A

Activity Recognition with Evolving Data Streams: A Review

Zahraa S. Abdallah, Faculty of Information Technology, Monash University, Australia
Mohamed Medhat Gaber, School of Computing & Digital Technology, Birmingham City University, UK
Bala Srinivasan, Faculty of Information Technology, Monash University, Australia
Shonali Krishnaswamy, AIDA Technologies Pte Ltd, Singapore

Activity recognition aims to provide accurate and opportune information on people's activities by leveraging sensory data available in today's sensory rich environments. Nowadays, activity recognition has become an emerging field in the areas of pervasive and ubiquitous computing. A typical activity recognition technique processes data streams that evolve from sensing platforms such as mobile sensors, on body sensors, and/or ambient sensors. This paper surveys the two overlapped areas of research of activity recognition and data stream mining. The perspective of this paper is to review the adaptation capabilities of activity recognition techniques in streaming environment. Categories of techniques are identified based on different features in both data streams and activity recognition. The pros and cons of the algorithms in each category are analysed and the possible directions of future research are indicated.

CCS Concepts: •**General and reference** → **Surveys and overviews;** •**Theory of computation** → **Streaming models;** *Machine learning theory;* •**Computing methodologies** → **Transfer learning;** •**Hardware** → **Sensor applications and deployments; Sensor devices and platforms;**

ACM Reference Format:
Zahraa Said Abdallah, Mohamed Medhet Gaber, Bala Srinivasan, and Shonali Krishnaswamy. Activity recognition in evolving data streams: A review. *ACM Computing Surveys* V, N, Article A (January YYYY), 41 pages.
DOI: http://dx.doi.org/10.1145/0000000.0000000

1. INTRODUCTION

Sensors are becoming more pervasive. They exist ubiquitously around us and are embedded in our phones, cameras, clothing, buildings, cars, and in all kinds of everyday objects. Massive amounts of data are generated from these sensors continuously. The availability of real time sensory information through these sensors has led to the emergence of research into "Activity Recognition" (AR). Activity recognition aims to provide accurate and opportune information based on people's activities and behaviours. Activity recognition has become an emerging field in the areas of pervasive sensory data processing and ubiquitous computing. Many applications have demonstrated the usefulness of activity recognition. These include applications in healthcare [Tentori and Favela. 2008; Do et al. 2013; Wu et al. 2008; Mohomed et al. 2008; Zhang et al. 2008; Sánchez et al. 2008], social networks [Miluzzo et al. 2008], environmental monitoring [Mun et al. 2009], surveillance, and emergency response [Zhang et al. 2008; Nait-Charif and McKenna 2004]

Activity recognition has been widely studied using different approaches and from various perspectives. Probabilistic, statistical and logical reasoning approaches have been applied to understand and predict various user activities. Additionally, machine learning approaches based on sensory data have also been leveraged for activity recog-

Author's addresses: Z. S. Abdallah and B. Srinivasan, School of Information Technology, Monash University, Melbourne, Australia; Mohamed Medhat Gaber, School of Computing & Digital Technology, Birmingham City University, UK; Shonali Krishnaswamy, AIDA Technologies Pte Ltd, Singapore.
 1539-9087/YYYY/01-ARTA $15.00
DOI: http://dx.doi.org/10.1145/0000000.0000000

nition. The premise underlying the use of machine learning in activity recognition is that activities can be recognised and even anticipated using prior knowledge of previously collected data representing different activities. Typical activity recognition process consists of three main components: namely, *data collection and preprocessing*, *modelling*, and finally *recognition*. The *data collection and preprocessing component* that gathers annotated sensory data evolves from diverse data sources such as on body wearable sensors, mobile sensors, and/or smart environment sensors. Then, raw sensory data is processed into features that help discriminate between activities. The *modelling component* uses the extracted features to train a baseline learning model that is then deployed to predict activities from new incoming sensory data by the *recognition component*.

State of the art activity recognition techniques rely strongly on prior knowledge to recognise activities based on models built from samples of the population. However, typical activity recognition techniques deal with data that is continuously streaming from various sensors. Sensory data is a data stream that contains unbounded data which arrives at high speed. Therefore, it is unrealistic in activity recognition to assume that data is static over time. Dynamic changes in activities that reflect variations in user's activities are expected and natural. Change of existing activities or emergence of novel activities occur in evolving activity data.

In this paper, we survey the area of research concerning activity recognition in data streams. The focus of this paper is on the research gaps and challenges that are faced by activity recognition approaches when dealing with data streams. We start the survey by discussing the scope of this survey across activity recognition and stream mining disciplines in Section 2. Within this scope, we first discuss different learning approaches for activity recognition in Section 3. Then, section 4 explores approaches to mine changes in data streams in general. Section 5 represents the research area comprising the intersection of stream mining and activity recognition approaches. Section 6 surveys key activity recognition systems applied in streaming settings. We further present in Section 7 the current research challenges and gaps, along with a comparison of key systems. Section 8 discusses future directions of research. The paper is summarised and concluded in Section 9.

2. SURVEY SCOPE

Activity recognition is a wide research area that has been investigated from many perspectives. A well investigated research perspective focuses on managing the process of data collection. This perspective concerns issues related to the kinds of sensors and sensing platforms. Yet, immense research has been directed towards learning methods for recognising activities from sensory data. The literature covers a wide variety of learning techniques applied for activity recognition. This research focuses on learning in activity recognition, yet, in streaming environments. Although the literature represents subsets of the research themes across activity recognition and stream mining disciples, this survey primarily focuses on the holistic approach that merges both activity recognition and stream mining.

In the light of the survey's scope, we introduce different learning techniques for activity recognition with data streams. To explain the scope, we first introduce the known term of *i.i.d.* in statistical and probability theory, which refers to data that is both independent and identically distributed. A key challenge with learning from sensory data in a streaming environment is that it requires learning beyond identically distributed and independent conditions. In machine learning, a collection of data D is defined as independent and identically distributed if all samples in $D : X_1, X_2, ...X_n$ follow the same distribution function f and are independent of each other.

Many state-of-the-art statistical and machine learning approaches rely on processing the data that follows the *i.i.d.* conditions. These approaches assume that data instances are independent and follow the same distribution. Thus, the prediction of new data relies strongly on prior knowledge. In the ubiquitous environment, either of the *i.i.d.* conditions or even both can be violated. The distribution condition is clearly challenged with the basic concepts of data streaming. In a streaming dynamic environment, we can not assume the fulfilment of the identical distribution condition, as a typical data stream evolves over time. There is an increasing interest in the domain of learning from data that is not identically distributed. Changes in data distribution leads to concept drifts that are formally defined using the Bayes rule as the change in the prior and/or the likelihood. A concept drift can be abrupt, gradual, incremental, or recurrent. The appearance of new concept corresponds to concept evolution. While, outliers can be considered as a new concept that only occurred once. A main difference between outliers and concept evolution is the recurrence of the concept and action required upon detecting. In concept evolution, we expect more points in the stream that represent the new concept and therefore the classification model is required to be adapted to include the novel concept. On the other hand, outliers are mostly sparse and not forming a concept. The action required for outliers is filtration rather than adaptation. Indeed, these methods still preserve data independence. Other areas of research focus on the violation of the independent data condition. These approaches deal with data that is dependent while the distribution is fixed. Stationary time series and Markov Chains are examples of approaches that challenge the dependency assumption.

Activity recognition data that represents a sequence of performed activities is intuitively dependent. Therefore, an efficient activity recognition system focuses on dealing with the dependency among data for predicting performed activities. Furthermore, in a streaming environment, activity recognition violates not only the assumption of independence but also identical distribution. Changes in activity recognition context can also be abrupt (e.g., change in walking pattern after an accident), gradual (e.g., change in walking pattern for toddlers), incremental (change in walking pattern during healing from an injury), or recurrent (e.g., repeated change of walking pattern according to situations). However, new activities performed frequently represent novel concepts or concept evolution (e.g., using a gym instrument for the first time). On the other hand, a new activity performed once can be an outlier (e.g., sudden fall or malicious behaviour). Figure 1 depicts the intersection of the two research areas of stream mining and activity recognition. In the following sections, we survey methods for both activity recognition and data stream mining as well as the intersection between them.

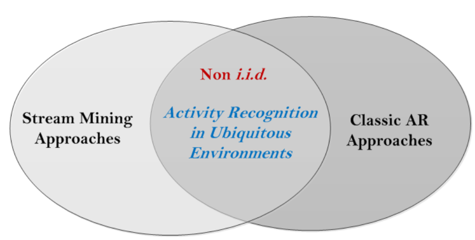


Fig. 1: Survey Scope

3. ACTIVITY RECOGNITION TECHNIQUES

We survey in this section the main body of research in activity recognition. Different from traditional probabilistic and logical approaches, other approaches that consider activity recognition from the machine learning perspective have been studied. In general, approaches in machine learning for activity recognition can be broadly divided into two major strands. The first strand concerns the underlying learning approach. That includes supervised, unsupervised, and semi-supervised learning. The second strand focuses on the dynamic capabilities of the recognition system beyond the learning phase. This includes personalisation and adaptation of a learning model and the concept of transfer learning.

3.1. Overview

Activity recognition is one of the emerging applications in the area of ubiquitous computing. Systems that can recognise human activities opened the door to many important applications such as:

— Health and wellness [Longstaff et al. 2010; Karantonis et al. 2006; Hong et al. 2010; Sánchez et al. 2008; Do et al. 2013]: Progressive research in activity recognition has provided the foundation for many applications in health and wellness. In recent years, fitness tracking applications have attracted much attention in activity recognition. Devices and applications, such as fitbit [1], monitor distances walked and corresponding burned calories. More advanced fitness applications aim at tracking Activities of Daily Living (ADL) [Hong et al. 2010]. Recognition and monitoring of the type and frequency of ADL is essential for creating what is known as an activity diary/log [Yang 2009]. These diaries help users to understand their personal lifestyle patterns and effect healthy changes (e.g. increase physical exercise, reduce number of hours sitting in front of the computer). Such activity recognition applications are important for preventing medicine and avoiding chronic illness such as cardiovascular diseases, diabetes and obesity. Other applications in health monitoring are concerned with the remote supervision of home based patients or at risk elderly people. Medical professionals believe that one of the best ways of early detection and prevention of emerging medical conditions is to recognise changes and abnormality in different activities [Lawton and Brody 1969].

— Activity-based crowdsourcing and surveillance [Bodor et al. 2003; Wilson and Atkeson 2005]: Recognising activities for crowds leads to interesting applications. A case of a large number of people running in a place where they normally walk or sit indicates a possible emergency or disaster [Lockhart et al. 2012]. Also, surveillance applications that are able to understand and model people's activities could predict intent and motive as people interact with the environment. Therefore, activity recognition applications aim to proactively detect abnormal behaviours in busy environments [Niu et al. 2004]. An example is a suspicious person who is spending longer than usual time on a train platform.

— Targeted advertising [Intille 2004; Sricharan et al. 2006]: Activity recognition in real time is an important component of applications that interact with users to deliver context relevant information and services. Applications in this category target conveying the right message at the right place and at the right time using activity recognition. Examples include personalised advertisements or discount deals in smart shopping scenarios. The term Know Your Customer (KYC) [Woodruff and Gardial 1996], which has been used by businesses, refers to understanding customer needs and provides them with satisfying services. Activity recognition contributes

[1] http://www.fitbit.com

to KYC analytics by inferring the general interests of customers and thus helps in providing them with relevant information and services.

Typical activity recognition process consists of three main components: data collection and preprocessing, modelling, and finally recognition. The flow of the learning process through different components in activity recognition is illustrated in Figure 2 The data collection and preprocessing component that gathers annotated sensory data evolves from diverse data sources such as on body wearable sensors, mobile sensors, and/or smart environment sensors. Then, raw sensory data is processed into features that help discriminate between activities. The modelling component uses the extracted features to train a baseline-learning model that is then deployed to predict activities from new incoming sensory data by the recognition component.

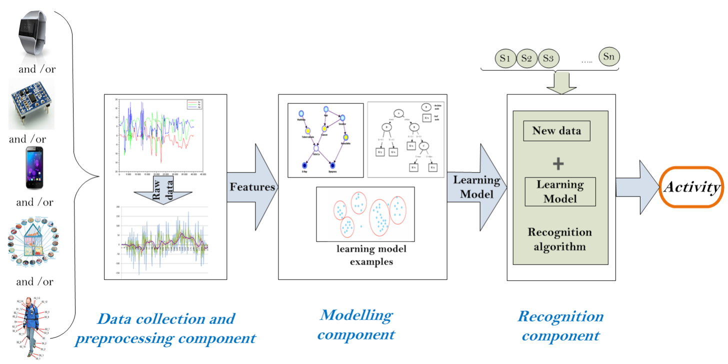


Fig. 2: Activity Recognition Components

Data collection and preprocessing component represents the very initial step of any activity recognition process. Data collection includes issues related to the sensing platform such as types and locations of the sensors. Both types and locations vary based on the aim of the recognition. For instance, recognising hand gestures may require accelerometer sensors attached to hands or fingers. Variously, spatiotemporal activities require collection of GPS data that could be from a user's device that they carry. Atomic activities such as sitting and walking may also be recognised using accelerometer data embedded in a mobile device or attached as wearable sensors. Another consideration for the data collection concerns the annotation or labelling of activities in the collected data. The process of data collection is followed by a preprocessing and feature extraction steps, which aim to prepare the raw collected data for the following component of modelling.

The collected data is then pre-processed in order to recognise the target activities. The pre-processing step in activity recognition strongly relies on sensors applied for data collection and aim of the recognition process (kind of activities). Raw data is processed in this step with feature selection and extraction methods to extract meaningful information that can distinguish between different activities. For example, accelerometer raw data (i.e. x, y and z component) needs to be transformed to a set of features that include the magnitude, mean, standard deviation, and number of peaks of the accelerometer readings along the three axes. Studies in activity recognition have applied various well-known feature extraction and selection techniques in the collection and preprocessing component that also includes data filtration from both noise and outliers [Hong et al. 2010; Kwapisz et al. 2011]

Popular features include mean, standard deviation, maximum, peak-to-peak, root-mean-square, and correlation between values of modality axes. Autoregressive modelling can also be applied on raw data to form augmented-feature vectors. Survey on features used for activity recognition can be found in [Bulling et al. 2014]. Section 3.2 discusses learning techniques related to modelling and recognition components in Activity Recognition process.

3.2. Learning Approaches

Typical activity recognition systems deploy supervised learning methods such as Naive Bayes, Decision Trees, Hidden Markov Models, Nearest Neighbour, Support Vector Machines, and different Boosting techniques for activity classification. Moving beyond fully supervised settings, researchers have started studying the feasibility of other machine learning techniques for activity recognition: unsupervised and semi-supervised learning.

Supervised learning is applied pervasively for activity recognition. Basically, labelled data is collected to train a static classification model for recognising a set of activities. The classification model built from labelled data is used to recognise the incoming unlabelled data. Figure 3 explains the supervised learning process. Supervised learning is categorised as either generative, discriminative or hybrid approach [Rubinstein et al. 1997]. Algorithms that follow the generative approach models the class conditional distribution. Naive Bayes is an effective generative approach that is applied pervasively for activity recognition such as in [Ravi et al. 2005]. Hidden Markov Models are also generative methods that have been successfully applied for recognising activities [Patterson et al. 2005; Ward et al. 2006]. Discriminative models on the other hand learn the boundaries between classes. Decision trees [Bao and Intille 2004; Logan et al. 2007] and nearest neighbour [Maurer et al. 2006; Lee and Mase 2002] are well-studied examples of the discriminative approach for activity recognition. Moreover, a hybrid approach is the one that combines the two approaches into a single classifier. In [Viola and Jones 2001], authors applied a modified version of AdaBoost that combines a set of static classifiers for recognising activities. They demonstrated the efficiency of combining discriminative and generative classifiers for smooth recognition of activities [Lester et al. 2005]. The hybrid approach discriminatingly selects useful features and learns an ensemble of static classifier to recognise different activities. In [Yuan and Herbert 2014], a hybrid classifier was developed that combines both threshold based and machine learning methods to select the most suitable classifier dynamically on the cloud. Artificial Neural Networks rely on a hybrid generative-discriminative approach. Authors in [Do et al. 2013] developed a system based on stream reasoning and Artificial Neural Networks for recognising activities from mobile phone sensors.

A wide range of supervised methods commonly used for activity classification was reviewed in [Preece et al. 2009; Peterek et al. 2014]. One main characteristic of these methods is the necessity of a significant amount of labelled data to build the classification model. The assumption that labelled data is consistently available is unrealistic. Due to many reasons, activities may vary while time evolves for the same person or across different individuals. For each activity, data is required to be collected for each user to attain an accurate recognition. However, the annotation process is a time consuming, error prone, and mostly tedious process. Therefore, researchers investigated an unsupervised learning approach for activity recognition to overcome the limitations of the supervised approaches.

