Log-likelihood Clustering Enabled Passive RF Sensing for Residential Activity Recognition

Wenda Li*, Bo Tan†, Yangdi Xu*, Robert. J. Piechocki*

*Department of Electronic and Electrical Engineering, University of Bristol, UK
†School of Computing, Electronics and Mathematics, Coventry University, UK
*{wenda.li, yangdi.xu, Robert. J. Piechocki}@bristol.ac.uk,†bo.tan@coventry.ac.uk

***Abstract*—Physical activity recognition is an important research area in pervasive computing because of its importance for e-healthcare, security and human-machine interaction. Among various approaches, passive Radio Frequency (RF) sensing is a well-tried radar principle that has potential to provides unique non-invasive human activity detection and recognition solution. However, this technology is still far from mature. This paper presents a novel HMM-log-likelihood matrix based feature characterizing for the Doppler shifts to break the fixed sliding window limitation in traditional feature extraction approaches. We prove the effectiveness of proposed feature extraction method by K-means & K-medoids clustering algorithms with experimental Doppler data gathered from a passive radar system. The time adaptive log-likelihood matrix-based approach outperforms the traditional Singular Value Decomposition (SVD), Principal Component Analysis (PCA) and physical feature based approaches, and reaches 80% in recognizing rate.**

***Index Terms*—Human Activity Recognition, Log-likelihood matrix, Doppler Radar, Passive Sensing**

I. Introduction

Human activity capturing and behavior modeling in residential environment draw increasing attention in communities because of its significant importance in healthcare, security and economic research. Physical activity recognition is one of the essential elements in these areas, particularly for its value in tackling challenges in modern healthcare problems, for example, Chronic Noncommunicable Diseases (NCDs) which reported causes two million deaths every year [1]. Technologies like computer vision, Micro-electromechanical (MEMS) based wearable sensors, and environmental sensors have been listed as potential solutions. Research project like SPHERE [2] dedicates in integrating above sensors in smart home context to model human behaviors for healthcare purpose.

Among those candidate technologies, we notice that videos [3]–[5] and wearable sensors [6]–[8] like accelerometers and gyros are most popular approaches. However, these approaches have obvious defects in the practical residential applications. Wearable sensors suffer from battery life limitation, uncomfortableness and oblivion in long-term monitoring applications. While deploying of vision based sensors are constrained by privacy issues and environmental illuminance conditions. In this context, researchers start to shift attention to passive RF sensing which is based on well-tried passive Radar principle [9]. Passive RF sensing is the option can potentially create a seamless, ever-present, contact-less human activity sensing approach in residential context without taking sensitive resident's image information.

Early attempts of the RF sensing focus on active Ultra-Wide-Band (UWB) radar [10]–[12]. Benefiting from the high operating carrier frequency and wider bandwidth, the active UWB radars have sensitive Doppler detection and high range resolution that are advantages of capturing human target movement. However, active UWB needs the extra RF signal source in given environment, which may cause Electromagnetic Compatibility (EMC) problem or license issue during deploying. Thus, researchers shift the vision from active UWB to passive sensing which uses the existing RF resources in the given environment. Regarding the residential area, WiFi signal is the most investigated passive RF source for sensing purpose because of its ubiquitous existence in contemporary residential space. Some works like [13] simply taking advantage of Received Signal Strength (RSS) measurement on IoT device to interpreting the signal disturbance caused by human activities. However, RSS is a coarse index for interpreting detailed human gesture. Then, the concept is extended to WiFi Channel State Information (CSI) based approaches [14]. CSI expands the single RSS parameter to Direction-of-Arrival (DoA) [15], Angle-of-Arrival (AoA) [16] and Doppler information [10] which are accurate enough to represent the human activity details.

Among these parameters, Doppler information is the only one that is differentiated from stationary clutters which make it an ideal for detecting moving objects. Thus, we can see many passive RF sensing systems using Doppler as the main parameter for human activity and gesture recognition. For example, work [17] presents a device-free Through-The-Wall (TTW) system for personnel sensing at standoff distances and [18] shows a solution for gait recognition using Doppler information with WiFi signals. Our pioneer works [19]–[21] and researchers in [12], [22] have also shown that the recorded Doppler shifts within a gesture or activity cycle can be potentially used for various application including non-invasive breathing detection and home area lifestyle monitoring, even through-wall.

In order to interpret human gesture or activity information from the Doppler information, a "good feature vector" need to be defined. In most research and practices, popular Doppler

feature extraction approaches include Physics Features [11], [21], [23], PCA features [12], [24], and Eigenvalue based features [25], [26]. However, these approaches need Sliding-Window (SW) algorithm which requires a pre-designed fixed window sliding on the Doppler sequence. This fact makes the SW based features are insufficient or need to be tuned for high accurate activity recognition for varying durations of different activities [27]. To overcome the limitation of SW algorithm, we propose the clustering methods on the basis of Hidden Markov Model (HMM) based log-likelihood matrices to adapting the varying activity durations problem. HMM is a widely used statistical model for sequence clustering problems like the DNA sequence and speech recognition [28]. In [22], HMM is used for micro-Doppler signatures classification in the SW category. In this work, we creatively use the HMM as middle-processing for log-likelihood feature extraction of each individual human gesture/activity Doppler sequence which presents strong time sequential characteristics.

The methodology described in this paper makes following contributions to the research community:

- **Time-Adapting Log-likelihood Feature:** The log-likelihood matrix derived from HMM training brings adjustable window size to improve the time diversity of different Doppler sequence. This operation brings more than 15% improvement in activity recognition rate compared with previous approaches.
- **Symmetric log-likelihood Metrics** In addition, we propose to use the symmetric log-likelihood matrix rather than the original log-likelihood matrix to further extend diversity in clustering process. This process further benefits the recognition performance by taking account the training quality of HMM.
- **Framework & Prove-of-Concept (PoC) System:** A framework which systematically connects functional blocks of RF signal processing, the HMM-based log-likelihood feature extracting and clustering are proposed. A Software Define Radio (SDR) based passive RF sensing platform is designed as the PoC system to prove those concepts. The recognition performance is supported by the lab experimental data.

