IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 13, 2020, accepted October 19, 2020, date of publication October 26, 2020, date of current version November 4, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3033589

Graph-Based Semi-Supervised Learning for Activity Labeling in Health Smart Home

YAN HU[1,2], BINGCE WANG[1], YUYAN SUN[3,4], JING AN[5], AND ZHILIANG WANG[1]

[1]School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China
[2]Beijing Engineering and Technology Center for Convergence Networks and Ubiquitous Services, University of Science and Technology Beijing, Beijing 100083, China
[3]Beijing Key Laboratory of IoT Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100195, China
[4]School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100195, China
[5]School of Marxism Studies, University of Science and Technology Beijing, Beijing 100083, China

Corresponding author: Yan Hu (huyan@ustb.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61802016 and Grant 61702506, in part by the Fundamental Research Funds for the Central Universities under Grant FRF-TP-20-012A2, in part by the National Key Research and Development Project under Grant 2017YFB0802805, and in part by the National Social Science Foundation of China under Grant 17ZDA331.

ABSTRACT Health Smart Home (HSH) is an important part of smart city. This technology provides a new kind of remote medical treatment, and can effectively alleviate the shortage of medical resources caused by aging population and help elderly people live at home more safely and independently. Activity recognition is the core of Health Smart Home. However, constructing activity recognition models usually requires a large amount of labeled data, which imposes a heavy burden on manual labeling. In this article, the authors propose an activity labeling approach based on a graph-based semi-supervised learning algorithm. This approach can divide the raw sensor event sequence without any label information into appropriate segments. Consecutive sensor events that occurred in a same activity are grouped into a same segment. In addition, this approach requires only a small number of manually labeled segments to complete the labeling of the remaining large number of unlabeled segments, thereby greatly reducing the burden of manual labeling. After that, all the labeled data can be further used for activity recognition in smart homes. Finally, a series of comprehensive experiments are conducted on freely available data sets to validate the effectiveness of the proposed activity labeling approach.

INDEX TERMS Smart city, health smart home, activity labeling, semi-supervised learning, label propagation algorithm.

I. INTRODUCTION

It is predicted that 70% of the world's population will live in cities by 2050, so cities need to become smart by using information and communications technologies (ICTs) to make city services more aware, interactive and efficient [1], thus to meet the various requirements of such a large urban population for work, education and daily life. Additionally, the average age of the world's population has been increased by the tremendous advancement in the field of medicine. The United Nations predicts that 22% of the world population will be over 65 years old by 2050 [2]. Therefore, it is an urgent task for most nations to develop the Health Smart Home technology, which is an important part of smart city, to improve the quality of life of elderly people. In modern society, the long distance between family members makes elderly people have to live alone and be autonomous. In addition, diseases like Alzheimer's become more prevalent with the increase of life expectancy. In order to avoid distress situations as much as possible, telemonitoring technologies should be used to detect significant changes in activities or habits of elderly people and ensure their safety in smart home environments.

Sensor technologies provide a promising solution for telemonitoring in smart home environments [3]. In the MIT (Cambridge, Massachusetts) project House_n, a variety of sensors are deployed in a flat for user activity monitoring [4]. Moreover, the system provides users with a series of Human-Machine Interfaces (HMI) to control their

The associate editor coordinating the review of this manuscript and approving it for publication was Zhibo Wang.

environment, which can promote the physical and mental health of elderly people. The "Aware Home Research Initiative" project developed by the Georgia Institute of Technology [5] constructs a two-floor smart home environment to satisfy the requirements of different generations of a family, i.e., children with mental disabilities and the elderly people. In this project, a series of video cameras, RFID tags and motion and environmental sensors are installed to help the handicapped children or the elderly people live independently and safely at home. In France, the researchers of both the AILISA [6] and PROSAFE [7] projects use presence infrared sensors to monitor activities of users and raise alarms in case of anomaly.

However, a large number of sensors deployed in a smart home environment undoubtedly produce a large amount of heterogeneous and multidimensional streaming sensor data. Realizing activity recognition based on a large amount of data is a big challenge. Besides, training activity recognition models usually requires a large number of labeled samples, which imposes a heavy burden on manual labeling. Therefore, realizing automatic labeling of training data for activity recognition in smart home environments is a promising research direction. In this article, we propose a novel activity labeling approach based on a graph-based semi-supervised learning algorithm. This approach can divide the raw sensor event sequence without any label information into appropriate segment. Additionally, it requires only a small number of manually labeled segments to realize the labeling of the remaining large number of unlabeled sensor events. This approach can effectively reduce the burden of manual labeling, and further improve the efficiency of activity recognition. The key contributions of this work are summarized as follows:

- We design a segmentation method for the raw sensor event sequence without any label information. Consecutive sensor events that occur in a same activity are grouped into a same segment as much as possible, which facilitates the feature extraction of different activities.
- We propose an activity labeling technique based on a graph-based semi-supervised label propagation algorithm. It only needs a small number of manually labeled segments to complete the labeling of the remaining large number of unlabeled training cases, thereby greatly reducing the burden of manual labeling.
- Comprehensive experiments are conducted on freely available datasets to validate the effectiveness of the proposed segmentation method for the sensor event sequence and the activity labeling technique.

The rest of the paper is organized as follows. The research literature about activity monitoring and activity recognition is presented in Section II. Section III presents the problem statement and an overview of the proposed approach. In Section IV, we elaborate on the proposed activity labeling approach, which mainly includes the segmentation of the sensor event sequence and the activity labeling based on the graph-based semi-supervised learning algorithm. Experimental setup and results are presented in Section V. Finally, we draw a conclusion in Section VI.

II. RELATED WORK

A. ACTIVITY MONITORING

Monitoring user actions and environmental changes is an essential prerequisite for human activity recognition. According to different monitoring devices, the activity monitoring technology mainly falls into two categories: video-based and sensor-based.

1) VIDEO-BASED ACTIVITY MONITORING

Video-based activity monitoring [8] uses cameras deployed in a smart home to track and record user actions continuously. Then, a variety of image processing algorithms are applied to the collected 2D or 3D images to realize activity recognition [9], [10]. Although the video-based technology is regarded as a very intuitive method, it still has some disadvantages. First, the video quality is easily affected by the viewing angle of a camera and the ambient light intensity, so it is difficult to maintain satisfactory video quality at different times of the day. Besides, transmitting unencrypted videos over networks can easily leak users' sensitive information [11]–[13]. In addition, processing and transmitting videos usually requires a lot of computational hardware resources, e.g., memory and bandwidth, thus further limiting the wider application of this technology. Fortunately, these problems can be alleviated to a certain extent by using the sensor-based monitoring technology. Therefore, the sensor-based monitoring technology is becoming more and more prevalent in the field of activity recognition.