The main goal of ***unsupervised learning*** techniques in activity recognition is to discover variation and likelihood among data. Figure 4 depicts an overview of the unsupervised learning process. There are few researchers who studied unsupervised learning for activity recognition such as [Lee et al. 2009; Wyatt et al. 2005; Li and

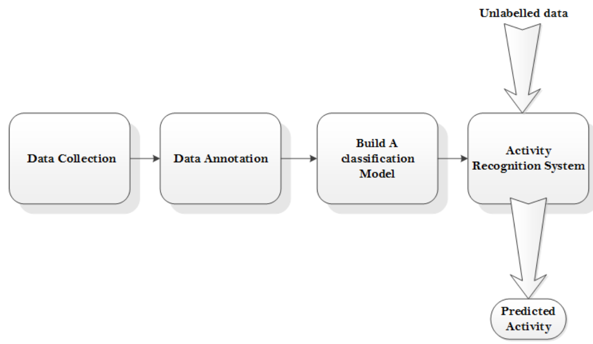


Fig. 3: Supervised Learning in Activity Recognition

Dustdar 2011; Huynh and Schiele 2006; Ye et al. 2014]. Lee et al.[Lee et al. 2009] used unsupervised learning for abnormality detection. To detect whether a pattern is registered or not, a probability model based on the past activity pattern is created. The Expectation-Maximisation (EM) algorithm is applied with the feature vectors to decide whether the activity has abnormal behaviour. In another study, Li and Dustdar [Li and Dustdar 2011] studied the feasibility of applying a specific type of unsupervised learning to high-dimensional, heterogeneous sensory input. The correspondence between clustering output and classification input is proposed as well. Although clustering is a promising approach in discovering data structure and patterns, at least a few labels have to be provided for performing the actual recognition of activities. Also, the feasibility of traditional clustering approach is questionable for high dimensional streaming data [Li and Dustdar 2011]. Furthermore, methods for unsupervised learning require a large pool of unlabelled data in order to find interesting patterns.

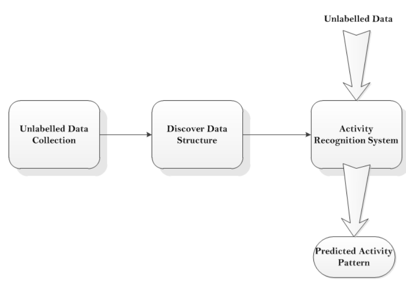


Fig. 4: Unsupervised Learning in Activity Recognition

In realistic conditions, large amounts of unlabelled data are easily collected while small set of labelled training data is available. In order to integrate the advantages of both supervised and unsupervised learning; the concept of ***semi-supervised learning*** is considered for activity recognition. A semi-supervised approach requires less labelled data for recognising a substantial amount of unlabelled data. Figure 5 represents an outline of the semi-supervised learning approach. The ability to use unlabelled data for enhancing the recognition system became an interesting topic for many researchers. Different approaches have been applied for semi-supervised learning that include self-learning, co-learning, multi-graph based and Multiple Eigenspaces. In the self-learning paradigm, a small amount of annotated data is used to build the classification model for later prediction of the unlabelled data. The predicted label with highest confidence is added to the training seed for rebuilding the classification model. Self learning has been successfully applied in many applications such as text analysis [Yarowsky 1995] and image processing [Li et al. 2007]. Continuing research is being conducted to study self-learning techniques for activity recognition. Longstaff, Reddy, and Estrin [Longstaff et al. 2010] investigated methods of further training classifiers after a user begins to use them using semi-supervised learning techniques.

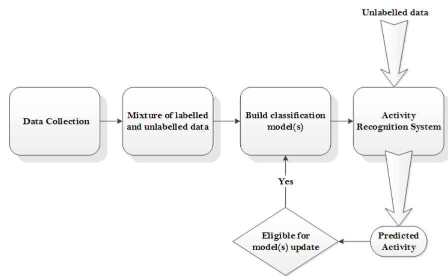


Fig. 5: Semi-supervised Learning in Activity Recognition

Co-learning was first developed by Blum and Mitchell [Blum and Mitchell 1998]. It uses two classifiers, each trained on different perspective of the data. Each classifier adds its most confidently predicted label to the training set and rebuilds the model. Guan et al. [Guan et al. 2007] applied co-learning with three classifiers from the same view of data. They also used majority voting to choose the most confident label to augment the training data. Stikic, Laerhoven, and Schiele [Stikic et al. 2008] applied various semi-supervised techniques for enhancing accuracy of recognising users' activities. Lee et al. [Lee and Cho 2014] presented a method to recognise a person's activities from sensors in a mobile phone using mixture-of-experts model.

Authors in [Stikic et al. 2009; Stikic et al. 2011] proposed multi-graph based methods that propagate information through a graph containing both labelled and unlabelled data. These methods deploy two different ways of combining multiple graphs based on feature similarity and time. Each node of the graph corresponds to an instance while every edge encodes the similarities between a pair of nodes as a probability value. Ali, King and Yang [Ali et al. 2008] implemented a Multiple Eigenspaces (MES) technique based on the Principal Component Analysis combined with Hidden Markov Models. The system is designed to recognise finger gestures with a laparoscopic gripper tool. Huynh and Schiele [Huynh and Schiele 2006] combined Multiple Eigenspaces with Support Vector Machines to recognise eight ambulation and daily activities.

Semi-supervised approaches can be categorised in general as transductive or inductive. The concept of transductive learning has been introduced in 1998 [Gammerman et al. 1998] in contrast to inductive learning. In transductive learning the aim is to predict the labels of unlabelled test data without building a model that maps between the input and output space. According to the transductive approach, there is no need to build a generic model to identify specific labels, as it will add unnecessarily computation overhead. Instead, in transductive learning, unlabelled test data is directly labelled based on the labelled training data.

The majority of semi-supervised methods applied for activity recognition are inductive. Induction approach is more applicable for activity recognition because of two reasons. First, the test data is not required to be available at the training time in inductive learning which is more applicable in activity recognition. In a typical activity recognition learning flow, labelled data exist in the training phase, while unlabelled data stream on later recognition/testing phase that mostly occurs online. It is computationally inefficient in streaming environment to keep both training and testing data available if applying induction methods. Second, the deployment domain may differ from the development domain. I.e., data that is used for training and testing is drawn from completely different sets (of users). Therefore, using induction approach to build a general classification model that maps between training and testing would be more suitable to avoid possible overfitting.

Despite the ongoing research on semi-supervised learning, most of the developed techniques rebuild the whole classification model upon predicting the most confident label. Yet, this is impractical for real time recognition in streaming settings. Also, the focus of aforementioned methods was mainly to minimise the labelled data required for building the initial classification model rather than improving the classifier itself, especially with confusing data [Longstaff et al. 2010].

Active learning is another approach in the semi-supervised learning category. Unlike self learning and co-learning, active learning requires user input to label data with the true label. According to Muslea, Minton, and Knoblock [Muslea et al. 2000], the main goal of active learning algorithm is to find the more profitable and less costly data to label. Kapoor and Horvitz [Kapoor and Horvitz 2008] compared different methods to decide upon the selection of profitable data. While Stikic et al. [Stikic et al. 2008] employed a multi-sensor approach to choose important data to be labelled. The data is selected for active learning based on two approaches. One approach chooses the data with the lowest classification confidence. The other approach chooses the data that causes a high degree of disagreement between two classifiers. Results showed improved performance when active learning is applied. State-of-the-art active learning techniques in activity recognition assume the retraining of models. This is not applicable in streaming settings. Alternatively, dynamic update to tune the classification model in real time are crucial in a streaming environment. Improving active learning methods to consider the value of label compared to the cost of interrupting the user is essential in the context of activity recognition. Also, estimating the time taken to provide true labels and their effect on the prediction accuracy in real time requires more investigations. Table I summarises basic learning approaches in activity recognition.

Table I: Learning Approaches in Activity Recognition

Learning Approach		**References**
Supervised	Generative	[Ravi et al. 2005; Patterson et al. 2005; Ward et al. 2006]
	Discriminative	[Bao and Intille 2004; Logan et al. 2007; Maurer et al. 2006; Lee and Mase 2002]
	Hybrid	[Viola and Jones 2001; Lester et al. 2005; Do et al. 2013]
Unsupervised		[Lee et al. 2009; Wyatt et al. 2005; Li and Dustdar 2011; Huynh and Schiele 2006; Ye et al. 2014]
Semi-Supervised	Self-learning	[Longstaff et al. 2010]
	Co-learning	[Stikic et al. 2008; Lee and Cho 2014]
	Multi-graph learning	[Stikic et al. 2009; Stikic et al. 2011]
	Multiple eigenspaces	[Ali et al. 2008; Huynh and Schiele 2006]

In this section, we have reviewed the three main approaches of activity recognition according to the underlying learning approaches. Traditional activity recognition systems are based on fully supervised learning approaches. However, these techniques require all data to be annotated which is impractical, especially when applied in a streaming environment. On the other hand, unsupervised learning finds pattens in unlabelled data. Unsupervised learning is not capable of finding the actual predicted label of activity without at least some labelled data presenting the ground truth. Therefore, semi-supervised learning is a recent trend in activity recognition that employs only a small set of labelled data for training. In semi-supervised learning, the recognition system continuously learns from unlabelled data either automatically with self learning or similar approaches or interactively with user input via an active learning approach.

3.3. Dynamic capabilities

In this section we aim to review the dynamic capabilities of the recognition system according to the objective of a system update. In the literature, two reasons urge model update beyond the learning phase: model *personalisation* to best fit a specific user; model *adaptation* by adding new activities or deleting abandoned ones. The wide concept of *transfer learning* is also presented to explain the different kinds of anticipated changes in activity recognition and the proposed techniques to handle these changes. We represent, in the following, the literature contribution to each category.

3.3.1. Model personalisation. It is hard to generate one learning model that fits all users in activity recognition. Due to many reasons, different individuals might perform the same activity but in different ways. A "walking" activity for one person, for instance, might seem to be "running" for another. The most accurate recognition results can be obtained if we train the learning model with the annotated data for a specific user. This assumption is invalid in a ubiquitous environment when labelled data is scarce. Therefore, continuous learning approach for tailoring the model to best fit a specific user is crucial for improving recognition accuracy. We define model personalisation as *the process of tuning a general model to represent a user's personalised way of performing different activities*.

The vast majority of activity recognition research did not consider the personalisation issue. Only few studies investigated the impact of training model on personalised data and compared it to training the model on general data collected from different users. The researchers showed the improved accuracy when deploying subject-specific data for training instead of the general model [Kwapisz et al. 2011; Weiss and Lockhart 2012]. Weiss and Lockhart [Weiss and Lockhart 2012] demonstrated the improved accuracy if a personalised model is deployed even using only a small amount of user-specific training data. Whereas Kwapisz et al. [Kwapisz et al. 2011] created models individually for each user, then deployed a personalised model for recognition.

Due to the sparsity of labelled data for a specific user, a more practical approach is investigated to tailor a general model to best fit a specific user. In [Zhao et al. 2011], the authors developed an algorithm that learns a binary decision tree model for one person from his labelled data, transfers its structure to another person, and automatically adapts its non-determinate nodes with the unlabelled samples of the new person. This accomplishes the cross-people knowledge transfer task. Pärkkä, Cluitmans, and Ermes [Pärkkä et al. 2010] proposed a similar approach based on a binary decision tree. In this method, the user's input is required to tune tree thresholds for a specific user. Moreover, it takes 3–10 minutes of new data with annotation and uses that for updating the thresholds in each node. The problem of activity recognition is more challenging with multi-dimensional data in a streaming environment. Therefore, the binary decision trees applied in the aforementioned studies are not efficient with large scale data and complicated scenarios. Gomes et al. [Gomes et al. 2012b] constructed a personalised activity recognition system that is deployed in a streaming environment. Despite the efficiency of the developed system, their model still required user specific annotated data to achieve personalisation. Similarly, Reiss and Stricker [Reiss and Stricker 2013] used a set of classifiers as a general model which is later updated with new labelled data from a specific user. In [Vo et al. 2013], authors adjusted the learning model from person A with a selected confident sample for another person B. The proposed algorithm is an integration of an SVM classifier and clustering approach for updating the model automatically. However, the proposed system has not been evaluated in a streaming setting. The deployment of activity recognition system in streaming environment imposes more challenges as the change of data distribution while a stream evolves may cause the model to drift away from the actual data distribution.

3.3.2. Model adaptation. The other dynamic feature of the recognition system concerns its ability to capture significant data changes. It is impractical to assume that there is always a static set of activities along evolving data streams. In the recognition system, the initial data represents a set of activities that is collected to train the primitive model. Then, the learning model is deployed for the actual recognition of incoming unlabelled sensory data. However, state-of-the-art approaches in activity recognition do not consider the appearance of new activities that did not exist in the initial training data [Yang et al. 2010]. New activities appear because of two common reasons. First, it is unrealistic to collect annotated sensory data for all kind of activities that exist in a domain. A typical activity recognition system contains only a few activities that are annotated with experts or with a means for assistance such as videos. Other activities may appear beyond the learning phase. Therefore, the set of activities may need to be extended later after deployment because of the scarcity of annotated data in the learning phase. Second, novel activities include also sudden activities. It is impractical to collect data representing sudden activities for training the model. An example of this kind of sudden activity is a sudden fall for an elderly person in the application of health care. Due to the difficulties of collecting annotated data for such sudden activities, other approaches have to be considered for detecting novel activities. Traditional methods solve this problem by rebuilding the entire model based on the training data of the new set of activities. This is impractical for real time activity recognition and especially in a streaming environment.

Developing a recognition system that can recognise new activities and assimilate it with existing model for further recognition is essential for real life recognition. The same concept applies for the removal of abandoned activities that are no longer relevant to a particular user. Model adaptation is a key criterion for the flexibility and accuracy of any activity recognition system with an evolving data stream that changes over time. There is no notion of adaptation/refinement of the learning models in the literature. Models do not detect activities that may emerge over a period of time (post the data collection) or changes in a user's patterns, which are both completely realistic in the context of a mobile user. The adaptation process needs to update the recognition model to recent changes in a real life user's activities in real time. Different from model personalisation that only tunes existing models, model adaptation enforces changes to core activities with incremental assimilation of new discovered activities or elimination of abandoned activities.

3.3.3. Transfer Learning. A major assumption in activity recognition is that the underlying concepts of both training and target (deployment) are the same. This assumption is invalid in many cases. Data collected for training in activity recognition may be different from data received in actual recognition in terms of distribution, domain and tasks. Transfer learning concerns "*developing systems that can leverage experience from previous tasks into improved performance in a new task which has not been encountered before*" [Cook et al. 2013]. The power of transfer learning is in the flexibility of mapping between training and deployment. Thus, we can reuse knowledge learned previously to solve a new problem faster or in a more efficient way.

We first define basic notations of transfer learning. The notations in this paper are consistent with Pan [Pan and Yang 2010] and Weiss [Weiss et al. 2016] definitions. Transfer learning adapts between training (D_T) and deployment domain (D_D). Each domain is defined by a feature space X and labels Y, where X_T, Y_T pairs represent the training space, X_D, Y_D depicts the deployment space. The marginal probability distribution of X is $P(X)$. While the predictive function learned from label pairs x_i, y_i is $f(.)$ where x_i in X and y_i in Y. The prediction function is also termed as conditional probability distribution of the domain $P(Y|X)$. Basic notations of transfer learning are illustrated in Figure 6.

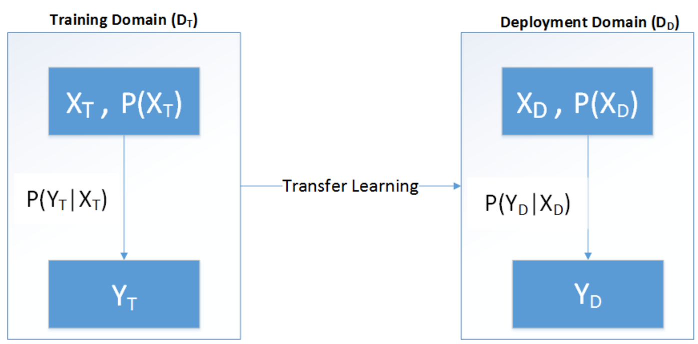


Fig. 6: Illustration of Transfer Learning

Transfer learning is categorised according to four aspects: alignment of feature spaces, distribution, what has been transferred and availability of labelled data. The first aspect focuses on the relationship between features in both training and deployment spaces. In this direction, transfer learning can be either heterogeneous [Harel and Mannor 2010] or homogenous [Chattopadhyay et al. 2012; Yao and Doretto 2010; Long et al. 2013]. In terms of notations, the case where $X_T \neq X_D$ is defined as heterogeneous transfer learning. Homogenous transfer learning is where the feature spaces are the same for both training and deployment, i.e., $X_T = X_D$. Transfer learning can also be categorised according to its aim in distribution correction. It can be either correcting for marginal distribution differences [Pan et al. 2011], conditional distribution differences [Yao and Doretto 2010] or differences in both [Chattopadhyay et al. 2012; Long et al. 2013]. Distribution correction mainly occurs in homogenous condition, where features in training and deployment domains are aligned. Heterogeneous methods, on the other hand, focuses on aligning X_T and X_D while assuming that distributions are the same. Work such as [Shi et al. 2010] addressed the distribution correction in addition to feature spaces alignment. Further studies are required to combine distribution correction methods with feature spaces alignment.

