Personicle: Personal Chronicle of Life Events

Laleh Jalali , Da Huo, Hyungik Oh, Mengfan Tang, Siripen Pongpaichet, Ramesh Jain
University of California, Irvine
{lalehj, huod, hyungiko, mengfant, spongpai, jain} @ics.uci.edu

ABSTRACT
This paper presents an overview and early results of our project to build a framework for integrating, aggregating, and analyzing heterogeneous personal data streams to build persona and use it to recognize evolving personal situation. Using increasing volumes of heterogeneous personal data streams, we recognize first movements and personal situations for providing individuals actionable information and insights. We demonstrate recognition of high-level life events through simultaneous use of asynchronous observations consisting of continuous GPS and accelerometer measurements. The main contributions of our framework will be to develop an architecture to integrate, store, and analyze data from heterogeneous data streams; detect low level physical activities using various unobtrusive sensors embedded in a mobile phone; and design a hierarchical classifier that identifies high-level life events using location context and physical activity and create a chronicle of life events called *Personicle*. We plan to collect test data sets and make them available publicly as a resource to other researchers.

1. INTRODUCTION
We are moving towards a *Universal Web* in which mobile communications, social technologies, Internet of Things, and sensors are connecting people and the physical world into one interconnected network [1]. Personal data plays a vital role in countless aspects of our everyday lives. Individuals are using their data for self-tracking as a form of accountability, to connect with other people with similar interests, and find relevant information and services. Physicians use health data to address public health issues, better diagnose illnesses, and develop new cures for diseases. Businesses are using a wide range of personal data to design new products, innovate new services, and improve targeted recommendations.

The promise of the internet of things (IoT) is extensive. The increasing adoption and decreasing cost of smartphones, sensors, and devices for monitoring physiological and physical activities has resulted in the generation of massive multisensory and multimodal observational data in many domains including health and wellness monitoring, safety monitoring, home rehabilitation, sensor surveillance for elderly care, etc. Mobile devices are becoming increasingly sophisticated and the latest generation of smart cell phones now incorporates diverse and powerful sensors. These sensors include GPS sensors, vision sensors (i.e., cameras), audio sensors (i.e., microphones), light sensors, temperature sensors, direction sensors (i.e., magnetic compasses), and acceleration sensors (i.e., accelerometers). The availability of these sensors in mass-marketed communication devices creates new opportunities for collecting, aggregating, and analyzing personal data. The information obtained from sensors varies in many respects. Methods to convert data to information and the reliability of the reported information could be entirely different for different sensors. Usually a sensor measures only a specific attribute of a user, which is one of many different independent or correlated attributes required to detect the overall state of user in the environment.

The act of observation performed by heterogeneous sensors creates an avalanche of data that must be integrated and interpreted in order to provide knowledge of the evolving situation. The primary challenges of analyzing personal data produced by a myriad of sensors are:

- Data is in the form of data streams.
- Most data is heterogeneous and multimodal.
- Data is used to derive high-level knowledge and evolving situation of an individual from low-level sensor observation.

As the number and ubiquity of sensors and mobile devices continue to grow, the need for computational methods to analyze the avalanche of heterogeneous sensor data for deriving situation awareness will grow and novel information processing architectures should be developed to enable easy handling of the produced data from different sources. In this paper we demonstrate such architecture and illustrate how to use objective data from heterogeneous sensors as a first step towards recognition of high-level events as they occur in daily life and create a detailed personal chronicle, called *Personicle*. More specifically, we investigate how recognition of low-level physical activities can be scaled to the recognition of high-level life events using data from desperate sources. Our experimental results suggest that it is feasible to recognize certain events such as office activities (*attending a meeting*), and everyday life events (*shopping*, *dining*) using sensors embedded in mobile phone and individual's calendar entries. In contrast to existing activity monitoring systems that typically require users to wear extra sensors, our platform leverages on-the-shelf smartphones. As a result, it is unobtrusive to users and enables the users to monitor and track their living routines on a daily basis, which is often difficult to realize in the conventional activity monitoring system. We demonstrate that it is possible to perform activity recognition with commonly available (nearly ubiquitous) equipment and yet achieve highly accurate results.

Interestingly, a large part of research in Human Activity Recognition (HAR) focuses on rather low-level and short-term activities [3-5]. However, in many applications ranging from healthcare and assisted living to modeling of human behavior, the analysis and recognition of high-level and longer-term activities is important. Let us briefly define the terms *low-level physical activity* and *high-level life events* as we understand them, since to the best of our knowledge there exists no generally accepted definition of these terms in the activity recognition and ubiquitous computing communities. As *low-level physical activities* we consider activities which can be characterized by a periodic sequence of body motions or postures such as *walking, sitting, standing, ascending stairs, descending stairs, jogging,* and *cycling*. *High-level life events*, on the other hand, are usually composed of a collection of low-level physical activities, and spans over a longer period of time, from several minutes to several hours. Such life events are *shopping, attending meeting, leisure, watching TV,* etc. Combining low-level physical activity with other user contextual information derives life events. Here by context we mean any common information about user herself or her environment that might be useful in determining the ongoing life events. For example in shopping event, user's location is mall/ store and activity is mostly walking and standing. Figure 1 shows translation process from raw data observations to low-level activity detection and finally high-level life event extraction and situation recognition.