2) SENSOR-BASED ACTIVITY MONITORING

According to different types of sensors used for activity monitoring, the sensor-based activity monitoring technology can also be divided into two categories: portable sensor-based and non-intrusive sensor-based. The portable sensor-based technology mainly employs Radio Frequency IDentification (RFID) technology and acceleration sensors to monitor user activities. RFID tags can be attached to different objects, so user activities that are highly related to these objects can be easily tracked. Patterson *et al.* [14] build a RFID-based smart kitchen, in which RFID tags are attached to dozens of objects, e.g., coffee machines, refrigerators, ovens, tablewares, dishwashers and cabinet doors. Accordingly, user activities can be monitored and analysed at different times of the day. Besides, acceleration sensors are also a type of portable sensors that are commonly used for activity monitoring. They are particularly sensitive to activities consisting of repeated actions, e.g., climbing stairs, walking, standing, and running. Zhang *et al.* [15] attach acceleration sensors to the hands of users, and then analyse the collected sensor data with a Back-Propagation (BP) neural network to realize daily activity recognition.

However, the portable sensor-based activity monitoring technology still has some shortcomings. First of all, it is inconvenient for most users, especially the elderly, to wear these sensors for most of the day. Besides that, portable sensors still face some technical problems, such as battery life, sensor size, water resistance and wearing comfort, thereby greatly limiting the wide application of this technology. To alleviate these problems, some researchers use smartphones to assist in daily activity monitoring [16].

Compared with portable sensors, non-intrusive sensors do not impose any burden on target users or interfere with users' daily life. This kind of sensors are not only low-cost, but also can be deployed in different locations in a smart home to monitor user actions at any time. van Kasteren *et al.* [17] build a smart home with a set of non-intrusive sensors. Reed switch sensors are used to monitor the door states of rooms, refrigerators, wardrobes and ovens. Besides, the mercury contact sensors attached to objects such as medicine boxes, tablewares and books are used to monitor movements of different objects. Moreover, a floating sensor embedded in a toilet is able to detect whether the user is using the bathtub or the toilet.

The Center for Advanced Studies in Adaptive Systems (CASAS) at Washington State University also builds a smart home environment with a variety of non-intrusive sensors to assist in the home medical technology [18], [19]. This project uses a group of passive infrared motion sensors to monitor whether a user appears in a certain area, and uses temperature sensors to perceive the ambient temperature in real time. In addition, object sensors are used to monitor whether the user is using an object, and water flow meters are used to calculate the amount of water used by the target user, and door sensors are used to monitor the opening and closing status of doors.

As discussed above, the activity monitoring technology mainly provides raw data for the follow-up activity recognition, so it is a necessary prerequisite for high-quality activity recognition. In the following of this section, we present some typical activity recognition technologies.

B. ACTIVITY RECOGNITION

User actions and environmental data collected by activity monitoring devices are then processed and analyzed by specific algorithms for activity recognition. Generally speaking, the collected sensor data can be regarded as a time series of sensor events. By dividing the sequence of sensor events, we can get a series of fixed-length time windows. Then, we can extract a feature vector from each window by applying statistical methods. In activity recognitions, the commonly used features include time and locations of sensor events, and the order of appearance of sensors in a time window [20]. Wu [21] devised a mixed feature extraction technique based on time segment coding. Time segments are Gray-encoded and combined with existing features to enrich the feature set, so as to improve the recognition accuracy. Moreover, some environmental data (e.g., locations, time and traffic routes) of target users are also combined with acceleration sensor data collected by smartphones [22], thereby enriching identifiable daily activities.

After feature extraction, a training set should be built based on a portion of manually labeled feature vectors. Then, a series of supervised model training algorithms can be applied to the training set for model training. The commonly used supervised training technology includes template matching, discriminant and generative methods. The template matching methods first compute the similarity between each pair of feature vectors and then determine the activity label of a new feature vector based on the labels of its nearest neighbors [23], [24]. The discriminant methods mainly use machine learning algorithms, such as decision tree and Artificial Neural Network (ANN), to identify different activity categories by looking for the boundaries between different classes of feature vectors. The Decision tree algorithm continuously selects features that can best distinguish different activities based on information gain [25]. ANN realizes activity recognition by modeling non-linear relationships between feature vectors and activity categories [26]–[28]. However, the high complexity of an objective function usually makes the parameter selection time-consuming and sometimes makes itself converge to a local minimum. Therefore, it is important to choose a reasonable network topology before model training [29]. The generative methods utilize probability models such as Naive Bayes Classifier [30] and Hidden Markov Model [31] to characterize the joint probability distribution of feature vectors and activity categories. Then, the association probabilities of a new feature vector and different activity categories can be estimated. The category with the maximum association probability is selected as the activity recognition result.

Based on the above analysis, we can see that the current activity recognition technologies generally require a large number of manually labeled training cases. However, due to the large amount of streaming sensor data generated in real-world smart home applications, it is undoubtedly an arduous task to label so much raw data for model training. In this work, we propose an activity labeling approach to support high-efficiency activity recognition. This approach utilizes a segmentation technique to group consecutive sensor events that occur in a same activity into a same segment, which facilitates the feature extraction of different activities. Additionally, this approach requires only a small number of manually labeled segments to realize labeling of the remaining large number of unlabeled segments, thereby greatly reducing the burden of manual labeling and improving the efficiency of activity recognition.

III. PROBLEM STATEMENT AND APPROACH OVERVIEW

In this section, we present the problem statement and an overview of the proposed activity labeling approach.

