Probabilistic Approach to Human Activity Recognition from Accelerometer Data

Walid Gomaa
Cyber-Physical Systems Lab,
Egypt-Japan University of Science and Technology,
and
Faculty of Engineering, Alexandria University,
Alexandria, Egypt.
walid.gomaa@ejust.edu.eg

***Abstract*—Due to the wide availability of personal mobile and wearable devices, it is now easier than ever before to obtain information from a wide range of sources, including the actions performed by the users or the environment in which a user is located. This is advantageous in a number of applications including, in particular, remote health monitoring and diagnostics, where the manner in which an action is performed or the conformance to some treatment or living regime may be relevant to treatment or health outcomes. The goal of human activity recognition is to determine the identity of the action the user is currently performing based on the data streamed from some sensor modality. In this work we consider the accelerometer signals streamed through a wearable IMU unit and use this data to recognize the user's activity. We consider typical activities of daily living (ADLs). We adopt a simple approach based on the probabilistic modeling of the streamed signals. Some of the training samples are taken as templates and are modeled by empirical distributions. Testing for new sample consists essentially of measuring the probabilistic distribution of the test sample against those of the templates using a chosen set of distance/dissimilarity measures. This approach has proven to be both very computationally efficient and effective regarding its predictive performance.**

***Index Terms*—human activity recognition, probability distribution, dissimilarity measure, Inertial Measurement Unit, accelerometer.**

I. Introduction

Automatic recognition of human activities has become a very substantial research topic. A wide range of digital applications depend mainly on *human activity recognition (HAR)* in their work such as monitoring patients and elderly people, surveillance systems, robots learning and cooperating, human robot interaction, and in the military. The idea of an automatic HAR system depends on collecting measurements from some appropriate sensors which are affected by selected human motion attributes. Then, depending on these measurements, some features are extracted to be used in the process of training activity models, which in turn will be used to recognize these activities later.

As a typical example of such activities are Activities of Daily Living (ADL) that people have the ability for doing on a daily basis like eating, moving, hygiene, dressing, etc. [9]. There are many methods and data acquisition systems which are based on different sensory readings for recognizing people's actions. Previously, heavy devices were used to collect accelerometer data such as data acquisition cards (DAQs) [10]. Later, smaller integrated circuits connected to PDAs were used for this aim, for instance, camera-based computer vision systems and inertial sensor-based systems. In computer vision, activities are recorded by cameras [14]. However, the drawback of the camera-based approach is that it may not work due to the absence of full camera coverage of all person's activities. In addition, cameras are meddlesome as people don't feel free being observed constantly.

Based on the *data acquisition* paradigm, HAR systems can be divided into two categories: ambient/surrounding fixed-sensor systems and wearable mobile sensor systems. In the first category, data are collected from distributed sensors attached to fixed locations in the activity surrounding environment (where users activities are monitored) such as surveillance cameras, microphones, motion sensors, etc. See [6] for applications in smart homes. Alternatively, the sensors are attached to interactive objects in order to detect the type of interaction with them, such as a motion sensor attached to the cupboard doors or microwave ovens (to detect opening and/or closing), or on water tap to feel turning it on or off, and so on. Although this method can detect complex actions efficiently, it has many limitations due to its fixed nature. Another limitation is that if the user wants to leave the place, she will not be observed from the fixed sensors and her activities won't be detectable. Privacy is also another issue, especially when considering video surveillance cameras or auditory sensors.

In the *wearable* based systems, the measurements are taken from mobile sensors mounted to human body parts like wrists, legs, waist, and chest. The typical sensory unit used is the Inertial Measurement Unit (IMU) which streams signals about the acceleration, Gyroscope rotational velocity, Gyroscope orientation, and the gravitational field. These readings are complete in the sense that they can in principle uniquely distinguish the typical human daily activities. These kinds of sensors are essentially embedded as MEMS sensors in wearable gadgets such as smart phones, smart watches, etc. Unlike fixed-sensor based systems, wearables are able to measure data from the user everywhere since it is not bounded

978-1-7281-5617-0/19/$31.00

by a specific place where the sensors are installed. Also, it is very easy to concentrate on directly measuring data of particular body parts efficiently without a lot of preprocessing that are needed, for example, in fixed depth cameras. Examples of wearables include smart watches, smart shoes, sensory gloves, hand straps, and clothing [4], [11]. Therefore, the disadvantages of wearable mobile sensors of being intrusive, uncomfortable, and annoying have vanished to a great extent, making this method of on-board sensing from smart mobile and wearable devices very suitable for HAR data acquisition.

