Proceedings of the 5th International Workshop on Wearable and Implantable Body Sensor Networks, in conjunction with
The 5th International Summer School and Symposium on Medical Devices and Biosensors
The Chinese University of Hong Kong, HKSAR, China. Jun 1-3, 2008

Layered Hidden Markov Models for Real-time Daily Activity Monitoring Using Body Sensor Networks

Jin He, Sheng Hu and Jindong Tan

Abstract— **This paper presents an inferring and training architecture for the long-term and continuously monitoring daily activities using a wearable body sensor network. Energy efficiency and system adaptation to subjects are two of the most important requirements of a body sensor network. This paper proposes a two-layered hidden Markov model (HMM) architecture for in-network data processing to achieve energy efficiency and model individualization. The bottom-layer HMM is used to preprocess the sensor readings locally at each wireless sensor node to significantly reduce the amount of data to be transmitted. The top-layer HMM is utilized to find the activity sequence from the result of this local preprocessing. This approach is energy efficient in that only the results of the decoding procedure in each node need to be transmitted rather than the raw sensor readings; therefore, the volume of transmitting data is significantly reduced.**

I. INTRODUCTION

In the current healthcare system, the failure of frequent and regular health monitoring is particularly problematic for the elderly whose health situation changes rapidly, sometimes with multiple co-morbidities. However, with the advances in wireless sensor and wearable computing technologies, real-time monitoring for daily human activities and physiological parameters has become practical. Meanwhile, body sensor networks (BSN) have become a significant application of sensor networks and have achieved much progress in these years. So the idea of using a BSN to monitor and classify human activities emerges upon this circumstance.

The human activity series can be considered as a Markov process, where the set of activities can be considered the set of states. In this paper, a layered classification system is developed to achieve individualized and energy efficient daily activity classification. The hidden Markov model and associated algorithms are applied to train and then decode the hidden states from observations. At the bottom-layer, each sensor uses an HMM to map the sensor readings to a sequence of inference. The inferences of the bottom-layer HMM are transmitted to host. The combination space of the inferences from the sensors is divided to different observation subspaces. The top-layer HMM is used to map the observation sequence to the hidden state sequence which represents the final classification result of the activity sequence.

Manuscript received January 15, 2008. This work was supported in part by the National Science Foundation under Grant ECS #0528967, CNS #0720781, and Army Research Laboratory under contract ARL W911NF-06-2-0029 and CERDEC W15P7T-06-P228, and Air Force Research Laboratory under a DURIP project AFOSR FA-9550-07-1-0500.

Jin He, Sheng Hu and Jindong Tan is with the Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, MI 49931, USA shengh@mtu.edu

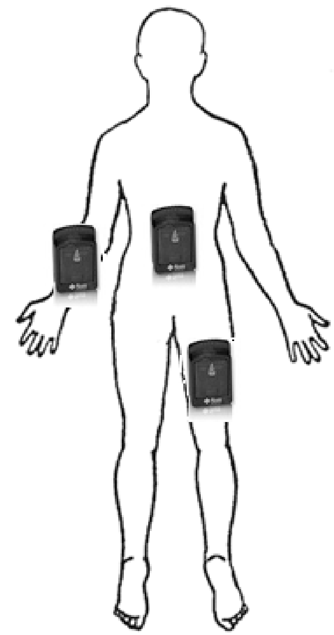

Fig. 1. The distribution of wireless sensor nodes

II. RELATED WORK

Acceleration has been used in various researches and applications such as the obesity monitoring, the Parkinson's disease monitoring, the soldier monitoring, the epilepsy seizure detection, the fall detection, and the general activity classification, etc. [1]. A wearable platform named LiveNet is reported for the long-term, ambulatory health monitoring [1]. An architecture for the activity classification is discussed using a hierarchical binary tree [2]. A new device is reported [3] for the measurements of duration, frequency, and intensity of various types of human physical activities. A single, waist-mounted tri-axial accelerometer unit is used in a real-time classification system [4]. In this work, a median filter and a low pass filter are applied to the sensor reading processing. Then the normalized signal magnitude area is computed and a threshold is set to make a primary decision. The postural orientations are used to determine whether a person's arm movement is similar to that of a person suffering from a seizure [5].

Lots of the previous works on activity classification [6], [7], [2], focus on the classification of only one specific activity. However, when a sequence of activities needs to be classified and monitored, the previous works neglect the interior correlation within these sequential activities. In this paper, the inferring result of the LHMM is a sequence of activity states which suits the online monitoring better.