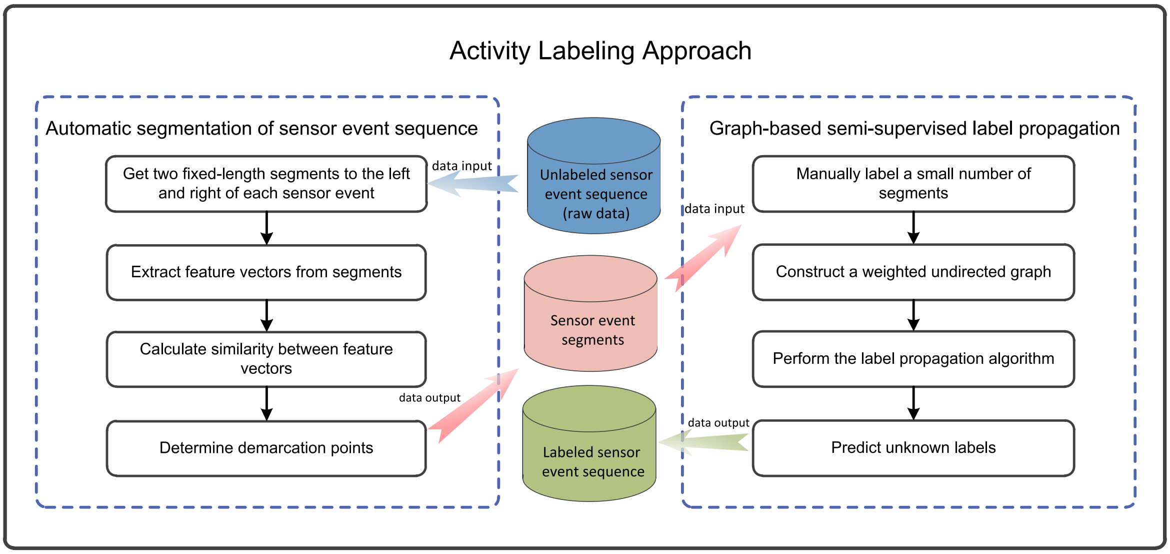

FIGURE 1. Approach overview.

A. PROBLEM STATEMENT

Suppose that there are M sensors deployed in a smart home for activity monitoring and a sequence of sensor events $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_T\}$ is collected over time. Each sensor event can be denoted by a quintuple $\boldsymbol{e} = (d, t, s, l, a)$, where d, t, s, l, a respectively denote the date, time, sensor, location, and activity category of the sensor event. Please note that the category information a is unknown in the raw sequence of sensor events. Therefore, a labeling technique should be devised to predict the unknown category information of different sensor events $\boldsymbol{e}$ before they can be used for model training.

B. APPROACH OVERVIEW

In order to address the problem stated above, we propose a novel activity labeling approach based on a graph-based semi-supervised learning algorithm. The entire procedure of this approach is illustrated in Figure 1, mainly including the segmentation of the sensor event sequence and the graph-based semi-supervised label propagation. The segmentation of the sensor event sequence aims to group consecutive events that occur in a same activity into a same segment. In this procedure, feature extraction, similarity calculation of feature vectors and determination of demarcation points should be performed in sequence to obtain a series of segments. The label propagation procedure requires only a small number of manually labeled segments to complete the labeling of the remaining large number of unlabeled segments. This procedure first builds a weighted undirected graph, which takes all labeled and unlabeled segments as nodes and the relationships between segments as edges. Then, a small number of manually annotated labels are propagated iteratively on the graph until the algorithm converges. The convergence results are used to predict the unknown category information of the remaining unlabeled segments, so as to realize the labeling of the entire training data set.

IV. ACTIVITY LABELING

In this section, we elaborate on the proposed activity labeling approach, which mainly includes the segmentation of the sensor event sequence and the graph-based semi-supervised label propagation algorithm.

A. SEGMENTATION OF SENSOR EVENT SEQUENCE

Segmentation of the sensor event sequence can group consecutive sensor events that occur in a same activity into a same segment, which facilitates the feature extraction of different activities. Moreover, the efficiency of labeling can be improved effectively, since all events in a same segment share a same activity label. In other words, once the label of the entire segment is determined, the label of each sensor event in the segment can be determined. In this subsection, a segmentation algorithm is designed to divide the unlabeled sensor event sequence, trying to make consecutive events that occur in a same activity category fall into a same segment. Generally, sensor groups triggered in similar activities have a higher similarity than sensor groups triggered in different activities. In other words, the greater the difference between two sensor groups, the more likely they are to belong to different activity categories. Based on this underlying principle, the segmentation algorithm is devised as follows.

Suppose a subsequence of sensor events $S_{sub} = \{\boldsymbol{e}_{i-\delta+1}, \ldots, \boldsymbol{e}_{i+\delta+1}\}$ consists of at least two different activities, which means that there exists at least one demarcation point in S_{sub}. If a sensor event $\boldsymbol{e}_i$ divides S_{sub} into two segments $S_{i1} = \{\boldsymbol{e}_{i-\delta+1}, \ldots, \boldsymbol{e}_i\}$ and $S_{i2} = \{\boldsymbol{e}_{i+1}, \ldots, \boldsymbol{e}_{i+\delta+1}\}$, and the similarity between S_{i1} and S_{i2} is lower than a pre-defined threshold, $\boldsymbol{e}_i$ is considered as a demarcation point, and S_{i1} and S_{i2} are regarded as two discriminative segments that belong to different activity categories. Here, the length of each discriminative segment is pre-set to a fixed value δ, and the similarity between two segments can be measured by the similarity between their feature vectors. We extract such a feature vector $\boldsymbol{x} = (c_1, c_2, \ldots, c_n)$ from each segment, where n is the total number of sensors deployed in a smart home environment, and c_k $(k = 1, \ldots, n)$ denotes the number of times the kth sensor appears in the current segment. Therefore, The similarity between S_{i1} and S_{i2} can be obtained by calculating the cosine similarity between their corresponding feature vectors $\boldsymbol{x}_{i1}$ and $\boldsymbol{x}_{i2}$ [32]. It is worth noting that if a sensor event is located near an actual demarcation point, it can yield two adjacent segments whose similarity may also be lower than the predefined threshold. Therefore, in a series of consecutive candidate demarcation points, only the sensor event that generates two adjacent segments with the lowest similarity value is selected as the final demarcation point. The Automatic Sequence Segmentation Algorithm (*ASSA* for short) is summarized by Algorithm 1. Lines 1 performs some initializations. Lines 2 to 9 traverse the entire sequence to find all candidate demarcation points that can generate two adjacent segments with a similarity value lower than a predefined threshold θ and put them into a set P. Then, all the longest subsequences of consecutive points in P are found and put into another set CP (Line 10). In each longest subsequence, the sensor event that yields two adjacent segments with the lowest similarity value is selected as a final demarcation point and put into the result set B (Lines 11 to 20). Finally, the result set B is returned (Line 21).