In this paper we develop an approach for activity recognition based on probabilistic modeling of the time series data and use metric-theoretic techniques to measure the distance/dissimilarity between two probability distributions. We focus of activities of daily living (ADLs) which are routine activities that people tend to do every day without needing assistance [1]–[3], [9]. We have used a public dataset of accelerometer data collected with a wrist-worn accelerometer [7], [8]. The data are streamed while performing 14 activities using several human subjects and with varying number of samples for the different activities. Table I provides a summary of the monitored activities along with each activity sample size. Each experiment consists of three time series data corresponding to the tri-axial accelerometer directions. To the best of our knowledge, most of the methods targeted at activity recognition are based either on the machine learning paradigm or time series analysis. So in the current paper we explore new direction that has shown computational efficiency with reasonable predictive performance. We hypothesize that the predictive performance can be boosted by including more sensory data such as the gyroscope data.

The paper is organized as follows. Section I is an introduction. Section II presents the dataset used for empirical verification. Section III describes our approach along with experimentation and results. Section IV concludes the paper with prospective future work.

II. Data Description

We have obtained our labeled accelerometer activity data from the UCI machine learning public repository. The dataset is: 'Dataset for ADL Recognition with Wrist-worn Accelerometer Data Set'. [1] A detailed description of the data can be found in [5].

The data are collected using a wrist-worn accelerometer with sampling rate 32Hz. The measurement range is $[-1.5, 1.5]$ (it is in g-force units). One data sample is a tri-axial acceleration: x-axis pointing toward the hand, y-axis pointing toward the left, and z-axis perpendicular to the plane of the hand. Acceleration data are collected for 14 activities with varying number of samples for each. Table I lists the activities along with the number of collected samples per each.

[1]https://archive.ics.uci.edu/ml/datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer#

TABLE I
List of activities and associated number of samples.

No.	Activity	Number of samples	No.	Activity	Number of samples
1	brush teeth	12	8	pour water	100
2	climb stairs	102	9	eat meat	5
3	comb hair	31	10	walk	100
4	drink glass	100	11	liedown bed	28
5	getup bed	101	12	standup chair	102
6	sitdown chair	100	13	descend stairs	42
7	use telephone	13	14	eat soup	3

III. Estimation Using Discrete Distributions

In this paradigm we use a variety of tools from the theory of probability and measure theory. Here we deal with the *statistical* properties of the raw signal without taking time into consideration. We basically rely on probabilistic modeling of the signals and measuring the distance among these probability distributions in order to classify the given activity signal.

Here we consider the points of a time series as generated from a discrete distribution. Hence, the empirical distribution of the time series is estimated using the discrete distribution that corresponds directly to the observed data. We then apply several distance measures among the discrete distributions in order to measure the *dissimilarity* among different time series samples and accordingly decide on activity membership. An important thing to notice is that the resolution of the accelerometer data is fine enough to provide a comparable discrete distributions among the time series samples. We mean by resolution both the sampling rate relative to the accelerometer range ($[-1.5, 1.5]$), the precision of measurements as real numbers, and the length of each sampled time series. In the following we briefly describe four distance measures that we have used to measure dissimilarity among discrete probability distributions.

The first is the *Hellinger distance* which is defined in terms of the Hellinger integral and can be computed for discrete distributions as follows [12]. Given two discrete distributions $P = (p_1, \dots, p_n)$ and $Q = (q_1, \dots, q_n)$, then

$$d_H(P,Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{n} \left(\sqrt{p_i} - \sqrt{q_i}\right)^2} \tag{1}$$