Transfer learning can also be categorised with respect to what has been transferred which includes four approaches: instance-based transfer, feature-based transfer, parameter-based transfer, and relational-based transfer.

— **Instance-based transfer** uses reweighted data in the training domain according to certain criteria into the deployment domain. Instance-based transfer performs well when features in both domains are homogenous. In [Chattopadhyay et al. 2012], for example, the weights of different sources in the training domain are computed based on the marginal distribution differences between the training and deployment domains. These weights are then changed according to the difference in the conditional distribution. The weights in [Yao and Doretto 2010] are adapted according to the performance of the algorithm in each boosting iteration.

— **Feature-based transfer** aims to adapt the feature space of the training domain to attain the best performance in the deployment domain. Feature-based transfer can be either symmetric [Pan et al. 2011] or asymmetric [Long et al. 2013]. In Symmetric feature transfer finds a common latent feature space between the two domains that improves the performance while correcting the marginal distribution in both domains. Asymmetric transfer, on the other hand, maps the features directly from the training domain through reweighting to closely match the deployment domain.

— **Parameter-based transfer** assumes common parameters across training and target domain. Therefore, prior knowledge of parameters would be transferred between training and deployment domains [Chattopadhyay et al. 2012].

— **Relational-based transfer** discovers relationships among the training data and transfers this knowledge to the target. This type is relatively new area with few work published such as [Li et al. 2012].

It is important to highlight the difference between homogeneous feature-based transfer and heterogeneous transfer. Homogenous feature based transfer is the case where the feature space is the same between the source and target, and the adaptation is performed on individual features in an attempt to align the distributions. Heterogeneous transfer is the case of feature spaces being different. In this case the adaptation between the source and target attempts to find a transformation to make both feature spaces common.

Cook [Cook et al. 2013] classified transfer learning in regards to the presence of labeled data. Transfer learning is defined as supervised, unsupervised and semi-supervised based on the availability of labeled data in the training domain. It can also be either informed or uninformed with respect to the presence or absence of labeled data in the deployment domain. Therefore, uninformed supervised transfer learning, for example, means labeled data is available only in the training domain while data in deployment domain is unlabeled. The majority of the methods are supervised, which assume the availability of data labels in the training domain. Few methods, such in [Zhu et al. 2011], are unsupervised. However, these methods are very specific to their related application, such as image classification in [Zhu et al. 2011], and are difficult to use in other applications. Transfer learning has been applied successfully to many machine learning applications such as text classification [Li et al. 2012], image classification [Zhu et al. 2011], human activity classification [Harel and Mannor 2010] and object recognition [Yao and Doretto 2010]. Figure 7 summarises transfer-learning categories.

Transfer learning is founded on an underlying assumption that the training domain and deployment domain are related. Therefore, knowledge transfer would bring benefit to the deployment domain. However, there is a situation when both domains are not related and hence transfer learning will be unsuccessful, and even worse, it might have a negative impact on the performance in some cases. This situation is often referred to as negative transfer learning. This field of research is relatively new [Yang et al. 2016]. Future research in transfer learning focuses more on investigating negative transfer and ways to measure and qualify it.

Definition related to transfer learning might vary with respect to the application domain. In the context of activity recognition, transfer-learning terms have different

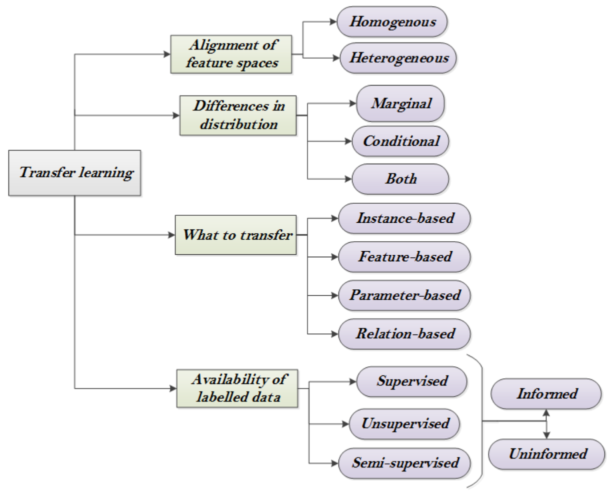


Fig. 7: Transfer Learning Approaches

domain specific meanings. This has been explained in details in cook [Cook et al. 2013]. Traditional approaches for activity recognition assume that training and deployments domains are typically identical. However, this assumption is not valid due to many differences that may exist between both domains. These differences can be because of changes in one or many aspects. We list some of the common reasons for change as follows:

— **Sensor modality [Kurz et al. 2011; Roggen et al. 2010]**: Sensors used for activity recognition can be on-body, mobile or ambient sensors. This includes changes of type, location and layout of sensors. These changes may affect data distribution, parameters and features. Thus, changes in sensor modality require instance-based, feature-based, parameter-based transfer to align both domains.
— **Time [Pan et al. 2007; Pan et al. 2011]**: Changes in activity times between training and deployment domains are very likely in activity recognition. An example of health information associated with user activities, when medical tests, treatments, immunization times for specific disease might be different between domains. Differences could be due to change in the medical plans, test times, immunization availability. Time changes will require feature and/ or parameter based transfer to map between training and deployment domain.
— **People [Hachiya et al. 2012]**: Typical activity recognition application uses different subsets of people in training and deployment spaces. Thus, different people perform the same activity or the same activity is performed in different ways. This essentially will require instance-based transfer, feature-based transfer and parameter-based transfer.
— **Sampling rate [Zhao et al. 2010]**: Change in sampling rate between domains is very common in activity recognition and requires mostly parameter transfer to align both domains.
— **Activities/labels [Zheng et al. 2008]**: Data provided in training could represent activities that are different from those of the deployment domain. The difference could be with respect to activities granularity, labels set, and/or label context. Consider an example of differences in activities granularity in smart home settings. In this example, labels of training domain might include cooking, washing the dishes, and watching, while labels of the deployment domain only include high-level activ-

ities such as walking, standing or sitting. Labels of activity could also have context bias that makes them different across domains. For example, lying activity in the hospital has different context meaning from lying in smart home context. Difference in activities/labels mostly require instance, feature based transfer and relational-knowledge transfer.

In activity recognition, these differences have a direct impact on various data aspects: distribution (marginal, conditional or both), alignment of feature spaces (Heterogeneous) and what to be transferred (Instance, features, parameters, relational-knowledge). An efficient learning algorithm will need to be able to transfer information between domains by addressing the aforementioned changes. As the number of difference increases, the less related training and deployment domains are and more effort is required to align them. Work in the literature focuses on addressing a single change between the domains, which is not realistic when multiple changes exist. Also, more attention needs to be given to negative transfer to be able to decide when it is better or worse to apply transfer learning on activity recognition data.

In the next section, basic concepts of the deployment domain of stream environment are presented to introduce the developed challenges. These challenges of activity recognition that arise from the streaming environment are described in Section 7.

4. MINING CHANGES IN DATA STREAMS

This section discusses various methods proposed to capture changes in data streams. We first differentiate between input and target domains in stream mining. The input data is received before time point t_0; while the target data is received between t_0 and t_1. Thus, the change is monitored between the two time points t_0 and t_1. In a streaming dynamic environment, changes are expected to occur between the input and target data. These changes might occur once or many times, gradually or suddenly [Gama et al. 2014]. Figure 8 represents an example of the change in data distribution between the input domain and target domain in 1-D data.

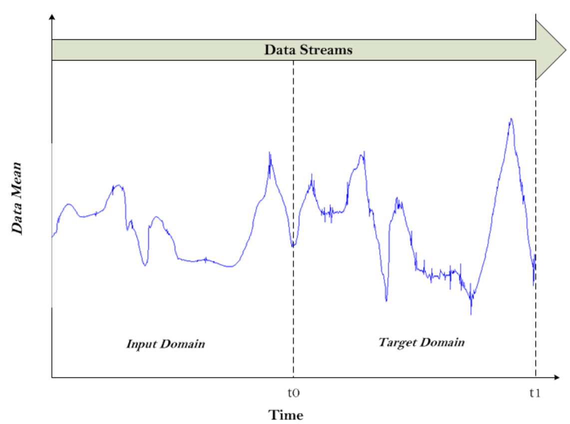


Fig. 8: Example of Change in Data Streams

Various types of change could be detected in streams. Concept drift is one of them that refers to the change in the distribution that occurs while the stream evolves. Another type is concept evolution which refers to the appearance of a new concept in the target domain that did not exist in the input domain. An extension of concept evolution is concept forgetting. Opposite to detecting novel concepts, concepts that are no longer relevant in the target domain require a forgetting technique for adapting the model to the recent changes in the stream. Outliers are considered as a rare change of the stream. Unlike concept evolution, outliers are not incorporated into the system or added to the model for enhancement.The goal of detecting outliers is to isolate irrelevant data and filter them out from real data. Noise can also be considered as a special case of outliers. For various kinds of change, adaptive learning techniques are applied in order to update the learner and cope with the evolution of data. Figure 9 illustrates categories of changes.

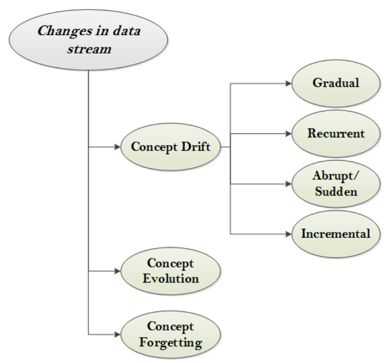


Fig. 9: Categories of Changes in Data Streams

Adaptive learning in stream mining can be categorised into two approaches. The first approach aims to update underlying concepts to cope with the most recent changes in data streams. In this approach, changes are not explicitly detected, yet the learner is updated periodically to accommodate for the expected changes in the data streams. While the other approach aims to observe data streams in order to detect the changes and then adapt the model upon the identified changes. In the following, we illustrate different types of change in data streams. Then, we discuss adaptive learning techniques that are applied for model update.

4.1. Tracking changes

Concept drift.

Figure 10(a) represents the type of change known as *concept drift* which is first identified in [Schlimmer and Granger Jr 1986]. It refers to the change in distribution between input and target domains. Klinkenberg and Renz [Klinkenberg and Renz 1998] specified three indicators for concept drift. The first one is based on the classifier performance metrics such as the accuracy of the classifier. While the second indicator is based on model properties such as model complexity. The last one concerns the change of data properties, i.e. data distribution. Examples of techniques that detect concept drift based on performance indicators are the FLORA family of algorithms [Widmer and Kubat 1996]. A FLORA algorithm monitors the accuracy and the coverage of the model of a rule based classifier. The algorithm adapts the window size dynamically according to the measured performance metric. This approach requires true labels provided by the user in order to measure the accuracy. Indeed, this input is impractical in streaming settings when data arrives at high speed and requires real time adaptation. Other relevant techniques in [Gama et al. 2004; Bouchachia 2011] have applied statistical evaluation to monitor the performance and adapt accordingly. Hulten, Spencer, and Domingos [Hulten et al. 2001] presented a system for concept adapting very fast decision trees - CVFDT. The adaptive Hoeffding tree monitors the quality of the previous model and adapts the model in terms of the splitting features in the tree. While Gaber and Yu [Gaber and Yu 2006] presents a STREAM-DETECT technique that capture change in data streams by monitoring data distribution using an online clustering deviation method.

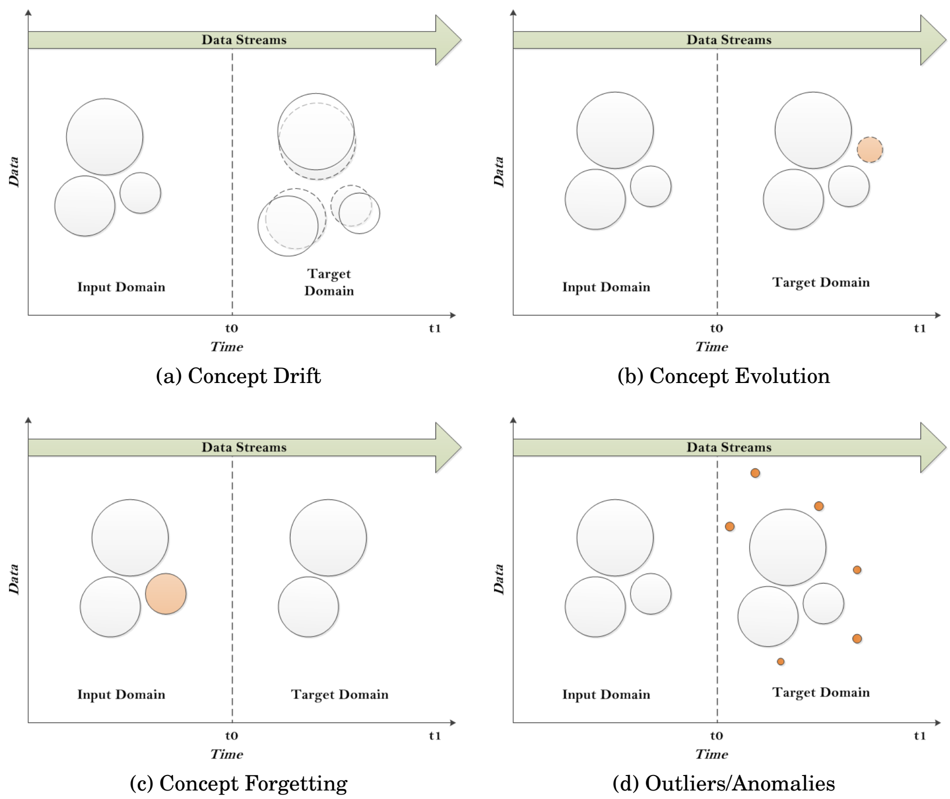


Fig. 10: Changes in Data Streams

The change is monitored between two time points t_0 and t_1. Detecting concept drift occurs over either a fixed window size, an adaptive window or combination of both. One major problem with the fixed window is the selection of the window size. VFDT [Domingos and Hulten 2000] and CVFDT [Hulten et al. 2001], for instance, applied a sliding window of fixed size to handle data streams. A small size of window in stable data causes exhaustion of resources and increases the probability of false alarms. While a long time window in fluctuating data might miss a crucial information and result in slow reaction to the encountered changes. Unlike the fixed window approach, an adaptive window is adjusted according to the detected changes. Thus, window size either expands or shrinks upon the change indicator. Widmer and Kubat [Widmer and Kubat 1996], for instance, applied a sliding window of adaptive size to monitor the incoming data. The assumption in this approach is that the most recent data is the most important data. Therefore, the model is updated based on the most recent data appearing in the most recent window. Klinkenberg and Joachims [Klinkenberg and Joachims 2000] proposed an SVM model that compares the error on various window sizes, then selects the size with minimum error. The third approach is based on the combination of two different window sizes; one is fixed and the other is adaptive. The fixed window stores the baseline historical information, while the sliding window captures the incoming data streams. Where statistical measures change between the two windows, the change is detected [Dries and Rückert 2009; Adä and Berthold 2013]. Bifet and Gavalda [Bifet and Gavaldà 2006; 2007] proposed an adaptive sliding window technique - ADWIN and ADWIN2 - that maintains the size of windows according to the rate of change. The main drawback of applying the combination of both windows is the resource constrains. Figure 11 presents the taxonomy of concept drift approaches.

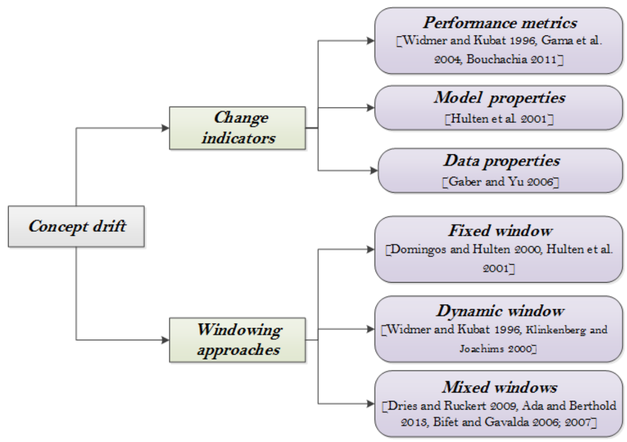
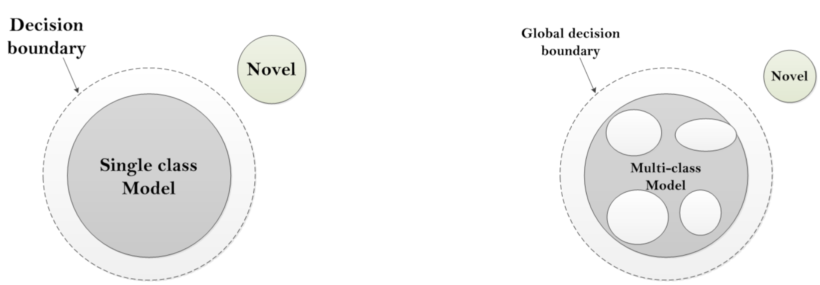


Fig. 11: Summary of Concept Drift Approaches

Concept evolution. The other type of change is *concept evolution*, which refers to the appearance of novel concept in the stream. Figure 10(b) represents an illustration of concept evolution. Detecting a novel concept is a challenging task in data streams, especially when dealing with data streams with concept drift. An efficient approach has to be able to distinguish between the drifting of an existing concept and the appearance of an entirely new concept. The appearance of a new concept is followed later in the stream by detection of the recurring data instances that belong to the novel concept. Concept evolution techniques aim to capture the arrival of novel concepts and incorporate the detected novel concept into the existing underlying concept. This integration allows further detection of recurring instances that belong to the novel concept. Different approaches have been developed to distinguish between existing and novel concepts.