Algorithm 1 *ASSA* Algorithm

Input: the raw sensor event sequence S, the length δ of each discriminative segment, the segment similarity threshold θ

Output: the result set B of demarcation points

1 $B, P, CP \leftarrow \emptyset$;
2 **for** $(i = \delta;\ i \leq S.length - \delta - 1;\ i{+}{+})$ **do**
3     $S_{i1} \leftarrow \{\boldsymbol{e}_{i-\delta+1}, \ldots, \boldsymbol{e}_i\}$, $S_{i2} \leftarrow \{\boldsymbol{e}_{i+1}, \ldots, \boldsymbol{e}_{i+\delta+1}\}$;
4     $\boldsymbol{x}_{i1} \leftarrow$ the feature factor of S_{i1}, $\boldsymbol{x}_{i2} \leftarrow$ the feature factor of S_{i2};
5     $sim_i \leftarrow \text{CosineSimilarity}(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2})$;
6   **if** $sim_i < \theta$ **then**
7        $P \leftarrow \{P, \boldsymbol{e}_i\}$;
8   **end**
9 **end**
10 put all the longest subsequences of consecutive events in P into CP;
11 **for** *each* $L \in CP$ **do**
12     $sim_d \leftarrow +\infty$, $index_d \leftarrow 0$;
13   **for** *each* $\boldsymbol{e}_i \in L$ **do**
14   **if** $sim_i < sim_d$ **then**
15           $sim_d \leftarrow sim_i$;
16           $index_d \leftarrow i$;
17   **end**
18   **end**
19     $B \leftarrow \{B, e_{index_d}\}$;
20 **end**
21 return B;

B. GRAPH-BASED SEMI-SUPERVISED LABEL PROPAGATION FOR ACTIVITY LABELING

After obtaining a set of demarcation points, we can split the entire sequence of sensor events into a series of segments. Afterwards, we can construct a weighted undirected graph by taking segments as nodes and relationships between segments as edges. If two segments have a high similarity, the edge connecting them will be assigned a high weight. In order to predict the category information of unlabeled sensor events through the graph-based label propagation algorithm, a small number of segments should first be manually labeled. Then, the label propagation algorithm spreads the label information associated with each node to its neighbors iteratively until achieving global convergence. Since a weighted undirected graph can be denoted by a matrix, we can analyze the label propagation process in terms of matrix operations.

Suppose that a total of m segments are obtained through the *ASSA* algorithm. The set of manually labeled segments is represented by $L = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots, (\boldsymbol{x}_l, y_l)\}$, where $\boldsymbol{x}_i$ and y_i $(i = 1, \ldots, l)$ are respectively the feature vector and the category label of the ith segment, and l is the number of manually labeled segments. Analogously, the set of unlabeled segments is represented by $U = \{(\boldsymbol{x}_{l+1}, y_{l+1}), (\boldsymbol{x}_{l+2}, y_{l+2}), \ldots, (\boldsymbol{x}_{l+u}, y_{l+u})\}$, where $\boldsymbol{x}_{l+j}$ and y_{l+j} $(j = 1, \ldots, u)$ are respectively the feature vector and the unknown label of the $(l + j)$th segment, and u is the number of unlabeled segments. Please note that the number of manually labeled segments is much smaller than the number of unlabeled segments, i.e., $l \ll u$ and $l + u = m$. Afterwards, we can build a weighted undirected graph $G = (V, E)$ based on the m segments, where V and E are the set of nodes and the set of edges, respectively. The node set V is denoted by $V = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_l, \boldsymbol{x}_{l+1}, \boldsymbol{x}_{l+2}, \ldots, \boldsymbol{x}_{l+u}\}$, and the weight of the edge connecting the ith and jth $(i, j = 1, \ldots, m)$ nodes is defined by the following equation:

$$W_{ij} = \begin{cases} \exp\Big(\dfrac{-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2}{2\sigma^2}\Big), & \text{if} \quad i \neq j, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

where the operator $\|\cdot\|_2$ calculates the 2-norm of a vector, and $\sigma > 0$ is a predefined standard deviation of a Gaussian function. The weights of all edges can form a m-by-m affinity matrix $\boldsymbol{W}$, based on which we can construct a probability transition matrix $\boldsymbol{P} = \boldsymbol{D}^{-1}\boldsymbol{W}$, where $\boldsymbol{D} = \text{diag}(d_1, d_2, \ldots, d_m)$ is a diagonal matrix with its (i, i)-element equal to the sum of the terms of the ith row of $\boldsymbol{W}$, i.e., $d_i = \sum_{j=1}^{m} W_{ij}$.

Next, in order to assign category labels to unlabeled nodes, a real-valued function $f\colon V \rightarrow \mathbb{R}$ should be computed based on the graph G, where $\mathbb{R}$ is the set of real numbers. Based on f, we can get such a classification rule $y_i = \text{sign}(f(\boldsymbol{x}_i))$ for the follow-up labeling. Intuitively, nodes that are close to each other are expected to have similar labels, which motivates the choice of the quadratic energy function:

$$E(f) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} W_{ij}\big(f(\boldsymbol{x}_i) - f(\boldsymbol{x}_j)\big)^2 = \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{D} - \boldsymbol{W})\boldsymbol{f}, \quad (2)$$

where $\boldsymbol{f}$ is a vector denoted by:

$$\boldsymbol{f} = (\boldsymbol{f}_l^{\mathrm{T}}, \boldsymbol{f}_u^{\mathrm{T}})^{\mathrm{T}}. \quad (3)$$

In the above equation, $\boldsymbol{f}_l$ and $\boldsymbol{f}_u$ are respectively the vectors of function values on the labeled and unlabeled data sets:

$$\boldsymbol{f}_l = \big(f(\boldsymbol{x}_1), f(\boldsymbol{x}_2), \ldots, f(\boldsymbol{x}_l)\big)^{\mathrm{T}}, \quad (4)$$

and

$$\boldsymbol{f}_u = \big(f(\boldsymbol{x}_{l+1}), f(\boldsymbol{x}_{l+2}), \ldots, f(\boldsymbol{x}_{l+u})\big)^{\mathrm{T}}. \quad (5)$$

The minimum energy function satisfies $(\boldsymbol{D} - \boldsymbol{W})\boldsymbol{f} = \boldsymbol{0}$ on the unlabeled data set and satisfies $f(x_i) = y_i$ on the labeled data set. To compute the solution of $(\boldsymbol{D} - \boldsymbol{W})\boldsymbol{f} = \boldsymbol{0}$ subject to $\boldsymbol{f}|_L = \boldsymbol{f}_l$ in terms of matrix operations, we split the weight matrix $\boldsymbol{W}$ (and similarly $\boldsymbol{D}$ and $\boldsymbol{P}$) into 4 blocks after the lth row and column:

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{W}_{ll} & \boldsymbol{W}_{lu} \\ \boldsymbol{W}_{ul} & \boldsymbol{W}_{uu} \end{pmatrix}, \quad (6)$$

Based on the split matrices, the solution is given by:

$$\boldsymbol{f}_u = (\boldsymbol{D}_{uu} - \boldsymbol{W}_{uu})^{-1}\boldsymbol{W}_{ul}\boldsymbol{f}_l = (\boldsymbol{I} - \boldsymbol{P}_{uu})^{-1}\boldsymbol{P}_{ul}\boldsymbol{f}_l. \quad (7)$$

Since $\boldsymbol{f}_l = (y_1, y_2, \ldots, y_l)^{\mathrm{T}}$ is already known, we can predict the unknown categories of unlabeled segments in U by substituting $\boldsymbol{f}_l$ into Equation (7).

