From Multimedia Logs to Personal Chronicles

Hyungik Oh
University of California, Irvine
Irvine, California
hyungiko@uci.edu

Ramesh Jain
University of California, Irvine
Irvine, California
jain@ics.uci.edu

ABSTRACT

Multimodal data streams are essential for analyzing personal life, environmental conditions, and social situations. Since these data streams have different granularities and semantics, the semantic gap becomes even more formidable. To make sense of all the multimodal correlated streams we must first synchronize them in the context of the application, and then analyze them to extract meaningful information. In this paper, we consider the problem of modeling an individual by using daily activity in order to understand their health and behavior. The first step is to correlate diverse data streams with atomic-interval, and segment a person's day into her daily activities. We collect the diverse data streams from the person's smartphone to classify every atomic-interval into a daily activity. Next, we use an interval growing technique for determining daily-activity-intervals and their attributes. Then, these daily-activity-intervals are labeled as the daily activities by using Bagging Formal Concept Analysis (BFCA). Finally, we build a personal chronicle, which is a person's time-ordered list of daily activities. This personal chronicle can then be used to model the person using learning techniques applied to daily activities in the chronicle and relating them to biomedical or behavioral signals. We present the results for this daily activity segmentation and recognition by using lifelogs of 23 participants.

CCS CONCEPTS

• Information systems → Mobile information processing systems; **• Human-centered computing → Ubiquitous computing**; **Mobile computing**; **Ambient intelligence**; **Smartphones**;

KEYWORDS

Activity recognition; activity segmentation; personal chronicle; lifelogging; multimodal data streams; event modeling; mobile sensing;

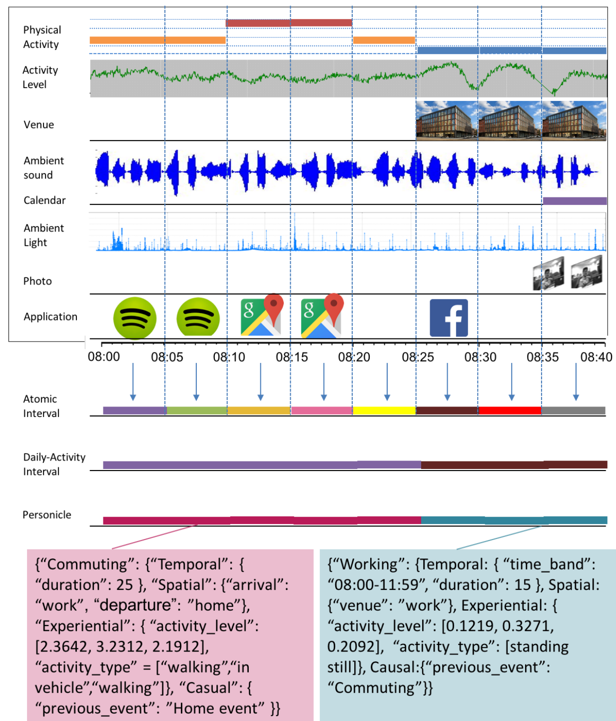

Figure 1: From multi-modal sensor data streams to atomic-interval, daily-activity-interval and chronicle of daily activities.

1 INTRODUCTION

Understanding the daily lives of human beings, what people have experienced, how people have spent their time, when and where they have been and whom they have been with, has long been the subject of scientific inquiry. This interest has led people in the field of multimedia to develop scientific approaches to monitoring and analyzing personal lifestyles and behavioral patterns. Multimedia researchers have tried to extract semantic level information from visual content so that they can analyze people's lives, and even environmental conditions and social situations. They also have analyzed real-time behavior data, which is collected via wearable devices, such as smartphones or smartbands, and social media, to understand more about personal lifestyles and behavioral patterns. However, recognizing people's daily lives at higher cognitive and more abstract levels (e.g., working, exercising, shopping, or relaxing) than low-level multimedia lifelogs (e.g., step count, GPS, venue, or physical activity), which makes inferring and predicting people's lifestyles more intuitive, remains relatively undeveloped.

Advances in sensor technology have increased the number of quantitative and qualitative multimedia lifelogs that are captured via wearable devices. Thus, we can now automatically aggregate and analyze heterogeneous multimedia data streams. Since these data streams have different granularity and semantics, the data

MM '17, October 23–27, 2017, Mountain View, CA, USA

ACM ISBN 978-1-4503-4906-2/17/10…$15.00
https://doi.org/10.1145/3123266.3123375

streams need to be correlated by synchronizing them in the context of the application. The synchronized data streams can then be raised up to higher-level forms, so-called daily activity, by analyzing relationships between the daily activity and their temporal, causal, spatial, experiential, informational, and structural aspects [35]. Finally, the personal chronicle of the daily activity can be generated by chronologically ordering the recognized results [21]. In this paper, we automatically recognize these daily activities using multimodal data streams from each individual's smartphone. Figure 1 shows the steps of our recognition approach: collecting multimedia lifelogs, synchronizing and segmenting the data streams, recognizing daily activities, and generating the personal chronicle.

We consider the problem of modeling an individual to ultimately help them with personalized health management. We believe that objectively understanding the daily activities of human beings has a strong potential to improve health research, given that these daily activities and the sequences of these high-level data abstractions contain their life experiences, behavioral patterns, and even their feelings. According to Kahneman et al., quantifying information about time usage and its frequency, as well as stress level, pleasure, and other affective states of each individual user, is potentially useful for health research [23]. More specifically, they tried to find this information by identifying each person's daily activity. Thus, the authors first conducted a survey categorizing people's common daily activities, and then described how to quantify them. Jain and Jalali's research on objective self models has also shown that analyzing the personal chronicle of daily activities can be used to build sophisticated models that will help the monitoring of individual health and building disease models [20]. They built a complete infrastructure for the objective self, but it has not yet had actual implementations and experimental validations. Thus, with the same goals in mind, we recognize Kahneman's common daily activities and generate personal chronicles of the daily activities in order to build objective self models.

