

Research Article

Feature Selections Using Minimal Redundancy Maximal Relevance Algorithm for Human Activity Recognition in Smart Home Environments

Hongqing Fang , Pei Tang, and Hao Si

College of Energy and Electrical Engineering, Hohai University, Nanjing, Jiangsu 211100, China

Correspondence should be addressed to Hongqing Fang; fanghongqing@gmail.com

Received 13 September 2020; Revised 17 October 2020; Accepted 9 November 2020; Published 29 November 2020

Academic Editor: Ivan Miguel Pires

Copyright © 2020 Hongqing Fang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, maximal relevance measure and minimal redundancy maximal relevance (mRMR) algorithm (under D-R and D/R criteria) have been applied to select features and to compose different features subsets based on observed motion sensor events for human activity recognition in smart home environments. And then, the selected features subsets have been evaluated and the activity recognition accuracy rates have been compared with two probabilistic algorithms: naïve Bayes (NB) classifier and hidden Markov model (HMM). The experimental results show that not all features are beneficial to human activity recognition and different features subsets yield different human activity recognition accuracy rates. Furthermore, even the same features subset has different effect on human activity recognition accuracy rate for different activity classifiers. It is significant for researchers performing human activity recognition to consider both relevance between features and activities and redundancy among features. Generally, both maximal relevance measure and mRMR algorithm are feasible for feature selection and positive to activity recognition.

1. Introduction

The aging of population and the increasing number of the elderly who chooses to live on their own [1–3] is an indisputable social reality. To implement the goals, smart home technology can play an important role to detect and analyze health events [4] and to provide corresponding medical assistant and caregiver for frail elderly and disabled people who are unable to live independently for a long period of time and in their home far away from hospital or their families, e.g., to remind them of time to take medicine, to see the doctor, to assist them in cutting off the water, turning off the oven, etc.

Actually, accurate assessment of human Activities of Daily Livings (ADLs) is the prerequisite for smart home to yield the correct service, whether it is for the elders or for the severe disabilities of health monitoring, or provide them with other relevant helps. Once the dangerous behavior is

detected, smart home itself can cope with it and eliminate as much of the inhabitant's risk as possible. Therefore, the accurate recognition of human activity in smart home is of great significance and gives a pattern for the realization of healthcare for solitary elderly or disabled as well, which is the most important process in incorporating ambient intelligence into smart environments [5–8].

Recently, human activity discovery and recognition has gained a lot of interest due to its enormous potential in context aware computing systems, including smart home environments. The primary objective of human activity recognition in smart home environment is to find the interesting patterns of behavior from gathered sensor data and to recognize such patterns. Currently, one of the primary challenges of human activity recognition is the choosing of machine learning algorithms which perform better in the same sequence of sensor data collected by smart home environment during the activity. In the last years, several

intelligent algorithms applied for human activity recognition in smart home have been reported. Singla, Crandall, and Cook et al. described the applications of some probability-based algorithms that include naïve Bayes (NB) classifier, Markov model (MM), and hidden Markov model (HMM) [9, 10] to train the partially labeled motion sensor events data to obtain the values of the prior parameters and then to validate the performances of the algorithms by testing the rest of labeled sensor data. Liu et al. presented a Bayesian network-based probabilistic generative framework based on Allen's temporal relations over primitive events to characterize the structural variabilities of complex activities [11]. Gayathri et al. proposed a statistical relational learning approach which augments ontology based activity recognition with probabilistic reasoning through Markov Logic Network (MLN) [12]. Kim et al. proposed a discriminative and generative probabilistic model, conditional random field (CRF), as a more flexible alternative to HMM [13]. Zhu et al. presented a two-layer CRF model to represent the action segments and activities in a hierarchical manner, which allows the integration of both motion and various context features at different levels and automatically learns the statistics that capture the patterns of the features [14]. Chen et al. introduced a knowledge-driven approach to continuous activity recognition based on multisensor data streams in smart homes [15]. Fahim et al. introduced a novel Evolutionary Ensembles Model (EEM) that values both minor and major activities by processing each of them independently, which is based on a Genetic Algorithm (GA) to handle the nondeterministic nature of activities [16]. Fleury et al. proposed support vector machine- (SVM-)based multimodal classification of ADLs in health smart homes [17]. Wen and Wang combined Latent Dirichlet Allocation (LDA) and AdaBoost to jointly train a general activity recognition model with partially labeled data [18]. Hong et al. composed a hybrid model of Bayesian networks and support vector machines to accurately recognize human activity [19].

Besides the suitable choosing of machine learning algorithms, another key point for human activity recognition in smart home is to select valid features from sensor events datasets collected in smart home environment. Usually, sensor events datasets include a large number of observed sensor events generated by various activities and any activity annotated in the dataset has various features, even redundant and irrelevant features [20]. However, these features are selected in one method in all tests, almost, and the influences of these features on the performance of classifiers are seldom addressed. Actually, the features which are irrelevant to activity recognition and redundant between initial features need to be removed prior to activity recognition. Furthermore, feature selection means to select the features subset which is the most favorable for activity recognition and compressing of data successfully.

Recently, minimal redundancy maximal relevance (mRMR) feature selection algorithm has been widely applied in many researching fields, which aims to achieve the best classification performance by reducing redundancy among the selected features and maximizing their relevance to the

target class. Mohamed et al. implemented the mRMR filter and a metaheuristic approach as a feature selection process for drug response microarray classification [21]. Che et al. presented a novel mutual information feature selection method based on the normalization of the maximum relevance and minimum common redundancy for nonlinear classification or regression problems [22]. Xu et al. proposed a new distributed monitoring scheme which integrates mRMR, Bayesian inference, and principal component analysis for plant-wide processes [23]. Li et al. provided a granular feature selection method with an mRMR criterion based on mutual information (MI) for multilabel learning [24]. Escalona-Vargas et al. proposed a method that uses mRMR as criteria to automatically select references for the frequency-dependent subtraction method to attenuate maternal and fetal magnetocardiograms of fetal magnetoencephalography recordings [25]. Mallik et al. developed a new framework of identifying statistically significant epigenetic biomarkers using mRMR criterion-based feature (gene) selection for multiomics dataset [26]. Tiwari calculated the weighted distance to improve the prediction performance of G-protein coupled receptors families and their subfamilies by using sequence derived properties, and the feature selection method based on fusion of mRMR and other supervised filter were provided [27]. Chen and Yan developed an optimized multilayer feedforward network by using mRMR-partial mutual information clustering integrated with least square regression to construct a soft sensor for controlling naphtha dry point [28]. Wang et al. presented a multiobjective evolutionary algorithm which employs the Pareto optimality to evaluate candidate feature subsets and finds compact feature subsets with both the maximal relevance and the minimal redundancy [29]. Morgado and Silveira proposed a multivariate procedure capable of selecting nonredundant subsets of features significantly faster than other similar methods to the diagnosis of Alzheimer's disease and related disorders which is inspired in mRMR algorithm [30]. Kamandar and Ghassemian used a modified mRMR as a criterion for feature extraction for hyperspectral images classification based on information theoretic learning [31]. Kandaswamy et al. extracted the best features using mRMR feature selection algorithm and used the random forest algorithm to predict extracellular matrix proteins [32]. Jin et al. proposed a novel method for health monitoring and anomaly detection for cooling fans in electronic products based on Mahalanobis distance with mRMR features [33]. Unler et al. presented a hybrid filter-wrapper feature subset selection algorithm based on particle swarm optimization for support vector machine classification. The filter model is based on the mutual information and is a composite measure of feature relevance and redundancy with respect to the feature subset selected [34]. Zdravevski et al. proposed a generic feature engineering method for selecting robust features from a variety of sensors, which is from the originally recorded time series and some newly generated time series (i.e., magnitudes, first derivatives, delta series, and fast Fourier transformation-(FFT-) based series); a variety of time and frequency domain features are extracted [35, 36].