The above method is mainly designed for binary classification, but can be extended for multi-class classification. Analogously, a graph $G = (V, E)$ is first constructed and then a m-by-C matrix $\boldsymbol{F} = (\boldsymbol{F}_1^{\mathrm{T}}, \boldsymbol{F}_2^{\mathrm{T}}, \ldots, \boldsymbol{F}_m^{\mathrm{T}})^{\mathrm{T}}$ with non-negative entries is built, where C is the total number of activity categories. The ith row of $\boldsymbol{F}$ is denoted by $\boldsymbol{F}_i = (\boldsymbol{F}_{i1}, \boldsymbol{F}_{i2}, \ldots, \boldsymbol{F}_{iC})$, which assigns a label $y_i = \arg\max_{1\le j\le C} \boldsymbol{F}_{ij}$ to the ith node $\boldsymbol{x}_i$. The matrix $\boldsymbol{F}$ can be initialized as:

$$\boldsymbol{F}(0) = \boldsymbol{Y}_{ij} = \begin{cases} 1, & \text{if } 1 \le i \le l \wedge y_i = j, \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

Clearly, $\boldsymbol{Y}$ is consistent with the initial labels on L. Afterwards, a label propagation matrix $\boldsymbol{S}$ is built based on the weight matrix $\boldsymbol{W}$:

$$\boldsymbol{S} = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{W}\boldsymbol{D}^{-\frac{1}{2}}, \quad (9)$$

where

$$\boldsymbol{D}^{-\frac{1}{2}} = \mathrm{diag}\Big(\frac{1}{\sqrt{d_1}}, \frac{1}{\sqrt{d_2}}, \ldots, \frac{1}{\sqrt{d_{l+u}}}\Big). \quad (10)$$

Here, the weight matrix $\boldsymbol{W}$ is normalized symmetrically, which is necessary for the convergence of the following iteration. The iterative equation for label propagation is derived as follows:

$$\boldsymbol{F}(t+1) = \alpha\boldsymbol{S}\boldsymbol{F}(t) + (1-\alpha)\boldsymbol{Y}, \quad (11)$$

where $\alpha \in (0, 1)$ is a parameter that specifies the relative amount of the information from its neighbors and its initial label information in each iteration step. Note that self-reinforcement is avoided since the diagonal entries of $\boldsymbol{W}$ are set to 0. The iteration process continues until $\boldsymbol{F}$ converges:

$$\boldsymbol{F}^* = \lim_{t\to\infty} \boldsymbol{F}(t) = (1-\alpha)(\boldsymbol{I} - \alpha\boldsymbol{S})^{-1}\boldsymbol{Y}. \quad (12)$$

Based on the converged matrix $\boldsymbol{F}^*$ and the following classification rule:

$$y_i = \arg\max_{1\le j\le C} \boldsymbol{F}^*_{ij}, \quad (13)$$

the unlabeled segments $(\boldsymbol{x}_{l+1}, \boldsymbol{x}_{l+2}, \ldots, \boldsymbol{x}_{l+u})$ in U can be automatically labeled.

V. EXPERIMENTS

In this section, we conduct experiments on a real-world dataset to validate the effectiveness of the proposed activity labeling approach.

A. DATASET

In this experiment, a freely available data set ''Human Activity Recognition from Continuous Ambient Sensor Data''[1] provided by the Center of Advanced Studies in Adaptive Systems (CASAS) at Washington State University [18] is used. The entire data set consists of 15 sub-data sets: CSH101 to CSH115, each of which records the daily activities of a volunteer in a smart home within a month. And about 100,000 to 120,000 sensor events are collected for each volunteer. In the CASAS project, motion sensors, door sensors, temperature sensors, light sensors and other kinds of sensors are deployed in different locations throughout a smart home for activity and environment monitoring. For simplicity, we only use the data generated by motion sensors and door sensors for experiments.

B. DATA PROCESSING

In order to obtain standardized data for experiments, common activities in these 15 sub-data sets are analyzed and similar activities are merged. For instance, the three activities ''Eating Breakfast'', ''Eating lunch'' and ''Eating dinner'' in the original data sets can be combined into the ''Eat'' activity. Finally, we obtain a total of 11 activity categories, i.e., ''Cook'', ''Eat'', ''Bathe'', ''Sleep'', ''Toilet'', ''Bed_Toilet_Transition'', ''Leaving'', ''Wash_Dishes'', ''Work_At_Table'', ''Personal_Hygiene'' and ''Other_Activity''. Specially, Figure 2 depicts the distribution of different activities in the first sub-data set CSH101.

C. EVALUATION METRIC

In order to evaluate the performance of the sequence segmentation technique, the label information provided by the experimental data sets is utilized. Please note that the label

[1] https://archive.ics.uci.edu/ml/datasets.php

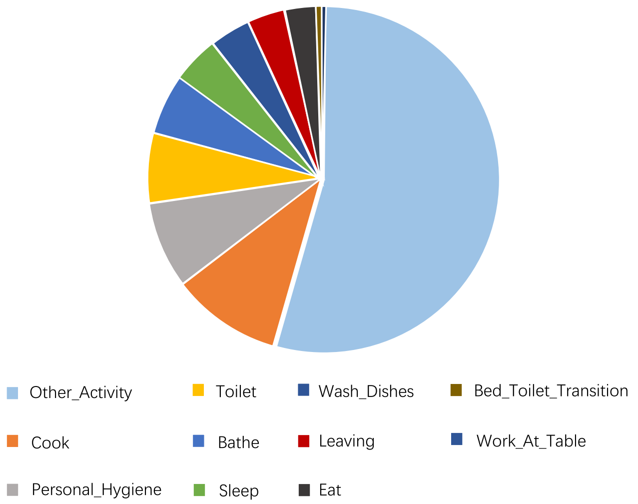

FIGURE 2. Distribution of activities in the CSH101 dataset.