The main goals of our proposed framework are: (i) providing an architecture to integrate, store, and analyze data from heterogeneous data streams; (ii) detect low-level physical activities using various unobtrusive sensors embedded in a mobile phone; (iii) design a hierarchical classifier that identifies high-level activities with simultaneous use of asynchronous observations consisting of GPS and accelerometer measurements to create a chronicle of life events named *Personicle*; (iv) tools to represent Personicle for a person and derive persona, and (v) visualizing and mining persona's to form societal models. In this paper we address only a partial set of these goals.

2. ARCHITECTURE

Translating low-level observation from personal data streams into high-level knowledge of a person's evolving situation is a perceptual task and a variety of enabling technologies should be incorporated. Figure 2 displays architecture of a multisensory evolving personal situation recognition system. This architecture has five layers: data ingestion, data processing, visualization, repository, and client. First, the data ingestion layer is composed of various sensors and sensor fusion component. The sensors ranges from smartphone, wearable activity tracking, calendar, to

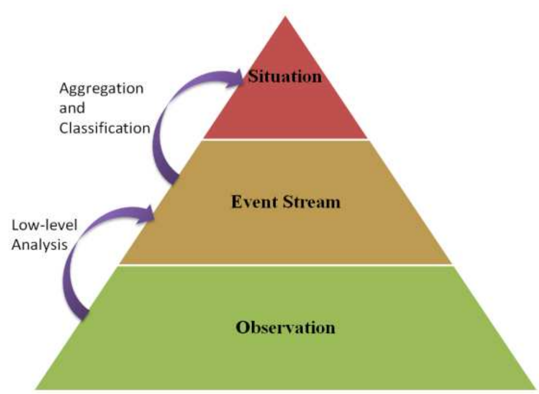

Figure 1. Translating low-level observation to knowledge

physiological sensors that track vital signs (e.g., heart rate, respiration rate, skin temperature, skin conductivity, ECG, etc.). More sensors can be added to improve the performance of our system. Client layer consists an expert and the person/user herself. User contributes to data collection and expert's knowledge is used to predefine risky situations for different application domains. For example in case of applying this framework to asthma management problem, medical expert defines a set of situations that might put asthmatic patient at risk (e.g. patient being vigorously active more than 15 minutes in a geographical location with high pollen count). System monitors patient continuously to detect these predefined situations. Also by analyzing history of patient's asthma attack over a period of time, correlations between attacks and local and global factors will be detected. In the following, major layers in the architecture are explained in more details.

2.1 Data Ingestion Layer

Sensor fusion is the combination of sensory data or data derived from disparate sources so that the resulting information is in some sense better than when these sources were used individually [2]. Using additional sensors, the amount of data to be processed will increase. This data can contain meaningful information and may even create a synergy effect when processed properly. With significant improvements in mobile sensor technology a new chapter in the era of self-aware devices and continuous data logging has just started. However, the sensor-packed smartphone can't be beneficial in isolation and shall be as a part of a larger platform. For life event recognition we leverage these types of sensors as they are providing unobtrusive and continuous monitoring, easily available in mass market, and aren't research only devices. Audio sensor in smartphone can be leveraged to detect complex events such as watching TV and sleep quality [26, 15]. High resolution cameras can be employed for food and calorie consumption logging [14]. GPS and acceleration sensors can also be used to track physical motion and location of users.

2.2 Processing Layer

Processing layer has two main components: Personicle creation; and Predictive data analytics. Personicle creation involves detecting low-level physical activities from activity tracking wearable sensors and unobtrusive sensors embedded in a mobile phone. Once user's motion is detected, it can be combined with other contextual information to create chronicle of life events. A chronicle means "a record or register of historical events in chronological order". Since we detect and store the abstracted life events of person through time, we call it Personicle. As data ingestion layer allows for accepting data streams from diverse sources, processing layer provides appropriate mapping and assimilation, as will discuss in the section 3 and 4, to convert this data to life events.

Life events thus detected are independent atomic elements of analysis for building persona in predictive data analytics component. Persona is the model of person. We will develop correlation and co-occurrence based environment in this component to detect recurring patterns for a person. This may result in insights such as sensitivity to a particular activity under specific climatic conditions or the effect of time and duration of exercise on one's sleep pattern and sleep quality. The goal is to provide a flexible and powerful analytics environment for personal data management.