Inspired by the ideas mentioned above, in this paper, maximal relevance measure and mRMR feature selection algorithm (under D-R and D/R criteria) have been applied to select features and to compose different features subsets based on the observed motion sensor events for human activity recognition in smart home environments. And then, the selected features subsets have been evaluated, and the activity recognition accuracy rates have been compared with two probabilistic algorithms: NB classifier and HMM.

The rest of the paper is organized as follows: Section 2 describes the smart apartment testbed, the data collection, and 13 features of observed sensor events for human activity recognition. Section 3 presents the concepts of information entropy, mutual information (MI), and minimal redundancy maximal relevance (mRMR) feature selection algorithm. Section 4 gives the training and testing activities, the features subsets selected through analyzing of maximal relevance measure, and mRMR feature selection algorithm. Finally, we present the comparison results of activity recognition accuracy rates with the selected features subsets and the performance measures of NB classifier as well as HMM. Section 5 summarizes the main contributions.

2. Smart Apartment Testbed and Data Collection for Human Activity Recognition

2.1. Smart Apartment Testbed and Data Collection. The smart apartment testbed for this research is located on Washington State University campus and is maintained as part of the Center for Advanced Studies in Adaptive Systems (CASAS) smart home project, which includes three bedrooms, one bathroom, a kitchen, and a living/dining room. The smart apartment is equipped with motion sensors distributed approximately 1 meter apart throughout the space on the ceilings, as shown in Figure 1. In addition, other sensors installed provide ambient temperature readings and custom-built analog sensors provide readings for hot water, cold water, and stove burner use. Sensor data is captured using a sensor network that was designed in-house and is stored in a Structured-Query-Language (SQL) database. After collecting data from the smart apartment testbed, the sensor events are annotated for ADLs, which are used for training and testing the activity recognition algorithms [3, 4, 7, 9, 10].

The data gathered by CASAS smart home is represented by the following parameters, which specify the number of features that are used to describe the observed sensor events. The generalized syntax of the dataset (Cairo Dataset, 2009) is

Date Time SensorID SensorValue <label>

An example of the dataset of Night_wandering activity is:

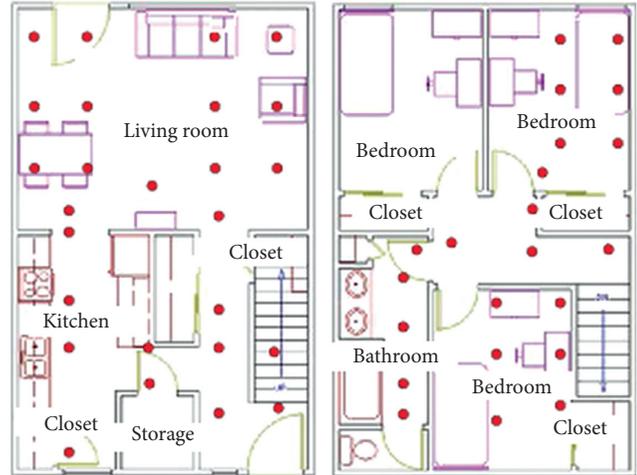
{

2009-06-10 03:20:59.08 M006 ON Night_wandering begin

2009-06-10 03:25:19.05 M012 ON

2009-06-10 03:25:19.08 M011 ON

2009-06-10 03:25:24.05 M011 OFF



● Motion sensor

FIGURE 1: The smart apartment testbed and sensors in the apartment to monitor motion.

```
2009-06-10 03:25:24.07 M012 OFF Night_wandering
end
}
```

This example shows that the observed sensor events correspond to the Night_wandering activity with concrete Date, Time, Sensor ID, Sensor Value, and activity label parameters.

2.2. Features of Observed Sensor Events. Considering the actual situation, each activity has 13 features of the observed sensor events:

- (1) The means of logical values of Sensor IDs of each activity's sensor events is f_1 .

Considering that the place where each activity happens is relatively stable, therefore, selecting the average of Sensor IDs means the focus area where the activity occurs. The equation is

$$\bar{S}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} S_{ik}, \quad (1)$$

where \bar{S}_i is the means of Sensor IDs of activity i , n_i is the number of sensors, and S_{ik} is the k^{th} Sensor ID of activity i .

- (2) The variance of all Sensor IDs triggered by the current activity, f_2 , is

$$S_i^2 = \frac{1}{n_i} \sum_{k=1}^{n_i} (S_{ik} - \bar{S}_i)^2. \quad (2)$$

- (3) Day of the week, which is converted into a value in the range of 0 to 6, is f_3 .

- (4) Previous activity, which represents the activity that occurred before the current activity, is f_4 .
- (5) Activity length, which represents the length of the current activity measured in a number of sensor events, is f_5 .
- (6) The logical value of the first Sensor ID triggered by the current activity is f_6 .
- (7) The logical value of the last Sensor ID triggered by the current activity is f_7 .
- (8) The duration of the current activity, which indicates the time interval between the last sensor and the first sensor triggered by the current activity, is f_8 .
- (9) The beginning time of the current activity is f_9 .
- (10) The ending time of the current activity is f_{10} .
- (11) Next activity, which represents the activity that occurred after the current activity, is f_{11} .
- (12) The mode value of the Sensor IDs triggered by the current activity is f_{12} .
- (13) The median value of the Sensor IDs triggered by the current activity is f_{13} .

Usually, the optimal features subset contains the least number of dimensions that contribute to higher recognition accuracy rate. Therefore, it is necessary to remove the remaining and unimportant features.

3. Minimal Redundancy Maximal Relevance Feature Selection Algorithm

3.1. Entropy. Information entropy is a more abstract mathematical concept which can be understood as the probability of the emergence of some specific information. Generally, the higher of the probability of a kind of information indicates that it is spread more widely or more highly cited. Therefore, information entropy is able to represent the value of information. The information source is the source of the message and the message sequence. For example, the simplest discrete information source is $X = \{x_1, x_2, \dots, x_n\}$, and $p(x_i)$ is the probability of a given x_i ; then, the entropy of X can be defined as

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i). \quad (3)$$

3.2. Conditional Entropy, Joint Entropy, and Mutual Information

3.2.1. Conditional Entropy. In information theory, the conditional entropy quantifies the amount of information needed to describe the outcome of a random variable X given that the value of another random variable Y is known. If X and Y are dependent on each other, then, in the condition of $\{Y = y\}$, the conditional entropy of X is

$$H(X|Y = y) = \sum_x p(x|y) \log \frac{1}{p(x|y)}. \quad (4)$$

If Y is given, then the conditional entropy of X is

$$\begin{aligned} H(X|Y) &= \sum_y p(y) H(X|Y = y), \\ &= \sum_{x,y} p(x,y) \log \frac{1}{p(x|y)}, \end{aligned} \quad (5)$$

where $p(x, y)$ is the joint probability of x and y .