In terms of the underlying concept, some approaches identified the underlying existing concept as a single model that represents only a single class, while others considered a multi-class underlying concept. Spinosa, Carvalho, and Gama [Spinosa et al. 2007] represented the underlying model by a single concept with all incoming data either part of the underlying model or novel concept, as explained in Figure 12(a). This approach assumes that there is only one “normal” class and any other classes are novel. The cluster based system, named OLINDDA, is based on three models: normal profile model, concepts that extend the normal profile, and novel concepts. Learning phases in OLINDDA are offline and online. The normal model is built in the offline phase, while the extension in the online phase detects minor changes in the normal model. A novel concept is detected when incoming data is located away from the normal model and also satisfies a specific validation criterion. A single model approach limits the capabilities of a concept evolution algorithm to differentiate only between existing and novel concepts without considering the presence of multiple existing concepts. Thus, other approaches have been developed to address the multi-class structure of an underlying concept, illustrated in Figure 12(b). The techniques developed in [Masud et al. 2011; Faria et al. 2013; Hayat and Hashemi. 2010] are examples of stream learning approaches for novelty detection with multi-class underlying concept. ECSMiner [Masud et al. 2011] applies an ensemble of classifiers to an incoming stream of equally sized data chunks for prediction. The global decision boundary is defined as the union of local decision boundaries for existing classes in the underlying model. The classifier model in the ensemble classifier is dynamically updated to detect instances that are outside the global boundary. If no classifier is able to predict the incoming data, then data is stored in short memory for further processing. The novel concepts are detected when data in a buffer maintains cohesion with other buffer data and separation from existing underlying concepts. This approach addresses the novelty detection in multi-class underlying concepts, yet it requires all data chunks to be labelled to define the new concept. Faria, Gama, and Carvalho [Faria et al. 2013] proposed the MINAS system for concept evolution which applies unsupervised learning approaches. MINAS classifies new incoming instances as known or unknown. Unknown instances are the ones located outside the decision boundaries. The declared unknown instances are stored in short time memory. Then, data in short memory is clustered in order to discover new concepts. Hayat and Hashemi [Hayat and Hashemi. 2010] proposed an approach that is based on the discrete cosine transform to build normal concepts of multi-classes with sub-clusters. The distance measure is also applied to distinguish between existing and novel concepts. The aforementioned approaches rely mainly on the distance measures that predict novel concept based on its location from the decision boundaries. The new concept is declared as novel if it is outside the global decision boundary which is the union of local boundaries of clusters. Although the underlying concept contains multi classes, creating a global decision boundary results in a similarity between multi-class models and single class models. It combines all existing concepts in one concept and defines the global boundary for the combined concepts. Thus, these approaches also did not address the appearance of novel concepts that might exist outside the local boundaries, yet inside the global boundary. Moreover, most of these approaches did not consider the labelling cost of data streams when identifying novel concepts.



(a) Single Class Model (b) Multi-class Model

Fig. 12: Underlying Concept Approaches for Concept Evolution

Concept evolution approaches are also categorised based on the action taken upon novelty detection. Yeung and Chow [Yeung and Chow 2002] and Yang et al. [Yang et al. 2002] presented instance based algorithms that consider whether the incoming data is sufficiently close or far from the underlying concepts based on some appropriate metric. Novelty detection in these techniques detect “filtered out” instances without tracking data for the detection of normal concepts. While approaches in [Faria et al. 2013; Hayat and Hashemi. 2010; Masud et al. 2011; Spinosa et al. 2007] extend the detection of “filtered out” instances by detecting the level of cohesion among these instances to form a novel concept. Studies in [Masud et al. 2011; Al-Khateeb et al. 2012] further integrated the novel concept with underlying ones in order to detect recurring instances that belong to novel concept. Figure 13 summarises various approaches in concept evolution.

The idea of concept evolution is also attached to concept forgetting. Dynamic environments with non-stationary distributions require the forgetfulness of the observations not consistent with the actual behaviour of the nature [Gama et al. 2010]. Both concept drift and concept evolution adapt the model to newer information, while concept forgetting abandons old information. Forgetting mechanisms are necessarily with data streams in order to preserve resources (i.e., memory) while dealing with the most recent observations for better accuracy. The system performance relies on its capabilities to learn changes and new concepts appear in the stream as well as forgetting outdated concepts that became a burden on the system [Kifer et al. 2004]. Concept forgetting, in this direction, is concerned about outdated concepts instead of observations. It might use the same forgetting mechanisms, applied for observations, yet at concept level. Whereas incremental learning manages the continuous learning of new concepts, decremental learning focuses on forgetting abandoned concepts[Cauwenberghs and Poggio 2001]. Figure 10(c) explains the concept forgetting in data streams.

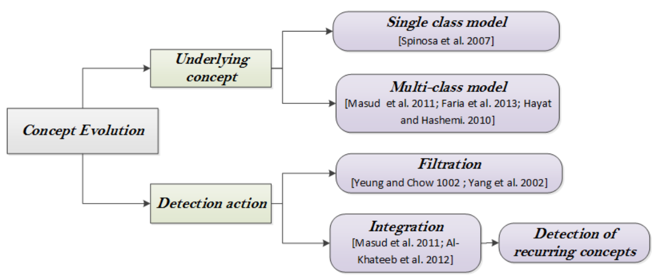


Fig. 13: Summary of Concept Evolution Approaches

Outliers. The other category of change that is explained in Figure 10(d) is *outliers/anomalies*. Many of the outliers are considered as noise, while others are of paramount importance such as credit card frauds. Concept evolution in data streams is closely related to the outlier detection. Outliers are defined as data instances which deviate from underlying concepts. However, this definition also applies for novel concepts. The main characteristic that differentiates between outliers and concept evolution is the cohesiveness among “filtered out” instances. The novel concept has to satisfy a validation criterion that concerns mainly the separation of short memory instances from underlying concepts as well as cohesion among these instances to form a novel concept. In streaming settings, outlier detection under realistic assumptions is an unsupervised learning approach because they are mostly rare to occur. Therefore it is not possible to train the model on them beforehand. Different metrics have been developed in literature to measure the deviation of incoming data from underlying concept that define anomalies. The deviation measure such as distance and distribution have been applied pervasively for outlier detection. There are many techniques that studied outlier detections in data streams such as [Angiulli and Fassetti 2007; Assent et al. 2012; Niennattrakul et al. 2010; Pokrajac et al. 2007]. Few studies have combined the detection of outliers with concept evolution such as in [Masud et al. 2011; Masud et al. 2013; Al-Khateeb et al. 2012]. A detailed survey of outlier detection approaches is presented in [Aggarwal 2013]

In a nutshell, evolving data streams may have encountered many kinds of change. That includes concept drift, concept evolution, and concept forgetting. Concept drift could be gradual, sudden, incremental, recurrent. In this section, we discussed various kinds of changes and different approaches to detect each. The main target of these approaches is to spot and identify the change. In the following section, we review techniques to learn from rapidly changing data streams. Thus, we study the problem of learning in an evolving streaming environment when changes occur.

4.2. Learning from changes

Adaptive learning. The evolving nature of data streams necessities the need of a learning approach that is capable of accommodating the anticipated changes. Continuous learning in evolving data streams refers to the well-known term of *incremental learning* or *adaptive learning*. There are two categories of adaptive learning; blind/implicit learning or informed/explicit learning [Gama et al. 2010]. Techniques for blind learning update the learner periodically without prior knowledge of the encountered changes. The adaptation occurs at fixed time intervals independent of the kind of change. Unlike the blind approach, explicit/informed learning techniques are triggered when change is detected. Therefore, informative learning based on the kind of change is performed whenever change occurs.

An adaptation process in the blind approach is incremental without prior explicit knowledge of the change itself. VFDT [Domingos and Hulten 2000] is a typical example of blind learning where the decision tree leaves are updated periodically according to a loss function. In contrast, informed learning requires explicit knowledge about change to trigger the adaptation of the learner. Masud et al. [Masud et al. 2011], for instance, applied an ensemble classifier for detecting concept evolution in data streams with concept drift. The proposed system incorporates a new learnt model into a learner for future prediction of recurring instances.

Blind learning has the advantage of the periodic update without relying on the detection of change and its corresponding performance. However, there are also some limitations with this kind of implicit learning. The adaptation may take different forms based on the kind of change. For example, the action required for tuning the learner in case of concept drift is different from the action required when a new concept has emerged. Moreover, performing updates based on fixed time intervals strongly relies on both the interval length and the rate of change. The response to change might be slow if the interval size is big especially in data streams with high rates of change. On the other hand, unnecessary updates might cause over use of resources in short time intervals with stable data streams. Thus, the trade-off between the cost of updates and the gain in performance has to be considered for choosing an adequate interval length [Gama 2012]. On the other hand, informed/explicit learning only responds to the detected changes. Based on the change detected, different actions are taken to respond to the specified changes. The main limitation of this approach is its dependency on the change detection. Learning only occurs when change is detected and is tightly related to the detection performance. Either an incapability of detecting changes or a high rate of false alarm in the detection technique might mislead the learning process. For example, whenever the change detection fails to detect the appearance of new concept, the new concept and its recurring instances will be either misclassified or classified as *unknown*.

Adaptive learning is also categorised based on the adaptation response. The adaptation reaction is either revolutionary or progressive. In the revolutionary approach, the learner is retrained and reconstructed with the most recent important data. The existing model is discarded and totally replaced with a new model from new data [Gama et al. 2014; Street and Kim 2001; Gama et al. 2004]. Alternatively, new data adapts existing models incrementally in the progressive approach in order to react to the changes in the stream. An example is the OLINDDA system [Spinosa et al. 2007] that applies online and incremental learning for extensions of existing concepts or discovering novel concepts.

Figure 14 summarises different approaches for adaptive learning.

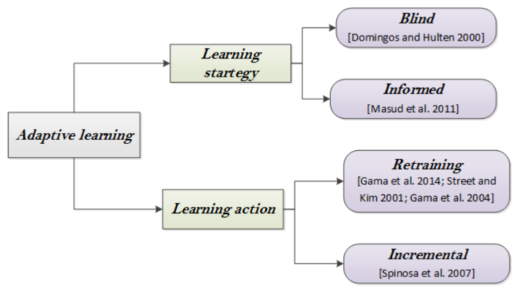


Fig. 14: Summary of Adaptive Learning Approaches

A special category of incremental learning is active learning. Active learning focuses on labelling only few data instances in order to enhance the learner accuracy. In data stream settings, it is impractical to assume the availability of labels while a stream evolves. Many approaches have been proposed for efficient active learning techniques in data streams, such as in [Widyantoro and Yen 2005; Masud et al. 2012; Huang and Dong 2007]. Studies in [Huang and Dong 2007; Fan et al. 2004; Lindstrom et al. 2013] developed an active learning approach based on the change detection method. When the system detects change in a data streams, that triggers active learning to inquire for true labels. Other approaches built on an assumption that part of the data streams is labelled while the other part is unlabelled [Li et al. 2012]. Recent work in [Zliobaite et al. 2014] presents a framework for active learning in data streams with concept drift. The developed system learns in batches while saving the labelling cost. The main target of active learning is to keep the trade-off between minimising the cost of labelling while maximising the gain in performance.

All in all, we reviewed state-of-the-art-techniques in data stream mining. Approaches for data stream mining focuses mainly on the non-identical distribution feature of data. Techniques for data stream mining include handling data streams, learning from data streams and capturing change in evolving data streams. Incoming data in activity recognition typically streams from sensors. In the following section, we discuss the intersection between stream mining and activity recognition areas of research. That also includes the mapping of terminologies between the two fields of study.

5. ACTIVITY RECOGNITION TECHNIQUES USING DATA STREAM MINING

This section focuses on the intersection of the two fields of research: stream mining and activity recognition. State of the art activity recognition techniques address the dependency nature of activity data, while approaches in stream mining focus on the non-identically distributed streaming data. Traditional activity recognition approaches are based on the assumption that the processed data is stationary. This assumption is violated when dealing with real time sensory data that is typically evolving over time. Data streams approaches deal with non-identically distributed data and also constraints imposed in the streaming nature, e.g. infinite number, high speed, concept drift etc. However, most of the stream mining approaches have not been used to address activity recognition. Activity recognition in data streams approach is an overlapping between stream mining and activity recognition when both *i.i.d.* conditions of independent and identically distributed are violated. The developed techniques for activity recognition in data streams have to be adaptable and flexible, to accommodate for the evolving nature of activities.

Some approaches in data stream mining might be seen as applicable for activity recognition [Fong et al. 2016; Žliobaitė et al. 2016]. Particularly, mining sequential data streams could be considered for activity recognition. This approach deals with sequential data streams as ordered chunks with or without a notion of time. Although techniques in this approach consider the dependency in data, two major core differences make these techniques incompatible with activity recognition. According to the previously explained approaches in Section 4, most techniques in sequential data processing aim to find interesting/frequent patterns instead of mining the sequential patterns in order to predict classes of incoming data. This is different from the main target of activity recognition which is the prediction of incoming activities with sub-targets that might also be included in some cases for finding the interesting patterns. In data stream mining, few studies have addressed the actual mining of sequential patterns in data streams. Yet, these techniques assume that the sequential data is in a transactional form which assumes that the stream is segmented and each segment is assigned to one well-defined class. This assumption is not valid for activity recognition when a stream of data representing a sequence of performed activities arrives with no boundaries in-between. Activities also can be interleaving, such se pauses while walking or recurrent, such as running while bouncing the ball in a basketball court. Although sequential stream mining is conceptually related to activity recognition, its techniques are not applicable for activity recognition because of the aforementioned core differences in terms of the learning target and domain definition. It is also important to emphasis the relationship between time series as a special type of data streams that has a time notation and activity recognition.

5.1. Time Series and Activity Recognition

Most techniques that are developed for time series focus on handling/simplifying time series representation (such as SAX) or aligning different time series in order to find similarities (such as DTW). Special characteristics of data streams, such as concept drift, have attracted less attention in the time series community. A variety of time series methods are applied for activity recognition. Some of these methods focus on transforming raw time series into an efficient representation such as [Wanigatunga et al. 2016] where authors applied symbolic aggregate approximation (SAX) to accelerometer data for activity pattern visualisation. In [Liu et al. 2015], the authors used a feature vector of shapelet, which is computationally expensive, hence cannot be applied in streaming environment. Converting raw time series into an efficient presentation is followed by applying a similarity measure for classification. K- nearest neighbour (KNN) is commonly used for time series classification. KNN shows effective performance in time series in general especially when using dynamic time warping (DTW). [Seto et al. 2015] applied a modified DTW for activity recognition to avoid feature extraction and domain knowledge. However, the model is overfitted and handles only atomic activities. [Chen et al. 2013] proposed a new DTW measure, DTW-D, that targets Semi supervised learning for many applications, including activity recognition. Subwindow Ensemble Model [Zheng et al. 2013] used an ensemble of classifiers trained on features made up of coarse summary statistics computed from different temporal scales. Kogeh [Hu et al. 2013] built an alignment-free time series classification framework that requires only weakly-labeled data. The framework applies a threshold-based approach to distinguish between patterns inside a single time series. It also reduces the computation of 1-NN by constructing data dictionary. The framework showed fast and accurate performance compared to traditional time series techniques for seven different applications including activity recognition.

Applying time series classification for activity recognition might encounter a number of challenges which are:

— **Segmentation:** Most literature on time series classification assumes that the beginning and ending points of the pattern of interest can be correctly identified [Hu et al. 2013]. This assumption is unrealistic for activity recognition in data streams when data arrives as a continuous sequence of multi-layered of concurrent and overlapped activities. Segmentation of a time series into a sequence of physical activity types rather than classifying an entire time series as a single activity type is essential. The quality of segmentation directly influences the recognition results.
— **Complex activities:** Activity patterns are complex, repetitive and represented with different lengths. Time series techniques work well when the task is to match the overall shape of a time series. However they perform poorly on activities with repetitive patterns [Zheng et al. 2013]. The basic machine-learning task involves classifying a time series as a single physical activity type [Zheng et al. 2013]. Using time series to classify complex activities could be challenging. Moreover, data used for time series are arranged to be of same length. For example, in the world's largest collection of time series datasets, the UCR classification archive, all datasets contain only equal-length data [Chen et al. 2015]. This assumption is not applicable for activity recognition.
— **Personalisation:** Time series representing activities can change from one person to another. Therefore, a classified time series activity can belong to more than one well-defined class. Some people tend to have pauses of standing while walking; others may walk as fast as jogging. The current solution to preprocess such data requires human intervention in order to keep only data that represent the activity of interest [Hu et al. 2013].
— **New activities:** In real-life streaming data, new activities that have not been trained with the classification model might evolve over time. Detecting new activities and assimilate them into the classification model will enhance the classification accuracy. We found no reference in the literature of using time series techniques for detecting new activities.