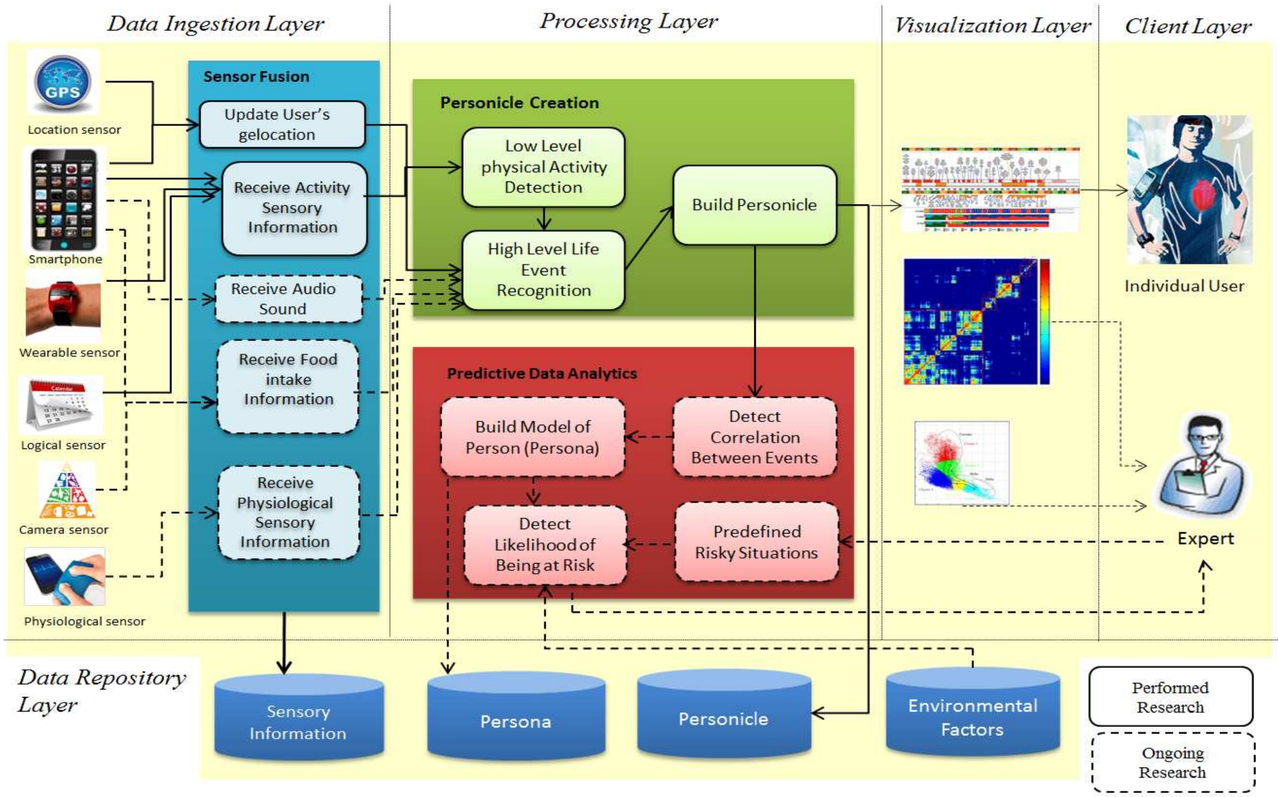

Figure 2. Architecture of multisensory personal evolving situation recognition system

Personicle and Persona such created have enormous potential to be used in medical diagnosis, population medicine, behavior modeling, and particularly in any field of knowledge that requires not only continuous monitoring but also acquiring insight about individual's life style. It doesn't take much imagination to figure out that Personicle is a game changer in health monitoring. Today's healthcare and medication system disregards individual's variability and promotes considerably more unnessacary medical testing and procedures. As the first use-case for our personal evolving situation recognition system, we chose asthma management. Section 4.3 describes this matter in more details.

2.3 Data Repository Layer

In data repository and management, we use the powerful NoSQL MongoDB[1] as our backend database to store raw data from different sources. Then the raw data is going through a preprocessing module to separate information content and unify temporal granularity of observations from disparate sources. Different sensors produce data at different time intervals, so it's important to have a temporal granularity conversion module. Data streams thus generated will be stored in database. Personicle as the output of processing layer will also be stored in form of an event steam. In order to provide rich queries base on Personicle, we are going to store Personicle to ElasticSearch[2] which is a flexible and powerful open source, real-time search engine.

[1] https://www.mongodb.org/

[2] http://www.elasticsearch.org/

3. LIFE EVENT RECOGNITION

High-level life event recognition has great potentials in areas such as medical diagnosis and human behavior modeling. However, most of the work in activity recognition has centered on identification of a specific type of activity in a particular scenario such as "opening/closing a window" or "ascending/descending stairs". Less effort has been applied to identification of more complex patterns of human activities and behavior, which extend over a long period of time. The explicit goal of this section is to enable the recognition of longer-term life events. Intuitively, recognition of low-level physical activities is a prerequisite to recognize more complex and high-level activities. We employed a multilevel representation that allows for explanation of multiple temporal granularities, by capturing different levels of temporal details. In the first level we use a classifier to extract low-level physical activity from raw accelerometer data. Also in the same level we process raw GPS measurements and map them to location categories. The result of these computations will feed as features to the second level where life event detection is performed.

Next we describe two phases of data acquisition. Then first level which includes low-level activity recognition and location category extraction is explained. Section 4 describes life event recognition and a use case of Personicle in healthcare domain.

3.1 Data Acquisition

An essential first step is to record an interesting and realistic dataset of high-level life events. Besides being tedious and time-consuming, the recording of high-level activities is a non-trivial task and data should be as realistic and representative as possible.