Hellinger distance is a proper metric and it is normalized by the factor $1/\sqrt{2}$ such that it is upper bounded by 1. The second distance metric is the *total variation distance*. Given two probability measures P and Q on the same σ-algebra $(\Omega, \mathcal{F})$, the total variation distance between P and Q can be computed as follows:

$$d_T(P,Q) = \sup\{|P(A) - Q(A)| \colon A \in \mathcal{F}\} \tag{2}$$

So it is the largest difference between the two measures for any given event in the σ-algebra. In the assumed discrete case, as also the support is finite, this can be reduced to the following formula:

$$d_T(P, Q) = \frac{1}{2}\sum_{i=1}^{n}|p_i - q_i| \tag{3}$$

Note that this reduces to the Manhattan distance. The third distance metric is the Cramér-von Mises distance, which computes the distance between the empirical cumulative distribution functions that are estimated from the given data [13]. This distance can be computed as follows:

$$\begin{aligned} d_C^2(P, Q) = \int &(P(\{y \in [-1.5, 1.5] \colon y \leq x\}) - \\ &Q(\{y \in [-1.5, 1.5] \colon y \leq x\}))^2 \mu(dx) \end{aligned} \tag{4}$$

where μ is the integration measure, which in our experiments taken to be the probability measure Q. The last distance metric that we adopt in our experimentation is the Kolmogorov distance. It is inspired by the Kolmogorov-Smirnov goodness-of-fit test [13]. It computes the supremum distance between the two given probability distributions evaluated over the sets in the σ-algebra defined over $[-1.5, 1.5]$:

$$\begin{aligned} d_K(P, Q) = \sup\{&|P(\{y \in [-1.5, 1.5] \colon y \leq x\}) - \\ &Q(\{y \in [-1.5, 1.5] \colon y \leq x\})| \colon x \in [-1.5, 1, 5]\} \end{aligned} \tag{5}$$

For each activity A, some of its collected samples s are randomly selected to act as a model of the underlying activity. We call these samples 'support samples'. They act as templates for the activity they are drawn from. The remaining samples are used for testing. In experimentation we use 65% of the samples as 'support samples'.

The offline modeling phase for the above techniques consists of the easy, and computationally efficient, task of estimating the discrete probability distributions of the given time series samples. The testing phase consists essentially of computing the distance, using one of the four distance metrics, between the test sample and all of the support samples of the 14 activities. For each time series sample we have three estimated empirical distributions for the tri-axial accelerometer directions. When a new test sample s is classified (to predict to which of the 14 activities this sample belongs), the following is done:

1) For each activity A, we find the distance between s and every member s' of the supporting samples of A. The test sample s has three empirical distributions associated with it for the tri-axial directions. Each such distribution is compared against the corresponding distribution in the supporting sample s'. So we have three scores. These scores are combined together by taking their summation.

$$d(s, s') = d_X(s_x, s'_x) + d_X(s_y, s'_y) + d_X(s_z, s'_z) \tag{6}$$

where d_X is either one of the four distance metrics described above.

2) Given that we have computed the distance scores of the given test sample s against all the supporting samples of activity A. We need to choose a single score representing the distance between the test sample s and activity A. We choose the minimum score:

$$d(s, A) = \min_{s' \in Supp(A)} d_X(s, s') \tag{7}$$

3) The decision is then taken by assigning s to the activity A' with minimum distance: $A' = argmin_A d(s, A)$.

Notice that this procedure is equivalent to a k-NN classifier with $k = 1$. Fig. 1 illustrates a comparison in prediction accuracy over the 14 activities using the four distance measures. From this figure it is clear that working with probability mass functions, as in Hellinger and total variation, outperforms working with cumulative distribution functions, as in Cramer-von Mises and Kolmogorov. Also working with an L_2 norm style as in the Hellinger case outperforms working with an L_1 norm style as in the total variation case. Of course, the former is smoother and more uniform in accounting for distances among the various parts of the two input signals. The best performance is achieved using the Hellinger distance, using the whole set of 14 activities its accuracy reaches about 80%.

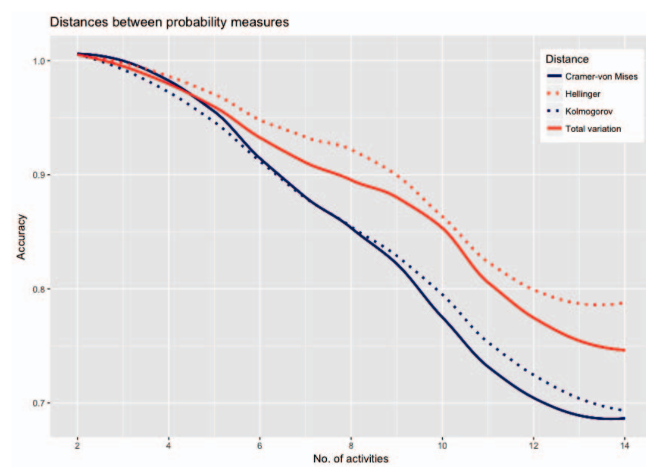

Fig. 1. Comparing the accuracy using different distance metrics between discrete empirical distributions.