To automatically quantify the daily activity of each individual, the recognition method should be unobtrusive and effortless, and user tracking should only use common devices. More importantly, we should not intervene in users' life patterns by pushing them to do something or putting them in specific situations in order to recognize their daily activity. However, one major technical challenge is that this sort of fully-automated tracking is not always a guarantee of high recognition accuracy [9]. Some daily activities might require more diverse features than current smartphone sensors, and some others might be user-dependent or subjective daily activities, which need user feedback for personalization. This paper describes how to overcome these challenges for fully-automated tracking and explores to what extent Kahneman's daily activities can be recognized.

Our approach begins with understanding daily-activity-intervals by classifying every atomic-interval into a daily activity. We propose the idea that daily activities in a time-line are similar to objects in two-dimensional pixel space in that both the daily activities and objects are determined by a correlation between the times/pixels. We first collect multimedia logs via each individual's smartphone. Then, the collected logs are used to segment a person's day into their daily activities. We use diverse data streams from the person's smartphone to classify every atomic-interval into a daily activity. Next, we use interval growing techniques for determining daily-activity-intervals and their attributes. Then, these intervals are labeled as the daily activities by using Bagging Formal Concept Analysis (BFCA). Finally, we build a personal chronicle represented as events.

We believe that recognizing atomic-level daily activities, which can be automatically recognized, is one important step towards higher-level activity recognitions. Our main contribution in the area of activity recognition is 1) revealing and quantifying these atomic-level daily activities with our automated and unobtrusive approach, and 2) increasing the possibility of automatically recognizing the higher cognitive daily activities, and thus 3) quantifying the personal chronicle of these daily activities, as in Figure 1.

2 RELATED WORK

Research on human behavior analysis is not a new area. It has been around for decades in many different forms. In 1945, Vannevar Bush's "Memex" vision had already presented a systematic approach, which organized a person's life-time knowledge, such as books, records and communication, by providing a user-authored data store, its linkages, and labels of the data to understand personal experiences [6, 16]. However, the capability to greatly develop this vision has recently become possible with advancements in technology [16]. The significant advances in computer storage, processing power, sensing technology, and network systems have encouraged researchers to participate in the field of human behavior recognition [20]. Classification techniques have also contributed to the recognition of higher-level semantics, such as physical activity [3, 17, 24, 36], more so sensor measurements.

There have been several data-driven studies analyzing the contexts or lifelogs of each individual user. A. K. Dey devised an architecture named Context Toolkit, which would allow the combination of data resulting in an abstraction that can be used to better understand how people experience the real-world [1, 10]. They provided higher-level contexts by aggregating and interpreting lower-level contexts in the conceptual framework. Since the Context Toolkit was introduced in 2001, the agent for human sensing has moved from the computer-based toolkit to mobile/wearable sensor-based loggers [5, 7, 12, 18]. With the trend of using the Internet-of-Things for data-driven studies, the so-called lifelogging, which is focused on a process of pervasively collecting, processing, and reflecting on an each individual's life experience data, has become more popular [16]. For example, Gordon Bell recorded many aspects of his everyday life by capturing a series of real-world images using the wearable camera, called SenseCam [14, 15], for the purpose of aiding recollection of past experiences. He expected that "total capture" of daily life would lead to "total recall" of our lives [4, 29]. P. Wang and A. F. Smeaton have also highlighted the importance of visual lifelogs because they identify various semantic concepts across individual subjects. They automatically identified high-level human activities such as eating, drinking, or cooking using SenseCam images, and data models [32–34].

Many human activity recognition systems have been based on situation specific capture. MIT's "PlaceLab" installed hundreds of sensors in all parts of a home seeking to automatically record activities [29]. Kasteren et al. collected location data and voice labeled annotations for each activity, such as breakfast, sleeping, or toileting, from the house. They constructed a probabilistic model using a hidden markov model (HMM) to predict future sensor readings [31]. Research on situation specific capture has drawn much attention in Activity of Daily Living (ADL) recognition. To automatically recognize ADL (e.g., toileting, grooming, bathing, showering or sleeping, etc.) for the purpose of preventative medical monitoring or building a smart home, researchers have set up low-cost sensors at critical locations in a home [8, 13, 19, 27, 30, 31] and then have predicted activities using naive bayes classifier [26, 30], HMM [31], ontologies and semantic reasoning [8], and Formal Concept Analysis [27], etc. Luštrek et al. used smartphone data, such as location (GPS), physical activity (accelerometer), and sound, and combined machine learning algorithms and symbolic reasoning to recognize high-level activities of a diabetic patient [26].

Another research group seeks to segment events on a lifelog of images. Doherty and Smeaton extract MPEG-7 features from images, such as accelerometer sensor values, light-level, ambient temperature, and passive infrared, and compare the similarity to those of adjacent images for the purpose of event segmentation [11]. There is another group who plans to recognize Kahneman's daily activities by analyzing taken photos from smartphones [2]. However, to the best of our knowledge, there is no approach for daily activity recognition that begins with understanding physical activity patterns by using non-visual smartphone lifelogs, and then gradually finding daily-activity-intervals in order to recognize daily activity. Moreover, we have not seen any approach to identify atomic-level daily activities to recognize higher cognitive and more abstract levels.

3 METHODOLOGY OVERVIEW

In this section, we describe our overall methodology for recognizing daily activity. We first explain what daily activity is, and then finalize the target corpus of the daily activity. Next, we categorize the daily activity corpus into three levels that describe their characteristics in terms of recognition possibility. Lastly, we provide the definitions of each daily activity. Since we ask our participants to label their daily activities with the exact name of that moment, we must synchronize the exact meaning of each daily activity. We refer to dictionaries, such as the Oxford and Cambridge English Dictionaries, and modify the meanings to match our contexts. We explained these definitions to each participant, and encouraged them to correctly label their daily activities according to the definitions.

We consider that daily activity is a brief name for each episode, such as "commuting to work" or "eating lunch", that can generally happen in the daily lives of human beings. Thus, we think that the continuous series of the daily activities can imply the person's lifestyle, behavioral patterns, and even their feelings. Kahneman et al. have also insisted that quantifying these daily activities would potentially be useful for research on human well-being. Furthermore, they have tried to categorize common daily activities by conducting a survey, and suggested 16 common daily activities. We refine our daily activities into Kahneman's daily activity corpus, which has already been verified for human well-being research [23].