Since we couldn't find an open source dataset for detecting complex activities using wearable sensors and mobile phone embedded sensors, we had to create our own dataset. This dataset that we have collected and will continue to collect can serve as a resource to other researchers. Our data collection has two phases: phase one, collects accelerometer measurements related to low-level physical activities; phase two, collects both accelerometer and GPS measurements related to more complex life events.

3.1.1 *Phase One*
In order to collect data for our supervised learning task, it was necessary to have a number of participant carry an Android-based smart phone while performing certain physical activities. These activities are *walking*, *jogging*, *cycling*, *sitting*, and *standing*. Data collection was controlled by an application we created that executed on the phone. This application, through a simple graphical user interface, records user's name, start and stop of data collection, and a user defined label for the activity being performed. These activities are chosen since they are basic motions that people usually have in their daily living. For example, user is usually in walking or standing position while shopping or in sitting position while dinning. Accelerometer data is collected every 1sec, and a non-overlapping window with 10-second duration is defined to create the number of examples shown in table 1. The last row in Table 1 shows the percentage of the total examples associated with each activity. Please note that in our work the terminology of sitting and standing are different from most activity recognition areas. Here by sitting/ standing we refer to user being in sit/stand position and NOT the process of sitting/standing.

Table 1. Number of examples per user and physical activity

ID	Walk	Jog	Cycle	Sit	Stand	**Total**
1	11	2	3	5	19	40
2	391	21	0	23	34	469
3	112	5	2	0	0	119
4	14	35	0	214	12	265
5	15	0	0	79	0	94
6	50	0	18	3	5	76
7	198	0	201	11	7	417
8	507	0	0	4	2	513
9	276	18	744	5	11	1054
10	82	3	0	83	24	192
Sum	1656	84	968	427	115	3240
%	51.1	2.6	29.6	13.1	3.6	

3.1.2 *Phase Two*
In the second phase, two participants from the initial group were volunteered to collect labeled life event data using the same Android application while they engage in their daily routines. This time in addition to raw accelerometer data, GPS measurements are captured as well. Initially, events such as *attending a meeting*, *shopping*, and *dining* are considered. During all these activities, user should place the phone in her front pocket. In activity recognition community there is no common definition or structure of human activities that would allow us to formulate a clear and common problem statement (how a specific activity should be characterized). So inspired from events happening in their natural form, we asked participants to perform activities as described in table 2. We explicitly started with the recording of these life events and later automatically labeled the low-level physical activities that were performed during these life events with the prediction model described in section 3.2. Each participant recorded each life event four times in different days and in natural environment. We call a recording of life event from start time to end time a scenario. The length of each scenario varies between 30 to 90 minutes. The total length of data is 874 minutes.

Table 2. Description of three life events

Life Event	*Scenario Definition*
Dinning	• Walking into a restaurant • Standing by the reception desk • Sitting behind the table (at least 10 minutes) • Walking out of the restaurant
Shopping	• Walking into a store/mall • Strolling through the store • Waiting in the line to pay • Walking out of the store
Meeting	• Enter meeting information in Google calendar • Walking to the meeting room • Sitting/Standing behind the desk

3.2 Low-level Activity Recognition
The topic of accelerometer-based activity recognition is not new. Bao & Intille [3] developed an activity recognition system to identify twenty activities using bi-axial accelerometers placed in five locations on the user's body. Additional studies have similarly focused on how one can use a variety of accelerometer based devices to identify a range of activities [4, 5]. Our work differs from most prior work in that we use a commercial mass-marketed device rather than a research-only device; we use a single device conveniently kept in the user's pocket rather than multiple devices distributed across the body. Also we are interested in physical activities such as walking, jogging, sitting, standing, and cycling that person is performing while engaging in their daily routines. Figure 3 plots raw accelerometer data for a typical user, for all three axes and for the activities concatenated together. It is clear that sitting and standing have distinctive patterns, based on the relative magnitudes of the x, y, and z, values. Walking and jogging activities involve repetitive motions and exhibit periodic behavior. Note that for most activities the y values have the largest accelerations. This is the contribution of the force of gravity, which causes the accelerometer to measure a value of 9.8 m/s^2 in the direction of the Earth's center. For all activities except sitting this direction corresponds to the y axis.

In order to obtain useful information from raw accelerometer sensor data we should be able to accurately discover the characteristics or *features* of the signal coming from a given sensor. To accomplish this we divided data stream into time

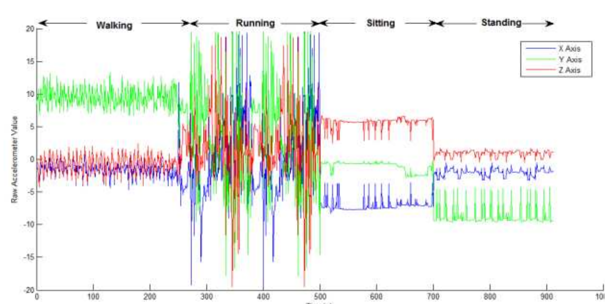

Figure 3. Plot of a 250-seconds of raw acceleration value for each activity

segments and then generated features that were based on the raw data readings contained within each time segment. We refer to the duration of each segment as the window length. Each segment can span from seconds to minutes and depending on the activity type, it provided sufficient time to capture several repetitions of the (repetitive) motions involved in that activity. Although we have not performed experiments to determine the optimal example duration value, we did compare the results for a 10-second, 30-second, and the 60-second window length and didn't find a major difference. In this paper we are extracting 22 features from both *time* domain and the *frequency* domain over each window. These features are

- Mean, Std. Deviation, Median, Range, Max, Min (for all three x, y, and z axes)
- Root Mean Square
- Average Resultant Acceleration: Average of the square roots of the sum of the values of each axis squared $\sqrt{x_i^2 + y_i^2 + z_i^2}$
- Spectral Energy: The sum of the squared discrete FFT component magnitudes of the signal. The sum was divided by the window length for normalization
- Information Entropy: The normalized information entropy of the discrete FFT component magnitudes of the signal.