3.2.2. Joint Entropy. Joint entropy is a measure of the uncertainty associated with a set of random variables. Supposing two random variables X and Y and each of them is given a limited value, then, the joint entropy is

$$H(X, Y) = - \sum_{x,y} p(x, y) \log p(x, y). \quad (6)$$

The joint entropy is a measure of the correlation of X and Y . If X and Y are independent, then, the joint entropy is

$$H(X, Y) = H(X) + H(Y). \quad (7)$$

3.2.3. Mutual Information. Mutual information (MI) is a quantity that measures the level of similarity as well as correlation of random variables [33, 37]. Supposing two random variables X and Y and Y contains some information of X , then MI between X and Y can be defined as

$$I(X; Y) = \sum_{x,y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}. \quad (8)$$

MI is typically defined as the measure of the mutual dependence of two random variables. A larger value of MI means a closer relationship between the two random variables which have larger correlation. If the value of MI is zero, it means that the two random variables are uncorrelated and independent. Therefore, in this paper, MI can be used to measure the similarity among features and the correlation between feature and activity.

3.3. Minimal Redundancy Maximal Relevance (mRMR) Algorithm. Although MI is widely applied in the feature selection fields, there still exist some deficiencies. Most of the feature selection algorithms only consider the relationship between features and classification categories but ignore the mutual influence among features. Instead, mRMR feature selection algorithm applied in this paper considers not only the amount of information provided by these features for categorical attributes but also the influence of interaction among features on classification [37].

MI can weigh the quantity of information between feature variables X and Y ; furthermore, it can measure how much information quantity that X can provide to Y to

classify activity as well. Therefore, MI can not only show the ability of each feature of identifying the activities, but also measure whether there is redundancy among features. According to the traits of MI, two different criteria can be extended to evaluate the features: redundancy and relevance.

3.3.1. Redundancy Measure. Redundancy measure utilizes the quantity of MI between features. If the value of MI is large, it means that there is a large amount of information duplication between the two features; i.e., there are redundancies between the two features. A lower value of redundancy measure indicates a better feature selection criterion. Utilizing redundancy measure is to find the feature which has the minimal value of MI among all features.

According to the idea that the smaller the value of redundancy of information between features is, the more beneficial it is to activity classification, which can be expressed by minimizing the MI among features, the minimal redundancy condition is

$$\min R(S), R = \frac{1}{|s|^2} \sum_{x_i, x_j \in S} I(x_i; x_j), \quad (9)$$

where $|s|$ is the number of features in features subset S and $I(x_i; x_j)$ is MI between feature i and j .

3.3.2. Relevance Measure. Relevance measure utilizes the value of MI between the feature and the target activity. If the value of MI is small, it indicates that there is a weak correlation between the feature and the target activity. On the contrary, the larger value of MI means that the feature has greater amount of information to classify the activity. Therefore, it is necessary to select the maximum value of MI between the features and the target activity, the maximal relevance criterion, which can be expressed as

$$\max D(S, c), D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c), \quad (10)$$

where c is the target activity and $I(x_i; c)$ is the MI between the feature i and the target activity c .

Feature selection based on the maximum value of relevance measure is to compose the optimal features subset by selecting MI to target activity.

3.3.3. Algorithm Designing. mRMR feature selection algorithm is based on the relevance measure and redundancy measure described above. It combines the relevance between the features and the target activity as well as the redundancy among the features [33, 37]. According to (9) and (10), the influences of relevance measure and redundancy measure have been taken consideration in feature selection, comprehensively. mRMR feature selection algorithm has two evaluation criteria, which optimizes D and R , simultaneously, as

$$\max \Phi_1(S, c), \Phi_1 = D - R, \quad (11)$$

$$\text{or } \max \Phi_2(S, c), \Phi_2 = \frac{D}{R}. \quad (12)$$

Supposing that there is a features subset S_m which is composed of m features, the next step is to extract the optimal $(m+1)^{\text{th}}$ feature from the features subset $\{S - S_m\}$ according to (11) or (12), through

$$\max_{x_i \in S - S_m} \left[I(x_i; c) - \left(\frac{1}{m} \right) \sum_{x_j \in S_m} I(x_i; x_j) \right], \quad (13)$$

or

$$\max_{x_i \in S - S_m} \left[\frac{I(x_i; c)}{(1/m) \sum_{x_j \in S_m} I(x_i; x_j)} \right]. \quad (14)$$

The incremental procedures of mRMR feature selection algorithm are as follows:

Step 1. In the original features set Ω , the optimal feature x_i can be selected by $I(x_i; c)$ and then put into the optimal features subset S ;

Step 2. In the features subset $\Omega_S = \Omega - S$, the next optimal feature x_j is selected which satisfies (11) or (12);

Step 3. Repeat Step 2 to find the optimal features subset S which meets the size requirement finally.

4. Experimental Results

4.1. Training and Testing Activities. A total of 10 activities were performed in the CASAS smart apartment by 2 volunteers to provide physical training data for NB classifier and HMM. These activities include both basic and more complex ADLs that are found in clinical questionnaires. These activities are the following:

- (1) Bed_to_toilet (activity 0): transition happens between bed and toilet in the night time (30 instances)
- (2) Breakfast (activity 1): the resident has breakfast (48 instances)
- (3) Bed (activity 2): this is the activity of sleeping in bed (207 instances)
- (4) C_work (activity 3): the resident works in the office space (46 instances)
- (5) Dinner (activity 4): the resident has dinner (42 instances)
- (6) Laundry (activity 5): the resident cleans clothes using the laundry machine (10 instances)
- (7) Leave_home (activity 6): the resident leaves smart home (69 instances)
- (8) Lunch (activity 7): the resident has lunch (37 instances)
- (9) Night_wandering (activity 8): the resident wanders during night time (67 instances)

- (10) R_medicine (activity 9): the resident takes medicine (44 instances)

The data have been collected in the CASAS smart apartment testbed for 55 days, resulting in total of 600 instances of these activities and 647, 485 collected motion sensor events. The 3-fold cross validation is applied in the data for NB classifier and HMM under the same conditions to ensure that the experimental comparison is fair [38].

4.2. Feature Selection with Maximal Relevance Measure. Firstly, maximal relevance (MR) measure of the 13 features has been analyzed. Usually, MI is employed as the criterion of evaluation. Since the feature with larger value of MI has the closer relationship to the target activity, therefore, the values of MI between features and activities must be calculated. By analyzing relevance measure, the ranks of the 13 features according to the values of MI are listed in Table 1, and then the 13 features have been sorted in a descending order successively, composed of 13 features subsets, respectively [39].

It can be found that feature f_{10} has the largest value of MI, which means that f_{10} is the most important feature to activity recognition and has the closest relationship to the target activity. On the contrary, f_3 has the smallest value of MI, which means that f_3 is not discriminatory to activity recognition and has the weakest link to the target activity.

Then, the 13 features subsets with MR measure are the following:

- Features subset 1: (f_{10})
- Features subset 2: (f_{10}, f_9)
- Features subset 3: (f_{10}, f_9, f_6)
- Features subset 4: (f_{10}, f_9, f_6, f_8)
- Features subset 5: ($f_{10}, f_9, f_6, f_8, f_7$)
- Features subset 6: ($f_{10}, f_9, f_6, f_8, f_7, f_1$)
- Features subset 7: ($f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}$)
- Features subset 8: ($f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4$)
- Features subset 9: ($f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}$)
- Features subset 10: ($f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}, f_5$)
- Features subset 11: ($f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}, f_5, f_2$)
- Features subset 12: ($f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}, f_5, f_2, f_{13}$)
- Features subset 13: ($f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}, f_5, f_2, f_{13}, f_3$)

The activity recognition accuracy rate can be divided into two categories: the individual activity recognition accuracy rate (IARAR) and the total activities recognition accuracy rate (TARAR). The definitions are

$$\text{IARAR} = \frac{\text{correct sample number of this activity}}{\text{the sample number of this activity}}, \quad (15)$$

$$\text{TARAR} = \frac{\text{total correct sample number}}{\text{total sample number}}.$$

The results of activity recognition accuracy rates with the 13 features subsets of NB classifier based on MR measures

are shown in Table 2, and the results of HMM are shown in Table 3, respectively.