Table 1: Kahneman's daily activity on concept levels

Level 1	Level 2	Level 3
Still	Working	Watching TV
Walking	Commuting	Preparing food
Running	Exercising	Socializing
Cycling	Religious event	Housework
Driving	Shopping	Intimate relations
Direct communication	Eating	Relaxing
Remote communication	Using toilet	Taking a break
On the smartphone	Home event	Sleeping

We classify Kahneman's common daily activity in three levels. The level definitions are as follows:

- **Level 1 (L1):** a daily activity which can be automatically recognized. It can be seen as the atomic-level.
- **Level 2 (L2):** a daily activity which has the possibility of automatic recognition in the near future using sensing technology, but can not yet be recognized.
- **Level 3 (L3):** a daily activity which is not possible to be automatically recognized, but is soon to be recognized once richer data is gathered. We also deem subjective or user-dependent daily activities as level 3.

Since there are limits and restrictions on smartphone-based recognition, such as the lack of sensor data, or difficulties in understanding user-dependent or subjective activities, we think that it is not possible to recognize all the daily activities at the current stage. Our approach is to focus on recognizing daily activities, which can be automatically recognized via smartphone first (atomic-level), and then gradually try to recognize the daily activities which have a high possibility of automatic recognition (L2). Once the daily activity is recognized, we start considering that activity is the atomic-level activity, and using it as an attribute for other daily activity recognitions. Table 1 is the classification of Kahneman's daily activities in these three levels. In this paper, we refine the target activities to L1 and L2 activities, and see to what extent L2 activities can be automatically recognized.

We define Kahneman's L2 activities based on their dictionary definition. People might have different definitions for each daily activity. Thus, we provide them with the following general definitions for correctly labeling their daily activities:

- **Working**: the activity of doing a job at the workplace (indoors)[1].
- **Commuting**: the activity of traveling regularly between work and home[1].
- **Exercising**: the activity of performing physical actions to make or keep your body healthy[1].
- **Religious event**: the activity occurring at religious places.
- **Shopping**: the activity of looking for things to buy in a shopping mall[1].
- **Eating**: the activity of taking food in a restaurant[2].
- **Using toilet**: the activity of going to the bathroom.
- **Home event**: the activity occurring in a structure in which a person lives, esp. a house or apartment[1].

[1]http://dictionary.cambridge.org/us/dictionary/english
[2]http://www.oed.com/

Table 2: Atomic-interval sample dataset. a1: still, a2: walking, a3: running, a4: bycle, a5: vehicle

atomic interval	activity level	activity type	venue type	...	app type
59	0	[a1]	building	...	-
60	1.15	[a1,a2,a1,a2]	route	...	fitness
61	1.99	[a3,a2,a1]	park	...	music
...	...	...	...	...	...
288	0	[a1]	building	...	music

4 LIFE LOGGING

Lifelogging signifies the process of gathering, processing, and storing data regarding personal life experiences [16]. We collect, process, and record a user's contextual information while the user is carrying their smartphone. As shown in Figure 1, each exclusive data receiver, which is responsible for the generation of each data stream, pulls or processes the collectable data independently using built-in smartphone sensors and different APIs. The agent is always running in the background of each smartphone, logging the data without any user interventions, and storing the derived results locally on the device for user-studies. We collect the following lifelogs:

- **time:** time_window (e.g., 20161028_59), time_band (e.g., 0: 00:00 - 03:59, 1: 04:00 - 07:59, 2: 08:00 - 11:59, 3: 12:00 - 15:59, 4: 16:00 - 19:59, 5: 20:00 - 23:59), week (e.g., 0: week, 1: weekend), long_time (e.g., 1477655468)
- **location:** latitude[3], longitude[3], venue_name[3] (e.g., [Cheesecake Factory, Starbucks, Yogurt Land]), venue _type[3] (e.g., [restaurant, cafe, food]), venue_likelihood[3] (e.g., [30%, 10%, 5%]), point_of_interest
- **activity:** activity_type[3] (e.g., [still, walking]), duration[3] (e.g., [250, 50]), activity_level (e.g., 0.4012)
- **phone oriented lifelog:**
  (1) **application:** count, name (e.g., [off, Facebook]), category[4] (e.g., [none, communication]), duration (e.g., [200, 100])
  (2) **photo:** count, concept[5] (e.g., [person, pasta, dish, man, woman])
  (3) **media:** play time
  (4) **sound setting:** silence, bell, vibration
  (5) **calendar:** event (e.g., birthday party), where (e.g., Cheesecake Factory), start_time, end_time

We collect not only low-level lifelogs, such as latitude and longitude, but also high-level semantics. For example, we provide venue name set (e.g., [Cheesecake Factory, Starbucks, Yogurt Land]), which is the exact names of a given GPS point, and the categories of that venue (e.g., [restaurant, cafe, food]). Considering one GPS point may contain multiple venues, we also provide the probabilities of being at each venue (e.g., [30%, 10%, 5%]). In addition, we analyze the places the user frequently visited, and provide the user's point of interests. Furthermore, we accumulate a sequence of the user's physical activity, calculate activity level, which is an average score of the physical activity set [28], and then provide these as high-level lifelogs.

Since these lifelogs are collected as data streams, and they have different granularities and semantics, we must synchronize the data streams by correlating them with a periodic time-interval. We define this periodic time-interval as atomic-interval. Atomic-interval is a 1 x N matrix having N kind of lifelogs collected for a given time-interval. Each row in Table 2 shows the atomic-interval. The numbers in the first column indicate the order of the atomic-interval of the day. The sequentially collected lifelogs, such as activity type, are chronologically collected in an array. Average value, such as activity level, calculated based on pre-defined weights and their amount. Semantic data, such as activity type, venue, photo concept, or application category, are gathered by trustworthy APIs. The length of the atomic-interval can be decided by the designer depending on the precision requirement of the application, and thus there can exist the following separated atomic-intervals per day if we assume the unit of interval as minute.

$$number\_of\_atomic\_intervals = \frac{24hours \times 60minutes}{time\_interval} \quad (1)$$

We organize these atomic-intervals as json format in Figure 1, and then store them in the mobile phone database. We also define daily-activity-interval as a length of the daily activity. This daily-activity-interval can be determined by using our interval growing technique. This technique analyzes the characteristics of sequential atomic-intervals, and groups similar atomic-intervals together to make the daily-activity-interval. This is also shown in Figure 1.