For each window, derived features and the label assigned to them is called a data record. Table 3 summarizes the performance of five classifiers for low-level physical activity recognition task. In most cases we can achieve high level of accuracy. Our results indicate that there isn't a significant change between different window lengths and none of the five learning algorithms consistently performs best, but the random forest does perform best overall. Confusion matrix associated with this algorithm is presented in Tables 4.

Table 3. Summary of classifier results (% of records correctly predicted) using 10-fold cross validation over different window lengths.

Classifier	*10-second*	*30-second*	*60-second*
Naïve Bayes	83.27	85.10	85.69
J48	93.70	94.95	94.73
Random Forest	96.11	95.87	97.26
Logistic Regression	87.36	83.40	82.63
Multilayer Perceptron	91.39	94.49	96.42

Although there are very few examples of jogging, we can still identify this activities quite well because it involves more extreme changes in acceleration. Also sitting and standing are predicated with high accuracy since the accelerometer data associated with these activities has a significant pattern as shown in figure 3.

3.3 Location Tracking and Place Categories

GPS tracking generates a huge amount of geographic data which is tricky to handle in its raw form, and requires extraction of

Table 4. Confusion matrix of Random Forrest for 10-second windows

Actual Class	***Predicted Class***				
	walking	cycling	jogging	sitting	standing
walking	**1383**	22	0	11	3
cycling	23	**739**	0	0	0
jogging	5	0	**18**	0	0
sitting	7	0	0	**424**	0
standing	6	0	1	0	**93**

activity locations. GPS readings are the input to our model—a typical trace consists of approximately one GPS reading per second; each reading is a point in 2D space. We use reverse geo-coding to assign location names to raw GPS measurements. Reverse geocoding is the process of back (reverse) coding of a point location (latitude, longitude) to a readable address or place name. Our approach does not return actual addresses, only estimates of what should be there based on the predetermined range. We use the same windowing technique discussed in section 3.2. So for each GPS measurement within a window, our algorithm returns a list of possible locations. We then look at the temporal sequence of recorded locations and use a set of decision rules based on distance and time to identify place name for each window. We have created a taxonomy of place categories from Foursquare [9] which then will be used to assign place category to each approximate location. In the next section we describe how we used place categories to determine life events such as shopping, dining, and meeting.

4. KNOWLEDGE DISCOVERY

The final goal of this system is to decompose the temporal sequence obtained from the sensors in real-time into concepts at different levels of abstraction or temporal granularity. The lowest analysis level is more sensitive to raw observations and we might need to re-train this level to obtain best results for each user. Since there isn't a significant change between different window lengths at first level, we chose windows of 10-second to create the new feature vector for second level. The feature vector for the second level results from the concatenation of low-level activity and location categories. Since life events duration expands from several minutes to several hours, classification function at the second level can produce the output at a lower temporal granularity. We chose 1 minute for window length of second level classifier.

4.1 Results

An important source of influence on the recognition performance is the classifier. In this section we evaluate and compare the following classification techniques: Support Vector Machine (SVM), Naïve Bayse (NB), Decision Tree (J48), and k-NN with k=3. For evaluation we compare two schemes, person-dependent vs. person-independent leave-one-scenario-out cross validation. For person dependent evaluation, for each participant, we leave one scenario of data recording out for testing and use the rest of the scenarios of the same participant for training. In this case, the model in the first level was also trained based on this specific participant's data. For the person-independent case, we train on all scenarios of one participant and test on all scenarios of the second participant. Also first level classifier was trained with all participants' data. In both cases, overall recognition performance is calculated as the average performance across all cross validations.

Figure 4 summarizes the recognition performance achieved using different classifiers. As can be seen from the figure, in person-dependent case, best results (precision= 94.7, recall= 84.5) achieved by SVM and NB exhibits the worst results (precision= 71.2, recall= 65.6). In person independent case, k-NN and J48 classifiers achieved better precisions. Despite the fact that results are not conclusive between different classifiers, person-dependent case lead to better overall performance.

Figure 5(a) shows Personicle visualization for a 24-hour period for one participant. In some cases that we didn't have enough contextual information, life event is either unknown or low-level physical activity was shown instead. We also aggregated our life event stream with *"transport"* activity from Moves[3] mobile application. As demonstrated in the figure, Personicle's life events are color coded. We used Google API timeline visualization for this purpose. Discarding unknown activities, distribution of Personicle for 24-hour period is shown as a pie chart in figure 5(b).