The rest of this paper is organized as follow: Section II presents the system design and experiment setup; the preprocessing for passive radar sensor is outlined in Section III; the descriptions of proposed HMM-based log-likelihood matrix are expressed in Section IV; Section V presents the recognition performance for the log-likelihood based clustering in contrast with other classical features; Conclusions are given in Section VI.

II. Framework, System and Experiment Design

The top-level block diagram of our passive activity recognition radar system is presented in Fig 1. Three main processing stages are included: RF IQ samples preprocessing, log-likelihood calculation and clustering. The data formats of input/output ports of each block are also given in top-level diagram and will be discussed in corresponding sections. The

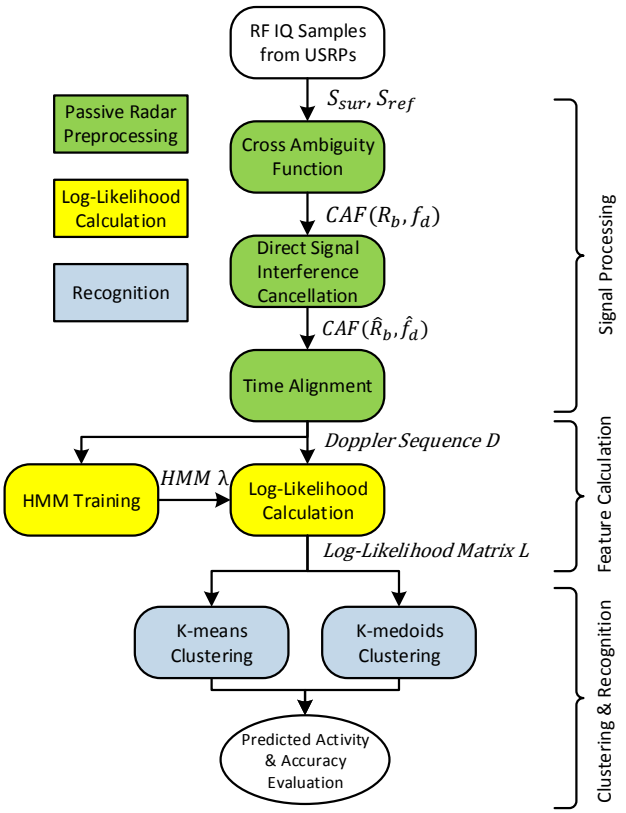


Fig. 1: Block diagram of our proposed passive radar activity recognition system

signal processing stage is the preprocessing of received RF IQ signals from RF front end which is NI USRPs in this paper. This stage outputs Doppler sequence corresponding to each activity sequentially. Afterwards, in feature calculation stage, we use HMM to train those Doppler sequences with the classical Baum-Welch algorithm. The idea is that for Doppler sequences captured from same activity shall receive similar log-likelihoods from HMMs trained from other Doppler sequences. Thus those log-likelihoods should be clustered into each activity groups (if the predict HMM is correct). The outputs at current stage are the log-likelihood matrix.

K-means and K-medoides clustering algorithms are then used to allocate activities to clusters according to the value in the log-likelihood matrix. Different from previous works [11], [12], [25], [29] which all use the supervised classifier (like SVM or decision tree), both K-means and K-medoides belongs to unsupervised learning that no labels are required from dataset. In this work, these clustering algorithms group the activities into several clusters on the basis of distance among the value in log-likelihood matrix. This approach shows an attractive solution for the data from weak labelling environment. Note that, labels were only used to evaluate the clustering and recognition performance at final stage.

All measurements were carried out in the lab at the University of Bristol, the experiment layout is shown in Fig 2. The monitoring area is 7 m x 5 m with equipments, chairs,

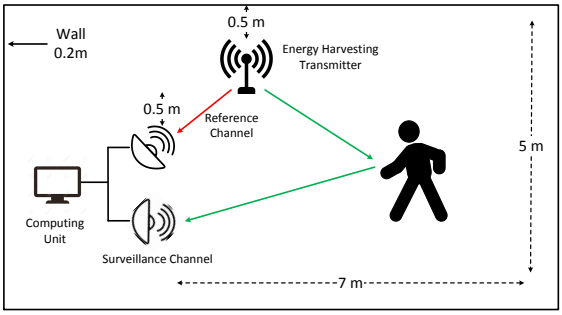


Fig. 2: Experiment Setup

and tables surrounded. A table is cleared of any other items and used to collect the Doppler data with an energy harvesting transmitter (TX91501 POWERCASTER) as the signal source which operates at 915 MHz ISM band (902-928 MHz). Our passive radar sensor contains two synchronized channels, one is the surveillance channel which aims to capture the reflected signal from surveillance area, the other one is reference channel which aims to receive the transmitted signal from a comparable stable channel - a direct channel in most cases. The location of the signal source was near to the wall in the room, whereas the system was located 0.5 m away from the source and faced to the monitoring area. This layout minimizes the influence of the system geometry to the profile of Doppler sequence since the monitoring area is within the range of surveillance antenna. Due to the nature of bistatic geometry, there will be no Doppler observed if the subject moving along the bistatic contour [30]. Thus, adding additional surveillance channels from different angles is necessary to cover blind spots from application aspect.

The passive radar system is implemented based on a Software Defined Radio (SDR) platform, where two NI USRP-2920s were used for surveillance and reference channels respectively. Each USRP was equipped with a directional PCB antenna (P2110-EVB) and connected to a laptop via an Ethernet port. For each channel, the bandwidth was set at 1 MHz to enable the real-time processing for the system. The system measurement rate was set at 5 Hz with a shifting window design to ensure high Doppler resolution [31]. Moreover, we processed and stored data for one minute for each measurement despite the difference in duration time of each activity.

In this pilot study, the dataset includes five basic activities (standing, walking, running jumping and turning) and one inactivity (standing). Descriptions of above activities are presented in Table I. To increase the diversification for the dataset, we conducted the measurement from four volunteers (three males and one female) at different ages (ranges from 25 to 38). The measurement from different people will increase the dissimilarity even for the same activity. So that the proposed concept can be verified with a realistic environment.

The total number of the dataset is inactivity (1) $\times$ (20 repetitions) + 5 activities (2-6) $\times$ (40 repetitions) = 220.