5 DAILY ACTIVITY RECOGNITION

5.1 Daily Activity Segmentation

Daily activity segmentation is the process of partitioning a day into multiple sets of daily-activity-intervals. We pull diverse data streams from a user's smartphone, synchronize each data stream by using atomic-intervals, and then segment a day with our interval growing technique when chronological atomic-intervals have similar patterns of physical activity. For these reasons, determining a length of the atomic-interval must be the first step. We have proven that a five-minute time interval can be a reasonable amount for the atomic-interval. We have tried to find situation transition moments by comparing the similarities of sequential five-minute atomic-intervals, and then proved that this amount of time can be a base unit of daily activity segmentation [22, 28]. Thus, we use five-minutes as the length of the atomic-interval, and then divide a day into 288 atomic-intervals. Most importantly, we assume that indications of the changes of physical activity pattern can be involved in the changes of other attributes, which can be considered as ending one daily activity and starting another. For example, let's say a user has been working at the office, and he has been sitting on the chair. After 10 minutes, if he moves towards the cafeteria for lunch, we recognize this change of physical activity, segment this moment, and make a daily-activity-interval by segmenting from the first atomic-interval to now. In other words, our daily activity segmentation focuses on the interval-growing technique appropriate for daily activity segmentation in which the relevant atomic-intervals are identified by the patterns of physical activities.

[3]https://developers.google.com/android/guides/overview
[4]https://play.google.com/store
[5]https://clarifai.com/

Algorithm 1 Solution for BIG

```
Input: current atomic-interval I_i, seed atomic-interval S_j
Output: daily-activity-interval set R;
Set S_j = I_i if i = 0 and j = 0, or S_j = ∅, and then
  set k = 0;
repeat
    Wait for next atomic-interval, I_i = I_{i+1};
    Extract activity level l_i, and total amount of moving
      time t_i from I_i;
    Extract activity level l_j, and total amount of moving
      time t_j from S_j;
    Calculate δ(i);
    Make a daily-activity-interval r_k by segmenting from
      I_j to I_i, increment k and j, set new seed atomic-interval
      S_j = I_i if δ(i) = 1;
until the system is terminated.
return R
```

Binary Interval Growing (BIG): More specifically, we apply our binary interval growing technique to determine whether consecutive atomic-intervals have similar patterns of physical activities. In order to compare the similarities, we classify each atomic-interval into the moving or the non-moving type of interval, and then deal with the atomic-intervals as one or the other. Algorithm 1 shows the procedures of how to segment atomic-intervals into daily-activity-intervals. We first set up a seed atomic-interval S_j, and then keep calculating $\delta(i)$ every five minutes to determine the similarity between sequential atomic-intervals. $\delta(i)$ can be represented by the following formula:

$$\delta(i) = \|f(S'_j) - f(I'_i)\|_2^2 \tag{2}$$

where S'_j is $\{l_j, t_j\}$, I'_i is $\{l_i, t_i\}$, $f(x)$ is a classification algorithm to classify the non-moving (0) or the moving (1) type of atomic-interval, and $\delta(i)$ is a distance between S_j and I_i. Thus, we segment atomic-intervals when $\delta(i)$ is equal to 1, and then make a daily-activity-interval by segmenting from I_j to I_i. For example, if the type of the seed atomic interval is non-moving, then $f(S'_j)$ is equal to 0. After 5 minutes, if the type of the current atomic-interval is also non-moving, $f(I'_i)$ will be 0, and thus $\delta(i)$ is also equal to 0. However, after another 5 minutes, if the type of the current atomic-interval is moving, $f(I'_i)$ will be 1, and we will finally get $\delta(i) = 1$. Then, we segment this moment, make a daily-activity-interval by segmenting from I_j to I_i, and repeat this process again.

5.2 Daily Activity Recognition

To recognize the daily activities, we now build a common daily activity model. Westermann et al. have built a common multimedia event model by identifying the global unique properties of each individual event. This model addresses several fundamental aspects of events, such as temporal, spatial, experiential, causal, structural, and informational aspects [35]. Specifically, Westermann et al. approach the common event modeling by understanding physical (e.g, event occurrence time stamp and interval), logical (e.g, temporal domain), and relative (e.g, temporal relationships to other events) relationships between each aspect and an event. We bring in these general aspects as the categories of our modeling attributes, and

Table 3: Simplified cross table defining relationships between daily activity and their attributes.

		Attribute		
		Walking (Experiential)	Medium time-duration (Temporal)	Work (Spatial)
Object	Working		X	X
	Using Toilet	X		X
	Commuting	X	X	

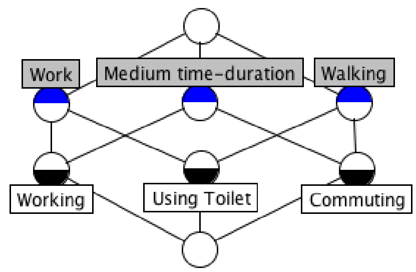

Figure 2: Sample concept lattice derived from Table 3.

modify the physical, logical, and relative components to match the daily activities.