Real time activity recognition might also be related to activity recognition in data streams. Techniques for real time recognition focus mainly on enhancing the response time by simplifying both data processing and model structure. In an early study, Karantonis et al. [Karantonis et al. 2006] implemented a real time classification to distinguish between activity and rest, before classifying further lower level activities. The classification technique is based on a hierarchical binary structure for broad classification at the top level. Sub-classification occurs at the lower levels. The resource aware techniques are performed onboard sensors and target real time recognition. CeneMe [Miluzzo et al. 2008] represents a system for real time recognition of contextual activity using off-the-shelf, sensor-enabled mobile phones. The algorithms used by the

CenceMe classifiers run on the phone and the backend server according to the split-level classification design. The backend server classification presents a higher level of contextual recognition of activities. The activity-recognition algorithm presented in [Pärkkä et al. 2010] is based on a binary decision tree classifier to automatically recognise physical activities on a portable device. The labelled data is provided to evaluate the system performance and update the decision tree threshold values with the user's own data.

Although the response time of these techniques seems to resemble one for stream mining techniques, none of the other performance characteristics of data streams have been addressed in these techniques. Particularly, they have no notion for handling infinite, high speed, and mostly unlabelled data streams. Moreover, these approaches lack the consideration of concept drift, outliers, or concept evolution that are known phenomena in data streams. These approaches are not flexible for personalisation and adaptation with the evolving activities.

Some terms in activity recognition have their corresponding meanings in stream mining, yet in different settings and with different contexts. A typical example of this is *outlier detection*. Hawkins [Hawkins 1980] defined an outlier in general as: "An outlier is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism". Outliers in data stream mining correspond to abrupt changes in data streams from the underlying concepts. Outliers can be referred to as noise, anomalies, and abnormalities. The term of outlier detection in activity recognition refers to *sudden activity* such as *fall detection*. Detection of sudden activities resemble outlier detection as they both aim to find the unusual events in data. The term of "sudden activity" or "fall detection" is pervasively discussed in activity recognition, especially for elderly people aids [Luštrek and Kaluža 2009; Mubashir et al. 2013; Bakar et al. 2016].

Novelties are considered as a special category of outliers, yet with aggregated data points [Aggarwal 2013]. However, in this survey we suggest a separation between the two concepts, outliers and novelties. By definition, outlier is an unusual event that is completely different from the underlying/known concepts. It is mostly expressed as a set of individual data points that correspond to noise, sensor failure, etc. Outliers cannot be aggregated under a single concept as they have different natures and different causes. On the other hand, novelty data represent novel concept which still have some commonality with existing underlying concepts. Though the novel concept is different from the underlying concept, they still preserve some common similarities. Examples of outliers are noise, credit card fraud, or sudden falls in activity recognition which are entirely different from the underlying concepts. While novelties are about discovering new "normal" concepts such as discovering a new category of credit card transaction or a new activity that the user started performing recently. There are two main differences between outliers and novelties. First, novelties represent novel concepts that have not been seen by the system before. Yet, novel concepts are neither completely different nor abnormal from the underlying concepts. Thus, new concepts might appear in the middle of existing ones. Moreover, novel concepts are represented by a set of aggregated instances, while outliers instances are mostly irregular that appear separately and most of the time with no relation to other detected outliers.

The term "novelties" is also referred to as concept evolution and novelty detection. In stream mining approaches, *concept evolution*, as introduced earlier in Section 4.1, is the process of monitoring the data streams in order to discover the appearance of novel concepts. The term of concept evolution is analogous to *detecting novel activity* in activity recognition. In contrast to concept evolution, some concepts became outdated and no longer relevant to the target domain. These concepts require an adaptive mechanism to forget the abandoned concepts. The term of *concept forgetting* has been presented for data streams in [Gama et al. 2010]. Concept forgetting is also relevant to activity recognition whereas activities are no longer performed by users. *Forgetting abandoned activities* aims to update the model continuously to reflect the most recent changes in data and remove outdated/abandoned activities.

Another term that is well-known in stream mining is *concept drift*. It primarily refers to the change in data distribution while a stream evolves. This change could be gradual or sudden that happens once or recurring. A typical data stream evolves over time. Thus it requires an effective approach to handle the drift and accommodate for the most recent changes in the stream. In activity recognition, the definition of concept drift is relevant, yet in a different context. The deviation between input and target domains in an activity stream occurs as activities are performed in a different way from one user to another. When the recognition model (target) is different from the incoming data (input), this deviation resembles concept drift for activity recognition. More precisely, approaches addressing concept drift in activity recognition are termed as *model personalisation*. The personalisation process targets the adaptation of the recognition model to fit data for a particular user.

The process of learning from aforementioned changes is done with *adaptive learning* in data streams which include incremental and active learning. The anatomy of adaptive learning in data streams is *model adaptation* in activity recognition. The goal of adaptive learning in data streams is the same of model adaptation in activity recognition, which is continuous learning to adapt to the most recent changes in the evolving data. The concept of model adaptation in activity recognition applies also to incremental and active learning.

Some of the issues that have been addressed in both stream mining and activity recognition, yet under different settings and with resembling meanings, are summarised in Table II. The table shows a subset of stream mining terminologies and its corresponding related terms in activity recognition.

Table II: Terminology Mapping between Stream Mining and Activity Recognition

Stream mining	**Meaning**	**Activity recognition**	**Meaning**
Learning from concept drift	*The detection and response of change in a data stream*	Model personalisation	*The tuning of the model to suit a personal way of performing activities*
Concept evolution	*The discovery of new concepts in the stream*	Detecting novel activity	*The discovery of new activities*
Outlier detection	*Detection of abnormal instances in data stream*	Sudden activity detection	*The detection of sudden changes in activity data*
Concept forgetting	*The decremental learning of outdated concepts*	Forgetting abandoned activity	*The decremental learning of abandoned activities*

Few studies have considered activity recognition in streaming environments. Krishnan and Cook [Krishnan and Cook 2014] developed an efficient technique for handling streaming data based on windowing technique. This system is based on the fact that different activities can be characterised by different window lengths. Sensors deployed in this study are binary motion sensors installed in a smart home environment. Another study that also deploys binary sensors in a smart home environment is presented in [Rashidi and Cook 2010]. In this study, authors applied a tilted time window to find sequential patterns from streaming data using multiple time granularity. The technique adapts window size, not the classifier model, for boosting the recognition accuracy. Unlike the aforementioned techniques, Gomes et al. [Gomes et al. 2012b; 2012a] have developed an on-board data stream mining technique for mobile activity recognition. The developed system predicts activities in the stream and adapts the model to fit a user's profile. Do et al. [Do et al. 2013] built a logic based framework for recognising basic and complex activities from mobile sensors. Recently, Lockhart and Weiss [Lockhart and Weiss. 2014] have presented the Actitracker system for mobile activity recognition. Actitracker builds a general/universal classifier that could be replaced by a personalised model for a particular user. The system collects data with fixed time windows and transmits data for backend server for processing. More details of these techniques and other state-of-the-art activity recognition techniques along with the research gaps are discussed in the following sections. Based on the aforementioned discussion of stream mining and activity recognition, the next section discusses in depth key techniques in activity recognition, Section 6, followed by analysis and discussion of research gaps, Section 7, and future direction in this area of research, Section 8.

6. KEY TECHNIQUES

In this section, we review key techniques in activity recognition. Krishnan and Cook [Krishnan and Cook 2014] proposed a sliding window based approach for recognising streaming activities from motion sensors events in a smart home environment. The learning classifier applies Support Vector Machine (SVM) for modelling activities. The main focus of this study is on the techniques for handling data streams. The study evaluated different windowing techniques for analysing a stream of activities. The performance is evaluated on a fixed size window with both time based weighting and mutual information weighting. It also includes the classification probabilities of activities that are previously recognised in the preceding window. The developed technique handles the 'other' class activity that does not correspond to any known activity classes. It explicitly trains the model offline for transitional activities and other unknown activities. Then, it gathers them all in one 'other' activity class and incorporates it with the model. While this study has been added to the field of activity recognition, research is still required to address other research gaps.

The approach addressed the important research of activity recognition in streaming settings. However, the developed approach did not allow personalisation or adaptation with the evolving streams. Thus, the classifier is built with training data, with no flexibility to be adapted and personalised post the deployment. The recognition of 'other activity' is not for detecting the emergence of a novel activity in the stream. The 'other activity' category adds more activities corresponding to transitional and unknown activities to the offline model during the training. Moreover, incorporating the 'other' activity into the model causes more confusion in recognising both known and unknown activities. Also, the system takes an average of 4 days to learn the different activity models for each technique. Moreover, a smart home environment requires installation of sensors in fixed laboratory settings which limits the application of activity recognition.

Another activity recognition system in smart home settings is presented in [Rashidi and Cook 2010]. Rashidi and Cook developed a system that recognises activities from unbounded input data: a stream rather than a transactional format. The aim of this study is to find the sequential patterns in the stream of data of a smart home environment. Sensors deployed for this study are also binary motion sensors. The proposed system applies a tilted window technique for handling data collection. The tilted window with an approximation approach applies a relaxed threshold in order to find the sequential patterns in the stream. The tilted window stores the records in order of time; the oldest records are kept at the highest granularity. For mining patterns in the stream, an extension of DVSM [Rashidi et al. 2011] method has been proposed. After mining the sequenced patterns, the system updates the tilted time window with the most recent patterns. The system combines sequential processing and data stream mining for activity recognition. Nevertheless, the proposed technique adapts only the tilted window rather than updating the recognition model. The main focus of this system is handling the stream with the adequate tilted window for recognising interesting patterns. Explicit recognition of sequenced activities in a data stream is not considered in this study. Furthermore, no techniques have been adapted for dealing with concept drift, concept evolution, and concept forgetting in data streams. Stream mining from smart home binary sensors with less constrained devices is less challenging than mobile devices with limited resources that require online learning and recognition in real time.

MARS [Gomes et al. 2012b; 2012a] stands for Mobile Activity Recognition System. It is an incremental system for predicting activities in data streams on the mobile device. The learning process in MARS is divided into two phases: training and recognition. In the training phase, a user performs activities and annotates them interactively while data is collected from mobile sensors. The collected annotated data is saved for building the model offline. In the recognition phase, the incoming unlabelled data is then classified based on the offline built model. The study compared the results of both static decision tree and incremental Naive Bayes for evaluating system performance. The proposed algorithms are light-weight thus can be deployed on mobile devices. Although, MARS has presented an early system that combines activity recognition with stream mining, some challenges still need to be addressed. The system assumes the availability of labelled data for each user. Each individual has to collect and annotate data representing the personalised activities for building the model. When new subject uses the system, the model has to be retrained with the data collected and annotated for this particular user. This assumption is impractical in streaming settings where the majority of data is unlabelled. Moreover, retraining the entire model for recognising the user specific activities is a time consuming process that may not be applicable in streaming settings. The developed system also lacks the description of techniques for handling the streaming nature of data. Online and incremental adaptation of the model to include entirely new activities or forget abandoned activities has not been considered in this study.

The system presented in [Do et al. 2013] applies logic based stream reasoning combined with an Artificial Neural Network (ANN) for recognising both complex and basic activities. The system comprises four components: Client, WebServer, data stream manager system (DSMS), and reasoning server. The client collects data from mobile sensors and trains the Artificial Neural Network for recognising basic activities. The collected data from a user's mobile phone (client) is transmitted to the WebServer which uses the GPS data to recognise the user's location. The WebServer sends all data to the server. It also facilitates the connection between the client and backend server. DSMS collects and stores data for the reasoning server component. The recognition of complex activities and ambiguity reasoning occur at the reasoning server component. The system tracks complex activities for suggesting a healthier user lifestyle. Despite the efficiency of the developed system, personalisation and adaptation are not considered in this study. Also, the reasoning on a backend server requires additional connectivity requirements and does not preserve a user's data privacy.

Actitracker [Lockhart and Weiss. 2014] is a recent system that is being developed for activity recognition application in the health domain. A universal model is built initially from general and impersonal data. Then, the user has the options of either to train the model with the personalised data or instead deploy the universal model. For model personalisation, the system retrains the model with sufficient amount of personalised annotated data collected for a particular user. Streaming data is collected from mobile sensors with a fixed size window of 10 seconds. The collected data is sent to a backend server for processing and classification. Various classification algorithms are applied: decision trees (C4.5 and Random Forest), instance-based learning (KNN), neural networks, Naive Bayes and Logistic Regression (LR). The Actitracker system only uses Random Forest models (which were shown to perform very well) [Weiss et al. 2014; act 2012]. Data is classified with a personalised random forest if available; with a universal random forest model otherwise. The developed system is available on the application store and ready to download for mobile devices. The data is shared with activity recognition researchers via publicly available datasets [Lockhart and Weiss. 2014]. This system uses a fixed window size technique to handle data streams. However, the personalised model is neither automated nor incremental. A collection of personalised activities requires not only annotation for each activity but also retraining the entire model to suit the personalised user. Training a model in such way is impractical in streaming settings that require automated and incremental approaches for 'adjusting' the model to fit a particular user. Moreover, the annotation process is time consuming, erroneous, and not applicable for a streaming environment. Therefore, selecting only the most profitable data with the minimum effect on performance is crucial for effective recognition. The system also lacks the consideration of discovering new activities or deleting irrelevant activities.

The technique developed in [Nguyen et al. 2015] aims to detect novel activities with limited labelled data. It requires model training on the available data that represents the new concepts. The technique uses a combination of feature-based and attribute-based learning to leverage the relationship between existing and new activities. The approach is also extended with adding a random sampling to address the imbalanced data problem. Three public datasets are used to evaluate the effectiveness of the technique in recognising new activities. This work is a pioneer in addressing the very important challenge of recognising novel activities. However, it has many limitations that require further development. The technique still requires labelled data to train the model beforehand on a new activity. Also, it lacks the capability to handle streaming data. The dynamics of the recognition includes adding new activities but not deleting abandoned ones.

The technique developed in [Abdallah et al. 2015] is an extension of [Abdallah et al. 2012] that addresses activity recognition in data streams. The system builds a flexible and lightweight framework that can be refined with the evolving activities along the stream. The technique addresses the personalisation of activity recognition system to fit the user personalised way of performing activities. The stream is captured by a fixed-size sliding window whilst detecting concurrent activities occurring in a single window. The developed technique applies incremental learning to refine the recognition model continuously with recent changes in the evolving activities. It also applies active learning for addressing the scarcity of annotated data by labelling only a small amount of data. The technique contains three components of modelling, recognition and personalisation across two phases of offline and online learning. The modelling component builds the initial base-line framework from historical data. The recognition component integrates the baseline framework with an ensemble classifier for recognising the incoming data stream. While the personalisation component updates the baseline framework continuously through batch active and incremental learning. The technique provides a lightweight solution that can be implemented on a mobile device for real time recognition. Although the developed technique addresses many research challenges, it still requires further extensions for model adaptation to add and remove activities while the stream evolves.

Some key techniques have been applied for personalised activity recognition yet in static environments. Although having an improved accuracy with personalised models, these techniques are incapable of handling the evolving nature of the incoming data streams. Moreover, the prediction relies on the static model built offline. This model cannot be pruned or expanded after deployment. We discuss in the following two key systems for personalised activity recognition.

Zhao et al. [Zhao et al. 2011] developed a cross-people motion activity recognition system. This developed smart phone based activity reporting system can accurately recognise the daily activities of stationary, walking, running, upstairs, and downstairs, and report the accumulated time of each activity. More significantly, the system is concerned about the calibration free and personalised problem. The algorithm learns a binary decision tree model for one person from his labelled samples, transfers its structure to another person and automatically adapts its non-determinate nodes with the unlabelled samples of the new person, thus accomplishing the cross-people knowledge transfer task. The system consists of two components, the first one is the TransDT layer (Transfer Decision Tree), and the other one is the EM Algorithm layer (Expectation Maximization Algorithm). TransDT is a binary decision tree learnt off-line and built on a well-labelled training set. In this layer, the system constructs a classifier as well as finds the attributes that can distinguish one class from others. The EM layer corresponds to transferring and adapting the TransDT model to a personalised one. When having collected sufficient unlabelled samples from a new user, the algorithm classifies them with the TransDT model and then uses the result as the initial condition of the EM algorithm. After the EM algorithm, the unlabelled samples are well labelled. So the system can update the non-terminal nodes of the TransDT model and thus form the personalised model.

Pärkkä et al. [Pärkkä et al. 2010] developed an activity recognition system based on a decision tree classifier that automatically recognises physical activities on a portable device, online. The model is personalised by updating the decision tree threshold values with a user's own data. The central device in data collection and activity recognition is a personal digital assistant (PDA). The application receives data over a Bluetooth connection, computes feature signals from raw data online, classifies the data in a second-by-second basis online, and stores the data on a memory card. The human movements are quantified with Nokia wireless motion bands using the 3-D accelerometer signal and Bluetooth connection for data transfer to the PDA. Only ankle sensor data are used for computing the feature signals. The time-domain features were computed from the most recent samples, and the frequency-domain features were computed from the same samples and one added zero for an efficient fast Fourier transform (FFT) implementation. The structure of the binary decision tree includes four nodes with one threshold value in each node. The personalised algorithm searches for an optimum decision boundary between the activities falling left and right in each node. It takes 3-10 minutes of new data with annotation and uses that for updating the thresholds in each node. The algorithm requires only a few comparisons and thus consumes very little battery power. Personalising the training model by adjusting the thresholds is a key contribution, however a long delay up to 10 seconds is required for updating the model.