III. Two-Layered Architecture for Activity Sequence Classification

A body sensor network could consist of many sensors from different parts of a human body. The placement of the wireless sensor nodes will affect the trained parameters of a model for classification but the structure and algorithm should not be affected. In this paper, a three-nods body sensor network is considered which are distributed respectively at the chest, the dominant wrist, and the non-dominant outer thigh, as shown in Figure 1. Each sensor node is embedded one tri-axial accelerometer. Therefore a total of nine signal streams are considered as the input of the system. In the previous work [8], a CHMM for daily activity classification with the observation space is defined as the readings of all of the sensor nodes and all of the raw data are transmitted to the host which takes charge of data processing. Consequently, a large communication bandwidth is required. Additionally, synchronization among the sensor networks needs to be satisfied all the time for data correlation. All the circumstances require the BSN to spend lots of energy in the data transmission. However, the BSNs have the requirement for longevity, and hence the power consumption becomes bottleneck for continuous and long-term monitoring using a BSN. In this section, a two-layered classification algorithm is proposed in which the large amount of raw data are preprocessed in the bottom-layer, and the in-network data are "compressed" a lot.

At the bottom-layer, data collection and signal processing are carried out locally to reduce the volume of the transmission data among sensor nodes. At the top-layer, processed inferential data are fused via the HMM again for the final activity classification. Figure 2 and Figure 3 show the effect of the bottom-layer HMM classification at one sensor node. Figure 4 and Figure 5 show the effect of the top-layer HMM classification at the host. These figures show a sample classification of taking 'sit', 'stand' and 'walk' respectively. It can be seen that the classification result of the top-layer HMM reflects this sequence of activities correctly, as Figure 5 shows. It should be pointed out that the bottom-layer classification process should be completed at one sensor node and could not be displayed on the host. But for our convenience to show the effect of the bottom-layer classification, we make the sensor nodes to transmit the raw data to the host to draw this figure.

In this work, the acceleration data are collected at a frequency of 20 Hz, which is enough to represent human daily activities. To achieve real-time processing, the collected data stream are divided into processing units and processed once at the end of each processing unit. We fix that each real-time processing unit has 128 sampling points, which represents $128 \times (1/20) = 6.4$ seconds. In terms of daily life, 6.4 seconds could be seen as a small delay, so we could say that the requirement of real-time processing is satisfied. Then one processing unit is divided into small windows and

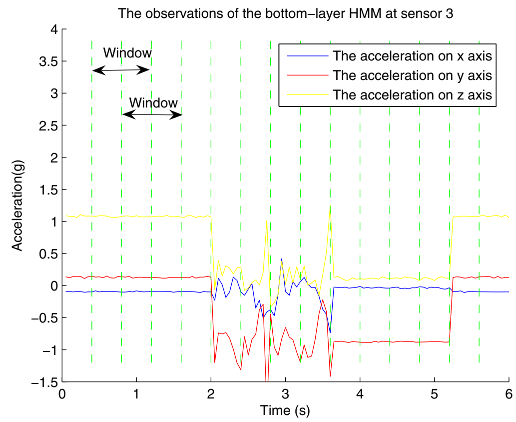

Fig. 2. The observations of the bottom-layer HMM at sensor 3.

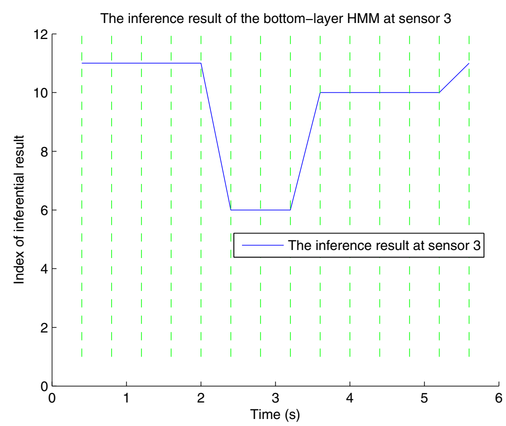

Fig. 3. The inference result of the bottom-layer HMM at sensor 3.

features are extracted from each window. The windows are 50% overlapped and each of them has 16 data points. So totally there are $(128/16) \times 2 - 1 = 15$ windows in one real-time processing unit. However, since the data in the first window are usually unstable, so they are discarded and the remaining 14 windows to infer a sequence of activities from the sequence of observations - features extracted from the 14 windows, as Figure 2 shows.

In our proposed algorithm we make an inference using a HMM firstly at the bottom layer at each node, then the inferences are transmitted to the host to make an final inference using HMM again at the top layer. So the bottom-layer classification is in fact a step of the local processing. The implementation of local processing makes the whole classification architecture more energy-efficient. This is the most important difference between the algorithm proposed in this paper and the one proposed in [9].