4.2 Asthma Management: A Potential Use Case of Personicle

Initially we plan to create Personicle for real-world application domains such as healthcare and long-term behavioral analysis. And in healthcare domain we chose asthma management problem. An estimated 235 million people worldwide suffer from asthma. There are more than 250,000 asthma-related fatalities each year [6]. Asthma is a long-term disease that has no cure; so managing asthma requires some life style changes. In our previous work [10] we suggested the concept of Focused Micro Blogs (FMBs). FMBs are on a focused topic and are targeted to a specific source. Since topic is fixed, the information in a FMB could be easily structured and parsed. Different smartphone apps can be a source of FMB to assist its owner in different situations. We have developed an asthma management application that collects user initiated information such as time, location, and severity of asthma attacks. This information is essentially an FMB that feeds to our system and represent an *asthma attack* life event. Also the app collects accelerometer and GPS measurements continuously in the background and has access to user's calendar information. Using a pre-trained life event recognition model described in section 3 and 4, we create Personicle for each user. It's worth mentioning that the data collection is extremely unobtrusive and doesn't impose any burden on user. Personicle thus created has enormous impact on managing the disease. Since this is a work in progress, asthma attack FMBs were not completed yet to be incorporated into our current Personicle visualization in figure 5.

For asthmatic patients the effect of meteorological factors such as pollen count, Air Quality Index, temperature, wind speed, and air pollutants is well known. As shown in our architecture, these factors combined with Personicle can be used to detect risky situations for asthmatic patients. In predictive data analytics component, for each user we build a model (persona) that reflects

[3] http://www.moves-app.com/

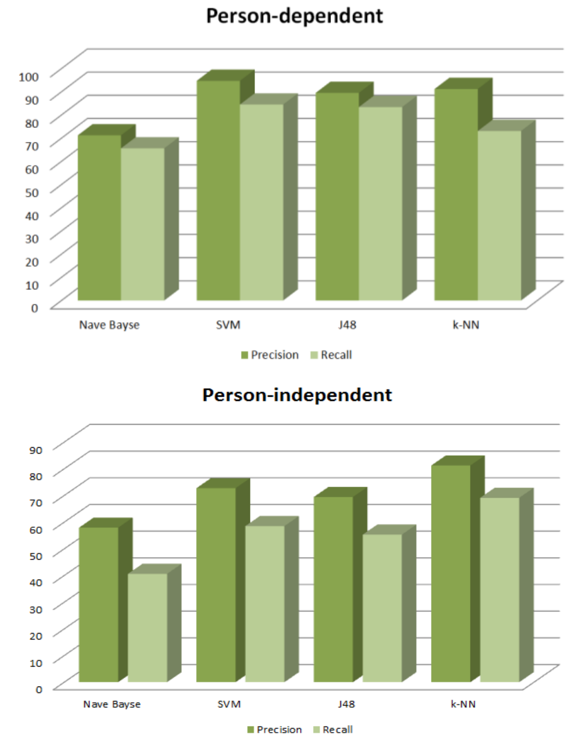

Figure 4. Life event recognition performance for different classifiers

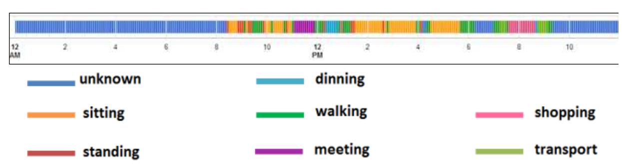

Figure 5(a). Personicle visualization

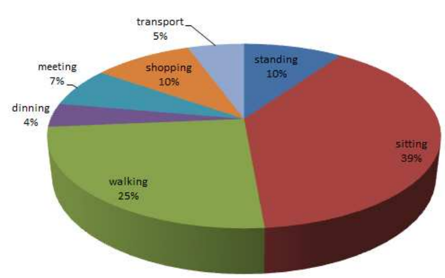

Figure 5(b). Distribution of life events in one day Personicle

the association between asthma attacks and local (physical activity) and global triggers. For example one might get asthma attack in spring and fall only by being exposed to high pollen count while someone else's asthma attack might occur while she is jogging or exercise outdoor in presence of certain air pollutions.

5. RELATED WORK

Until a few years ago, computational facilities limited data collection and processing to very narrow aspects of the world. A good example is the bold movement of Enterprise Data Warehouses towards the end of the last century. These systems collected data related to limited aspects of operations of an enterprise and tried to gain insights and understanding to create business intelligence. Around the same time one saw beginnings of stream processing and Complex Event Processing (CEP) that dealt with a number of data streams in an organization. In less than two decades, the situation has changed dramatically. Now we are creating a planetary data warehouse that involves massive number of heterogeneous data streams that must be analyzed in real time for predicting evolving situations and controlling them. Several data streaming management systems (DSMSs) [11 - 13] provide a declarative language like SQL, but complex event recognition is more naturally expressed using the operational flow of an imperative language. Thus, complex event processing (CEP) systems provided more promising approach and gained a lot of interest. Commercial vendors, such as IBM [16], Oracle [17], Tibco [18], Coral8 [19] and StreamBase [20], have built event processing systems with functionalities of both DSMS and CEP systems. Nonetheless, in traditional CEP systems, events are defined as a "change of state," when a measurement exceeds a predefined threshold for a specific value, e.g. when body temperature exceeds *100.8 °F* an event triggers. This approach is suitable for recognizing anomalies or burst detection but lacks the capability to semantically interpret and characterize data to recognize an evolving situation. Our framework focuses on occurrences that spans over a time interval and we intend to define and detect personal situation by considering a hierarchy of operators that can be unambiguously interpreted on life events and in general on Personicle.

