

Unsupervised Work Knowledge Mining through Mobility and Physical Activity Sensing

Le T. Nguyen

Electrical and Computer Engineering
Carnegie Mellon University, Silicon Valley
le.nguyen@sv.cmu.edu

Joy Zhang

Electrical and Computer Engineering
Carnegie Mellon University, Silicon Valley
joy.zhang@sv.cmu.edu

Abstract—Knowledge of working professionals gained through years of experience is invaluable for any organization. Extracting this knowledge allows an organization to optimize internal processes and facilitate training of new hires. Therefore, there has been a significant research effort in developing techniques for automated knowledge mining at workplaces. However, research in the past have been focused mainly on extracting knowledge of stationary professions such as office workers, who perform most of their day-to-day tasks at their desk. In this work, we propose an approach for mining work knowledge of physically active professions such as nurses, firefighters, waiters, housekeepers or janitors. We leverage the advances of mobile sensing to extract knowledge from workers with high level of mobility and physical activity patterns. We demonstrate the feasibility of the proposed approach on a real-world scenario of a janitor as a study subject. We show that using data collected from mobile devices carried by a janitor throughout their work, we are able to extract knowledge rules that describe generalized patterns of janitor’s behavior. We expect the proposed method to be applied to other fields to mine knowledge from workplaces.

I. INTRODUCTION

Years of experience in a profession makes a worker more effective and efficient at completing the work. This experience has a high value to organizations, especially when new hires need to be trained to replace experienced workers or when internal workflows need to be optimized to increase the productivity [3]. Therefore, it is crucial for organizations and operational researchers to capture and store the knowledge of workers [27].

A common technique for work knowledge extraction is to have employees manually maintain a knowledge base using tools such as wikis [16]. To automate the extraction process, knowledge can be mined directly from a text corpus available in organizations such as email communication [12]. Additionally, installing a logging application on a computer enables tracking of employee’s interactions with digital media (e.g., slides, documents, etc.) [3].

The traditional knowledge extraction techniques work well for stationary and knowledge professionals such as office workers, who perform their day-to-day tasks at a desk using computers. However, many new challenges arise when we aim to extract knowledge from non-stationary professions such as hospital nurses, fire fighters, waiters or janitors:

- **Diversity of physical activities:** In addition to cognitive activities (e.g., writing a report on a computer), employ-

ees perform a variety of physical activities, which cannot be easily captured through the existing approaches.

- **Mobility:** Workers frequently move to different locations to perform their daily tasks.
- **Non-intrusiveness:** In order to capture workers’ natural behavior patterns, the knowledge mining process should not interfere with their day-to-day workflow.
- **Privacy-preserving:** Often employees (e.g., nurses) are handling sensitive and personal information. The knowledge mining technique should respect the privacy of the target individuals and of other people in the environment.

To address above challenges, we propose a knowledge mining approach built on top of recent advances in mobile sensing [11], [28] and the well-established corpus-based knowledge extraction techniques [17], [6], [8]. Our approach uses smartphones carried by professionals to capture their mobility and physical activities. We then transform low-level sensor readings from smartphones into a text-like symbolic representation. Thus, we can leverage the existing corpus-based techniques to extract knowledge from professionals’ physical behavior patterns. We expect the proposed approach to be used in combination with traditional approaches to extract work knowledge from both cognitive and physical activities. The extracted knowledge can be further used by operational researchers to optimize the workflow.

Our contribution in this work is three-fold:

- 1) We show the feasibility of using smart phones that professionals carry with them to observe their activities without disturbing the professionals’ normal work patterns.
- 2) We propose an unsupervised approach for high-level knowledge mining from low level sensor signals. Specifically, we convert mobile sensor readings to a text-like symbolic representation. By applying text-processing techniques on the symbol-based time-series we extract high-level knowledge rules.
- 3) Through the experiments we demonstrate that the extracted knowledge rules capture generalized work patterns and can be used for analyzing and predicting an employee’s behavior.

The rest of the paper is organized as following: we first review related work in Section II. In Section III we introduce an knowledge mining scenario and discuss applications uti-

lizing the extracted knowledge. In Section IV we introduce the proposed approach for knowledge extraction. We conduct multiple experiment to demonstrate the effectiveness of extracted knowledge rules and present the findings in Section V. Finally, we conclude our work and discuss future directions in Section VI.

II. RELATED WORK

Corpus-based knowledge extraction has been a well-studied research area. Many techniques were developed to automatically extract knowledge from online text corpora such as Wikipedia or Wikihow [25].

In workplaces, similar techniques were used to mine knowledge from internal corpora such as email communications or reports [12]. These techniques were often combined with having employees manually maintaining knowledge wiki systems [16].

In addition to mining knowledge from a corpus, many approaches were proposed to automatically extract procedural knowledge from employees’ activities. Computer applications were used to log user interactions with digital content (e.g., slides, documents, etc.) [3]. By tracking these interactions, researchers were able to identify prominent behavioral patterns, which form the basis of efficient task solving knowledge.

The above-mentioned approaches focus on extracting knowledge of stationary professions, where employees performed most of their tasks at a desk using computers. To extract knowledge of non-desk based tasks (e.g., how to operate a certain machine), manual shadowing is commonly used technique [18]. A new employee would shadow an experienced employee for a certain time period to learn necessary skill and knowledge for completing a task [18].

Some approaches have been proposed to automate knowledge extraction from physical work activities by deploying cameras in the environment [22], [9]. The knowledge can be then extracted by leveraging computer vision techniques to capture the human behavioral patterns. For many occupations such as policemen, firefighters or medics, it is intractable to deploy cameras in all environments where the professionals could be. Moreover, sensors deployed in the environment raise privacy concerns as they record data not only of the target subject, but also of others not related to the knowledge extraction process.