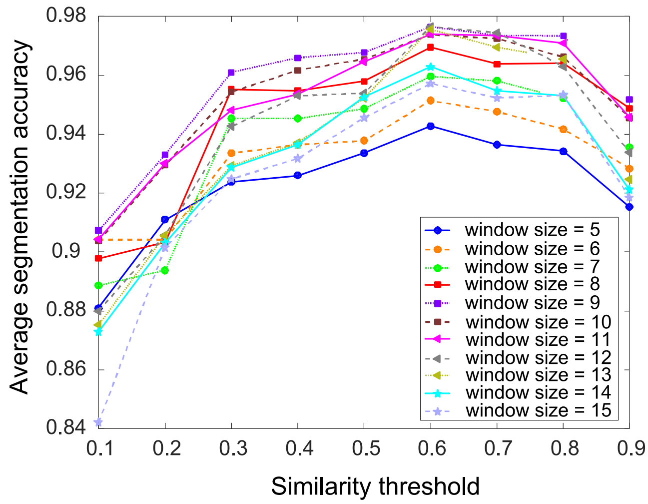

FIGURE 3. Impact of similarity threshold on segmentation accuracy.

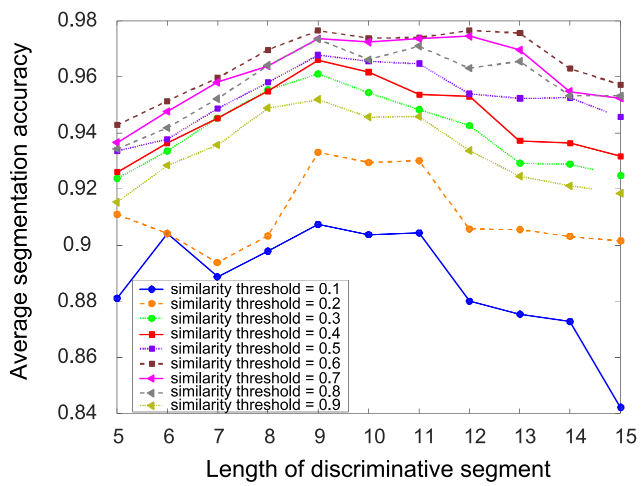

FIGURE 4. Impact of length of each discriminative segment on segmentation accuracy.

information is generally not available in real-world applications. Then, a majority-vote method is used to determine the category label of a segment. In other words, the most frequently occurring label in the current segment is used as the label of the segment. Afterwards, we calculate the proportion of sensor events with the same label as the segment to which they belong to measure the average Segmentation Accuracy (*SA*) of the sequence segmentation algorithm:

$$SA = \frac{1}{m}\sum_{i=1}^{m}\frac{|\{\boldsymbol{e}_{S_i}\}|}{|S_i|} \times 100\%, \tag{14}$$

where the operator $|\cdot|$ calculates the cardinality of a set, m is the total number of segments obtained by the segmentation algorithm, S_i is the ith segment, which can be seen as an ordered set of sensor events, $\{\boldsymbol{e}_{S_i}\}$ denotes all sensor events that belong to S_i and have the same label as S_i.

In addition, the accuracy of labeling is evaluated in this experiment. The Labeling Accuracy (*LA*) is defined as the proportion of correctly labeled events among all sensor events:

$$LA = \frac{|\{\boldsymbol{e}_c\}|}{|S|} \times 100\%, \tag{15}$$

where S is the entire sequence of unlabeled sensor events, and $\{\boldsymbol{e}_c\}$ represents the set of correctly labeled sensor events in S.

D. PERFORMANCE COMPARISON

The parameters that need to be predetermined in the *ASSA* algorithm are the length δ of each discriminative segment and the segment similarity threshold θ. Different values of the two parameters affect the segmentation accuracy, so we use different parameter values for experiments and choose the most suitable values that can achieve the best segmentation accuracy.

Take the CSH101 dataset as an example. Figures 3 and 4 respectively show the impacts of different values of the similarity threshold and the length of each discriminative segment on the segmentation accuracy. In Figure 3, the x-axis and the y-axis represent the similarity threshold and the average segmentation accuracy, respectively. The legend shows the different lengths of each discriminative segment corresponding to different polylines. We can see that when the length of a discriminative segment is fixed, the segmentation accuracy increases first and then decreases as the similarity threshold increases from 0.1 to 0.9, and reaches a maximum value when the similarity threshold is about 0.6.

In Figure 4, the x-axis and the y-axis represent the length of each discriminative segment and the segmentation accuracy, respectively. The similarity thresholds corresponding to different polylines is shown in the legend. Obviously, when the similarity threshold is fixed, the segmentation accuracy reaches a maximum value when the length of each discriminative segment is about 9. Therefore, we choose 0.6 and 9 as the optimal values of the similarity threshold and the length of each discriminative segment. The experimental results also show that the proposed sequence segmentation technique can achieve a high segmentation accuracy. δ and θ are two important parameters in the *ASSA* algorithm. They have different

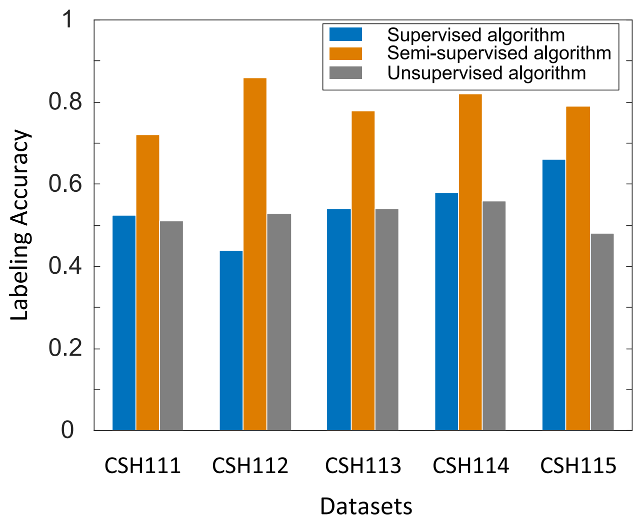

FIGURE 5. **Accuracy of activity labeling by different models.**

optimal values in different smart home environments. In our experiments, we evaluate the impacts of the two parameters on the accuracy of the proposed *ASSA* algorithm and set them as the values that can yield the highest segmentation accuracy. Now, we have to use the try-and-error method for parameter selection in different smart home environments. In the future work, we will explore new methods to automatically select optimal values of these parameters.