TABLE I: Description of Measured Activities

	Activity	Description
(1)	Standing	Standing still in front of the surveillance antenna for 10 seconds at distance with 1 m, 2 m, 3 m and 4 m. This represents low-level body movements.
(2)	Walking	Walking in the area between two random points, forward and backward. This represents long middle-level body movements.
(3)	Running	Running in the area between two random points, forward and backward. This represents long high-level body movements.
(4)	Jumping	Jumping in front of the antenna at distance with 1 m, 2 m, 3 m and 4 m. This represents short high-level body movements.
(5)	Turning	Turning the body into different directions in front of the antenna at distance with 1 m, 2 m, 3 m and 4 m. This represents short middle-level body movements.
(6)	Standing	Standing up from a chair in front of the antenna at distance with 1 m, 2 m, 3 m and 4 m. This represents another type of short middle-level body movements.

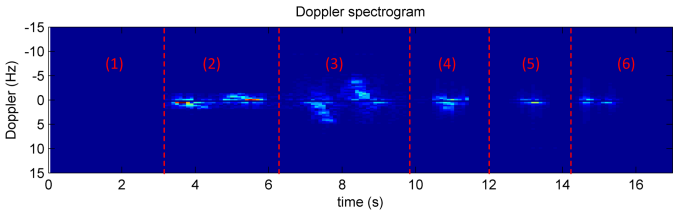


Fig. 3: Example of Doppler sequence from our passive radar sensor: activity (1) standing still, (2) Walking, (3) Running, (4) Jumping, (5) Turning and (6) Standing

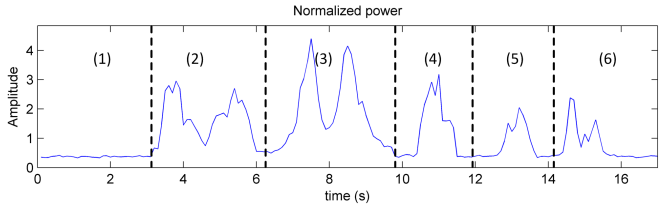


Fig. 4: Example of Doppler power intensity for time alignment: activity (1) standing still, (2) Walking, (3) Running, (4) Jumping, (5) Turning and (6) Standing

III. Preprocessing for Passive Radar Sensor

A. Cross Ambiguity Function (CAF)

In passive radar, CAF is an effective tool to obtain the range and Doppler information by taking Fast Fourier Transform (FFT) of cross-correlated signals from surveillance channel $S_{sur}(t)$ and reference channel $S_{ref}(t)$. However, there are two inherent limits of CAF. Firstly, long integration time of bandwidth signal will contain a significant amount of samples that the FFT process becomes impractical to be implemented. Secondly, the FFT algorithm expands all frequency components which may bury the desired Doppler shift in the wide frequency spectrum. In another hand, the segmentation technology can solve the computational issue by dividing a long sampled signal into group of short signal for the FFT transformation. This operation will lead to faster CAF processing. One of such technology is known as the batching

process [32] by splitting the baseband signal into several batches (each contains small but equal portion of signal). The equation for CAF with batching process can be presented as:

$$CAF(R_b, f_d) = \sum_{k=0}^{n_b-1} \int_{T_{i-1}}^{T_i} S^i_{sur}(t) \tag{1}$$

$$S^i_{ref}{}^*(t - kT_B - \frac{R_b}{c}) e^{j2\pi f_c f_d t} dt \tag{2}$$

where R_b and f_d represents the range and Doppler information respectively, $S^i_{sur}(t)$ and $S^i_{ref}(t)$ are down converted baseband signals from surveillance and reference channel with batching length T_B and n_b is the number of batches, c donates the light speed, f_c is the carrier frequency and $[*]$ indicates the complex conjugate. The range resolution is defined as $\Delta R_b = c/2B$ where B is the bandwidth of the RF signal (normally 20 MHz to 40 MHz for typical indoor communication/transmission system). This gives range resolution from 7.5 m to 3.75 m which obviously is too coarse for activity recognition purpose. On the other hand, the Doppler resolution is defined as $\Delta f_d = \frac{1}{T_i}$ which can be adjusted by the length of integration time.

B. Direct Signal Interference (DSI)

The major limitation of CAF is that the signal source to surveillance antenna signal leakage which contains much higher energy compare to the target's refections, thus, generates a strong interference at the zero Doppler bin. This strong interference could bury the target reflection and significant reduce the detection sensitivity. To remove this interference, the CLEAN algorithm [33] in introduced. The principle is to iteratively subtract scaled self-ambiguity function CAF_{self} surface. CAF_{self} is calculated from reference channel $S_{ref}(t)$ signal. The cleaned CAF $CAF^i(\hat{R}_b, \hat{f}_d)$ at i_{th} iteration can be then calculated as:

$$CAF^i(\hat{R}_b, \hat{f}_d) = CAF^{i-1}(R_b, f_d) - \alpha^i CAF_{self}(R_b - f_i, f_d) \tag{3}$$

where scaling factor α^i is the maximum absolute value of the shift factor related to the location of α^i and f_i is the corresponding phase shift factor. Note that, the batching process is also applicable in the DSI cancellation. Afterwards, the dominant peak caused by DSI will be removed and the desired Doppler peak can be detected. Finally, a Doppler sequence $D(f_d, t)$ is generated by combining a group of columns with the maximum Doppler peak from each CAF [19]. An example of the Doppler sequence measured from six activities is shown in Fig 3.

C. Time Alignment

One challenge in this system is to determine the starting point of an event. As there is no sign to label the start point of an activity, the system should be capable of automatically identifying the inactive and active periods. Traditional methods such as manifold alignment and dynamic time wrapping (DTW) [34] transform the general problem to high-dimensional vectors and therefore require high computational power which is not ideal for our system. In comparison, we transform this challenge into a pulse detection problem by calculating the power intensity PI of Doppler sequence:

$$PI(t) = \sum_{i=1}^{n_b} D(f_{d_i}, t) \tag{4}$$

We plot the corresponding Doppler power in Fig 4. As can be seen, the peaks for middle and high-level body movements are clear and distinguishable. However, for low-level body movement, there are few patterns in Doppler sequence that can be hardly detected. For this reason, we treat the low-level body movement as inactivity period and other levels as activity periods. We set the reference level at 10% of the waveform amplitude to detect the start point and duration of activity. Overall 97.3% (214 out of 220) accuracy has been observed between the activity and inactivity period. The average measured durations of each activity are shown in Table II.