We build the common daily activity model by using Formal Concept Analysis (FCA) based on these general aspects of events. FCA is one powerful technique when data sources are limited, and even when they are uncertain, due to its specialty for discovering implicit information based on pre-defined binary relationships between object and attributes. FCA can be applied for daily activity recognition as follows. One daily activity D can be represented by a triplet $T = (D, A, R)$, where A is a set of attributes, and R is the binary relationships between D and A, $R \subseteq D \times A$. Once each daily activity is defined by the triplet, the triplet can then be converted into a cross table (e.g., Table 3). Then, all possible formal concepts (X_i, Y_i), where $X_i \subseteq D_i$, and $Y_i \subseteq A_i$, are extracted from the cross table, and then are set up as nodes in the concept lattice, which is a graphical representation of the partially ordered knowledge. The hierarchy of formal concepts can be constructed by the following relations:

$$(X_1, Y_1) \leq (X_2, Y_2),\ if\, X_1 \subseteq X_2 \leftrightarrow Y_1 \supseteq Y_2 \tag{3}$$

X_i and Y_i satisfy the following relations:

$$X_i^{'} = \{a_i \in A_i \mid \forall d_i \in X_i, (d_i, a_i) \in R_i\} \tag{4}$$

$$Y_i^{'} = \{d_i \in D_i \mid \forall a_i \in Y_i, (d_i, a_i) \in R_i\} \tag{5}$$

Table 3 shows the simplified relationships between daily activity and their attributes. In order to build the FCA model, formal concepts should be derived from the cross table, such as (*Working*, {*Medium time-duration, Work*}), (*Using Toilet*, {*Walking, Work*}), (*Commuting*, {*Walking, Medium time-duration*}), ({*Working, Using Toilet*}, *Work*), ({*Working, Commuting*}, *Medium time-duration*), and ({*Using Toilet, Commuting*}, *Walking*). These formal concept pairs become each node in the concept lattice, and their hierarchy is determined by formula (3). Figure 2 shows the concept lattice reflects the partially ordered knowledge between each node. The top node and the bottom node indicate ({*Working, Using Toilet, Commuting*}, ∅), and (∅, {*Walking, Medium time-duration, Work*}), respectively. To navigate the concept lattice to obtain the expected results, depth first search is carried out with input attributes. For example, if input attributes are *Medium time-duration* and *Work* in Figure 2, these two nodes will indicate one daily activity, *Working*.

Basically, FCA finds an expected result depending on the structural similarity between an input attribute set and pre-defined attribute sets. Thus, different kinds of input attributes can significantly affect the structural similarities. Because of this, it is necessary to estimate what attributes are important keys to separating each different daily activity, and find all unique daily activity structures composed of those attributes. Moreover, we also need an effective method for estimating missing data while maintaining accuracy, considering that we recognize the daily activities in real-time on smartphone, and the smartphone status will not always be in the best condition. Lastly, given that the amount of actual user data is not always enough to train a powerful model, we also need to come up with how we can make a strong learner by using a group of weak learners. We believe that an ensemble classifier that consists of many concept lattice bags, and its voting process to obtain a majority result from all the recognitions, helps to overcome these challenges. We suggest Bagging Formal Concept Analysis (BFCA), which applies the ensemble approach to FCA, in order to solve those challenges. Bagging Formal Concept Analysis (BFCA) consists of the following steps:

(1) Categorize all the labeled daily-activity-intervals, which obtained from 23 participants for two weeks, by each daily activity.
(2) Make n number of classifiers where n is the number of the recognizable daily activity, make m number of bags per classifier, and bootstrap training data for each bag.
(3) In each bag, use one third random attributes $\frac{p}{3}$, where p is the number of total attributes, and extract all unique relationships between the labeled daily activity and their randomly picked attributes.
(4) Build the cross table in each bag by using those unique relationships, and generate the concept lattice. This concept lattice only determines whether the given input attribute set can be the labeled daily activity.
(5) When an input attribute set is given, which is an unlabeled daily-activity-interval, we navigate all the concept lattices for each daily activity classifier, calculate the possibility of being each daily activity, and then choose the highest possibility among the results.

Given that FCA requires discrete attributes, we convert our time-series values C, such as activity level, or time duration of daily-activity-intervals, into discrete space, such as w-dimensional space $\{high, medium, low\}$, by a vector $\bar{C} = \bar{c_1}, \bar{c_2}, ..., \bar{c_i}$. We use a discretization technique, SAX (Symbolic Aggregate ApproXimation), which reduces the time series of arbitrary length n into the w-dimensional space by the following equation [25]:

$$\bar{c_i} = \frac{w}{n} \sum_{j=\frac{n}{w}(i-1)+1}^{\frac{n}{w}i} c_j \qquad (6)$$

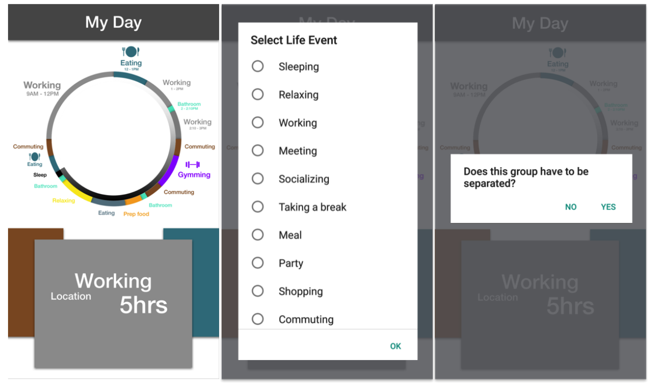

Figure 3: The system running for daily activity segmentation and recognition.

5.3 From Activities to Events

We proceed to create events with all facets by using all collectible data sources from multiple devices. We insist that an event is just a single unit in itself, but it can form the chronicle once it is stored in the database. Thus, we store all the recognized daily activities in the database as events with as many data sources as possible, such as *Personicle* in Figure 1, and quantify the chronicle. This personal chronicle can then be used to model the person by using learning techniques and relating them to biomedical or behavioral signals. In the current version, we use only the signals from smartphones, but pulling heterogeneous signals from multiple devices, and then analyzing a person with all the facets of the events will be an important topic for further research.

6 EXPERIMENTAL VALIDATION

We implemented an android application to test our segmentation and recognition methods. As shown in Figure 3, we asked 23 participants to give feedback on the results of their segmentations as well as label their daily activities for each segmented result during an average of two weeks. We stored all the collected lifelogs in their smartphone database, and then gathered these databases after the experiment had been completed. The total number of collected daily-activity-intervals was 35,967.

6.1 Segmenting User's Day

We assume that segmentation moments can mostly be affected by their adjacent atomic-intervals since atomic-intervals are on a one-dimensional time-line. Thus, interval growing based approaches, which compare contiguous atomic-intervals, must show better performance for the daily activity segmentation than those of statistical methods using all collected lifelogs. To verify the performance of BIG, we compare the BIG results to 1) ground truth, which was obtained by participants' feedback, and 2) the results to those of statistical techniques, such as clustering (k-means), and thresholding (otsu). We use the jaccard coefficient, which has the obvious advantage of similarity evaluation between two sets of binary data, for verifying the performance. The jaccard coefficient is calculated

Table 4: Overall segmentation results of 23 participants.

