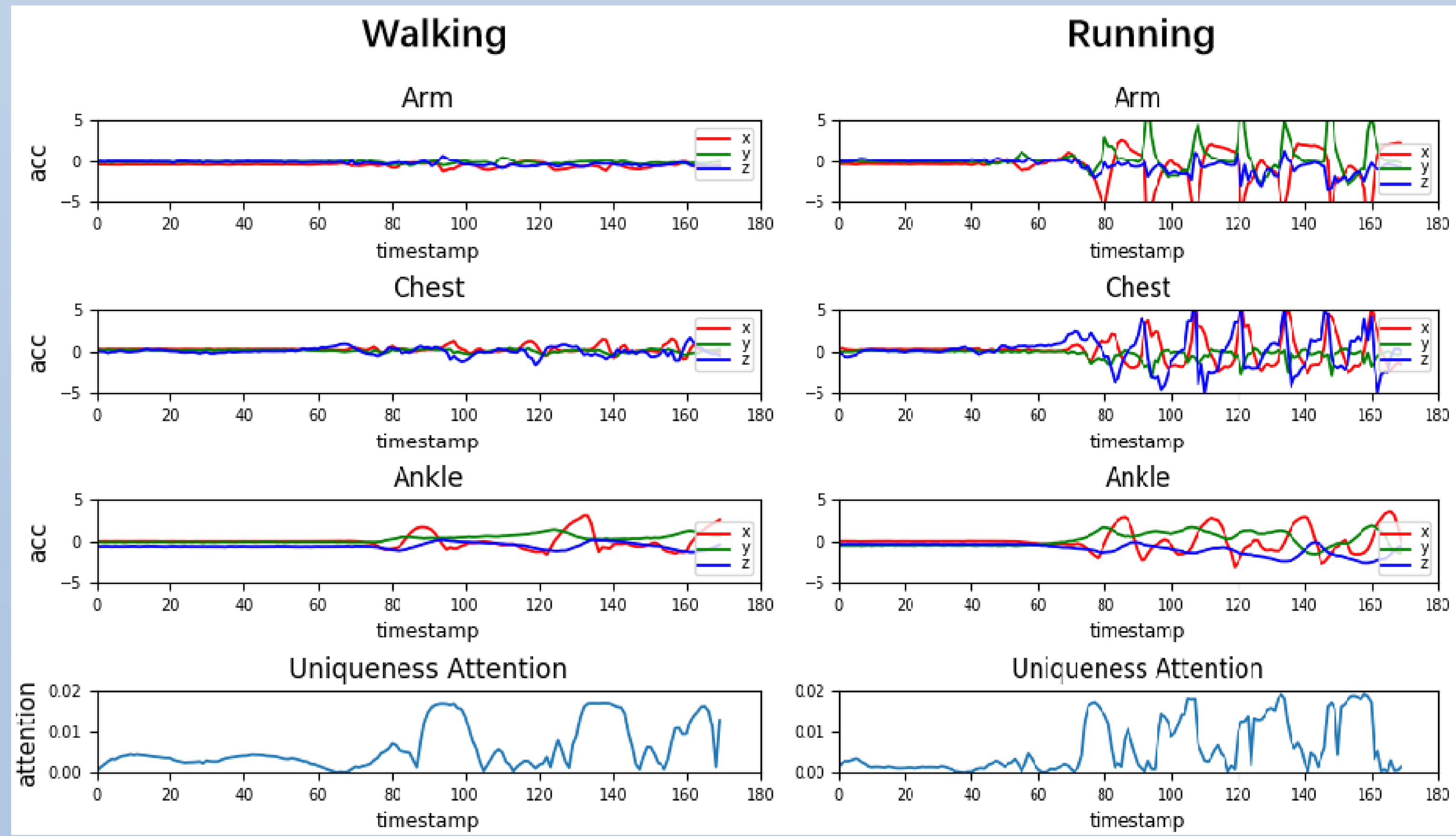


**Fig. 3.** Walking and running examples on PAMAP2. In either of these two subgraphs, the first three coordinates diagrams describe the acceleration information of the subjects' arms, chest and ankle respectively. The last coordinates diagram describes the corresponding uniqueness attention.

## Conclusion

Based on our understanding on the relationship between atom motion and human activity, we propose a model: LSTM with Uniqueness Attention.

- This model outperforms other model on two dataset, which indicates its effectiveness.
- This model is better than these two LSTM benchmarks suggests that the uniqueness attention mechanism does enable our model to focus on more important parts of the input.
- By visualizing our uniqueness attention, we show the characteristics of this attention, and explain how similar are these characteristics to human intuition.

**Other contact Information:**
WeChat: isteincn

**ManuscriptID**
**342**