7. RESEARCH CHALLENGES

Upon discussing the wide range of approaches proposed for activity recognition, we present the research gaps that have not been addressed or partially addressed in the literature. One of the main limitations that encounters the recognition system is the *static nature of the classification model*. Studies demonstrated that deployment of a personalised model for activity recognition outperforms the system of applying a general model [Weiss and Lockhart 2012]. It can easily be seen that models that are built for general activity recognition would need to be tuned and adapted to suit the evolving data of activities. Each individual has his/her own personalised way of performing activities. Thus, the significance of personalising the training model in real time became crucial in activity recognition in order to improve the recognition accuracy for a specific user. Personalisation is the process of tuning the general model to represent a user's particular way of performing various activities. Personalisation updates the current existing model without changing the core activity types. Therefore, the personalisation process only 'tunes' current activities for a particular user in an incremental manner. Few studies in activity recognition considered the personalisation issue. Most of these studies addressed the personalisation by retraining of the entire model with a particular user's annotated data. Few studies considered the incremental and automated approach for personalisation. Moreover, a data stream is an unbounded flow of data that is mostly unlabelled. Personalisation methods in activity recognition require annotation of recent data to update the model. However, *the scarcity of labelled data* as well as the *time consuming process of labelling each data instance* are in conflict with the streaming nature of activity recognition of sensory data. Alternatively, the recognition of activities needs to consider batch labelling for only the most profitable data in batch active learning approach. Thus, the time and resource consumptions are only limited for the most profitable data. Handling the streaming nature of data is crucial for processing and responding to a streaming activity in almost real time.

Moreover, in a real life application, activities that a user performs evolve over time. Therefore, the set of activities represented in the model has to be updated to reflect the change in the performed activities. That includes adding new recognised activities and also understanding activities that are no longer relevant to that particular user. Model adaptation is a key criterion of the robustness of any activity recognition system. There is no notion of adaptation/refinement of the classifier models in the literature. Models do not include activities that may emerge over a period of time (post the data collection) or changes in a user's patterns, which are both completely realistic in the context of a mobile user. The adaptation process needs to update the recognition model over time to reflect changes in a real life user's activities in real time.

The deployment of inexpensive sensors for collecting data for activity recognition coupled with stream mining techniques has resulted in the development of the wide variety of applications. Sensors could be either onboard the mobile device itself, body worn or embedded in a smart environment. Using mobile devices with limited resources for recognition of streaming activities became a hot topic of research in pervasive and ubiquitous computing. Two approaches have been applied for mobile based recognition systems. Firstly, the phone transmits the data to the backend sever with the server applying the activity recognition model and transmitting the results back to the phone. The second method involves implementing the activity recognition model directly on the mobile device. Given the computational power of these everyday devices, this is certainly a feasible option. One key advantage of this method is that it removes the need for a server and therefore save transmission time and allows real time prediction. Today's mobile devices make the solution perfectly scalable, and ensures the user's privacy, since the sensory data is kept locally on the device.

Time and accuracy are well-known critical factors used in judging the performance of any real-time activity recognition technique. Building a classifier that accurately recognises physical activities on a mobile device in real time is a key research issue.

7.1. Comparison of the State-of-the-art Activity Recognition Techniques

In Table III, we present a comparison of the different activity recognition systems addressing the aforementioned research issues. It is demonstrated that there are many significant missing research gaps that need to be addressed for the development of an efficient activity recognition system. Although the pervasive techniques in activity recognition are deployed in a static environment, we focus in this comparison on the missing gap in recent approaches for streaming environment.

Data collection platforms vary from one system to another. Data is collected from either mobile, body worn, or smart environment sensors. The processing platform is related to the recognition application and response time. Processing data on a backend server or on a PC implies two factors. The model is not light-weight. Therefore it has to be implemented on a high performing device, such as a PC. The other factor is the real time recognition as the implementation on a backend server or PC causes a delay in transmitting data back to the user and thus does not maintain a real time recognition.

Few of these systems considered real time adaptation for activity recognition. Recognition systems use a static training model which is built offline to recognise incoming data. The technique in [Abdallah et al. 2015] has the ability to expand the model after it is already deployed. When new activity was urged or any current activity abandoned, current techniques could not be adapted accordingly. Reflecting the change in activities in real time is also a crucial research issue. Only an approach in [Krishnan and Cook 2014] and [Nguyen et al. 2015] considered the discovery of 'other' or 'new' activity. The recognition of these activities is different from recognising novel activities from an evolving stream. In [Krishnan and Cook 2014], the technique has to build a static model offline that represents other transitional and unknown activities. It explicitly trains the model offline for transitional activities and other unknown activities. Then, it gathers them all in other activity class and incorporates it with the model. The technique developed in [Nguyen et al. 2015] also requires labelled data representing the novel activity. Therefore, these approaches are not detecting the emerging of novel activities from data stream with no labelled data representing the novel concepts.

Two approaches for model personalisation are presented in Table III. The first is a complete reconstruction of the learning model with the personalised data, while the other aims only on tuning the model to suit a particular user. There is a little focus on the personalisation of the learning models in activity recognition. Retraining the model is not applicable for streaming environments. In the other incremental learning approach, offline models in [Pärkkä et al. 2010] and [Zhao et al. 2011] are refined by tuning the model with incoming data. However, they both deployed in non-streaming environments. Moreover, the Pärkkä et al. technique performs the personalisation offline with long delays. Therefore, none of these systems is representing a real time approach to resolve the personalisation issue in a streaming environment. The approach in [Abdallah et al. 2015] addresses the personalisation of activity recognition model yet in streaming environment. The personalisation in this approach is more applicable for streaming data as it applies an incremental learning approach for personalisation instead of retraining the entire model as in [Gomes et al. 2012a].

The scarcity of labelled data requires incorporating active learning with recognition for labelling only a small amount of data that is most informative. Relying on less labelled data has not been addressed by the majority of the systems. The system developed in [Abdallah et al. 2015] proposed an approach that incorporates active learning based on the recognition confidence. However, the rate of triggered active learning inquiries still requires further improvement as it increases significantly with highly fluctuating data especially with numerous transition activities.

8. FUTURE RESEARCH DIRECTIONS

The research work surveyed in this paper stimulates different areas of research directions for extensions. In this section, we provide pointers to some of the key research directions that can be outlined as follows.

— Enabling *personalisation* in activity recognition benefits the research area for accurate and efficient recognition. With an efficient personalisation, activity recognition becomes widely available and accurate across users, as there is no need to train the recognition system on each user's way of performing activities. General models are automatically customised for a user's personalised activities at real time and on limited resource devices. The capability of personalisation supports an improved and accurate recognition across many applications. These applications include personalised advertisements, lifestyle monitoring, and elderly people monitoring.
— The capturing of novel activities and dynamic *adaptation* with evolving activities enables the development of reliable systems for activity recognition that can learn new knowledge and forget outdated knowledge. With the adaptation technique, there is no need to define a set of activities to recognise. The automatically adaptable technique is a self-learner that monitors the stream and understands the evolving data. This approach is a significant contribution in activity recognition as it creates less dependency on the historical annotated data. This future direction of this feature is to enable an independent technique that requires almost no prior knowledge to recognise activities. The feature can be extended for detecting and recognising activities on the fly with evolving streams. The dynamic adaptation feature can incrementally create the model with detected activities once they arrive and without prior training.
— Building a *holistic approach* for activity recognition that integrates the personalisation of existing activities with the adaptation for novel detected and forgotten activities. The holistic approach creates a single framework for detecting both major and minor changes in activities while a stream evolves.
— A big challenge in activity recognition is to collect sufficient amount of labeled data to train the learning model. The *annotating process* is expensive, time-consuming and erroneous. An incremental and active learning approach that addresses this problem has been previously developed as aforementioned. The initial model in the developed approach is built with a small amount of labeled data. The model is continuously accumulated through incremental learning. Active learning asks only for labels of informative samples. This approach opens the door for a wider range of applications in activity recognition as it allows learning from unlabeled data, which is pervasively available. The research in this direction is still new and promising as it aims at less supervision in the recognition process.
— An accurate recognition of activities can be combined with a *context aware* framework for recognising higher level and more complex activities. A future activity recognition tiered approach has to leverage all available information for context aware recognition. Building a context aware technique for activity recognition enables a wide range of applications that concern not only the activity of the user, but also the context of surrounding environment, to be able to provide the user with accurate and opportune information.
— Context aware in adaptive activity recognition should consider the *dynamics of both activities and context* in streaming environment. For example, when reasoning about the mode of transportation as a context of an activity, the recognition of the context has to assume that the activity is not static, i.e., the user can be walking, sitting or standing in the train. Extending the adaptation definition to include activities and their contexts is important in many application. One example is the detection of transport congestion when detecting the dynamics of activities while changing the mode of transpiration.
— *Activity prediction* is a natural extension of activity recognition. Unlike activity recognition that reasons about the current activity, activity prediction forecasts the upcoming activities in the timeline. Moreover, combining the adaptation capabilities

of personalisation and detection of novel activities with activity prediction provide interesting and challenging areas of research.

— Benefiting from the dynamic and adaptive recognition, the recognition technique could be applied to a wide range of users, to provide crowdsourcing information based on their performed activities. Aggregating on the recognised activities from users' devices on a higher level platform, such as the cloud, enables a wider perspective of understanding the performed activities based on different criteria, such as location based activity recognition on the cloud.

It is worth noting that some steps have been taken towards realising the aforementioned directions. However, the research done is still short in providing reliable solutions that address the different research challenges collectively and cohesively. It is expected that with the rise in the *Internet of Things (IoT)* in the coming years, activity recognition will be a core component in many IoT applications [Perera et al. 2014]. Thus, fulfilling the needs of such applications would require robust activity recognition systems.

9. SUMMARY AND CONCLUSION

In this survey, we reviewed state-of-the-art approaches for the two overlapped research areas concerning activity recognition in stream mining. We first introduced the area of activity recognition in general. We reviewed the variety of sensors applied for activity recognition along with the different kinds of activities to be predicted. Then, the known term of *i.i.d.* in statistical and probability theory that refers to independent and identically distributed- ness is introduced. A key challenge with learning from sensors in a streaming environment is that it requires learning beyond identically distributed and independent sensory data. We then illustrated the position of this survey from the two active areas of research in techniques concerning the distribution of data (stream mining) and techniques concerning data dependency (activity recognition).

Activity recognition techniques along with different approaches are depicted in detail. We survey two main categories of activity recognition. The first category focuses on the learning approaches; the other one concerns the dynamic capabilities beyond the learning process. Most of the developed techniques in activity recognition are based on supervised learning, however few studies apply unsupervised and semi-supervised approaches given the limited availability of labelled data. Beyond learning, different approaches for adapting the initial model are discussed. This includes adaptation for model personalisation or for adding/removing activities. The dynamic capabilities are related to the concept of transfer learning in terms of transfer between the static domain of modelling to the dynamic streaming deployment domain.

Deployment of activity recognition in streaming settings imposes many challenges. In order to understand the limitation and constraints imposed with stream mining for activity recognition, we survey state-of-the-art techniques for stream mining. Stream mining techniques correspond to the distribution oriented approach. Data that evolves in the stream is not identically distributed. In the literature, many approaches have addressed for the handling of high speed and infinite stream of data. We reviewed in this paper different techniques developed for handling streaming data. Unlike traditional machine learning approaches, new approaches have been developed to address the streaming nature of data. Various kinds of changes may encounter data streams over time. We presented in this paper the methods of both tracking these changes and learning from these changes.

Upon surveying the main approaches for both activity recognition and stream mining, we presented the overlap between the two research areas. We first presented a conceptual mapping between the two areas to demonstrate the connection. We then discussed techniques that considered activity recognition with data streams. The research gaps based on the discussed state-of-the-art approaches were illustrated. Then, the key techniques in the literature that target subset of the research gaps were presented in detail along with a comparison between them. Finally, the paper concluded with a description of the contribution of this survey in addressing the discussed research gaps with linkage to the research future directions.

REFERENCES

2012. Actitracker application. https://actitracker.com/. (2012). Accessed: 2017-8-28.

Zahraa Said Abdallah, Mohamed Medhat Gaber, Bala Srinivasan, and Shonali Krishnaswamy. 2012. StreamAR: Incremental and Active Learning with Evolving Sensory Data for Activity Recognition. In *Tools with Artificial Intelligence (ICTAI), 2012 IEEE 24th International Conference on*, Vol. 1. 1163–1170. DOI:http://dx.doi.org/10.1109/ICTAI.2012.169

Zahraa Said Abdallah, Mohamed Medhat Gaber, Bala Srinivasan, and Shonali Krishnaswamy. 2015. Adaptive mobile activity recognition system with evolving data streams. *Neurocomputing* 150, Part A (2015), 304 – 317. DOI:http://dx.doi.org/10.1016/j.neucom.2014.09.074

Iris Adä and Michael R Berthold. 2013. EVE: a framework for event detection. *Evolving systems* 4, 1 (2013), 61–70. DOI:http://dx.doi.org/10.1007/s12530-012-9067-0

Charu C. Aggarwal. 2013. *Outlier Analysis*. Springer New York. http://books.google.com.au/books?id=900CkgEACAAJ

Tahseen Al-Khateeb, Mohammad M Masud, Latifur Khan, Charu C Aggarwal, Jiawei Han, and Bhavani M Thuraisingham. 2012. Stream Classification with Recurring and Novel Class Detection Using Class-Based Ensemble.. In *ICDM*. 31–40.

Aziah Ali, Rachel C King, and Guang-Zhong Yang. 2008. Semi-supervised segmentation for activity recognition with multiple eigenspaces. In *Medical Devices and Biosensors, 2008. ISSS-MDBS 2008. 5th International Summer School and Symposium on*. IEEE, 314–317.

Fabrizio Angiulli and Fabio Fassetti. 2007. Detecting Distance-based Outliers in Streams of Data. In *Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management (CIKM '07)*. ACM, New York, NY, USA, 811–820. DOI:http://dx.doi.org/10.1145/1321440.1321552

Ira Assent, Philipp Kranen, Corinna Baldauf, and Thomas Seidl. 2012. AnyOut: Anytime Outlier Detection on Streaming Data. In *Proceedings of the 17th International Conference on Database Systems for Advanced Applications - Volume Part I (DASFAA'12)*. Springer-Verlag, Berlin, Heidelberg, 228–242. DOI:http://dx.doi.org/10.1007/978-3-642-29038-1_18

UABUA Bakar, Hemant Ghayvat, SF Hasanm, and SC Mukhopadhyay. 2016. Activity and anomaly detection in smart home: A survey. In *Next Generation Sensors and Systems*. Springer, 191–220.

Ling Bao and Stephen S. Intille. 2004. Activity Recognition from User-Annotated Acceleration Data. In *Pervasive Computing*. Lecture Notes in Computer Science, Vol. 3001. Springer Berlin Heidelberg, 1–17. DOI:http://dx.doi.org/10.1007/978-3-540-24646-6_1

Albert Bifet and Ricard Gavaldà. 2006. Kalman Filters and Adaptive Windows for Learning in Data Streams. In *Proceedings of the 9th International Conference on Discovery Science (DS'06)*. Springer-Verlag, Berlin, Heidelberg, 29–40. DOI:http://dx.doi.org/10.1007/11893318_7

Albert Bifet and Ricard Gavaldà. 2007. Learning from Time-Changing Data with Adaptive Windowing. In *Proceedings of the SIAM International Conference on Data Mining*. 443–448. DOI:http://dx.doi.org/10.1137/1.9781611972771.42

Avrim Blum and Tom Mitchell. 1998. Combining Labeled and Unlabeled Data with Co-training. In *Proceedings of the Eleventh Annual Conference on Computational Learning Theory (COLT' 98)*. ACM, New York, NY, USA, 92–100. DOI:http://dx.doi.org/10.1145/279943.279962

Robert Bodor, Bennett Jackson, Nikolaos Papanikolopoulos, and Human Tracking. 2003. Vision-based human tracking and activity recognition. In *Proceedings of the 11th Mediterranean Conference on Control and Automation*. Kostrzewa Joseph, 18–20.

Abdelhamid Bouchachia. 2011. Incremental learning with multi-level adaptation. *Neurocomputing* 74, 11 (2011), 1785–1799.

Andreas Bulling, Ulf Blanke, and Bernt Schiele. 2014. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)* 46, 3 (2014), 1–33.

Gert Cauwenberghs and Tomaso Poggio. 2001. Incremental and decremental support vector machine learning. *Advances in neural information processing systems* (2001), 409–415.