At the bottom layer, the features in one window contains the maximum, the minimum, the mean and the variance of acceleration on each axis, and then thresholds are used to divide the observation space. Finally the top-layer HMM maps the sequence of features into a sequence of hidden states. Then a set of hidden states includes 'intense movement',

Fig. 4. The observations of the top-layer HMM.

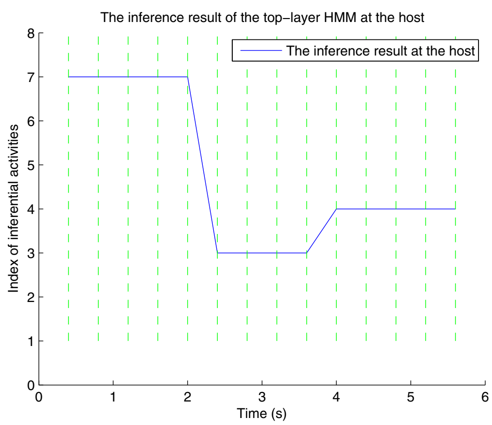

Fig. 5. The inference result of the top-layer HMM at the host.

'medium movement', 'delicate movement' and 'stable'. And 'stable' is further divided to 'X axis towards earth', 'X axis towards sky', 'Y axis towards earth', 'Y axis towards sky', 'X and Y axis are parallel to earth'. So we divide the hidden states space to 8 hidden states in sum. The Viterbi algorithm is used to map the sequence features to a sequence of hidden states. These sequence are transmitted to the host if the changes of the hidden state are detected.

At the top-layer, if no 'new' sequence of hidden states from any node is received then it just uses the 'old' recorded sequence. Anyhow, it uses the inferential result sequence of the bottom layer as its own observation sequence. And at the top layer, the HMM is again utilized to map the inferential result sequence of the bottom layer to the sequence of hidden states that represents the activity sequence, which is the final classification result. Figure 6 shows the two-layered architecture. The bottom-layer HMMs work on one single sensor node respectively, and the inferential states of the bottom-layer HMMs are transmitted to form the observations of the top-layer HMM. Compared with the single, standard HMM method in which HMM is only used as a top-layer classification tool, this layered HMM method could consider more on the correlation information hidden in a single signal channel but sacrifice some of the correlations between different signal channels.

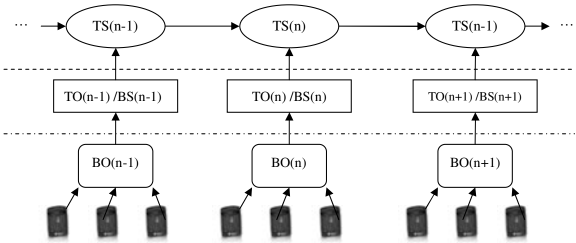

Fig. 6. Architecture of the two-layer HMM classification system. In the upper tier, TS is short for the state of the top-layer HMM; the middle tier are the states of the bottom-layer HMM (BS) that are also the observations of the top-layer HMM (TO); in the lower tier, the acceleration signals are extracted into the observations of the bottom-layer HMM(BO).

Energy is saved since the signal processing is locally implemented and the number of packets transmitted during a processing unit is much less then the method proposed in [9], which sends all the collected data to a host. The data throughput is 12 $\times$ 4 byte = 48bytes, where '12' represents the number of features in each window and '4' means a float number to express for each data. Moreover, compared to other window based method which also extract features from a window and then transmits the value of the features to a host, our algorithm has the advantage of only transmitting the bottom-layer HMM inference, which are only some indices, then only one byte, supporting 256 observations, of data should be transmitted for one transmission (may be different according to the number of hidden states to classify). Then we could see the ratio of the number of bits used for one transmission here and in the algorithms transmitting all the data to the host is 1byte / 48byte $\approx$ 2%. This means if the throughputs are the same in these two scenarios, then our algorithm could save 98% of the total energy compared with the algorithms sending all the collected data to a host.

IV. Experiment

In this experiment, The SunSPOT nodes manufactured by Sun Microsystems, Inc [10] are employed in the experiment. The Sun SPOT Device is a small, wireless, battery powered experimental platform which supports IEEE 802.15.4, and is embedded several useful sensors. One of them is tri-axial accelerometer. It can be configured into two modes, one is 2G sensing range and the other one is 6G, here we configured it into 2G mode, since the activities in our experiments would not be beyond this limit.