Since we are building a chronicle of life events, activity recognition is an inevitable part of our system. Complex activities are composed of a collection of simple activities and may consider contextual information such as time, interaction between people, and interaction with objects. Tao Gu et al. [27] build activity models by mining a set of emerging patterns from a sequential activity trace and used them to recognized sequential, interleaved, and concurrent activities. Hamid et al. [21] represent activities as bags of n-grams, cluster them into classes of activities, and characterize these classes by frequently occurring sequences. The patterns they discover on a set of 150 days of a person's indoor location traces are coarse and relatively difficult to interpret. In a more office and desktop-centered setting, [23] combine device usage with calendar data and time of day/time of week information to infer a user's availability. The work of Dearman et al. [24] shows that it is possible to utilize Yelp reviews to identify potential activities. As opposed to defining activity classes a priori they parsed for nearby verb-noun pairs to extract activity descriptions (e.g. buy book, appreciate art) as they are encountered in the textual data. There is a significant amount of work that uses location sensors to extract high-level information about a person's activities. Routinely visited locations such as home, work, or school can indicate pursued activities such as leisure, working, or picking up someone [25]. These works show that location is a powerful cue to the high-level structure of daily life. However, location is often not enough to identify daily routines reliably, as different activities can be performed at the same location. As an example at a mall, many people are shopping or they might be dinning. Similarly, in an office room one might work, hold meetings and even give presentation. Therefore, the work that we describe in section 4 is complementary to these approaches and the use of accelerometers allows detection of more fine-grained activities and can also account for different activities performed at the same location. Literature is abundant with activity recognition techniques from video signals [28-30]. Our work is different since we are using heterogeneous multi-modal sensors in real-world scenarios.

6. CONCLUSION AND FUTURE WORK

In this paper we presented a multisensory evolving personal situation recognition system with focus on creating Personicle, a chronicle of life events. Different layers of system's architecture for personal sensor data collection and transformation is demonstrated. A multilevel representation is employed which allows for explanation of multiple temporal granularities, by capturing different levels of temporal details. The experimental results suggest that recognition of high-level life events can be achieved through simultaneous use of observations consisting of GPS and accelerometer measurements. As a use case we considered asthma management to explore the power of Personicle in transformation of raw data to personal knowledge in a specific health application domain. Where our system really makes a difference is by gathering anonymous asthma data to help understanding the causes and external correlations of asthma and thereby, giving researchers visibility to find correlation between higher asthma rates in one specific vicinity, time, date, pollutant, and climate. The ability to gather this type of data, especially in real-time is unprecedented and we hope it would have a great impact in advancing asthma research.

Personicle is an event stream which demonstrated as a timeline of all events happened in person's life. The purpose of the timeline is to provide a visual tool for looking at events across a relatively long period of time and identify patterns and interrelationships involving a broad range of factors. Identification of patterns is particularly important when attempting to understand the needs of user and offering motivating insights for promoting a healthy lifestyle. The emerging field of visual analytics [7] focuses on integrating human judgment and visual representations to process heterogeneous and dynamic volumes of information.

It has already been demonstrated that insights from individuals can be aggregated for the society to gain societal insights. For example Google Flu Trends (GFT)[4] is a web service provided by Google uses aggregated Google search data to estimate flu outbreak. Though there are a lot of critiques why GFT failed [22], the fact remains that we need platforms to collect and correlate accurate individual data and more importantly, aggregate these data for gaining community insights. One solution relies on the power of FMBs. By using mobile phone applications designed for specific purposes (e.g. asthma management app), individuals can contribute information to a situation recognition system [8] and in return get a sense of situation at societal level.

7. REFERENCES

[1] D. Marc, R. Martinez, C. Kalaboukis, "Rethinking Personal Information – Workshop Pre-read." *Invention Arts and World Economic Forum*, June 2010.

[2] W. Elmenreich, "Sensor fusion in time-triggered systems." In *PhD Thesis*, Vienna University of Technology, Vienna, Austria, 2002.

[3] B. Lin, and S. S. Intille, "Activity recognition from user-annotated acceleration data." In *Pervasive Computing*

[4] https://www.google.org/flutrends/

Lecture Notes in Computer Science, Volume 3001, pp. 1-17. Springer Berlin Heidelberg, 2004.

[4] G. Lei, A. K. Bourke, J. Nelson, “Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems.” In *Medical engineering & Physics*, 2014.

[5] T. Choudhury, S. Consolvo, B. Harrison, J. Hightower, A. LaMarca, L. LeGrand, A. Rahimi, A. Rea, G. Borriello, B. Hemingway, P. Klasnja, K. Koscher, J. A. Landay, J. Lester, D. Wyatt, D. Haehnel, “The mobile sensing platform: An embedded activity recognition system.” In *IEEE Pervasive Computing*, Volume 7, Issue 2, pp. 32-41, 2008.