After applying the sequence segmentation algorithm to the raw sequence of sensor events, a series of segments are obtained. Then, a small set of segments should be manually labeled before the label propagation algorithm is performed. Generally, the accuracy of the activity labeling method increases with the increase of the number of manually labeled samples. In the experiments, we manually label the segments of the first day, based on which to label the remaining unlabeled segments. Empirically, a one-day dataset usually contains almost all daily activities of a family, and manually labeling a one-day dataset does not impose much burden on humans but can achieve a good performance of activity labeling, which is shown in the following experiments. Additionally, since the *ASSA* algorithm has tried to group consecutive sensor events that occur in a same activity into a same segment, the burden of manual labeling can be further reduced. Therefore, it is feasible to manually label a small subset of the total raw training data in real-world applications.

Next, we compare the labeling accuracy of the following three different algorithms: the proposed graph-based semi-supervised labeling propagation algorithm, a supervised algorithm based on Support Vector Machine (SVM) [33] and an unsupervised algorithm based on transfer learning [34]. Among different supervised machine learning algorithms, SVM is a commonly used one, so we use it as the baseline algorithm to investigate the effectiveness of our method. In the future work, we will try to compare the performance of our method with other machine learning algorithms such as HMM and Conditional Random Field. Additionally, in order to verify the superiority of our method, which relies on a small number of manually labeled samples, we compare it with the transfer learning method, which does not require any labeled data. The supervised algorithm uses 1 day of manually labeled data to train a SVM model, which can further predict the unknown category information of the remaining training data. The unsupervised algorithm utilizes the transfer learning technique to borrow label information from other data sets. To be specific, the first 10 sub-data sets CSH101 to CSH110 are used as source domains, and the remaining 5 sub-data sets CSH111 to CSH115 are regarded as target domains. The unsupervised labeling algorithm mainly relies on the label information extracted from source domains and the mapping relationships between the source and target domains, so it does not require any label information from the target domain. In addition, label information from multiple source domains should be fused for labeling in the target domain.

The labeling accuracy of the three algorithms on the CSH111 to CSH115 data sets is depicted in Figure 5. The x-axis and the y-axis represent different data sets and the labeling accuracy, respectively. The experimental results tell us that the labeling accuracy of the unsupervised algorithm and the supervised algorithm is comparable. The slight difference is due to different user habits and sensor distributions in different smart environments. Compared with the two algorithms, the semi-supervised labeling propagation algorithm can achieve the highest labeling accuracy on all the 5 data sets. Additionally, its labeling accuracy is always above 72%, indicating that the semi-supervised algorithm based on a small number of manually labeled data has an excellent ability of labeling.

VI. DISCUSSION AND CONCLUSION

In this article, we proposed an activity labeling approach in Health Smart Home based on the semi-supervised learning technique. This approach can automatically group consecutive sensor events that occur in a same activity into a same segment. Moreover, it only needs a small number of manually labeled segments to achieve high-accuracy labeling of the remaining unlabeled training data. Accordingly, the proposed approach can effectively reduce the huge burden of manual labeling in real-world applications.

However, the presented work still has some limitations. First, the fixed parameter values (e.g., the similarity threshold and the length of each discriminative segment) adopted in the segmentation algorithm will undoubtedly affect the sequence segmentation inaccuracy to a certain extent. Second, this approach still requires some manually labeled data, which reduces the efficiency of activity labeling. In the future work, we will try to solve these problems by exploring new methods to automatically select optimal values of algorithm parameters and incorporating transfer learning into our method, thereby further improving the accuracy and efficiency of activity labeling. Third, in this work, we only compare our method with one machine learning algorithm, i.e., SVM. In the future work, we will try to compare our method with more other machine learning algorithms, e.g., HMM and Conditional Random Field.

REFERENCES

[1] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, "An information framework for creating a smart city through Internet of things," *IEEE Internet Things J.*, vol. 1, no. 2, pp. 112–121, Apr. 2014.

[2] 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 Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 274–283, Mar. 2010.

[3] M. Chan, D. Estève, C. Escriba, and E. Campo, "A review of smart homes—Present state and future challenges," *Comput. Methods Programs Biomed.*, vol. 91, no. 1, pp. 55–81, 2008.

[4] S. S. Intille, "A new research challenge: Persuasive technology to motivate healthy aging," *IEEE Trans. Inf. Technol. Biomed.*, vol. 8, no. 3, pp. 235–237, Sep. 2004.

[5] G. D. Abowd, E. D. Mynatt, and T. Rodden, "The human experience [of ubiquitous computing]," *IEEE Pervas. Comput.*, vol. 1, no. 1, pp. 48–57, Jan. 2002.

[6] G. LeBellego, N. Noury, G. Virone, M. Mousseau, and J. Demongeot, "A model for the measurement of patient activity in a hospital suite," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 92–99, Jan. 2006.

[7] S. Bonhomme, E. Campo, D. Estève, and J. Guennec, "PROSAFE-extended, a telemedicine platform to contribute to medical diagnosis," *J. Telemed. Telecare*, vol. 14, no. 3, pp. 116–119, Apr. 2008.

[8] R. Poppe, "A survey on vision-based human action recognition," *Image Vis. Comput.*, vol. 28, no. 6, pp. 976–990, Jun. 2010.

[9] C. Wu, A. H. Khalili, and H. Aghajan, "Multiview activity recognition in smart homes with spatio-temporal features," in *Proc. 4th ACM/IEEE Int. Conf. Distrib. Smart Cameras (ICDSC)*, Aug. 2010, pp. 142–149.

[10] M.-C. Chang, N. Krahnstoever, S. Lim, and T. Yu, "Group level activity recognition in crowded environments across multiple cameras," in *Proc. 7th IEEE Int. Conf. Adv. Video Signal Based Surveill.*, Aug. 2010, pp. 56–63.

[11] P. Klasnja, S. Consolvo, T. Choudhury, R. Beckwith, and J. Hightower, "Exploring privacy concerns about personal sensing," in *Proc. 7th Int. Conf. Pervasive Comput.*, Nara, Japan, May 2009, pp. 176–183.

[12] Y. Pu, T. Xiang, C. Hu, A. Alrawais, and H. Yan, "An efficient blockchain-based privacy preserving scheme for vehicular social networks," *Inf. Sci.*, vol. 540, pp. 308–324, Nov. 2020.

[13] C. Hu, W. Li, X. Cheng, J. Yu, S. Wang, and R. Bie, "A secure and verifiable access control scheme for big data storage in clouds," *IEEE Trans. Big Data*, vol. 4, no. 3, pp. 341–355, Sep. 2018.

[14] D. J. Patterson, D. Fox, H. Kautz, and M. Philipose, "Fine-grained activity recognition by aggregating abstract object usage," in *Proc. 9th IEEE Int. Symp. Wearable Comput. (ISWC)*, Oct. 2005, pp. 44–51.