Algorithm	Segmentation accuracy			
	Best J_c	Worst J_c	Average J_c	Stdev
BIG	**0.9583**	**0.7896**	**0.9050**	**0.0432**
Clustering	0.7841	0.5803	0.6564	0.0601
Thresholding	0.8556	0.4370	0.5863	0.1467

as follows:

$$J_c(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{7}$$

where A is the ground truth and B is the algorithmic result. We see if BIG can be uniformly applied in all the users by achieving relatively higher results than those of other algorithms. For this reason, we handle each user's experimental data separately, calculate each user's jaccard coefficient, and then compare the best, worst, and average results as in Table 4.

We prove the BIG's performance by comparing the results to the statistical approaches. From the results in Table 4, we can see that the average accuracy of BIG is higher than those of the others. Even the worst result of BIG is almost similar to or a little less than other techniques' best results. Furthermore, the standard deviation of BIG shows that each user's accuracy is nearly the same; however, we can also see a 0.1687 difference between the best and worst accuracy. There are two reasons. Given that we depend on the API results for physical activity prediction (e.g., still, walking), some incorrect API results may lead to incorrect segmentation results. More specifically, the API returns "walking" or "in vehicle" activity when a user slightly shakes his legs or has minute-long movements. We found that the user who obtained the worst result in BIG had many of these cases, and thus these unexpected cases resulted in the incorrect segmentations results. The different awareness of segmentation moments between users and us also caused the incorrect results. We handle 5 minutes time-intervals, and thus we don't consider the short changes as segmentation moments. For example, if a user walks only for 5 or 10 seconds, and then immediately starts a non-moving activity, we consider this as one continuous non-moving segment given 5 minutes length of granularity. However, some of the users who participated in our experiment gave feedback many of these moments were segmentable moments.

The lower result of clustering and the thresholding technique show that reflecting past physical activity patterns for current segmentation moments can cause a bed effect on segmenting results. For example, if a user is a very active person, those techniques will not segment small movements even though these are a sufficient amount for the daily activity segmentation.

6.2 Recognizing Daily Activity

With the identified daily-activity-intervals, we now try to recognize L2 daily activities. 23 participants had labeled L2 daily activities on these daily-activity-intervals.

In the lifelogs, we observed that at times our system was unexpectedly killed by OS, which made the data discontinuous. At other instances, participants did not label their segmented results, or the participants' phones ran out of battery. We tried to avoid these exceptional cases by immediately restarting the system when it was terminated by OS, or asking the user to label the segments with the pop-up messages in the system. However, there were still many non-labeled and non-consecutive segments. We first clean these unclear data in order to precisely verify the performance of recognitions, and thus the total number of considered daily-activity-intervals are 15,087 samples of 35,967. And then we split these samples into 30% training dataset, and 70% test dataset to show that the model of BFCA can be robust despite the relatively small training dataset.

Table 5: F-measure (%) for combination of attribute sets. D_1: Commuting, D_2: Eating, D_3: Exercising, D_4: HomeEvent, D_5: ReligiousEvent, D_6: Shopping, D_7: UsingToilet, and D_8:Working. S_1: Temporal + Experiential, S_2: Temporal + Spatial, S_3: Spatial + Experiential, S_4: S_1 + Spatial, S_5: S_4 + Causal, S_6: S_5 + Structural aspect.

			Attribute set combination					
		# sample	S_1	S_2	S_3	S_4	S_5	S_6
Daily Activity	D_1	393	66.7	66.7	55.5	75.6	**90.4**	76.6
	D_2	404	28.2	71.9	43.2	70.7	77.8	**79.6**
	D_3	15	0	**100**	**100**	**100**	**100**	**100**
	D_4	10698	60.6	94.7	65.6	91.8	**96.6**	**96.6**
	D_5	588	0	**98.5**	**98.5**	97	76.4	**98.5**
	D_6	53	0	40	22.2	25	**44.4**	40
	D_7	28	56.3	0	38.5	9.5	**81.2**	55.2
	D_8	2908	6.9	69.5	44.9	81.8	**90.3**	89.1

In order to maximize the recognition performance, we assume that each daily activity has a specific combination of the common event attribute sets that most represent the daily activity. This means that all the aspects of the common event model (e.g., temporal, spatial, experiential, structural, informational, and causal aspects) are not vital elements for every daily activity recognition. For example, according to the definitions in Section 3, the "Commuting" activity, which refers to the activity of traveling regularly between work and home, can be recognized by only using spatial (e.g., work or home), structural (e.g., L1 daily activity, such as going, or still), and causal (e.g., the relations between current and previous daily activity) aspects. To verify this idea, we experimented with the different combinations of the common event model aspects, and figured out the best combinations by calculating their accuracy. We roughly use 10 bags of concept lattice for this experiment, and thereby calculate their f-measures to see the weighted harmonic accuracy between precision and recall. From the results in Table 5, we can see that some combinations of the attributes have better results than those of others, such as S_5 for D_1, D_6, D_7 and D_8, and S_6 for D_2. It shows that unnecessary information results in the confusion of modeling, and thus we use the specialized combination sets for each daily activity modeling.

We now try to find the best number of concept lattice bags, which also can maximize the recognition performance. First, we train separate BFCA models on different numbers of bags, which are from 1 to 1000, by using the selected attribute sets. Then, we experiment with the daily activity recognition on those trained models, respectively, by using the same test dataset. Finally, we calculate their f-measures to see what numbers of bags would return the best recognition accuracy. Figure 4 shows the variations of accuracy on the different number of bags. In our results, we

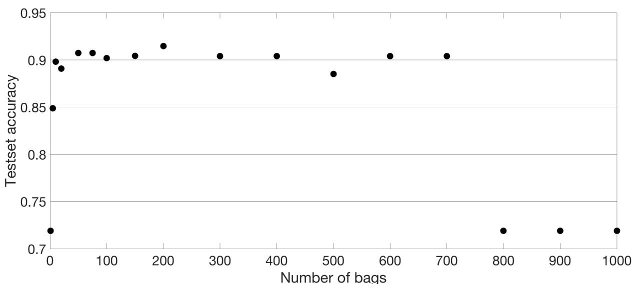

Figure 4: The Variations of BFCA accuracy on different number of concept lattice bags.