Three SunSPOTs are distributed respectively at the chest, the dominant wrist, and the non-dominant outer thigh, as shown in Figure 1. Therefore, a total of nine signal streams of acceleration data are collected as the inputs of the system. The sampling rate is set as 25Hz. TDMA (Time division multiple access) is used as the Medium Access Control (MAC) protocol in this small sensor network.

In this work, individualization are achieved by only re-training the bottom-layer HMM at each individual sensor and hold the top-layer HMM when it is applied on different human bodies. The bottom-layer HMM is trained by a half-hour period of data. The system is tested by 5 subjects with 19 different activity sequences (each is in a 6.4 seconds' interval) repeated 10 times have been carried out. The 19 activity series are: (1)lie, (2)sit, (3)stand, (4)walk,(5)run, (6)brush teeth, (7)sit with legs vibrating, (8)lie to sit, (9)sit to stand, (10)stand to walk, (11)walk to run, (12)stand to run, (13)run to walk, (14)walk to stand, (15)run to stand, (16)stand to brush teeth, (17)brush teeth to stand, (18)sit to sit with legs vibrating, (19)sit with legs vibrating to sit. The confusion matrix is shown in the following table.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	50	0	0	0	2	2	0	0	0	0	0	0	0	0	1	0	3	0
3	0	0	50	2	0	0	0	0	0	2	0	0	0	0	0	0	2	0	0
4	0	0	2	45	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0
5	0	0	0	3	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	2	47	4	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	1	44	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	48	0	8	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	46	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	42	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	48	0	2	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	2	0	48	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	48	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50
Acc	1	1	1	0.90	0.96	0.94	0.88	1	1	0.96	0.92	0.84	0.96	1	0.96	0.98	0.96	0.94	1

Here, we caculate that the overall accuracy of the experiments is 95.78%. We also compare the performance of our model with that of a single, standard HMM. We implement a CHMM to achieve the classification of the same activity set. Also the same features as used in LHMM are extracted from each sensor and one-layer standard HMM is used to map the feature to the activity sequence. The same test sequences are used too. The overall accuracy of this CHMM based classification system is 92.31%. We see that the accuracy of the LHMM based system proposed in this paper is higher than the CHMM based system in the same condition.

V. Conclusion

This paper presents a two-layered architecture for daily activity classification using a wearable body sensor network. At the bottom layer, features are extracted from the acceleration data using a window-based method. The HMM is trained and used to map the feature sequence into a sequence of hidden states which represents the movement status of a single sensor node. The inferential result sequence of a sensor node is transmitted to the host if there is a change of the hidden state. At the host, the top-layer HMM is trained and used to map the observations (the inferential result sequence of the bottom-layer HMM) to a hidden state sequence which is the activity classification result. The two-layered classification architecture is energy efficient in that communication power consumption is reduced by local processing and rare transmission. The accuracy of this LHMM based system is higher than the CHMM based system in the same condition.

References

[1] M. Sung, C. Marci, and A. Pentland, "Wearable feedback systems for rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 2, no. 17, 2005.

[2] M. J. Mathie, B. G. Celler, N. H. Lovell, and A. C. F.Coster, "Classification of basic daily movements using a triaxial accelerometer," *Medical and Biological Engineering and Computing*, vol. 42, no. 5, pp. 679–687, 2004.

[3] K. Zhang, P. Werner, M. Sun, F. X. Pi-Sunyer, , and C. N. Boozer, "Measurement of human daily physical activity," *Obesity Research*, vol. 11, no. 1, pp. 33–40, 2003.

[4] D. M. Karantonis, M. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelrometer for ambulatory monitoing," *IEEE transactions on information technology in biomedicine*, vol. 10, no. 1, 2006.

[5] T. R. Burchfield and S. Venkatesan, "Accelerometer-based human abnormal movement detection in wireless sesor networks," in *the 1st International Workshop on Systems and Networking Support for Healthcare and Assisted Living Environments*, (San Juan,Puerto Rico,USA), June 2007.

[6] M. J. Mathie, A. C. F. Coster, N. H. Lovell, and B. G. Celler, "Detection of daily physical activities using a triaxial accelerometer," *Medical and Biological Engineering and Computing*, vol. 41, pp. 296–301, May 2003.

[7] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Proceedings of the Second International Conference on Pervasive Computing*, (Vienna, Austria), 2004.

[8] J. He, H. Li, and J. Tan, "Real-time daily activity classification with wireless sensor networks using hidden markov model," in *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3192 – 3195, Aug 2007.

[9] N. Oliver, A. Garg, and E. Horvitz, "Layered representations for learning and inferring office activity from multiple sensory channels," *Computer Vision and Image Understanding*, vol. 96, pp. 163–180, November 2004.

[10] http://sunspotworld.com/.