Rita Chattopadhyay, Qian Sun, Wei Fan, Ian Davidson, Sethuraman Panchanathan, and Jieping Ye. 2012. Multisource domain adaptation and its application to early detection of fatigue. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 6, 4 (2012), 1–18.

Yanping Chen, Bing Hu, Eamonn Keogh, and Gustavo EAPA Batista. 2013. DTW-D: time series semi-supervised learning from a single example. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 383–391.

Yanping Chen, Eamonn Keogh, Bing Hu, Nurjahan Begum, Anthony Bagnall, Abdullah Mueen, and Gustavo Batista. 2015. The UCR Time Series Classification Archive. (July 2015). www.cs.ucr.edu/~eamonn/time_series_data/.

Diane Cook, KyleD. Feuz, and NarayananC. Krishnan. 2013. Transfer learning for activity recognition: a survey. *Knowledge and Information Systems* 36, 3 (2013), 537–556. DOI:http://dx.doi.org/10.1007/s10115-013-0665-3

Thang M. Do, Seng W. Loke, and Fei Liu. 2013. HealthyLife: An Activity Recognition System with Smartphone Using Logic-Based Stream Reasoning. In *Mobile and Ubiquitous Systems: Computing, Networking, and Services*. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol. 120. Springer Berlin Heidelberg, 188–199. DOI:http://dx.doi.org/10.1007/978-3-642-40238-8_16

Pedro Domingos and Geoff Hulten. 2000. Mining high-speed data streams. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 71–80. DOI:http://dx.doi.org/10.1145/347090.347107

Anton Dries and Ulrich Rückert. 2009. Adaptive concept drift detection. *Statistical Analysis and Data Mining* 2, 56 (2009), 311–327. DOI:http://dx.doi.org/10.1002/sam.v2:5/6

Wei Fan, Yi-an Huang, Haixun Wang, and Philip S Yu. 2004. Active mining of data streams. In *Proceedings of the 2004 SIAM International Conference on Data Mining*. SIAM, 457–461.

Elaine R. Faria, João Gama, and André C. P. L. F. Carvalho. 2013. Novelty Detection Algorithm for Data Streams Multi-class Problems. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC '13)*. ACM, New York, NY, USA, 795–800. DOI:http://dx.doi.org/10.1145/2480362.2480515

Simon Fong, Kexing Liu, Kyungeun Cho, Raymond Wong, Sabah Mohammed, and Jinan Fiaidhi. 2016. Improvised methods for tackling big data stream mining challenges: case study of human activity recognition. *The Journal of Supercomputing* (2016), 1–33.

Mohamed Medhat Gaber and Philip S Yu. 2006. Detection and classification of changes in evolving data streams. *International Journal of Information Technology & Decision Making* 5, 04 (2006), 659–670.

João Gama. 2012. A survey on learning from data streams: current and future trends. *Progress in Artificial Intelligence* 1, 1 (2012), 45–55.

João Gama, Pedro Medas, Gladys Castillo, and Pedro Rodrigues. 2004. Learning with drift detection. In *Advances in Artificial Intelligence–SBIA 2004*. Springer, 286–295.

João Gama, Pedro Pereira Rodrigues, Eduardo J Spinosa, and André Carlos Ponce Leon Ferreira de Carvalho. 2010. *Knowledge discovery from data streams*. Chapman & Hall/CRC Boca Raton.

João Gama, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. 2014. A Survey on Concept Drift Adaptation. *Comput. Surveys* 46, 4, Article 44 (2014), 37 pages. DOI:http://dx.doi.org/10.1145/2523813

Alexander Gammerman, Volodya Vovk, and Vladimir Vapnik. 1998. Learning by transduction. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 148–155.

Joao Bartolo Gomes, Shonali Krishnaswamy, Mohamed Medhat Gaber, Pedro A. C. Sousa, and Ernestina Menasalvas. 2012a. MARS: A Personalised Mobile Activity Recognition System. In *Proceedings of the 2012 IEEE 13th International Conference on Mobile Data Management (Mdm 2012) (MDM '12)*. IEEE Computer Society, Washington, DC, USA, 316–319. DOI:http://dx.doi.org/10.1109/MDM.2012.33

João Bártolo Gomes, Shonali Krishnaswamy, Mohamed M. Gaber, Pedro A. C. Sousa, and Ernestina Menasalvas. 2012b. Mobile Activity Recognition Using Ubiquitous Data Stream Mining. In *Proceedings of the 14th International Conference on Data Warehousing and Knowledge Discovery*. Springer-Verlag, Berlin, Heidelberg, 130–141. DOI:http://dx.doi.org/10.1007/978-3-642-32584-7_11

Donghai Guan, Weiwei Yuan, Young-Koo Lee, A. Gavrilov, and Sungyoung Lee. 2007. Activity Recognition Based on Semi-supervised Learning. In *Embedded and Real-Time Computing Systems and Applications, RTCSA. 13th IEEE International Conference on*. 469–475. DOI:http://dx.doi.org/10.1109/RTCSA.2007.17

Hirotaka Hachiya, Masashi Sugiyama, and Naonori Ueda. 2012. Importance-weighted least-squares probabilistic classifier for covariate shift adaptation with application to human activity recognition. *Neurocomputing* 80 (2012), 93–101.

Maayan Harel and Shie Mannor. 2010. Learning from multiple outlooks. *28th international conference on machine learning.* (2010), 401–408.

Douglas M Hawkins. 1980. *Identification of outliers*. Chapman and Hall, London. http://books.google.com.au/books?id=P8ZPAQAAIAAJ

Morteza Zi Hayat and Mahmoud Reza Hashemi. 2010. A DCT based approach for detecting novelty and concept drift in data streams. In *Soft Computing and Pattern Recognition (SoCPaR), 2010 International Conference of*. 373–378. DOI:http://dx.doi.org/10.1109/SOCPAR.2010.5686734

Yu-Jin Hong, Ig-Jae Kim, Sang Chul Ahn, and Hyoung-Gon Kim. 2010. Mobile health monitoring system based on activity recognition using accelerometer. *Simulation Modelling Practice and Theory* 18, 4 (2010), 446–455.

Bing Hu, Yanping Chen, and Eamonn Keogh. 2013. Time series classification under more realistic assumptions. In *Proceedings of the 2013 SIAM International Conference on Data Mining*. SIAM, 578–586.

Shucheng Huang and Yisheng Dong. 2007. An active learning system for mining time-changing data streams. *Intelligent Data Analysis* 11, 4 (2007), 401–419.

Geoff Hulten, Laurie Spencer, and Pedro Domingos. 2001. Mining Time-changing Data Streams. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '01)*. ACM, New York, NY, USA, 97–106. DOI:http://dx.doi.org/10.1145/502512.502529

Tâm Huynh and Bernt Schiele. 2006. Towards less supervision in activity recognition from wearable sensors. In *Wearable Computers, 2006 10th IEEE International Symposium on*. IEEE, 3–10.

Stephen S Intille. 2004. Ubiquitous computing technology for just-in-time motivation of behavior change. *Stud Health Technol Inform* 107, Pt 2 (2004), 1434–1437.

Ashish Kapoor and Eric Horvitz. 2008. Experience Sampling for Building Predictive User Models: A Comparative Study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 657–666. DOI:http://dx.doi.org/10.1145/1357054.1357159

Dean M Karantonis, Michael R Narayanan, Merryn Mathie, Nigel H Lovell, and Branko G Celler. 2006. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *Information Technology in Biomedicine, IEEE Transactions on* 10, 1 (2006), 156–167. DOI:http://dx.doi.org/10.1109/TITB.2005.856864

Daniel Kifer, Shai Ben-David, and Johannes Gehrke. 2004. Detecting change in data streams. In *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*. VLDB Endowment, 180–191.

Ralf Klinkenberg and Thorsten Joachims. 2000. Detecting Concept Drift with Support Vector Machines. In *ICML*. 487–494.

Ralf Klinkenberg and Ingrid Renz. 1998. Adaptive Information Filtering: Learning Drifting Concepts. In *Workshop Learning for Text Categorization AAAI98/ICML-98*. Press, 33–40.

Narayanan C. Krishnan and Diane J. Cook. 2014. Activity Recognition on Streaming Sensor Data. *Pervasive Mobile Computing* 10 (2014), 138–154. DOI:http://dx.doi.org/10.1016/j.pmcj.2012.07.003

Marc Kurz, Gerold Hölzl, Alois Ferscha, Alberto Calatroni, Daniel Roggen, and Gerhard Tröster. 2011. Real-time transfer and evaluation of activity recognition capabilities in an opportunistic system. In *The Third International Conference on Adaptive and Self-Adaptive Systems and Applications*. 73–78.

Jennifer R. Kwapisz, Gary M. Weiss, and Samuel A. Moore. 2011. Activity Recognition Using Cell Phone Accelerometers. *ACM SigKDD Explorations Newsletter* 12, 2 (2011), 74–82. DOI:http://dx.doi.org/10.1145/1964897.1964918

M. Powell Lawton and Elaine M. Brody. 1969. Assessment of Older People: Self-Maintaining and Instrumental Activities of Daily Living. *The Gerontologist* 9, 3 (1969), 179–186. DOI:http://dx.doi.org/10.1093/geront/9.3_Part_1.179

Min-Seok Lee, Jong-Gwan Lim, Ki-Ru Park, and Dong-Soo Kwon. 2009. Unsupervised clustering for abnormality detection based on the tri-axial accelerometer. In *ICCAS-SICE, 2009*. 134–137.

Seon-Woo Lee and Kenji Mase. 2002. Activity and Location Recognition Using Wearable Sensors. *IEEE Pervasive Computing* 1, 3 (2002), 24–32. DOI:http://dx.doi.org/10.1109/MPRV.2002.1037719

Young-Seol Lee and Sung-Bae Cho. 2014. Activity recognition with android phone using mixture-of-experts co-trained with labeled and unlabeled data. *Neurocomputing* 126 (2014), 106 – 115. DOI:http://dx.doi.org/10.1016/j.neucom.2013.05.044

Jonathan Lester, Tanzeem Choudhury, Nicky Kern, Gaetano Borriello, and Blake Hannaford. 2005. A Hybrid Discriminative/Generative Approach for Modeling Human Activities. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI'05)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 766–772. http://dl.acm.org/citation.cfm?id=1642293.1642416

Fei Li and Schahram Dustdar. 2011. Incorporating Unsupervised Learning in Activity Recognition. In *Proceedings of the 4th AAAI Conference on Activity Context Representation: Techniques and Languages (AAAIWS'11-04)*. AAAI Press, 38–41. http://dl.acm.org/citation.cfm?id=2908613.2908620

Fangtao Li, Sinno Jialin Pan, Ou Jin, Qiang Yang, and Xiaoyan Zhu. 2012. Cross-domain co-extraction of sentiment and topic lexicons. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*. Association for Computational Linguistics, 410–419.

Li-Jia Li, Gang Wang, and Li Fei-Fei. 2007. OPTIMOL: automatic Online Picture collecTion via Incremental MOdel Learning. In *Computer Vision and Pattern Recognition, IEEE Conference on*. 1–8. DOI:http://dx.doi.org/10.1109/CVPR.2007.383048

Peipei Li, Xindong Wu, and Xuegang Hu. 2012. Mining recurring concept drifts with limited labeled streaming data. *ACM Transactions on Intelligent Systems and Technology (TIST)* 3, 2 (2012), 1–29.

Patrick Lindstrom, Brian Mac Namee, and Sarah Jane Delany. 2013. Drift detection using uncertainty distribution divergence. *Evolving Systems* 4, 1 (2013), 13–25.

Li Liu, Yuxin Peng, Ming Liu, and Zigang Huang. 2015. Sensor-based human activity recognition system with a multilayered model using time series shapelets. *Knowledge-Based Systems* 90 (2015), 138–152.

Jeffrey W. Lockhart, Tony Pulickal, and Gary M. Weiss. 2012. Applications of Mobile Activity Recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 1054–1058. DOI:http://dx.doi.org/10.1145/2370216.2370441

Jeffrey W. Lockhart and Gary M. Weiss. 2014. The benefits of personalized smartphone-based activity recognition models. In *Proceedings of the 2014 SIAM International Conference on Data Mining*. 614–622. DOI:http://dx.doi.org/10.1137/1.9781611973440.71

Beth Logan, Jennifer Healey, Matthai Philipose, Emmanuel Munguia Tapia, and Stephen Intille. 2007. A Long-term Evaluation of Sensing Modalities for Activity Recognition. In *Proceedings of the 9th International Conference on Ubiquitous Computing (UbiComp '07)*. Springer-Verlag, Berlin, Heidelberg, 483–500. http://dl.acm.org/citation.cfm?id=1771592.1771620

Mingsheng Long, Jianmin Wang, Guiguang Ding, Jiaguang Sun, and Philip S Yu. 2013. Transfer feature learning with joint distribution adaptation. In *Proceedings of the IEEE International Conference on Computer Vision*. 2200–2207.

Brent Longstaff, Sasank Reddy, and Deborah Estrin. 2010. Improving activity classification for health applications on mobile devices using active and semi-supervised learning. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 4th International Conference on-NO PERMISSIONS*. IEEE, 1–7.

Mitja Luštrek and Boštjan Kaluža. 2009. Fall detection and activity recognition with machine learning. *Informatica (Slovenia)* 33, 2 (2009), 197–204.

Mohammad M. Masud, Tahseen M. Al-Khateeb, Latifur Khan, Charu Aggarwal, Jing Gao, Jiawei Han, and Bhavani Thuraisingham. 2011. Detecting Recurring and Novel Classes in Concept-Drifting Data Streams. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining (ICDM '11)*. IEEE Computer Society, Washington, DC, USA, 1176–1181. DOI:http://dx.doi.org/10.1109/ICDM.2011.49

Mohammad M Masud, Qing Chen, Latifur Khan, Charu C Aggarwal, Jing Gao, Jiawei Han, Ashok Srivastava, and Nikunj C Oza. 2013. Classification and adaptive novel class detection of feature-evolving data streams. *Knowledge and Data Engineering, IEEE Transactions on* 25, 7 (2013), 1484–1497. DOI:http://dx.doi.org/10.1109/TKDE.2012.109

Mohammad M. Masud, Jing Gao, L. Khan, Jiawei Han, and Bhavani Thuraisingham. 2011. Classification and Novel Class Detection in Concept-Drifting Data Streams under Time Constraints. *Knowledge and Data Engineering, IEEE Transactions on* 23, 6 (2011), 859–874. DOI:http://dx.doi.org/10.1109/TKDE.2010.61

Mohammad M Masud, Clay Woolam, Jing Gao, Latifur Khan, Jiawei Han, Kevin W Hamlen, and Nikunj C Oza. 2012. Facing the reality of data stream classification: coping with scarcity of labeled data. *Knowledge and information systems* 33, 1 (2012), 213–244.

Uwe Maurer, Asim Smailagic, Daniel P. Siewiorek, and Michael Deisher. 2006. Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions. In *Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks (BSN '06)*. IEEE Computer Society, Washington, DC, USA, 113–116. DOI:http://dx.doi.org/10.1109/BSN.2006.6

Emiliano Miluzzo, Nicholas D. Lane, Kristóf Fodor, Ronald Peterson, Hong Lu, Mirco Musolesi, Shane B. Eisenman, Xiao Zheng, and Andrew T. Campbell. 2008. Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of the CenceMe Application. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys '08)*. ACM, New York, NY, USA, 337–350. DOI:http://dx.doi.org/10.1145/1460412.1460445

Iqbal Mohomed, Archan Misra, Maria Ebling, and William Jerome. 2008. HARMONI: Context-aware Filtering of Sensor Data for Continuous Remote Health Monitoring. In *Proceedings of the 2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications (PERCOM '08)*. IEEE Computer Society, Washington, DC, USA, 248–251. DOI: http://dx.doi.org/10.1109/PERCOM.2008.110

Muhammad Mubashir, Ling Shao, and Luke Seed. 2013. A Survey on Fall Detection: Principles and Approaches. *Neurocomputing* 100 (2013), 144–152. DOI: http://dx.doi.org/10.1016/j.neucom.2011.09.037

Min Mun, Sasank Reddy, Katie Shilton, Nathan Yau, Jeff Burke, Deborah Estrin, Mark Hansen, Eric Howard, Ruth West, and Péter Boda. 2009. PEIR, the Personal Environmental Impact Report, As a Platform for Participatory Sensing Systems Research. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*. ACM, New York, NY, USA, 55–68. DOI: http://dx.doi.org/10.1145/1555816.1555823

Ion Muslea, Steven Minton, and Craig A. Knoblock. 2000. Selective Sampling with Redundant Views. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*. AAAI Press, 621–626. http://dl.acm.org/citation.cfm?id=647288.721119

Hammadi Nait-Charif and Stephen J. McKenna. 2004. Activity Summarisation and Fall Detection in a Supportive Home Environment. In *Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 4 - Volume 04 (ICPR '04)*. IEEE Computer Society, Washington, DC, USA, 323–326. DOI: http://dx.doi.org/10.1109/ICPR.2004.127

Le T Nguyen, Ming Zeng, Patrick Tague, and Joy Zhang. 2015. Recognizing new activities with limited training data. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 67–74.