Table 6: Confusion matrix of the BFCA. D_1: Commuting, D_2: Eating, D_3: Exercising, D_4: Home Event, D_5: Religious Event, D_6: Shopping, D_7: Using Toilet, and D_8: Working.

		Predicted (%)							
		D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8
Targeted (%)	D_1	**95.8**	0	0	4.2	0	0	0	0
	D_2	0	**97.8**	0	0	0	2.2	0	0
	D_3	0	0	**100**	0	0	0	0	0
	D_4	0	4.3	0	**95.7**	0	0	0	0
	D_5	0	2.9	0	0	**97.1**	0	0	0
	D_6	0	16.7	0	0	0	**66.7**	16.7	0
	D_7	5.3	0	0	0	0	0	**94.7**	0
	D_8	5.6	9.3	0	0	0	0	0	**85**

can see that the accuracy for under 700 bags is nearly the same; however, the accuracy rapidly decreased by 0.7191 once bags are over 800. Basically, the higher the number of bags, the higher the recognition performance in ensemble technique. However, a large number of bags in BFCA can confuse the voting process given that these bags can make all the classifiers robust. Therefore, among the good results under 800 bags, we choose the best accuracy, 0.9147 (bags=200).

Then, we build the confusion matrix to see the specific results of each daily activity recognition. In Table 6, we can see that 5 minutes length of granularity results in an ambiguous segmentation boundary between "Commuting" activity and "Home Event" activity (4.2%). We also can see that randomly picked $\frac{p}{3}$ attributes cause confusion in the daily activity modeling. For example, "Home Event" activity can be considered as "Eating" activity (4.3%), and "Shopping" activity can be classified either as "Eating" (16.7%) or "Using Toilet" activity (16.7%), if spatial aspects are missed. However, the overall accuracy of all the daily activity recognition (>90%) shows that using the randomly picked attributes, and a certain number of concept lattice bags can minimize the misclassification of daily activities. This is proven in Table 7.

As shown in Table 7, BFCA has greatly improved the recognition performance compared to the FCA. FCA only depends on the structural similarity between an input attribute set and pre-defined relations. Thus, it sometimes recognizes multiple daily activities if they have similar structures to the pre-defined relations. This issue is a critical problem, which can cause lower performance, given that FCA does not have any statistical methods to choose the most probable result. The result of BFCA shows that applying a statistical method to FCA, such as the ensemble approach, can be one solution to overcome the problem.

Table 7: Accuracy for the daily activity recognition on 23 participants.

	Precision	Recall	F-measure
FCA	0.2522	0.5114	0.3378
BFCA	**0.9098**	**0.9204**	**0.9151**
Decision Tree	0.6604	0.6826	0.6643
Random Forest	0.7358	0.7464	0.7411
Support Vector Machine	0.6981	0.7081	0.7031

Since BFCA brings the idea from random forest, which uses the ensemble technique bagged by decision trees, we also compare BFCA to random forest. In Table 7, we can see that BFCA has better results than the random forest. Basically, our dataset is imbalanced data because some daily activities occupy the better part of the day. For example, the "Sleeping", "Home Event", and "Working" activities used to be the majority of the daily activities. Moreover, these daily activities mostly share similar lifelogs to each other, and thus the decision tree and random forest must have difficulty clearly classifying them. This can also explain why the support vector machine, which is one of the most powerful classification algorithms, has lower recognition accuracy than BFCA.

7 CONCLUSION

Kahneman, who is a Nobel Prize winner, showed the importance of daily activities in human life experiences. This paper builds towards the research to develop techniques for objectively and automatically understanding the daily lives of human beings via common wearable devices. Specifically, this paper focuses on recognizing human daily activity to understand their lifestyle and behavior patterns for the purpose of building objective self model. Thus, it describes the methodology behind automatically recognizing daily activity with the goal to build a personal chronicle. We develop a logging application that runs on Android device, collects data, and converts the data into personal chronicle. Using the chronicle, one may proceed to determine individual models using machine learning techniques. Such models may play very important role in applications for health and behavior modification. We begin with synchronizing multimodal data streams by using atomic-intervals, and then use an interval growing technique for determining daily-activity-intervals and their attributes. Next, we use the common event model and BFCA to classify each daily activity. Lastly, the daily activities are stored in the database and consist of the chronicle of daily activities. Results obtained across different FCA and classification algorithms show the potential of such an approach for recognizing daily activities. Further research would allow for increasing the number of detectable atomic-level daily activities by combining more heterogeneous and higher cognitive multimedia logs, and thus recognizing more various daily lives.

REFERENCES

[1] Gregory D Abowd, Anind K Dey, Peter J Brown, Nigel Davies, Mark Smith, and Pete Steggles. 1999. Towards a better understanding of context and context-awareness. In *International Symposium on Handheld and Ubiquitous Computing*. Springer, 304–307.

[2] Rami Albatal, Cathal Gurrin, Jiang Zhou, Yang Yang, Denise Carthy, and Na Li. 2013. Senseseer mobile-cloud-based lifelogging framework. In *Technology and Society (ISTAS), 2013 IEEE International Symposium on*. IEEE, 144–146.

[3] Ling Bao and Stephen S Intille. 2004. Activity recognition from user-annotated acceleration data. In *International Conference on Pervasive Computing*. Springer, 1–17.

[4] C Gordon Bell and Jim Gemmell. 2009. *Total recall: How the e-memory revolution will change everything*. Dutton.

[5] Gregory Biegel and Vinny Cahill. 2004. A framework for developing mobile, context-aware applications. In *Pervasive Computing and Communications, 2004. PerCom 2004. Proceedings of the Second IEEE Annual Conference on*. IEEE, 361–365.

[6] Vannevar Bush et al. 1945. As we may think. *The atlantic monthly* 176, 1 (1945), 101–108.