From Tables 2 and 3, the 13 features impact on the classifiers' recognition accuracy rate can be found differently. Similarly, for NB classifier and HMM, TARARs vary from different features subsets. Generally, with the increasing of size of features subset, activity recognition accuracy rate increases. TARAR tends to be stable as the size of features subset becomes larger finally. However, the trend is not monotonic. Concretely, for NB classifier, TARAR does not improve with features subset 7 (88.7%) which introduces feature f_7 in features subset 6 (88.8%). It means that not all the features are positive to activity recognition obviously and different combination of features has different effect on activity recognition accuracy rate.

It also can be found that the maximal value of TARAR of NB classifier is slightly higher than that of HMM. With features subset 12, the TARAR of NB classifier reaches the maximal value 90.3%, while, for HMM, the maximal value of TARAR is 88.0% with features subset 13.

Furthermore, it can be seen that the trends of IARARs of NB classifier and HMM are similar as well; i.e., with the increasing of the number of features in features subset, IARAR increases until a specific number. Again, the trends are not monotonic. For example, for activity 1, IARARs of NB classifier and HMM have risen along with the sizes of features subset. For NB classifier, the optimal value of IARAR is 100.0%, with features subset 8, and then drop down with the increasing of size of features subset. For HMM, IARAR reaches the optimal value of 93.8% with features subset 6 and then drop down with the increasing of size of features subset. Moreover, from IARARs of activity 1, it can be seen that the performances of NB classifier and HMM are different by introducing the same feature, such as feature f_4 , compared with features subsets 7 (97.9%) and 8 (100.0%); NB generates a positive result with this feature; however, the recognition performance of HMM does not improve, from 89.6% decreasing to 83.3%. Even with same features subset, the recognition accuracy rates may be quite different from the performances of classifiers, which can be found much more from the recognition accuracy rates for each activity. For example, the performances of NB classifier and HMM differ greatly for activities 0 with features subset 1 (0.0% vs. 90.0%) and activity 2 with features subset 3 (88.4% vs. 39.1%).

Furthermore, from Table 2, it is indicated that the relatively better features subset is different for each activity; e.g., for activity 3, features subset 10 yields the best result, and the accuracy rate is 76.1%, while features subset 2 generates the best result of activity 2, and the accuracy rate is 100.0%, for NB classifier. With features subset 10, IARARs of activities 3, 4, 5, and 6 are better than or equal to those of the other features subsets; the proportion is 40% for all activities. However, with features subsets 12 and 13, the IARARs of activities 0 through 2, and 4, and 6 through 8 are better than or equal to those of features subset 10; the proportions are 70%. Therefore, the optimal results of TARARs (90.3%) are of features subsets 12 and 13. Relatively, features f_3 and f_{13} are not discriminatory to activity recognition with NB classifier.

TABLE 1: The values of mutual information (MI) and the ranks of the 13 features.

Feature	1	2	3	4	5	6	7	8	9	10	11	12	13
MI	1.1879	0.6885	0.0472	1.1734	0.8628	1.4903	1.2003	1.3213	1.5084	1.5350	1.1675	1.1832	0.5304
Rank	6	11	13	8	10	3	5	4	2	1	9	7	12

TABLE 2: The recognition accuracy rates with all features subsets of NB classifier based on maximum relevance selection.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.000	0.433	0.433	0.300	0.300	0.567	0.600	0.567	0.600	0.600	0.633	0.667	0.667
1	0.792	0.729	0.917	0.938	0.917	0.979	0.979	1.000	0.979	0.979	0.979	0.979	0.979
2	0.700	0.676	0.884	0.918	0.937	0.947	0.952	0.957	0.952	0.957	0.957	0.961	0.966
3	0.000	0.022	0.630	0.587	0.696	0.717	0.717	0.739	0.739	0.761	0.739	0.739	0.739
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.000	0.000	0.300	0.200	0.200	0.300	0.300	0.300	0.200	0.300	0.300	0.300	0.200
6	0.000	0.058	0.072	0.884	0.899	0.899	0.899	0.928	0.971	0.986	0.986	0.986	0.986
7	1.000	0.919	1.000	1.000	1.000	1.000	1.000	0.973	0.973	0.973	0.973	0.973	0.973
8	0.881	0.687	0.687	0.701	0.716	0.851	0.806	0.821	0.806	0.806	0.806	0.821	0.821
9	0.000	0.045	0.136	0.818	0.818	0.886	0.886	0.886	0.909	0.886	0.864	0.864	0.864
Total	0.535	0.528	0.680	0.827	0.843	0.888	0.887	0.893	0.895	0.900	0.898	0.903	0.903

TABLE 3: The recognition accuracy rates with all features subsets of HMM based on maximum relevance selection.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.900	0.900	0.900	0.700	0.700	0.767	0.767	0.767	0.767	0.700	0.733	0.667	0.667
1	0.000	0.000	0.104	0.521	0.896	0.938	0.896	0.833	0.771	0.792	0.750	0.750	0.750
2	0.000	0.005	0.391	0.667	0.797	0.860	0.870	0.884	0.913	0.918	0.928	0.942	0.961
3	0.000	0.174	0.739	0.804	0.804	0.783	0.783	0.804	0.761	0.761	0.761	0.761	0.761
4	0.000	0.000	0.000	0.524	0.929	0.976	0.976	0.976	0.976	1.000	1.000	0.976	0.976
5	0.000	0.000	0.000	0.200	0.300	0.300	0.300	0.300	0.200	0.400	0.400	0.400	0.200
6	1.000	0.928	0.942	0.986	0.986	0.986	1.000	0.986	0.986	0.986	0.986	0.986	0.971
7	0.000	0.000	0.000	0.514	0.973	0.973	0.973	0.946	0.973	0.919	0.919	0.973	0.946
8	0.090	0.149	0.179	0.313	0.313	0.701	0.701	0.701	0.731	0.776	0.776	0.806	0.821
9	0.000	0.386	0.432	0.750	0.932	0.568	0.568	0.659	0.909	0.932	0.864	0.864	0.864
Total	0.170	0.212	0.405	0.643	0.790	0.837	0.838	0.843	0.865	0.873	0.872	0.878	0.880

For HMM, it can be found from Table 3 that features subset 10 generates the highest proportion of optimal IARARs for all activities, 40%. Concretely, with features subset 10, IARARs of activities 4, 5, 7, and 9 are better than or equal to those of the other subsets. However, features subset 13 yields the optimal TARAR of 88.0%, which is slightly higher than that of features subset 10, 87.3%. Again, it also can be concluded that feature f_3 is not discriminatory to activity recognition with HMM.

4.3. Feature Selection with mRMR Algorithm. mRMR feature selection algorithm can be adopted to sort the 13 features and reduce the number of dimensions of features subset. This is done by analyzing the relevance of the features to the target activities on the original features set to remove the irrelevant features and keep the features which have strong correlations to the target activities and then analyzing the redundancy of the selected features to reduce the number of dimensions of features subset.