[6] “World Health Organization, Chronic Diseases and Health Promotion.” *http://www.who.int/respiratory/asthma/en/.*

[7] D. A. Keim, F. Mansmann, J. Schneidewind, J. Thomas, H. Ziegler, “Visual Analytics: Scope and Challenges.” In *Visual Data Mining Lecture Notes in Computer Science,* Volume 4404, pp. 76-90, 2008.

[8] M. Gao, V.K. Singh, R. Jain, “Eventshop: from heterogeneous web streams to personalized situation detection and control.” In *Proceedings of the third Annual ACM Web Science Conference,* ACM, 2012.

[9] Place Categories, http://aboutfoursquare.com/foursquare-categories/.

[10] R. Jain, L. Jalali, M. Fan., “From health-persona to societal health.” In *Proceedings of the 22nd international conference on World Wide Web companion,* pp. 1329-1334, 2013.

[11] T. Bernhardt, A. Vasseur, “Esper: Event stream processing and correlation.” In *http://www.onjava.com/lpt/a/6955.*

[12] D.J. Abadi, D. Carney, U. Cetintemel, M. Cherniack, C. Convey, S. Lee, M. Stonebraker, N. Tatbul, S. Zdonik, “Aurora: a new model and architecture for data stream management.” In *The VLDB Journal,* Volume 12, Issue 2, pp. 120-139, 2003.

[13] S. Chandrasekaran, O. Cooper, A. Deshpande, M. J. Franklin, J. M. Hellerstein, W. Hong, S. Krishnamurthy, S. R. Madden, F. Reiss, M. A. Shah, “TelegraphCQ: continuous dataflow processing.” In *Proceedings of the 2003 ACM SIGMOD international conference on Management of data,* pp. 668-668, ACM, 2003.

[14] K. Aizawa, “Multimedia FoodLog: Diverse Applications from Self-Monitoring to Social Contributions.”, In *ITE Transactions on Media Technology and Applications,* Volume 1, Issue 3, pp. 214-219, 2013.

[15] D. Majoe, P. Bonhof, T. K. Trachsel, J. Gutknecht, L. Widmer, “Stress and sleep quality estimation from a smart wearable sensor.” In fif*th International Conference on Pervasive Computing and Applications,* pp. 14-19, 2010.

[16] A. Biem, E. Bouillet, H. Feng, A. Ranganathan, A. Riabov, O. Verscheure, H. Koutsopoulos, C. Moran, “IBM infosphere streams for scalable, real-time, intelligent transportation services.” In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data,* pp. 1093-1104, ACM, 2010.

[17] Oracle Event Processing, http://www.oracle.com/technetwork/middleware/complex-event-processing/overview/index.html.

[18] Tibco Complex Event Processing, http://www.tibco.com/products/event-processing/default.jsp

[19] J. Morrell, S. D. Vidich, “Complex Event Processing with Coral8”, White paper, 2008.

[20] Stream Base Technical Documentation, *http://www.streambase.com.*

[21] R. Hamid, S. Maddi, A. Johnson, A. Bobick, I. Essa, C. Isbell, “Unsupervised discovery and characterization of activities from event-streams.” In arXiv preprint arXiv:1207.1381, 2012.

[22] D. M. Lazer, R. Kennedy, G. King, A. Vespignani, “The Parable of Google Flu: Traps in Big Data Analysis.” In *http://dash.harvard.edu/handle/1/12016836,* 2014.

[23] E. Horvitz, P. Koch, C. M. Kadie, A. Jacobs, “Coordinate: Probabilistic Forecasting of Presence and Availability.” In *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence,* pp. 224-233, ACM, 2002.

[24] D. Dearman and K. N. Truong, “Identifying the activities supported by locations with community-authored content.” In *Proceedings of the 12th ACM international conference on Ubiquitous computing,* pp. 23-32, ACM, 2010.

[25] L. Liao, D. Fox, and H. Kautz, “Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields.” In *The International Journal of Robotics Research,* Volume 26, Issue 1, pp. 119-134, 2007.

[26] Y. Zhan, T. Kuroda, “Wearable sensor-based human activity recognition from environmental background sounds.” In Journal of Ambient Intelligence and Humanized Computing, Volume 5, Issue 1, pp. 77-89, 2014.

[27] T. Gu, Z. Wu, X. Tao, H. K. Pung, J. Lu, “epsicar: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition.”, In *Pervasive Computing and Communications,* IEEE, 2009.

[28] N. Bingbing, G. Wang, and P. Moulin. "Rgbd-hudaact: A color-depth video database for human daily activity recognition." *Consumer Depth Cameras for Computer Vision*. Springer London, pp. 193-208, 2013.

[29] J. M. Chaquet, E. J. Carmona, and A. F. Caballero. "A survey of video datasets for human action and activity recognition." *Computer Vision and Image Understanding*, Volume 117, Issue 6, pp. 633-659, 2013.

[30] R., Neil, and I. Reid. "A general method for human activity recognition in video." *Computer Vision and Image Understanding*, Volume 104, Issue 2, pp. 232-248, 2006.