TABLE II: Average Duration of Each Activity

Activity	walking	running	jumping	turning	standing
Duration (s)	2.5	2.6	1.4	0.8	1.2

IV. HMM Training and Log-likelihood

After the preprocessing on the baseband signal, we can obtain a set of Doppler sequences from different activities. The problem in this paper is to allocate those Doppler sequences into K different groups. This is normally handled by using classifier like Support Vector Machine (SVM) or decision tree based on feature vectors. As discussed before, those feature vectors from [11], [12], [22], [23], [26] are all on the basis of the fixed sliding window algorithm. While in this work, we aim to use the original Doppler sequences instead of their segmentations from a fixed window. This is not a straightforward task since the Doppler sequences are varying in time length (shown in Table II). For this reason, we propose to use the HMM as the representation of Doppler sequences. Different from previous HMM work [22], we train one HMM for each individual Doppler sequence. The trained HMM will contain the activity information, not rely on activity start/end points inside a Doppler sequence. So that the limitation of fixed data length in HMM training process can be avoided. The outcome of Doppler sequence tests by HMM is the log-likelihood value which indicates how well the Doppler sequence fits the model. The details are present in this section.

A. Training with Hidden Markov Model

HMM is a popular statistical models for recognition/categorization. In the graphical model of HMM, there are two types of node: observations and hidden nodes. The observation nodes contain the input data which can either be continuous or discrete. The hidden states are discrete and characterized by a joint probability distribution. An HMM $\boldsymbol{\lambda}$ model follows the first order Markov assumption and can be defined as following [35], [36]:

- A finite set of hidden states $\boldsymbol{X} = \{X_1, X_2, ...X_m\}$.

TABLE III: State transition matrix and initial state vector for walking

Walking	State Transition Matrix					Initial
	State 1	State 2	State 3	State 4	State 5	Vector
State 1	0.12	0.23	010	0.30	0.26	0.70
State 2	0.02	0.04	0.05	0.40	0.48	0
State 3	0.01	0.37	0.22	0.23	0.16	0
State 4	0.18	0.28	0.22	0.26	0.06	0.30
State 5	0.33	0.36	0.27	0.03	0	0

- A state transition matrix $\boldsymbol{A}$ represents the probability from state X_i to state X_j as:
$$a_{ij} = p\left[q_{t+1} = X_j \mid q_t = X_i\right] \tag{5}$$
where q_t is the state at time t, $a_{ij} \geq 0$, $\sum_{j=1}^{N} a_{ij} = 1$ with $1 \leq i$ and $j \leq m$.
- A set of observations $\boldsymbol{O}$, here it represents Doppler sequences as $\boldsymbol{O} = \{D_1, D_2, ..., D_N\}$
- Emission matrix $\boldsymbol{B} = \{b(o \mid X_i)\}$ indicates the probability of the emission of output $o \in V$ at the state X_i. V is a continuous set with the probability density function $b(o \mid X_i)$
- Initial state probability vector $\boldsymbol{\pi}$
$$\pi_i = p\left[q_1 = X_i\right] \tag{6}$$
where $\pi_i \geq 0$ and $\sum_{i=1}^{N} \pi_i = 1$ with $1 \leq i \leq m$

For convenience, an HMM can be written with a triplet defines as $\boldsymbol{\lambda} = (\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi})$. And the hidden states $\boldsymbol{X}$ in this work represent the relationship between two adjacent columns within a Doppler sequence at time t and $t+1$.

We train an m-states HMM $\boldsymbol{\lambda}_i$ for each sequence D_i, $1 \leq i \leq N$ from the dataset. Firstly, two matrices $\boldsymbol{A}_{0i}$ and $\boldsymbol{\pi}_{0i}$ are randomly initialized as the input. This is because we have no priori knowledge about the model of any Doppler sequence. Secondly, Expectation Maximization (EM) is used as the optimization tool for calculating initial state vector and state transition matrix. An expectation (E) step is to create a function for expectation of current estimation, and a maximization (M) step is to compute the maximized expected log-likelihood found in E step. In this work, ten iterations are selected to find the best parameters that fit the Doppler sequence. Finally, we have the optimized matrix $\boldsymbol{A}_{1i}$ and $\boldsymbol{\pi}_{1i}$ as the parameters for HMM $\boldsymbol{\lambda}_i$. Moreover, the size of m is decided by performing a preliminary test over the recognition accuracy with state number from 2 to 7. We observe that HMMs with high state number are very easy to receive "NaN" value when calculating the covariance matrix. From the test, $m = 5$ states are found to be the optimized configuration which has the best performance 2% higher than others. An example of the transition matrix and initial vector for a walking sequence is shown in Table III. We use the Murphy's HMM toolbox [37] in this work.

B. Log-likelihood Calculation

After HMM training for each Doppler sequence, a log-likelihood matrix $L = \{L(D_i \mid D_j)\}$ is generated to present the similarities between each Doppler sequence and models. For each log-likelihood value L_{ij} is calculated as:

$$L_{ij} = \sum_{t=1}^{T_{D_i}} p(D_{i,t} \mid \boldsymbol{\lambda}_j) \tag{7}$$

where $p(D_{i,t} \mid \boldsymbol{\lambda}_j)$ is the forward probability of a testing Doppler sequence D_i with HMM $\boldsymbol{\lambda}_j$ and T_{D_i} is the length of sequence. Note that, since the inactivity period ((1) standing still activity) can be recognized by using time alignment method (described in Section II), thus the rest 200 sequences (5 activities) will be considered in this log-likelihood matrix L. Moreover, for the purpose of convenience, we use log-likelihood value instead of the likelihood as it is the direct output of HMM.