According to (11) to get the value of information of different features, the results of iterations are listed in Table 4 with D-R criterion. At the beginning, the MI equation has been applied to get the first optimal feature and the result is f_{10} . The next interaction, still according to (11), feature f_6 is the optimal feature among the other remaining 12 features. And then, the ranks of each feature have been obtained successively.

Finally, the order of the sorted features is $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9, f_2, f_{13}$, and f_3 . Again, it can be found that feature f_{10} is the most important feature and feature f_3 is not discriminatory to activity recognition.

Similarly, the information values of different feature groups can be obtained according to (12), and the results of iterations are listed in Table 5 under D/R criterion.

The order of the sorted features is $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8, f_1, f_9$, and f_3 .

Although the orders of the sorted features with D-R and D/R criteria are different, the first feature still is f_{10} , the most important feature, the second feature is f_6 , and the last

TABLE 4: The information values of feature groups under the D-R criterion.

Feature													
Iteration	1	2	3	4	5	6	7	8	9	10	11	12	13
0	1.188	0.688	0.047	1.173	0.863	1.490	1.200	1.321	1.508	1.535	1.168	1.183	0.530
1	0.476	0.213	-0.160	-0.009	0.441	0.774	0.310	0.146	-1.989	—	-0.032	0.607	0.134
2	0.376	0.230	-0.075	0.281	0.434	—	0.306	0.343	-0.594	—	0.261	0.363	0.151
3	0.520	0.288	-0.042	0.430	—	—	0.440	0.338	-0.041	—	0.453	0.520	0.195
4	0.425	0.268	-0.030	0.488	—	—	0.451	0.428	0.206	—	0.518	—	0.198
5	0.474	0.299	-0.023	0.452	—	—	0.456	0.458	0.238	—	—	—	0.216
6	—	0.173	-0.022	0.486	—	—	0.454	0.479	0.330	—	—	—	0.196
7	—	0.206	-0.019	—	—	—	0.484	0.482	0.326	—	—	—	0.207
8	—	0.207	-0.016	—	—	—	—	0.498	0.365	—	—	—	0.202
9	—	0.196	-0.025	—	—	—	—	—	0.364	—	—	—	0.188
10	—	0.196	-0.040	—	—	—	—	—	—	—	—	—	0.187
11	—	—	-0.040	—	—	—	—	—	—	—	—	—	0.190

TABLE 5: The information values of feature groups under the D/R criterion.

Feature													
Iteration	1	2	3	4	5	6	7	8	9	10	11	12	13
0	1.188	0.688	0.047	1.173	0.863	1.490	1.200	1.321	1.508	1.535	1.168	1.183	0.530
1	1.668	1.449	0.227	0.992	2.047	2.080	1.349	1.124	0.431	—	0.974	2.054	1.339
2	1.463	1.502	0.385	1.314	2.013	—	1.342	1.350	0.717	—	1.288	1.443	1.397
3	1.777	1.719	0.530	1.579	—	—	1.578	1.344	0.973	—	1.634	1.784	1.580
4	1.556	1.635	0.613	1.712	—	—	1.603	1.479	1.158	—	1.797	—	1.596
5	1.663	1.769	0.674	1.626	—	—	1.612	1.531	1.187	—	—	—	1.686
6	1.510	—	0.659	1.811	—	—	1.715	1.620	1.321	—	—	—	1.689
7	1.588	—	0.695	—	—	—	1.773	1.619	1.310	—	—	—	1.733
8	1.586	—	0.732	—	—	—	—	1.644	1.351	—	—	—	1.694
9	1.664	—	0.753	—	—	—	—	1.727	1.462	—	—	—	—
10	1.659	—	0.667	—	—	—	—	—	1.445	—	—	—	—
11	—	—	0.671	—	—	—	—	—	1.487	—	—	—	—

feature is f_3 , which is not discriminatory to activity recognition.

According to the orders of the sorted features with D-R and D/R criteria, 13 features subsets can be obtained through adding one feature to previous features subset each time successively.

The 13 features subsets under D-R criterion are the following:

- Features subset 1: (f_{10})
- Features subset 2: (f_{10}, f_6)
- Features subset 3: (f_{10}, f_6, f_5)
- Features subset 4: (f_{10}, f_6, f_5, f_{12})
- Features subset 5: ($f_{10}, f_6, f_5, f_{12}, f_{11}$)
- Features subset 6: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_1$)
- Features subset 7: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4$)
- Features subset 8: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7$)
- Features subset 9: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8$)
- Features subset 10: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9$)
- Features subset 11: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9, f_2$)
- Features subset 12: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9, f_2, f_{13}$)
- Features subset 13: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9, f_2, f_{13}, f_3$)

And the 13 features subsets under D/R criterion are the following:

- Features subset 1: (f_{10});
- Features subset 2: (f_{10}, f_6)
- Features subset 3: (f_{10}, f_6, f_5)
- Features subset 4: (f_{10}, f_6, f_5, f_{12})
- Features subset 5: ($f_{10}, f_6, f_5, f_{12}, f_{11}$)
- Features subset 6: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_2$)
- Features subset 7: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4$)
- Features subset 8: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7$)
- Features subset 9: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}$)
- Features subset 10: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8$)
- Features subset 11: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8, f_1$)
- Features subset 12: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8, f_1, f_9$)
- Features subset 13: ($f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8, f_1, f_9, f_3$)

The selected features subsets have been validated by NB classifier and HMM to recognize activities. TARARs of NB classifier and HMM, with each features subset, are listed in Table 6, respectively.

It can be found that although D-R and D/R criteria are different, the features subsets which have the same number

TABLE 6: Comparison results of TARARs under two mRMR evaluation criteria (D-R and D/R).

Features subset		1	2	3	4	5	6	7	8	9	10	11	12	13
NB	D-R	0.535	0.695	0.870	0.870	0.870	0.890	0.885	0.902	0.903	0.900	0.898	0.903	0.903
	D/R	0.535	0.695	0.870	0.870	0.870	0.880	0.877	0.892	0.888	0.887	0.903	0.903	0.903
HMM	D-R	0.170	0.255	0.447	0.523	0.638	0.745	0.770	0.832	0.837	0.873	0.872	0.878	0.880
	D/R	0.170	0.255	0.447	0.523	0.638	0.747	0.768	0.817	0.837	0.858	0.853	0.878	0.880

of features are similar, even the same, e.g., if the number of features is little than 6. Therefore, it can be observed that the results of the same classifier with the same number of features are very slightly different, under the two criteria. If the number of features is smaller than 6 or larger than 11, the results of TARARs are the same.

Moreover, the performances of NB classifier and HMM are different even with the same features subset. Generally, with the increasing of size of features subset, TARAR increases until the size reaches a certain value. Still, the trend is not monotonic. For example, the optimal results are with features subsets 9, 12, and 13 under D-R criterion (90.3%), and the optimal result is with features subset 13 under D/R criterion (88.0%), for NB classifier. For HMM, the optimal result is with features subset 13 under D-R (88.0%) and D/R (88.0%) criteria. The results show that NB classifier and HMM yield relatively higher TARARs with features subset 5 through 13 and TARARs of NB classifier are slightly higher than those of HMM with features subsets 8 through 13, but much higher with 1 through 7, under the same criterion.

Basically, it can be concluded that both D-R and D/R criteria can be applied to sort the features and then to compose features subset effectively.