[15] L. Zhang, Z. Kuang, K. Li, K. Wei, Z. Wang, S. Zhang, and R. Wang, "Human activity behavior recognition based on acceleration sensor and neural network," *Mod. Electron. Technique*, vol. 42, no. 16, pp. 71–74 and 78, 2019.

[16] J. Zhang, Z. Wu, F. Li, J. Luo, T. Ren, S. Hu, W. Li, and W. Li, "Attention-based convolutional and recurrent neural networks for driving behavior recognition using smartphone sensor data," *IEEE Access*, vol. 7, pp. 148031–148046, 2019.

[17] T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, "Accurate activity recognition in a home setting," in *Proc. 10th Int. Conf. Ubiquitous Comput. (UbiComp)*, 2008, pp. 1–9.

[18] D. Cook, "Learning setting-generalized activity models for smart spaces," *IEEE Intell. Syst.*, vol. 27, no. 1, pp. 32–38, Jan. 2012.

[19] D. J. Cook, N. C. Krishnan, and P. Rashidi, "Activity discovery and activity recognition: A new partnership," *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 820–828, Jun. 2013.

[20] Y. Tong, "Research on activity recognition in smart home based on conditional random fields," Ph.D. dissertation, Dept. Comput. Sci. Technol., Dalian Maritime Univ., Dalian, China, 2015.

[21] J. Wu, "Smart home activity recognition and behavior analysis based on environmental sensors," Ph.D. dissertation, Dept. Inf. Commun. Eng., Xiamen Univ., Xiamen, China, 2016.

[22] C. Bettini, G. Civitarese, and R. Presotto, "CAVIAR: Context-driven active and incremental activity recognition," *Knowl.-Based Syst.*, vol. 196, May 2020, Art. no. 105816.

[23] X. Yan, "Weighted KNN classification algorithm based on mean distance of category," *Comput. Syst. Appl.*, vol. 23, no. 2, pp. 128–132, 2014.

[24] C. Du, "Railway engineering staff behavior recognition method based on machine learning," *Comput. Syst. Appl.*, vol. 28, no. 7, pp. 199–205, 2019.

[25] Z. W. Yan He and K. Wang, "A user-independent behavior recognition model based on multi-classifier fusion," *J. Xi'an Univ. Posts Telecommun.*, vol. 21, no. 5, pp. 50–54, 2016.

[26] H. D. Mehr, H. Polat, and A. Cetin, "Resident activity recognition in smart homes by using artificial neural networks," in *Proc. 4th Int. Istanbul Smart Grid Congr. Fair (ICSG)*, Apr. 2016, pp. 1–5.

[27] J. Suto and S. Oniga, "Efficiency investigation of artificial neural networks in human activity recognition," *J. Ambient Intell. Humanized Comput.*, vol. 9, no. 4, pp. 1049–1060, Aug. 2018.

[28] S. Oniga and J. Suto, "Activity recognition in adaptive assistive systems using artificial neural networks," *Elektronika ir Elektrotechnika*, vol. 22, no. 1, pp. 68–72, Feb. 2016.

[29] Y.-S. Lee and S.-B. Cho, "Activity recognition with Android phone using mixture-of-experts co-trained with labeled and unlabeled data," *Neurocomputing*, vol. 126, pp. 106–115, Feb. 2014.

[30] X. Yang, A. Dinh, and L. Chen, "Implementation of a wearerable real-time system for physical activity recognition based on naive bayes classifier," in *Proc. Int. Conf. Bioinf. Biomed. Technol.*, 2010, pp. 101–105.

[31] D. Trabelsi, S. Mohammed, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "An unsupervised approach for automatic activity recognition based on hidden Markov model regression," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 3, pp. 829–835, Jul. 2013.

[32] D. Chen, Y. Shen, B. Xie, and Y. Ma, "A measure model of similarity for finding the best coach," *J. Northeastern Univ. (Natural Sci.)*, vol. 35, no. 12, pp. 1697–1700, 2014.

[33] N. Cristianini and J. Shawetaylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[34] D. Cook, K. D. Feuz, and N. C. Krishnan, "Transfer learning for activity recognition: A survey," *Knowl. Inf. Syst.*, vol. 36, no. 3, pp. 537–556, Sep. 2013.

YAN HU was born in 1988. She received the B.S. degree in automation from Xi'an Jiaotong University, Xi'an, Shaanxi, China, in 2011, and the Ph.D. degree in computer science from the University of Chinese Academy of Sciences, Beijing, China, in 2017.

Since 2017, she has been an Assistant Professor with the University of Science and Technology Beijing, Beijing. Her main research interests include smart city, the Internet of Things, and security of industrial control systems.

BINGCE WANG was born in 1998. He received the B.S. degree in Internet of Things from the University of Science and Technology Beijing, Beijing, China, in 2020. He is currently pursuing the Ph.D. degree in software engineering with Peking University, Beijing. His main research interests include the Internet of Things and machine learning.

YUYAN SUN was born in 1982. He received the B.S. degree in computer science from Peking University, Beijing, China, in 2004, and the M.S. and Ph.D. degrees in computer science from the University of Chinese Academy of Sciences, Beijing, in 2007 and 2016, respectively.

From 2007 to 2012, he was a Research Assistant with the Institute of Software, Chinese Academy of Sciences. Since 2016, he has been an Assistant Professor with the Beijing Key Laboratory of IoT Information Security, Institute of Information Engineering, Chinese Academy of Sciences, and the School of Cyber Security, University of Chinese Academy of Sciences. His main research interests include the Internet of Things and security of industrial control systems.

JING AN was born in 1984. She received the B.S. degree in management from Yunnan University, Kunming, Yunnan, China, in 2006, and the Ph.D. degree in ethnology from the Minzu University of China, Beijing, China, in 2012.

From 2013 to 2016, she was a Postdoctoral Researcher with the Chinese Academy of Social Sciences. From 2016 to 2018, she was an Assistant Professor with the University of Science and Technology Beijing, China, where she has been an Associate Professor, since 2018. Her main research interests include cyber security and cyber space governance.

ZHILIANG WANG was born in 1956. He received the B.S. degree in industrial automation from Yanshan University, Qinhuangdao, Hebei, China, in 1982, and the M.S. and Ph.D. degrees in electronic engineering from the Harbin Institute of Technology, Harbin, Heilongjiang, China, in 1985 and 1988, respectively.

From 1988 to 1991, he was a Postdoctoral Researcher with Zhejiang University. Since 1991, he has been a Professor and a Ph.D. Supervisor with the University of Science and Technology Beijing, China. His main research interests include smart city, the Internet of Things, and robot technology.