IV. Conclusion

In this paper we have studied an approach to human activity recognition based on probabilistic modeling of the underlying tri-axial accelerometer signal that describes the activity. The signal is assumed to be streamed from a wearable device. Our approach is based on representing the time series as a probability distribution, which of course loses the temporal succession of the streaming. Despite of this loss, the approach has shown effectiveness in the predictive performance as well as computationally. This can be attributed in part to the stationarity of the underlying time series as well as to a hypothesized short-term correlations in the given studied activities. These claims need to be verified using more datasets as well as more thorough investigation of the time series.

For each activity a set of samples from the training set are chosen at random to act as templates of the activity, we called them ‘support samples’. Then, at test/operation time a new test sample is then compared (using their representations as probability distributions) against each template set of each activity to decide its activity membership. The comparison is based on some typical dissimilarity metrics. It turned out that the Hellinger distance gives the best stable performance.

In the future we will investigate more diverse methods within the same paradigm of using probability distributions as representations of the signals along with multitude of distance metrics. We will experiment as well with other modes of sensory data, and in particular, the gyroscope.

Acknowledgment

This work is funded by the Information Technology Industry Development Agency (ITIDA), Information Technology Academia Collaboration (ITAC) Program, Egypt – Grant Number (PRP2019.R26.1 - A Robust Wearable Activity Recognition System based on IMU Signals).

References

[1] M. G. Abdu-Aguye and W. Gomaa. Robust Human Activity Recognition Based on Deep Metric Learning. *The 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 2019.

[2] M. G. Abdu-Aguye and W. Gomaa. VersaTL: Versatile Transfer Learning for IMU-Based Activity Recognition using Convolutional Neural Networks. *The 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 2019.

[3] M. G. Abdu-Aguye, W. Gomaa, Y. Makihara, and Y. Yagi. On The Feasibility of On-Body Roaming Models in Human Activity Recognition. *The 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 2019.

[4] S. J. M. Bamberg, A. Y. Benbasat, D. M. Scarborough, D. E. Krebs, and J. A. Paradiso. Gait analysis using a shoe-integrated wireless sensor system. *IEEE transactions on information technology in biomedicine*, 12(4):413–423, 2008.

[5] B. Bruno, F. Mastrogiovanni, and A. Sgorbissa. Laboratorium - Laboratory for Ambient Intelligence and Mobile Robotics, DIBRIS, University of Genova, via Opera Pia 13, 16145, Genova, Italia (IT), Version 1 released on Feb 11th, 2014.

[6] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria. Human motion modelling and recognition: A computational approach. In *Automation Science and Engineering (CASE), 2012 IEEE International Conference on*, pages 156–161. IEEE, 2012.

[7] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria. Human motion modelling and recognition: A computational approach. In *CASE*, pages 156–161. IEEE, 2012.

[8] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria. Analysis of human behavior recognition algorithms based on acceleration data. In *ICRA*, pages 1602–1607. IEEE, 2013.

[9] S. Katz, A. Chinn, and L. Cordrey. Multidisciplinary studies of illness in aged persons—ii: a new classification of functional status in activities of daily living. *Journal of chronic diseases*, 9(1):55–62, 1959.

[10] J. Mantyjarvi, J. Himberg, and T. Seppanen. Recognizing human motion with multiple acceleration sensors. In *Systems, Man, and Cybernetics, 2001 IEEE International Conference on*, volume 2, pages 747–752. IEEE, 2001.

[11] U. Maurer, A. Rowe, A. Smailagic, and D. P. Siewiorek. ewatch: a wearable sensor and notification platform. In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, pages 4–pp. IEEE, 2006.

[12] F. L.-J. Miescke. *Statistical Decision Theory: Estimation, Testing, and Selection (Springer Series in Statistics) by F. Liese (2008-06-11)*. Springer, 1616.

[13] H. Rieder. *Robust Asymptotic Statistics: Volume I (Springer Series in Statistics)*. Springer, softcover reprint of the original 1st ed. 1994 edition, 7 2012.

[14] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea. Machine recognition of human activities: A survey. *IEEE Transactions on Circuits and Systems for Video technology*, 18(11):1473–1488, 2008.