Tables 7 and 8 (under D-R criterion) show that the optimal IARARs of activities 0, 3, 5, 6, and 9 of HMM are higher than those of NB classifier, while the optimal IARARs of the activities 1, 2, 7, and 8 of NB classifier are higher than those of HMM. The optimal IARARs of activity 4 are the same of the two classifiers. Generally, the trends of the two classifiers are nearly the same with the increasing of size of features subset. Again, the trends are not monotonic. Moreover, it can be observed that different features have different effects on activity recognition. For example, features subset 5 means introducing feature f_{11} to features subset 4; then, IARARs have been improved of some activities, such as activities 0, 2, 3, 4, and 7 through 9; however, IARARs of other activities are degraded or not improved, of NB classifier, as shown in Table 7.

HMM has a similar result. It also can be found that, according to the different performance traits of NB classifier and HMM, the adding of the same feature has different effects on IARARs. For example, by introducing of the feature f_5 to features subset 2 to compose features subset 3, IARAR of activities 0 has been improved in NB classifier apparently but degraded in HMM. By introducing of feature f_{11} to features subset 4 to compose features subset 5, IARAR of activity 9 of NB classifier has been degraded but improved in HMM.

Furthermore, from Table 7, it is also shown that, with features subset 13, IARARs of activities 0, 1, 2, 4, and 6 are better than or equal to those of the other subsets; the proportion is 50% for all activities. With features subset 9, the proportion is 40% (activities 1, 5, 6, and 9). However, with features subset 9, IARARs of activities 1, 3, 5, 6, 7, and 9 are better than or equal to those of features subset 13. Therefore, features subset 9 yields the optimal TARAR of 90.3%, which is the same as that of features subsets 12 and 13. Still, features f_3 is not discriminatory to activity recognition with NB classifier.

It also can be found from Table 8 that, with features subset 13, IARARs of activities 2 and 8 are better than or equal to those of the other features subsets; the proportion is 20% for all activities, while features subset 13 gives the optimal TARAR of 88.0%, which is slightly higher than that of features subset 12. Moreover, with features subset 12, IARARs of activities 0, 1, and 3 to 7 and 9 are better than or equal to those of the features subset 13; the proportion is 80%. Again, it also can be concluded that feature f_3 is not discriminatory to activity recognition with HMM.

Under the D/R criterion, Tables 9 and 10 show the activity recognition accuracy rates with each features subset of NB classifier and HMM. From the experimental results, it can be found that the performances of NB classifier and HMM with the features subsets obtained under D/R criterion are similar to those of D-R criterion, respectively.

For example, for NB classifier, the optimal TARARs are 90.3% with features subsets 9, 12, and 13 under D-R criterion, which are the same to those of features subsets 11, 12, and 13 under D/R criterion. And for HMM, the optimal TARAR is 88.0% with features set 13 under D-R criterion, which is the same as that under D/R criterion.

From Table 9, for NB classifier, it is indicated that, with features subsets 12 and 13, IARARs of activities 1, 4, 6, and 9 are better than or equal to those of the other features subsets; the proportion is 40%, for all activities. However, with features subset 11, IARARs of activities 1, 3, 5, 6, and 9 are better than or equal to those of features subsets 12 and 13; the proportion is 50%. Features subsets 11, 12, and 13 yield the same optimal TARARs of 90.3%. Again, features f_3 and f_{13} are not discriminatory to activity recognition with NB classifier.

It also can be found from Table 10, for HMM, that with features subset 13, IARARs of activities 1, 2, 4, and 8 are better than or equal to those of the other features subsets. Features subset 13 yields the optimal TARAR of 88.0%, which is slightly higher than that of features subset 12. Actually, with features subset 12, IARARs of activities 0, 1, 3,

TABLE 7: The recognition accuracy rates with each features subset of NB classifier under the D-R criterion.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.000	0.000	0.400	0.400	0.400	0.667	0.667	0.667	0.600	0.600	0.633	0.667	0.667
1	0.792	0.979	0.938	0.958	0.958	0.958	0.958	0.958	0.979	0.979	0.979	0.979	0.979
2	0.700	0.899	0.913	0.913	0.913	0.908	0.899	0.937	0.947	0.957	0.957	0.961	0.966
3	0.000	0.848	0.848	0.761	0.761	0.761	0.783	0.783	0.804	0.761	0.739	0.739	0.739
4	1.000	0.976	1.000	1.000	1.000	1.000	1.000	1.000	0.976	1.000	1.000	1.000	1.000
5	0.000	0.100	0.300	0.600	0.600	0.600	0.400	0.500	0.600	0.300	0.300	0.300	0.200
6	0.000	0.072	0.913	0.928	0.957	0.957	0.957	0.957	0.986	0.986	0.986	0.986	0.986
7	1.000	1.000	0.973	0.973	0.973	0.973	0.946	0.973	0.973	0.973	0.973	0.973	0.973
8	0.881	0.881	0.821	0.806	0.821	0.836	0.851	0.851	0.806	0.806	0.806	0.821	0.821
9	0.000	0.045	0.864	0.864	0.795	0.886	0.886	0.886	0.886	0.886	0.864	0.864	0.864
Total	0.535	0.695	0.870	0.870	0.870	0.890	0.885	0.902	0.903	0.900	0.898	0.903	0.903

TABLE 8: The recognition accuracy rates with each features subset of HMM under the D-R criterion.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.900	0.900	0.800	0.800	0.833	0.867	0.833	0.833	0.700	0.700	0.733	0.667	0.667
1	0.000	0.000	0.292	0.271	0.250	0.229	0.083	0.500	0.458	0.792	0.750	0.750	0.750
2	0.000	0.005	0.184	0.256	0.483	0.657	0.768	0.860	0.908	0.918	0.928	0.942	0.961
3	0.000	0.870	0.935	0.957	0.978	0.935	0.891	0.804	0.826	0.761	0.761	0.761	0.761
4	0.000	0.000	0.524	0.524	0.714	0.952	1.000	0.952	0.929	1.000	1.000	0.976	0.976
5	0.000	0.000	0.400	0.700	0.400	0.700	0.500	0.500	0.600	0.400	0.400	0.400	0.200
6	1.000	0.986	1.000	1.000	1.000	0.971	0.971	0.986	0.986	0.986	0.986	0.986	0.971
7	0.000	0.000	0.297	0.703	0.892	0.973	0.892	0.946	0.838	0.973	0.919	0.973	0.946
8	0.090	0.119	0.239	0.373	0.388	0.701	0.701	0.701	0.701	0.731	0.776	0.806	0.821
9	0.000	0.205	0.614	0.705	0.886	0.773	0.886	0.909	0.955	0.932	0.864	0.864	0.864
Total	0.170	0.255	0.447	0.523	0.638	0.745	0.770	0.832	0.837	0.873	0.872	0.878	0.880

TABLE 9: The recognition accuracy rates with each features subset of NB classifier under the D/R criterion.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.000	0.000	0.400	0.400	0.400	0.700	0.700	0.700	0.733	0.600	0.600	0.667	0.667
1	0.792	0.979	0.938	0.958	0.958	0.938	0.958	0.958	0.958	0.958	0.979	0.979	0.979
2	0.700	0.899	0.913	0.913	0.913	0.908	0.903	0.947	0.947	0.961	0.961	0.961	0.966
3	0.000	0.848	0.848	0.761	0.761	0.783	0.783	0.783	0.783	0.804	0.804	0.739	0.739
4	1.000	0.976	1.000	1.000	1.000	0.976	1.000	0.976	0.976	0.952	0.952	1.000	1.000
5	0.000	0.100	0.300	0.600	0.600	0.600	0.400	0.400	0.400	0.500	0.600	0.300	0.200
6	0.000	0.072	0.913	0.928	0.957	0.971	0.942	0.942	0.942	0.971	0.986	0.986	0.986
7	1.000	1.000	0.973	0.973	0.973	0.946	0.946	0.946	0.946	0.946	0.946	0.973	0.973
8	0.881	0.881	0.821	0.806	0.821	0.836	0.851	0.851	0.821	0.791	0.806	0.821	0.821
9	0.000	0.045	0.864	0.864	0.795	0.750	0.750	0.773	0.750	0.727	0.864	0.864	0.864
Total	0.535	0.695	0.870	0.870	0.870	0.880	0.877	0.892	0.888	0.887	0.903	0.903	0.903

4, 5, 6, 7, and 9 are better than or equal to those of the features subset 13; the proportion is 80%. Again, features f_3 and f_{13} are not discriminatory to activity recognition with HMM.