Vit Niennattrakul, Eamonn Keogh, and Chotirat Ann Ratanamahatana. 2010. Data Editing Techniques to Allow the Application of Distance-Based Outlier Detection to Streams. In *Proceedings of the 2010 IEEE International Conference on Data Mining (ICDM '10)*. IEEE Computer Society, Washington, DC, USA, 947–952. DOI: http://dx.doi.org/10.1109/ICDM.2010.56

Wei Niu, Jiao Long, Dan Han, and Yuan-Fang Wang. 2004. Human activity detection and recognition for video surveillance. In *Multimedia and Expo. ICME '04. IEEE International Conference on*, Vol. 1. 719–722. DOI: http://dx.doi.org/10.1109/ICME.2004.1394293

Sinno Jialin Pan, James T Kwok, Qiang Yang, and Jeffrey Junfeng Pan. 2007. Adaptive localization in a dynamic wifi environment through multi-view learning. In *AAAI*. 1108–1113.

Sinno Jialin Pan, Ivor W Tsang, James T Kwok, and Qiang Yang. 2011. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks* 22, 2 (2011), 199–210.

Sinno Jialin Pan and Qiang Yang. 2010. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22, 10 (2010), 1345–1359.

Juha Pärkkä, Luc Cluitmans, and Miikka Ermes. 2010. Personalization Algorithm for Real-time Activity Recognition Using PDA, Wireless Motion Bands, and Binary Decision Tree. *Information Technology in Biomedicine, IEEE Transactions on* 14, 5 (2010), 1211–1215. DOI: http://dx.doi.org/10.1109/TITB.2010.2055060

Donald J. Patterson, Dieter Fox, Henry Kautz, and Matthai Philipose. 2005. Fine-Grained Activity Recognition by Aggregating Abstract Object Usage. In *Proceedings of the Ninth IEEE International Symposium on Wearable Computers (ISWC '05)*. IEEE Computer Society, Washington, DC, USA, 44–51. DOI: http://dx.doi.org/10.1109/ISWC.2005.22

C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos. 2014. Context Aware Computing for The Internet of Things: A Survey. *IEEE Communications Surveys Tutorials* 16, 1 (2014), 414–454. DOI: http://dx.doi.org/10.1109/SURV.2013.042313.00197

Tom Peterek, Marek Penhaker, Petr Gajdo, and Pavel Dohnlek. 2014. Comparison of Classification Algorithms for Physical Activity Recognition. In *Innovations in Bio-inspired Computing and Applications*. Advances in Intelligent Systems and Computing, Vol. 237. Springer International Publishing, 123–131. DOI: http://dx.doi.org/10.1007/978-3-319-01781-5_12

Dragoljub Pokrajac, Aleksandar Lazarevic, and Longin Jan Latecki. 2007. Incremental local outlier detection for data streams. In *Computational Intelligence and Data Mining, CIDM. IEEE Symposium on*. IEEE, 504–515. DOI: http://dx.doi.org/10.1109/CIDM.2007.368917

Stephen J Preece, John Y Goulermas, Laurence PJ Kenney, Dave Howard, Kenneth Meijer, and Robin Crompton. 2009. Activity identification using body-mounted sensorsa review of classification techniques. *Physiological measurement* 30, 4 (2009), 1–33.

Parisa Rashidi and Diane J. Cook. 2010. Mining Sensor Streams for Discovering Human Activity Patterns over Time. In *IEEE 10th International Conference on Data Mining (ICDM)*. 431–440. DOI: http://dx.doi.org/10.1109/ICDM.2010.40

Parisa Rashidi, Diane J Cook, Lawrence B Holder, and Maureen Schmitter-Edgecombe. 2011. Discovering activities to recognize and track in a smart environment. *Knowledge and Data Engineering, IEEE Transactions on* 23, 4 (2011), 527–539. DOI:http://dx.doi.org/10.1109/TKDE.2010.148

Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L. Littman. 2005. Activity Recognition from Accelerometer Data. In *Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence - Volume 3 (IAAI'05)*. AAAI Press, 1541–1546. http://dl.acm.org/citation.cfm?id=1620092.1620107

Attila Reiss and Didier Stricker. 2013. Personalized Mobile Physical Activity Recognition. In *Proceedings of the 2013 International Symposium on Wearable Computers (ISWC '13)*. ACM, New York, NY, USA, 25–28. DOI:http://dx.doi.org/10.1145/2493988.2494349

Daniel Roggen, Alberto Calatroni, Mirco Rossi, Thomas Holleczek, Kilian Forster, Gerhard Troster, Paul Lukowicz, David Bannach, Gerald Pirkl, Alois Ferscha, J. Doppler, C. Holzmann, M. Kurz, G. Holl, R. Chavarriaga, H. Sagha, H. Bayati, M. Creatura, and J. del R Millan. 2010. Collecting complex activity datasets in highly rich networked sensor environments. In *Networked Sensing Systems (INSS), 2010 Seventh International Conference on*. 233–240. DOI:http://dx.doi.org/10.1109/INSS.2010.5573462

Dan Y. Rubinstein, Trevor Hastie, and others. 1997. Discriminative vs Informative Learning. In *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining (KDD-97)*. AAAI Press, 49–53.

Dairazalia Sánchez, Monica Tentori, and Jesús Favela. 2008. Activity Recognition for the Smart Hospital. *IEEE Intelligent Systems* 23, 2 (2008), 50–57. DOI:http://dx.doi.org/10.1109/MIS.2008.18

Jeffrey C Schlimmer and Richard H Granger Jr. 1986. Incremental learning from noisy data. *Machine learning* 1, 3 (1986), 317–354.

Skyler Seto, Wenyu Zhang, and Yichen Zhou. 2015. Multivariate Time Series Classification Using Dynamic Time Warping Template Selection for Human Activity Recognition. In *Computational Intelligence, 2015 IEEE Symposium Series on*. IEEE, 1399–1406.

Xiaoxiao Shi, Qi Liu, Wei Fan, S Yu Philip, and Ruixin Zhu. 2010. Transfer learning on heterogenous feature spaces via spectral transformation. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*. IEEE, 1049–1054.

Eduardo J. Spinosa, André C. P. L. F. Carvalho, and João Gama. 2007. OLINDDA: A Cluster-based Approach for Detecting Novelty and Concept Drift in Data Streams. In *Proceedings of the 2007 ACM Symposium on Applied Computing (SAC '07)*. ACM, New York, NY, USA, 448–452. DOI:http://dx.doi.org/10.1145/1244002.1244107

M.S. Sricharan, V. Vaidehi, and P.P. Arun. 2006. An activity based mobility prediction strategy for next generation wireless network. In *Wireless and Optical Communications Networks, 2006 IFIP International Conference on*. 5–10. DOI:http://dx.doi.org/10.1109/WOCN.2006.1666596

Maja Stikic, Diane Larlus, Sandra Ebert, and Bernt Schiele. 2011. Weakly supervised recognition of daily life activities with wearable sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 12 (2011), 2521–2537.

Maja Stikic, Diane Larlus, and Bernt Schiele. 2009. Multi-graph based semi-supervised learning for activity recognition. In *Wearable Computers, 2009. ISWC'09. International Symposium on*. IEEE, 85–92.

Maja Stikic, Kristof Van Laerhoven, and Bernt Schiele. 2008. Exploring Semi-supervised and Active Learning for Activity Recognition. In *Proceedings of the 2008 12th IEEE International Symposium on Wearable Computers (ISWC '08)*. IEEE Computer Society, Washington, DC, USA, 81–88. DOI:http://dx.doi.org/10.1109/ISWC.2008.4911590

W Nick Street and YongSeog Kim. 2001. A streaming ensemble algorithm (SEA) for large-scale classification. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 377–382.

Monica Tentori and Jesus Favela. 2008. Activity-Aware Computing for Healthcare. *Pervasive Computing, IEEE* 7, 2 (2008), 51–57. DOI:http://dx.doi.org/10.1109/MPRV.2008.24

Paul A. Viola and Michael J. Jones. 2001. Rapid Object Detection using a Boosted Cascade of Simple Features. In *2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. 511–518. DOI:http://dx.doi.org/10.1109/CVPR.2001.990517

Quang Viet Vo, Minh Thang Hoang, and Deokjai Choi. 2013. Personalization in Mobile Activity Recognition System Using-Medoids Clustering Algorithm. *International Journal of Distributed Sensor Networks* (2013), 12. DOI:http://dx.doi.org/10.1155/2013/315841

Amal A Wanigatunga, Paul V Nickerson, Todd M Manini, and Parisa Rashidi. 2016. Using symbolic aggregate approximation (SAX) to visualize activity transitions among older adults. *Physiological Measurement* 37, 11 (2016), 1981.

Jamie A. Ward, Paul Lukowicz, Gerhard Troster, and Thad E. Starner. 2006. Activity Recognition of Assembly Tasks Using Body-Worn Microphones and Accelerometers. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 28, 10 (2006), 1553–1567. DOI:http://dx.doi.org/10.1109/TPAMI.2006.197

Gary M Weiss and Jeffrey W Lockhart. 2012. The impact of personalization on smartphone-based activity recognition. In *AAAI Workshop on Activity Context Representation: Techniques and Languages*.

Gary M Weiss, Jeffrey W Lockhart, Tony T Pulickal, Paul T McHugh, Isaac H Ronan, and Jessica L Timko. August 2014. Actitracker: A Smartphone-based Activity Recognition System for Improving Health and Well-Being. In *SIGKDD Exploration Newsletter, New York, NY, USA*.

Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. 2016. A survey of transfer learning. *Journal of Big Data* 3, 1 (2016), 1–40.

Gerhard Widmer and Miroslav Kubat. 1996. Learning in the presence of concept drift and hidden contexts. *Machine learning* 23, 1 (1996), 69–101.

Dwi H Widyantoro and John Yen. 2005. Relevant data expansion for learning concept drift from sparsely labeled data. *Knowledge and Data Engineering, IEEE Transactions on* 17, 3 (2005), 401–412.

Daniel H Wilson and Chris Atkeson. 2005. Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors. In *Pervasive computing*. Springer, 62–79.

Robert B Woodruff and Sarah Gardial. 1996. *Know your customer: new approaches to customer value and satisfaction*. Blackwell Business Cambridge, MA.

Wanhong Wu, Jiannong Cao, Yuan Zheng, and Yong-Ping Zheng. 2008. WAITER: A Wearable Personal Healthcare and Emergency Aid System. In *Proceedings of the 2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications (PERCOM '08)*. IEEE Computer Society, Washington, DC, USA, 680–685. DOI:http://dx.doi.org/10.1109/PERCOM.2008.115

Danny Wyatt, Matthai Philipose, and Tanzeem Choudhury. 2005. Unsupervised Activity Recognition Using Automatically Mined Common Sense. In *Proceedings of the 20th National Conference on Artificial Intelligence - Volume 1 (AAAI'05)*. AAAI Press, 21–27. http://dl.acm.org/citation.cfm?id=1619332.1619338

Jun Yang. 2009. Toward Physical Activity Diary: Motion Recognition Using Simple Acceleration Features with Mobile Phones. In *Proceedings of the 1st International Workshop on Interactive Multimedia for Consumer Electronics (IMCE '09)*. ACM, New York, NY, USA, 1–10. DOI:http://dx.doi.org/10.1145/1631040.1631042

Jun Yang, Hong Lu, Zhigang Liu, and Péter Pál Boda. 2010. Physical activity recognition with mobile phones: challenges, methods, and applications. In *Multimedia Interaction and Intelligent User Interfaces*. Springer, 185–213.

Liu Yang, Liping Jing, Jian Yu, and Michael K Ng. 2016. Learning transferred weights from co-occurrence data for heterogeneous transfer learning. *IEEE transactions on neural networks and learning systems* 27, 11 (2016), 2187–2200.

Yiming Yang, Jian Zhang, Jaime Carbonell, and Chun Jin. 2002. Topic-conditioned Novelty Detection. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '02)*. ACM, New York, NY, USA, 688–693. DOI:http://dx.doi.org/10.1145/775047.775150

Yi Yao and Gianfranco Doretto. 2010. Boosting for transfer learning with multiple sources. In *Computer vision and pattern recognition (CVPR), 2010 IEEE conference on*. IEEE, 1855–1862.

David Yarowsky. 1995. Unsupervised Word Sense Disambiguation Rivaling Supervised Methods. In *Proceedings of the 33rd Annual Meeting on Association for Computational Linguistics (ACL '95)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 189–196. DOI:http://dx.doi.org/10.3115/981658.981684

Juan Ye, Graeme Stevenson, and Simon Dobson. 2014. USMART: An Unsupervised Semantic Mining Activity Recognition Technique. *ACM Trans. Interact. Intell. Syst.* 4, 4, Article 16 (Nov. 2014), 27 pages. DOI:http://dx.doi.org/10.1145/2662870

Dit-Yan Yeung and C. Chow. 2002. Parzen-window network intrusion detectors. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, Vol. 4. 385–388. DOI:http://dx.doi.org/10.1109/ICPR.2002.1047476

Bingchuan Yuan and John Herbert. 2014. Smartphone-based Activity Recognition Using Hybrid Classifier. In *Proceeding of the 4th International Conference on Pervasive and Embedded Computing and Communication Systems*.

Daqing Zhang, Mossaab Hariz, and Mounir Mokhtari. 2008. Assisting Elders with Mild Dementia Staying at Home. In *Proceedings of the 2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications (PERCOM '08)*. IEEE Computer Society, Washington, DC, USA, 692–697. DOI:http://dx.doi.org/10.1109/PERCOM.2008.119

Zhongtang Zhao, Yiqiang Chen, Junfa Liu, and Mingjie Liu. 2010. Cross-mobile elm based activity recognition. *International Journal of Engineering and Industries* 1, 1 (2010), 30–38.

Zhongtang Zhao, Yiqiang Chen, Junfa Liu, Zhiqi Shen, and Mingjie Liu. 2011. Cross-people Mobile-phone Based Activity Recognition. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Three (IJCAI'11)*. AAAI Press, 2545–2550. DOI:http://dx.doi.org/10.5591/978-1-57735-516-8/IJCAI11-423

Vincent Wenchen Zheng, Sinno Jialin Pan, Qiang Yang, and Jeffrey Junfeng Pan. 2008. Transferring Multi-device Localization Models using Latent Multi-task Learning. In *AAAI*, Vol. 8. 1427–1432.

Yonglei Zheng, Weng-Keen Wong, Xinze Guan, and Stewart Trost. 2013. Physical Activity Recognition from Accelerometer Data Using a Multi-scale Ensemble Method. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI'13)*. AAAI Press, 1575–1581. http://dl.acm.org/citation.cfm?id=2891460.2891682

Yin Zhu, Yuqiang Chen, Zhongqi Lu, Sinno Jialin Pan, Gui-Rong Xue, Yong Yu, and Qiang Yang. 2011. Heterogeneous Transfer Learning for Image Classification. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI'11)*. AAAI Press, 1304–1309. http://dl.acm.org/citation.cfm?id=2900423.2900630

I. Zliobaite, A. Bifet, B. Pfahringer, and G. Holmes. 2014. Active Learning With Drifting Streaming Data. *Neural Networks and Learning Systems, IEEE Transactions on* 25, 1 (2014), 27–39.

Indrė Žliobaitė, Mykola Pechenizkiy, and João Gama. 2016. An overview of concept drift applications. In *Big Data Analysis: New Algorithms for a New Society*. Springer, 91–114.

Received Dec 2016; revised ; accepted

Table III: Comparison of Key Techniques for Activity Recognition

Criterion	Krishnan and Cook [Krishnan and Cook 2014]	Rashidi and Cook [Rashidi and Cook 2010]	Gomes et al. [Gomes et al. 2012b]	Do et al. [Do et al. 2013]	Lockhart and Weiss [Lockhart and Weiss. 2014]	Zahraa et al. [Abdallah et al. 2015]	Nguyen et al.[Nguyen et al. 2015]
Aim	Evaluate sliding window approaches for handling streaming data in activity recognition	Find sequential activity patterns in data streams using multiple time granularities	Personalise activity recognition model in data streams	Integrate stream reasoning with activity recognition techniques for recognising basic and complex activities	Build a personalised activity recognition system for real time recognition	Build an adaptive framework for personalised activity recognition in data streams	Recognise new activities using limited labelled data
Learning approach	Stream	Stream	Stream	Stream	Stream	Stream	Static (supervised)
Data collection platform	Smart home (binary motion sensors)	Smart home (binary motion sensors)	Environmental and on body sensors	Mobile	Mobile	Many (Generic)	Many (Generic)
Processing platform	PC	PC	Mobile phone	Mobile & Backend server	Backend server	Mobile	Not provided
Handling sensory data	Sliding window	Tilted window	Not provided	Sliding window	Fixed size sliding window	Fixed size sliding window	Fixed size sliding window
Classification technique	SVM	DVSM	Incremental NB, DT	ANN	Random Forest	Hybrid similarity measure approach	Feature-based and Attribute-based learner
Personalisation approach	No	No	Retrain	No	Retrain	Incremental	No
Discovering new activities	Partially	No	No	No	No	No	Partially
Forgetting abandoned activities	No	No	No	No	No	No	No
Incorporate active learning	No	No	No	No	No	Yes	No