In addition, we propose to apply a 'symmetrizing' process over the log-likelihood matrix before the clustering. We know that the log-likelihood value L_{ij} presents the similarity between Doppler sequence D_i and D_j, which is calculated by testing the D_i with HMM $\boldsymbol{\lambda}_j$ trained from D_j. However it has not actually take account the sequence D_j. In other words, it does not consider how good the HMM $\boldsymbol{\lambda_j}$ is trained by Doppler sequence D_j but assumes all sequences have same training quality. Thus we believe its necessary to consider the quality of HMM for the log-likelihood matrix. Following, three symmetric log-likelihood matrices used in this paper are presented.

1) Log-likelihood Symmetric (L_S): The most straightforward approach for 'symmetrizing' is simply to summing up L_{ij} and L_{ji} for the point L_{ij}. This symmetric matrix L_S can be defined as follow:

$$L_S^{ij} = \frac{1}{2}\left[L_{ij} + L_{ji}\right] \tag{8}$$

2) Log-likelihood B-Pair (L_{BP}): Paper [38] extends the idea of L_S and proposes the matrix L_{BP} to better evaluate the log-likelihood value. L_{BP} also takes account the L_{ii} when calculate the sequence D_i. Since the log-likelihoods in diagonal of the matrix represents the log-likelihood generated with the sequence itself which can be used as the reference for the point L_{ij}. The matrix L_{BP} is defined as:

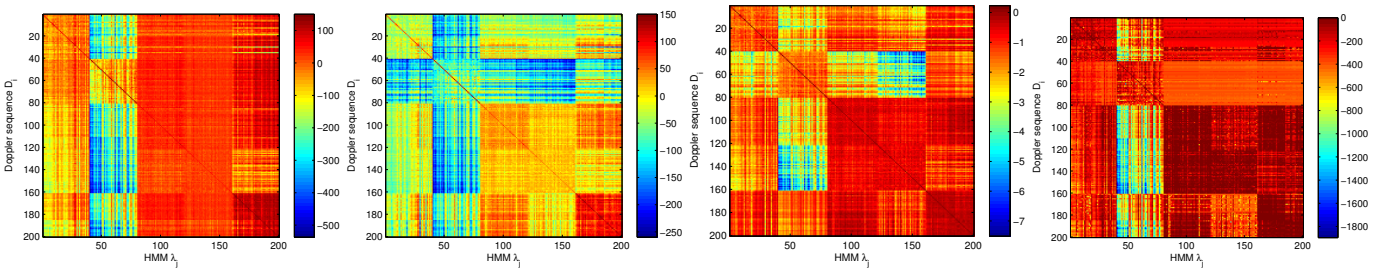
$$L_{BP}^{ij} = \frac{1}{2}\left[\frac{L_{ij}}{L_{ii}} + \frac{L_{ji}}{L_{jj}}\right] \tag{9}$$

3) Log-likelihood Kullback-Leibler (L_{KL}): Despite L_{BP} normalizes the log-likelihood L_{ij} with its corresponding reference at L_{ii}. However, the L_{BP} matrix uses linear normalization is quite simple and insufficient for the log-likelihood value. In comparison, we purpose to use the Kullback-Leibler (KL) number for L_{ij} and L_{ii}. The advantage is that it is more suitable for the log-likelihood value which provides better diversity in distance matrix in the clustering process. The matrix L_{KL} is defined as:

$$L_{KL}^{ij} = -\frac{1}{2}\left[\left|L_{ij}\left[ln\frac{L_{ij}}{L_{jj}}\right]\right| + \left|L_{ii}\left[ln\frac{L_{ii}}{L_{ji}}\right]\right|\right] \tag{10}$$

where $||$ means the absolute value, the $-$ sign is to reverse the matrix since the values on diagonal are 0s and should be the highest value in a log-likelihood matrix.

To better present the idea of log-likelihood matrix, we plot the log-likelihood matrix in Fig. 5 (a) L matrix, (b) L_S matrix, (c) L_{BP} matrix and (d) L_{KL} matrix. we get



(a) (b) (c) (d)

Fig. 5: log-likelihood matrix from (a) L, (b) L_S, (c) L_{BP} and (d) L_{KL}

TABLE IV: Differences in log-likelihood matrix

Log-likelihood	Mean Value	Max Value	STD
L	-32.26	150.56	27.28
L_S	-32.26	150.56	11.05
L_{BP}	-1.56	0.23	0.33
L_{KL}	-292.32	0	77.96

together all Doppler sequences from same activity to allow easier understanding of the log-likelihood matrix with orders as: Doppler sequence 1-40 are (2) Walking, 41-80 are (3) Running, 81-120 are (4) Jumping, 121-160 are (5) Turning, and 161-200 are (6) Standing from a chair. Same orders apply to HMMs. As the Fig 5 shows, all matrices show the clear border between different activity groups at both horizontal and vertical axis. This indicates that HMMs trained from one type of activity outputs similar log-likelihoods to other type sequence and allows the feasibility of activity recognition. As expected, the most significant value of each row is at the diagonal of the log-likelihood matrix. It can be seen the clear line in each log-likelihood matrix due to the self-testing that generate the highest similarity. We also observed that L_S matrix is very similar to L_{BP} matrix with some difference in magnitude. This is because that the log-likelihood values at diagonal L_{ii} are similar to each other, which makes the volume of $\left(\frac{L_{ij}}{L_{ii}} + \frac{L_{ji}}{L_{jj}}\right)$ close to $(L_{ij} + L_{ji})$. In addition, for L matrix (Fig. 5 (a)), we can see the log-likelihoods in activity (3), (4) and (5) are generally higher than (1) and (2), even with the HMMs generated from (1) and (2). This indicates that the training quality is various for different group activity. However, in case of activity recognition, the diversity in log-likelihood values actually shows advances in recognition performance. Therefore, this uncertainty in log-likelihood does not effect the clustering process in this work.

Afterwards, the mean, maximum and standard deviations (STD) of four log-likelihood matrices are shown in Table IV. As can be seen, L and L_S matrices have a very similar structure with the same value in mean and max. This is reasonable because the L_S symmetric the value of L, as a result, it reduces the STD value but no effect to mean and max value (max value is at diagonal). L_{KL} matrix has the lowest max value at 0 due to the value on diagonal $ln\frac{L_{ii}}{L_{ii}}$ gives 0 and the highest value in STD, While L_{BP} has the lowest value in both max and STD value.