[7] Scott Carter, Jennifer Mankoff, and Jeffrey Heer. 2007. Momento: support for situated ubicomp experimentation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 125–134.

[8] Liming Chen, Chris D Nugent, and Hui Wang. 2012. A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering* 24, 6 (2012), 961–974.

[9] Eun Kyoung Choe, Saeed Abdullah, Mashfiqui Rabbi, Edison Thomaz, Daniel A Epstein, Felicia Cordeiro, Matthew Kay, Gregory D Abowd, Tanzeem Choudhury, James Fogarty, et al. 2017. Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications. *IEEE Pervasive Computing* 16, 1 (2017), 74–84.

[10] Anind K Dey. 2001. Understanding and using context. *Personal and ubiquitous computing* 5, 1 (2001), 4–7.

[11] Aiden R Doherty and Alan F Smeaton. 2008. Automatically segmenting lifelog data into events. In *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS'08. Ninth International Workshop on*. IEEE, 20–23.

[12] Denzil Ferreira, Vassilis Kostakos, and Anind K Dey. 2015. AWARE: mobile context instrumentation framework. *Frontiers in ICT* 2 (2015), 6.

[13] James Fogarty, Carolyn Au, and Scott E Hudson. 2006. Sensing from the basement: a feasibility study of unobtrusive and low-cost home activity recognition. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*. ACM, 91–100.

[14] Jim Gemmell, Gordon Bell, and Roger Lueder. 2006. MyLifeBits: a personal database for everything. *Commun. ACM* 49, 1 (2006), 88–95.

[15] Jim Gemmell, Gordon Bell, Roger Lueder, Steven Drucker, and Curtis Wong. 2002. MyLifeBits: fulfilling the Memex vision. In *Proceedings of the tenth ACM international conference on Multimedia*. ACM, 235–238.

[16] Cathal Gurrin, Alan F Smeaton, Aiden R Doherty, et al. 2014. Lifelogging: Personal big data. *Foundations and Trends® in Information Retrieval* 8, 1 (2014), 1–125.

[17] Jin-Hyuk Hong, Julian Ramos, and Anind K Dey. 2016. Toward personalized activity recognition systems with a semipopulation approach. *IEEE Transactions on Human-Machine Systems* 46, 1 (2016), 101–112.

[18] https://www.google.com/fit/. [n. d.]. Google Fit. ([n. d.]).

[19] Stephen S Intille, Kent Larson, and Chuck Kukla. 2002. Just-in-time context-sensitive questioning for preventative health care. In *Proceedings of the AAAI 2002 Workshop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care*. AAAI Press Menlo Park, CA.

[20] Ramesh Jain and Laleh Jalali. 2014. Objective self. *IEEE MultiMedia* 21, 4 (2014), 100–110.

[21] Laleh Jalali, Da Huo, Hyungik Oh, Mengfan Tang, Siripen Pongpaichet, and Ramesh Jain. 2014. Personicle: personal chronicle of life events. In *Workshop on Personal Data Analytics in the Internet of Things (PDA@ IOT) at the 40th International Conference on Very Large Databases (VLDB), Hangzhou, China*.

[22] Laleh Jalali, Hyungik Oh, Ramin Moazeni, and Ramesh Jain. 2016. Human Behavior Analysis from Smartphone Data Streams. In *International Workshop on Human Behavior Understanding*. Springer, 68–85.

[23] Daniel Kahneman, Alan B Krueger, David A Schkade, Norbert Schwarz, and Arthur A Stone. 2004. A survey method for characterizing daily life experience: The day reconstruction method. *Science* 306, 5702 (2004), 1776–1780.

[24] 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.

[25] Jessica Lin, Eamonn Keogh, Li Wei, and Stefano Lonardi. 2007. Experiencing SAX: a novel symbolic representation of time series. *Data Mining and knowledge discovery* 15, 2 (2007), 107–144.

[26] Mitja Luštrek, Božidara Cvetković, Violeta Mirchevska, Özgür Kafalı, Alfonso E Romero, and Kostas Stathis. 2015. Recognising lifestyle activities of diabetic patients with a smartphone. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2015 9th International Conference on*. IEEE, 317–324.

[27] Sangeeta Mittal, Krishna Gopal, and SL Maskara. 2013. A versatile lattice based model for situation recognition from dynamic ambient sensors. *International Journal on Smart Sensing and Intelligent Systems* 6, 1 (2013), 403–432.

[28] Hyungik Oh, Laleh Jalali, and Ramesh Jain. 2015. An intelligent notification system using context from real-time personal activity monitoring. In *Multimedia and Expo (ICME), 2015 IEEE International Conference on*. IEEE, 1–6.

[29] Abigail J Sellen and Steve Whittaker. 2010. Beyond total capture: a constructive critique of lifelogging. *Commun. ACM* 53, 5 (2010), 70–77.

[30] Emmanuel Munguia Tapia, Stephen S Intille, and Kent Larson. 2004. Activity recognition in the home using simple and ubiquitous sensors. In *International Conference on Pervasive Computing*. Springer, 158–175.

[31] Tim Van Kasteren, Athanasios Noulas, Gwenn Englebienne, and Ben Kröse. 2008. Accurate activity recognition in a home setting. In *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 1–9.

[32] Peng Wang and Alan F Smeaton. 2012. Semantics-based selection of everyday concepts in visual lifelogging. *International Journal of Multimedia Information Retrieval* 1, 2 (2012), 87–101.

[33] Peng Wang, Alan F Smeaton, Yuchao Zhang, and Bo Deng. 2014. Enhancing the detection of concepts for visual lifelogs using contexts instead of ontologies. In *Multimedia and Expo Workshops (ICMEW), 2014 IEEE International Conference on*. IEEE, 1–6.

[34] Peng Wang, Lifeng Sun, Shiqiang Yang, Alan F Smeaton, and Cathal Gurrin. 2016. Characterizing everyday activities from visual lifelogs based on enhancing concept representation. *Computer Vision and Image Understanding* 148 (2016), 181–192.

[35] Utz Westermann and Ramesh Jain. 2007. Toward a common event model for multimedia applications. *IEEE MultiMedia* 14, 1 (2007).

[36] 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*. ACM, 1–10.