Table 11 shows the comparison results of the optimal IARARs and TARARs of NB classifier and HMM under the three evaluation criteria (MR, D-R, and D/R) and the required minimal size of features subset, respectively.

It indicates that although the features subsets are different under the three evaluation criteria, however, the optimal results of IARARs and TARARs are similar to the

same classifier. For example, for NB classifier, the optimal IARARs of activities 0, 2, 4, 6, 7, and 8 under MR criterion are the same as those under D-R criterion; furthermore, the optimal IARARs of activities 2, 4, 6, 7, and 8 are the same even under the three evaluation criteria, and even the optimal TARARs are the same.

Similarly, for HMM, the optimal IARARs of activities 0, 2, and 6 through 8 are the same even under the three evaluation criteria. Besides, the optimal TARARs are also the same with the three evaluation criteria. Additionally, it can be found that the results of the required minimal size of

TABLE 10: The recognition accuracy rates with each features subset of HMM under the D/R criterion.

Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.900	0.900	0.800	0.800	0.833	0.833	0.800	0.800	0.800	0.667	0.600	0.667	0.667
1	0.000	0.000	0.292	0.271	0.250	0.688	0.417	0.625	0.625	0.667	0.500	0.750	0.750
2	0.000	0.005	0.184	0.256	0.483	0.652	0.758	0.865	0.879	0.918	0.937	0.942	0.961
3	0.000	0.870	0.935	0.957	0.978	0.935	0.913	0.804	0.804	0.826	0.826	0.761	0.761
4	0.000	0.000	0.524	0.524	0.714	0.738	0.857	0.857	0.833	0.881	0.952	0.976	0.976
5	0.000	0.000	0.400	0.700	0.400	0.500	0.400	0.500	0.500	0.600	0.600	0.400	0.200
6	1.000	0.986	1.000	1.000	1.000	0.986	0.986	0.986	0.986	1.000	0.986	0.986	0.971
7	0.000	0.000	0.297	0.703	0.892	0.919	0.919	0.919	0.865	0.892	0.811	0.973	0.946
8	0.090	0.119	0.239	0.373	0.388	0.522	0.522	0.522	0.716	0.746	0.806	0.806	0.821
9	0.000	0.205	0.614	0.705	0.886	0.886	0.932	0.955	0.932	0.909	0.909	0.864	0.864
Total	0.170	0.255	0.447	0.523	0.638	0.747	0.768	0.817	0.837	0.858	0.853	0.878	0.880

TABLE 11: Optimal recognition accuracy rates under three evaluation criteria (MR, D-R, and D/R) and the required minimal size of features subset (in brackets).

Activity	Classifier	Criterion	0	1	2	3	4	5	6	7	8	9	Total
NB	MR		0.667 (12)	1.000 (8)	0.966 (13)	0.739 (10)	1.000 (1)	0.300 (3)	0.986 (10)	1.000 (1)	0.881 (1)	0.909 (9)	0.903 (12)
	D-R		0.667 (6)	0.979 (2)	0.966 (13)	0.848 (2)	1.000 (1)	0.600 (4)	0.986 (9)	1.000 (1)	0.881 (1)	0.886 (6)	0.903 (9)
	D/R		0.733 (9)	0.979 (2)	0.966 (13)	0.848 (2)	1.000 (1)	0.600 (4)	0.986 (11)	1.000 (1)	0.881 (1)	0.864 (3)	0.903 (11)
HMM	MR		0.900 (1)	0.938 (6)	0.961 (13)	0.804 (4)	1.000 (10)	0.400 (10)	1.000 (1)	0.973 (5)	0.821 (13)	0.932 (5)	0.880 (13)
	D-R		0.900 (1)	0.792(10)	0.961 (13)	0.978 (5)	1.000 (7)	0.700 (4)	1.000 (1)	0.973 (6)	0.821 (13)	0.955 (9)	0.880 (13)
	D/R		0.900 (1)	0.750 (12)	0.961 (13)	0.978 (5)	0.976 (12)	0.700 (4)	1.000 (1)	0.973 (12)	0.821 (13)	0.955 (8)	0.880 (13)

features subset according to the optimal IARAR of each activity and the optimal TARAR are also similar to the same classifier. For example, for NB classifier, the required minimal sizes of features subset of activity 1 to 5 and 7 to 8, under D-R criterion, are the same as those under D/R criterion. Moreover, for activities 2, 4, 7, and 8, the required minimal sizes of features subset are the same under the three evaluation criteria. For HMM, the required minimal sizes of features subset of activities 0, 2, 6, and 8 are the same under the three evaluation criteria; the proportion is 40%. Besides, the required minimal sizes of features subset of TARARs are also the same under the three evaluation criteria.

Obviously, maximal relevance measure and mRMR feature selection algorithm (under D-R and D/R criteria) are effective to feature selection for human activity recognition.

5. Conclusions

This paper has applied maximal relevance measure and minimal redundancy maximal relevance (mRMR) algorithm (under D-R and D/R criteria) to select features and to compose different features subsets based on the observed sensor events for human activity recognition in smart home environments. And then, the selected features subsets have been validated by NB classifier and HMM to

recognize human activities. Through the experimental results, it is shown that not all the features are beneficial to activity recognition, as estimated. Different combinations of features lead to different activity recognition results. Furthermore, even the same features subset has different effect on the activity recognition accuracy rate for different activity classifiers. It can be found that feature f_{10} (the ending time of the current activity) is the most important feature to activity recognition and feature f_3 (day of week) is not discriminatory to activity recognition. Therefore, the suitable features subset must be selected in advance, and the selection of unsuitable feature sets increases the computational complexity and degrades the human activity recognition accuracy rate. Moreover, it is significant for activity recognition to consider both relevance between features and activities and redundancy between features. Generally, maximal relevance measure and mRMR algorithm are beneficial to feature selection and positive to activity recognition of NB classifier and HMM.

Data Availability

The data were collected from the smart home testbed located on Washington State University campus, which can be downloaded from Dr. Cook's homepage.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

Thanks are due to Dr. Diane J. Cook whose positive and constructive comments helped to enhance the quality and presentation of this paper. Also, the authors gratefully acknowledge the data preparation work that was carried out by Dr. Cook's graduate students.