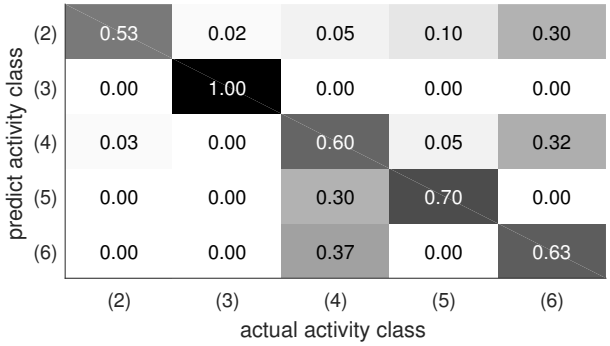


Fig. 6: Confusion matrix for log-likelihood Matrix L with K-means clustering

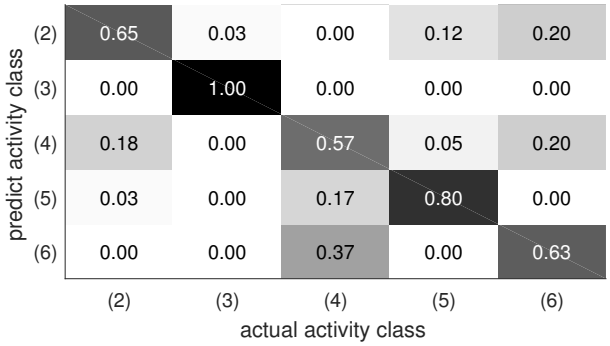


Fig. 7: Confusion matrix for log-likelihood Matrix L with K-medoids clustering

In the case where dataset is very large, the generation of log-likelihood matrix is not practical. It is possible to build a small log-likelihood matrix with a certain amount of dataset, then compute the log-likelihood for new sequence separately. Alternatively, the K-nearest neighbors (K-NN) algorithm could also be used to predict the activity group for new Doppler sequence by searching the closest data point. This enables the proposed log-likelihood approach to be used in a real application.

V. Recognition Performance & Comparison

In this section, the activity recognition performance base on K-means and K-medoids are presented. K-means and K-

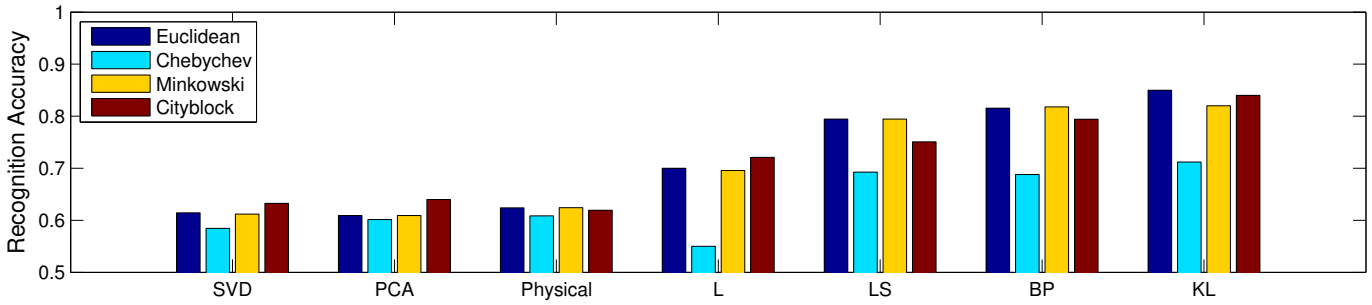


Fig. 8: Overall Accuracy Obtained by K-means Clustering with Different Distance Metric

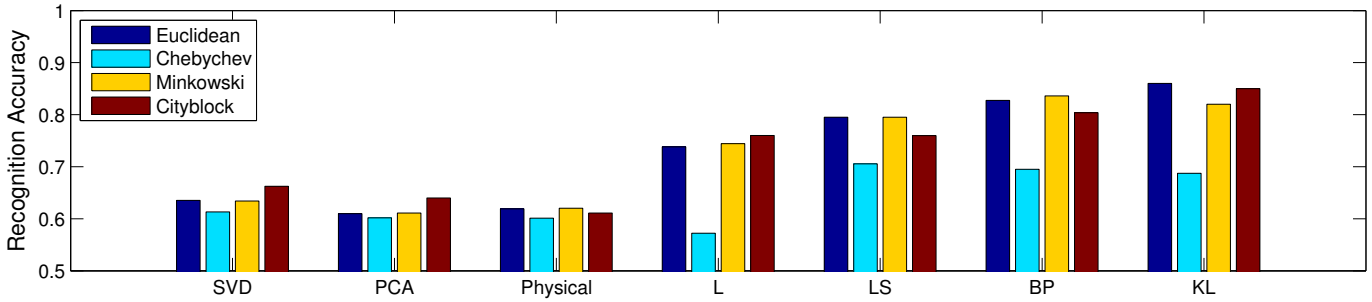


Fig. 9: Overall Accuracy Obtained by K-medoids Clustering with Different Distance Metric

medoids are two popular clustering algorithms which aim to partition observations (log-likelihood values) into clusters (activity classes) with no requirement on labels from the dataset. They exam the minimum distance between each observations and clusters and use for grouping. We compare the proposed log-likelihood matrix with other classical feature extraction methods including SVC, PCA, and physical meaning features. The results illustrate that the robustness of proposed log-likelihood matrix outperforms classical features more than 15% in recognition accuracy.

Here we provide two examples of confusion matrices for the K-means and K-medoids clustering on log-likelihood matrix L as shown in Fig 6 and 7 respectively. As can be seen from the confusion matrix, the accuracy of walking and running are better than others as expected due to the more significant diversity in log-likelihood values. Especially, the (3) running activity reaches 100% in both algorithms. While the (4) jumping has the lowest accuracy with lots of error recognition into (6) standing due to they have similar body upward movements.