References

- [1] L. Liu, E. Stroulia, I. Nikolaidis, A. Miguel-Cruz, and A. Rios Rincon, "Smart homes and home health monitoring technologies for older adults: a systematic review," *International Journal of Medical Informatics*, vol. 91, pp. 44–59, 2016.
- [2] M. R. Alam, M. B. I. Reaz, and M. A. M. Ali, "A review of smart homes-past, present, and future," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1190–1203, 2012.
- [3] D. J. Cook, "How smart is your home?" *Science*, vol. 335, no. 6076, pp. 1579–1581, 2012.
- [4] G. Sprint, D. J. Cook, R. Fritz, and M. Schmitter-Edgecombe, "Using smart homes to detect and analyze health events," *Computer*, vol. 49, no. 11, pp. 29–37, 2016.
- [5] 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, 2013.
- [6] T. Luor, H. Lu, H. Yu, and Y. Lu, "Exploring the critical quality attributes and models of smart homes," *Maturitas*, vol. 82, no. 4, pp. 377–386, 2015.
- [7] J. Dahmen, B. Thomas, D. Cook, and X. Wang, "Activity learning as a foundation for security monitoring in smart homes," *Sensors*, vol. 17, no. 4, p. 737, 2017.
- [8] H. M. S. Hossain, M. A. A. H. Khan, and N. Roy, "Active learning enabled activity recognition," *Pervasive and Mobile Computing*, vol. 38, no. 2, pp. 312–330, 2017.
- [9] G. Singla, D. Cook, and M. Schmitter-Edgecombe, "Tracking activities in complex settings using smart environment technologies," *International Journal of BioSciences, Psychiatry and Technology*, vol. 1, no. 1, pp. 25–35, 2009.
- [10] A. S. Crandall and D. J. Cook, "Coping with multiple residents in a smart environment," *Journal of Ambient Intelligence and Smart Environments*, vol. 1, no. 4, pp. 323–334, 2009.
- [11] L. Liu, S. Wang, G. Su, Z.-G. Huang, and M. Liu, "Towards complex activity recognition using a Bayesian network-based probabilistic generative framework," *Pattern Recognition*, vol. 68, pp. 295–309, 2017.
- [12] K. S. Gayathri, K. S. Easwarakumar, and S. Elias, "Probabilistic ontology based activity recognition in smart homes using markov logic network," *Knowledge-Based Systems*, vol. 121, pp. 173–184, 2017.
- [13] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *IEEE Pervasive Computing*, vol. 9, no. 1, pp. 48–53, 2010.
- [14] Y. Zhu, N. M. Nayak, and A. K. Roy-Chowdhury, "Context-aware activity modeling using hierarchical conditional random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 7, pp. 1360–1372, 2015.
- [15] L. Chen, C. D. Nugent, and H. Wang, "A knowledge-driven approach to activity recognition in smart homes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 6, pp. 961–974, 2012.
- [16] M. Fahim, I. Fatima, S. Lee, and Y.-K. Lee, "EEM: evolutionary ensembles model for activity recognition in smart homes," *Applied Intelligence*, vol. 38, no. 1, pp. 88–98, 2013.
- [17] 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, 2010.
- [18] J. Wen and Z. Wang, "Learning general model for activity recognition with limited labelled data," *Expert Systems with Applications*, vol. 74, pp. 19–28, 2017.
- [19] J.-H. Hong, J. Ramos, and A. K. Dey, "Toward personalized activity recognition systems with a semipopulation approach," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, pp. 101–112, 2016.
- [20] C. Debes, A. Merentitis, S. Sukhanov, M. Niessen, N. Frangiadakis, and A. Bauer, "Monitoring activities of daily living in smart homes: understanding human behavior," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 81–94, 2016.
- [21] N. S. Mohamed, S. Zainudin, and Z. Ali Othman, "Meta-heuristic approach for an enhanced mRMR filter method for classification using drug response microarray data," *Expert Systems with Applications*, vol. 90, pp. 224–231, 2017.
- [22] J. Che, Y. Yang, L. Li, X. Bai, S. Zhang, and C. Deng, "Maximum relevance minimum common redundancy feature selection for nonlinear data," *Information Sciences*, vol. 409–410, pp. 68–86, 2017.
- [23] C. Xu, S. Zhao, and F. Liu, "Distributed plant-wide process monitoring based on PCA with minimal redundancy maximal relevance," *Chemometrics and Intelligent Laboratory Systems*, vol. 169, pp. 53–63, 2017.
- [24] F. Li, D. Miao, and W. Pedrycz, "Granular multi-label feature selection based on mutual information," *Pattern Recognition*, vol. 67, pp. 410–423, 2017.
- [25] D. Escalona-Vargas, E. R. Siegel, P. Murphy, C. L. Lowery, and H. Eswaran, "Selection of reference channels based on mutual information for frequency-dependent subtraction method applied to fetal biomagnetic signals," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 1115–1122, 2017.
- [26] S. Mallik, T. Bhadra, and U. Maulik, "Identifying epigenetic biomarkers using maximal relevance and minimal redundancy based feature selection for multi-omics data," *IEEE Transactions on NanoBioscience*, vol. 16, no. 1, pp. 3–10, 2017.
- [27] A. K. Tiwari, "Prediction of G-protein coupled receptors and their subfamilies by incorporating various sequence features into Chou's general PseAAC," *Computer Methods and Programs in Biomedicine*, vol. 134, pp. 197–213, 2016.
- [28] C. Chen and X. Yan, "Optimization of a multilayer neural network by using minimal redundancy maximal relevance-partial mutual information clustering with least square regression," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 6, pp. 1177–87, 2015.
- [29] Z. Wang, M. Li, and J. Li, "A multi-objective evolutionary algorithm for feature selection based on mutual information with a new redundancy measure," *Information Sciences*, vol. 307, pp. 73–88, 2015.
- [30] P. M. Morgado and M. Silveira, "Minimal neighborhood redundancy maximal relevance: Application to the diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 155, pp. 295–308, 2015.
- [31] M. Kamandar and H. Ghassemian, "Linear feature extraction for hyperspectral images based on information theoretic

- learning,” *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 4, pp. 702–706, 2013.
- [32] K. K. Kandaswamy, G. Pugalenth, K.-U. Kalies, E. Hartmann, and T. Martinetz, “EcmPred: prediction of extracellular matrix proteins based on random forest with maximum relevance minimum redundancy feature selection,” *Journal of Theoretical Biology*, vol. 317, pp. 377–383, 2013.
- [33] X. Jin, E. W. M. Ma, L. L. Cheng, and M. Pecht, “Health monitoring of cooling fans based on Mahalanobis distance with mRMR feature selection,” *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 8, pp. 2222–2229, 2012.
- [34] A. Unler, A. Murat, and R. B. Chinnam, “MR2PSO: a maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification,” *Information Sciences*, vol. 181, no. 20, pp. 4625–4641, 2011.
- [35] E. Zdravevski, P. Lameski, V. Trajkovik et al., “Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering,” *IEEE Access*, vol. 5, pp. 5262–5280, 2017.
- [36] E. Zdravevski, B. Risteska Stojkoska, M. Standl, and H. Schulz, “Automatic machine-learning based identification of jogging periods from accelerometer measurements of adolescents under field conditions,” *PLoS ONE*, vol. 12, no. 9, p. e0184216, 2017.
- [37] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.
- [38] H. Fang, L. He, H. Si, P. Liu, and X. Xie, “Human activity recognition based on feature selection in smart home using back-propagation algorithm,” *ISA Transactions*, vol. 53, no. 5, pp. 1629–1638, 2014.
- [39] P. Tang, H. Fang, and H. Si, “Maximal relevance feature selection for human activity recognition in smart home,” in *Proceedings of the 2018 Chinese Control and Decision Conference. (CCDC)*, pp. 4264–4268, Shenyang, China, June 2018.