Afterwards, the overall recognition performance for K-means and K-medoids with all four log-likelihood matrices are shown in Table V. To avoid the uncertainty during the clustering process, each matrix has been repeatedly tested for 10 times so that the effect of random selection on the initial centroids can be minimized. As can be seen from the table, most recognition accuracies are above 80%, while the highest accuracy is obtained with L_{KL} distance with the K-medoids algorithm. Moreover, the overall performance of K-medoids is slightly better than K-means due to the reason that K-medoids minimize the sum of dissimilarities between the clustered point and the labeled point.

TABLE V: Overall Recognition Accuracy for K-means and K-medoids with four Matrices

	L	L_{LS}	L_{BP}	L_{KL}	**Overall**
K-means	**0.70**	**0.79**	**0.82**	**0.85**	**0.79**
K-medoids	**0.73**	**0.80**	**0.82**	**0.86**	**0.80**

A. Accuracy Comparison between Different Features

Following statements are three popular features which have been widely used in many radar-based activity recognition [11], [12], [21], [23], [24], [26]. Those features are generated from the same dataset to compare with the proposed log-likelihood matrix regarding the recognition performance.

1) Singular Value Decomposition: SVD is a tool to reduce the data dimension and has been used in analysis micro Doppler signatures [26]. Given the matrix D of Doppler sequence, the SVD can be carried out as $D = USV^T$ where U and V are the matrices with left and right singular vectors, S is the diagonal matrix with singular values of D. We output the eigenvalue from both U and V as the features.

2) Principal Component Analysis: PCA is known as a powerful technique in feature extraction. It breaks the data into its component vectors based on the eigenvectors of covariance matrix which has been widely used in radar application [12]. The matrix contains only the variation from the mean of the Doppler sequence is noted as D^T. Then the covariance of DD^T is calculated and the corresponding eigenvectors W are computed. Afterwards, a truncated eigenvectors W_L is used to further reduce the data dimension. Then the matrix is projected as $Y_L^T = D^T W_L$, where Y_L^T is the PCA coefficients.

3) Features with Physical Meaning: Another type of feature extraction is to obtaining the physical meaning from Doppler sequence [11], [21], [23]. The following six features are selected according to our previous work [21]: (1) the duration of the activity, (2) the maximum upper Doppler shift of the activity, (3) the maximum lower Doppler shift of the activity, (4) the peak-to-peak bandwidth of Doppler, (5) the mean power of the activity and (6) the standard deviation of the power of the activity.

As mentioned before, above feature extraction methods require a fixed time window. For this reason, a 2.5s time window has been applied which has been determined as the optimized window length after we test the performance from 0.5s to 4s. Also we further test the clustering algorithm with four distance metrics including Euclidean, Chebychev, Minkowski and Cityblock. The test on those distance metrics can further reveal the robust of log-likelihood matrix over the previous approaches. The recognition accuracy for both K-means and K-medoids with SVD, PCA, Physical features and four log-likelihood matrices are shown in Fig 8 and Fig 9 respectively. As can be seen, the overall accuracy of log-likelihood matrix clustering outperform the previous approaches. For SVD, PCA and physical features, they have the accuracies around 60%, while the accuracies of log-likelihood matrix are mostly above 80%. There are two main reasons for the robust of log-likelihood matrix. Firstly, as mentioned before, the log-likelihood matrix avoids the sliding-window algorithm so that the time-varying feature of Doppler spectrogram are also included. Secondly, unlike previous feature extraction methods (SVD, PCA and physical feature) which all features for one Doppler sequence are generated from its own. The log-likelihood values for each Doppler sequence are generated from the HMMs from other Doppler sequences. This means the log-likelihood matrix also takes into account the relationship between every two Doppler sequences, as a result it provides more diversity in distance matrix. In addition, the proposed "symmetrizing" process including L_S, L_{BP} and L_{KL} show more improvements in recognition accuracy when compared to the original log-likelihood matrix L. Especially, the L_{KL} gives the best performance with average 5% higher than the rest. This is benefited by considering the training quality of HMMs as the normalization with self log-likelihoods at L_{ii} during the "symmetrizing" process. Moreover, the only degradation performance for log-likelihood matrix is with the Chebychev distance metric. This is due to the Chebychev metric picks the maximum difference between two vectors as distance. However, this is insufficient for vectors with large amount of elements due to the diversity cannot be fully presented. But for L_S, L_{BP} and L_{KL}, they are still better than SVD, PCA and physical features.

These results indicate that the proposed log-likelihood matrix can largely improve the recognition performance over the previous approaches. Motivated by this result, we envision that the idea of use log-likelihood as the representation for Doppler pattern can be extended to other machine learning methods or classifiers like SVM, classification tree, etc.

VI. Conclusions

This paper presents a novel approach for passive RF sensing based activity recognition which breaks the limitation of traditional sliding window feature extraction techniques, provides adaptive capability in recognizing the daily activities. Our HMM-based clustering copes directly with this fundamental challenge by using the HMM-log-likelihood as the measurement of similarity between Doppler sequences. We apply this novel feature construction method in our passive radar system which uses K-means and K-medoids clustering for categorizing different activities. The experimental results show that the HMM-log-likelihood based feature outperform traditional SVD, PCA features and represent robustness to varying activity durations. However, RF sensing based activity capturing is still a promising area with open topics. Our future work will focus on tackling other challenges in passive indoor RF sensing like multiple signal sources and multiple users problems on the basis of time-adaptive feature characterization. We also plan to deploy this system in the SPHERE house [2] for out-of-lab trials in coming future with even larger number of random subjects.

References

[1] Physical activity. [Online]. Available: http://www.who.int/topics/physical_activity/en/

[2] P. Woznowski, A. Burrows, T. Diethe, X. Fafoutis, J. Hall, S. Hannuna, M. Camplani, N. Twomey, M. Kozlowski, B. Tan, N. Zhu, A. Elsts, A. Vafeas, A. Paiement, L. Tao, M. Mirmehdi, T. Burghardt, D. Damen, P. Flach, R. Piechocki, I. Craddock, and G. Oikonomou, "Sphere: A sensor platform for healthcare in a residential environment," *Designing, Developing, and Facilitating Smart Cities: Urban Design to IoT Solutions*, pp. 315–333, 2017. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-44924-1-14

[3] S. Sadanand and J. J. Corso, "Action bank: A high-level representation of activity in video," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, June 2012, pp. 1234–1241.

[4] O. Oreifej and Z. Liu, "Hon4d: Histogram of oriented 4d normals for activity recognition from depth sequences," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, June 2013, pp. 716–723.

[5] B. Yao and L. Fei-Fei, "Modeling mutual context of object and human pose in human-object interaction activities," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2010, pp. 17–24.

[6] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys Tutorials*, vol. 15, no. 3, pp. 1192–1209, Third 2013.

[7] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, "Sensor-based activity recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 790–808, Nov 2012.

[8] A. Fleury, M. Vacher, and N. Noury, "Svm-based multimodal classification of activities of daily living in health smart homes: Sensors, algorithms, and first experimental results," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 274–283, March 2010.

[9] B. Tan, K. Woodbridge, and K. Chetty, "Awireless passive radar system for real-time through-wall movement detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2596–2603, October 2016.

[10] J. L. Geisheimer, E. F. Greneker III, and W. S. Marshall, "High-resolution doppler model of the human gait," in *AeroSense 2002*. International Society for Optics and Photonics, 2002, pp. 8–18.

[11] Y. Kim and H. Ling, "Human activity classification based on micro-doppler signatures using a support vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1328–1337, 2009.

[12] J. Bryan, J. Kwon, N. Lee, and Y. Kim, "Application of ultra-wide band radar for classification of human activities," *IET Radar, Sonar & Navigation*, vol. 6, no. 3, pp. 172–179, 2012.

[13] Z. Chi, Y. Yao, T. Xie, Z. Huang, M. Hammond, and T. Zhu, "Harmony: Exploiting coarse-grained received signal strength from iot devices for human activity recognition," in *2016 IEEE 24th International Conference on Network Protocols (ICNP)*, Nov 2016, pp. 1–10.

[14] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-invasive detection of moving and stationary human with wifi," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2329–2342, 2015.

[15] C. Yang and H. r. Shao, "Wifi-based indoor positioning," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 150–157, March 2015.

[16] F. Adib and D. Katabi, "See through walls with wifi!" *SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 75–86, Aug. 2013.

[17] K. Chetty, G. E. Smith, and K. Woodbridge, "Through-the-wall sensing of personnel using passive bistatic wifi radar at standoff distances," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1218–1226, April 2012.

[18] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 363–373.

[19] W. Li, B. Tan, and R. J. Piechocki, "Non-contact breathing detection using passive radar," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[20] W. Li, B. Tan, R. J. Piechocki, and I. Craddock, "Opportunistic physical activity monitoring via passive wifi radar," in *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, Sept 2016, pp. 1–6.

[21] W. Li, Y. Xu, B. Tan, and R. J. Piechocki, "Passive wireless sensing for unsupervised human activity recognition in healthcare," in *2017 International Wireless Communications and Mobile Computing Conference (IWCMC)*, Aug 2015.

[22] M. O. Padar, A. E. Ertan, and Ç. ĝatay Candan, "Classification of human motion using radar micro-doppler signatures with hidden markov models," in *Radar Conference (RadarConf), 2016 IEEE*. IEEE, 2016, pp. 1–6.

[23] B. alyan and S. Z. Grbz, "Micro-doppler-based human activity classification using the mote-scale bumblebee radar," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 10, pp. 2135–2139, Oct 2015.

[24] J. Park and S. H. Cho, "Ir-uwb radar sensor for human gesture recognition by using machine learning," in *2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Dec 2016.

[25] D. P. Fairchild and R. M. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-doppler signatures," *IET Radar, Sonar Navigation*, vol. 8, no. 5, pp. 425–434, June 2014.

[26] F. Fioranelli, M. Ritchie, and H. Griffiths, "Performance analysis of centroid and svd features for personnel recognition using multistatic micro-doppler," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 725–729, May 2016.

[27] E. Keogh and J. Lin, "Clustering of time-series subsequences is meaningless: implications for previous and future research," *Knowledge and information systems*, vol. 8, no. 2, pp. 154–177, 2005.

[28] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[29] M. O. Padar, A. E. Ertan, and . . Candan, "Classification of human motion using radar micro-doppler signatures with hidden markov models," in *2016 IEEE Radar Conference (RadarConf)*, May 2016, pp. 1–6.

[30] M. Jackson, "The geometry of bistatic radar systems," in *IEE Proceedings F (Communications, Radar and Signal Processing)*, vol. 133, no. 7. IET, 1986, pp. 604–612.

[31] B. Tan, K. Woodbridge, and K. Chetty, "A real-time high resolution passive wifi doppler-radar and its applications," in *Radar Conference (Radar), 2014 International*. IEEE, 2014, pp. 1–6.

[32] F. Ansari and M. R. Taban, "Implementation of sequential algorithm in batch processing for clutter and direct signal cancellation in passive bistatic radars," in *2013 21st Iranian Conference on Electrical Engineering (ICEE)*, May 2013, pp. 1–6.

[33] K. Kulpa, "The clean type algorithms for radar signal processing," in *Microwaves, Radar and Remote Sensing Symposium, 2008. MRRS 2008*, Sept 2008, pp. 152–157.

[34] S. Masood, M. P. Qureshi, M. B. Shah, S. Ashraf, Z. Halim, and G. Abbas, "Dynamic time wrapping based gesture recognition," in *Robotics and Emerging Allied Technologies in Engineering (iCREATE), 2014 International Conference on*, April 2014, pp. 205–210.

[35] P. Smyth, "Clustering sequences with hidden markov models," in *Advances in neural information processing systems*, 1997, pp. 648–654.

[36] M. Bicego, V. Murino, and M. A. Figueiredo, "Similarity-based clustering of sequences using hidden markov models," in *International Workshop on Machine Learning and Data Mining in Pattern Recognition*. Springer, 2003, pp. 86–95.

[37] Hidden markov model (hmm) toolbox for matlab. [Online]. Available: https://www.cs.ubc.ca/ murphyk/Software/HMM/hmm.html

[38] A. Panuccio, M. Bicego, and V. Murino, "A hidden markov model-based approach to sequential data clustering," *Structural, Syntactic, and Statistical Pattern Recognition*, pp. 734–743, 2002.