In recent years, new opportunities arise with the rapid adoption of smartphones. As people carry their mobile devices everywhere and all the time, they become a new platform for continuous sensing and information inference. A large amount of work has been focused on inferring user’s activity [28] and location [15] through mobile sensing. The inferred information have been used for understanding user’s behavioral patterns and for future tasks prediction [20].

In this work, we propose using mobile sensing to extract knowledge of a professional in the real world. We leverage existing natural language processing techniques to infer high-level meanings of workers’ activities (e.g., “picking up trash can and dump trash”) from the low-level sensor information

such as the raw accelerometer readings. We then extract knowledge rules that describe user behavioral patterns in the form of IF-THEN rules such as *IF it is Monday 8am THEN vacuum the kitchen floor*. We show the effectiveness of the extracted knowledge rules by predicting patterns in a dataset collected in real-world settings.

III. REAL-WORLD KNOWLEDGE MINING

In this section, we first introduce a real-world scenario to explain the intuition of the knowledge extraction process. We then discuss how the extracted knowledge can be further utilized and describe the capabilities and limitations of using mobile sensing for knowledge extraction.

A. Knowledge Mining Scenario

The goal of the proposed approach is to extract the knowledge of a worker with high level of mobility and physical activity. We will explain the knowledge mining process using a scenario with a janitor working at a university as a target subject. His job is to clean and maintain a building with more than 30 rooms including the kitchen, classrooms, offices etc. Everyday, he enters almost every room in a building and performs a variety of tasks shown in Table I. For the study, the janitor is asked to carry a mobile phone with him (as he does anyway) while performing everyday tasks.

Cleaning	Maintenance
Cleaning toilets, urinals	(Un-)locking buildings
Cleaning mirrors, sinks	Restocking bathrooms
Cleaning floors (mopping, sweeping)	Movement of equipment
Carpet cleaning and vacuuming	Internal deliveries
Cleaning meeting rooms	Room setups
Wiping white boards	Watering plants
Cleaning tables in cubicles	Simple repairs
Clearing lunch room/kitchen	Refilling coffee machines
Emptying trash and recycling bin	etc.
Cleaning air-conditioner vents	
Litter picking	
Spot cleaning (generally spills)	
Sanitization	
etc.	

TABLE I
LIST OF JANITOR’S TASKS

The mobile phone runs a sensing application as a service [28] in the background. Since the janitor typically works from 8am to noon (4 hours), the application automatically starts the data collection at 7:30am and ends at 12:30pm. At the end of each workday the data is automatically uploaded to a server for further analysis. The janitor does not need to interact with the mobile application at any time.

We shadow the janitor for several days to obtain the ground truth information, which are used to explain the intuition of the knowledge extraction process. While shadowing the janitor, we manually log his activities with starting and ending time, room number and type of activity.

From the interviews and by shadowing the janitor, we observe a complex set of *constraints* influencing the janitor’s work patterns. For example, the janitor should vacuum (1) each room at least once a week, (2) preferably early morning and (3) when a room is not occupied. Considering that there are more

than 30 rooms, the janitor needs to keep track of which room he already vacuumed, which room needs vacuuming more than once a week and which room is usually occupied at which time.

The goal of the janitor is to finish a large amount of context-dependent tasks such as vacuuming rooms in a short time period (4-hours work days). In order to achieve the goal the janitor needs to plan his workflow. The janitor’s work planning can be framed as an optimization problem with the goal of reducing the total amount of work time given a set of constraints.

One of the challenges is that the experienced janitor often cannot enumerate all the constraints, even though he is subconsciously aware of them. The knowledge the janitor possesses but cannot explicitly describe is known as *tacit knowledge* [24], which has been shown to be invaluable for many professions [27]. In this work, we propose an approach to extract both explicit and tacit knowledge of the janitor by first observing his mobility and activity patterns, which are then used to extract rules representing the knowledge of the janitor.

Mobility patterns often provide cues about the janitor’s activity. Figure 1 shows a floor plan of the workplace with activities performed by the janitor at specific locations. The location-related cues can be observed through the janitor’s task of collecting trash. He typically enters each room for only a few second to empty the trash cans. After collecting the trash in all rooms, he walks out of the building to throw it into a trash container. Thus, by observing the janitor’s location traces, we can estimate whether he is currently performing a certain tasks such as collecting trash.

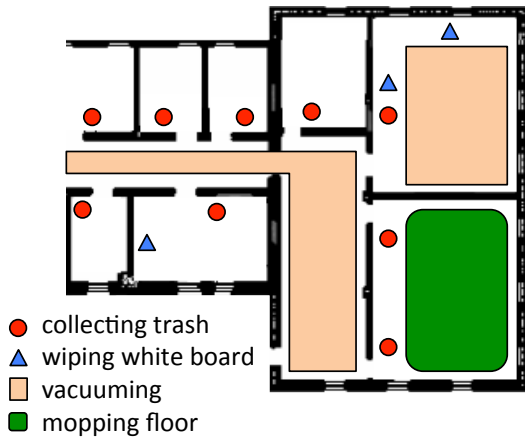


Fig. 1. Floor plan of a workplace with locations of performed activities.

Janitor’s activities often depend on the time of the day and day of the week. Figure 2 shows a timeline of a janitor’s workday. A janitor typically starts his work early in the morning to avoid disturbing the faculty members and students. He typically starts with loud tasks such as vacuuming or task that requires him to enter many office rooms such as collecting trash. In the later part of the workday, the janitor focuses on

the restocking and maintenance tasks, which are performed in public areas, such as kitchen or restrooms.

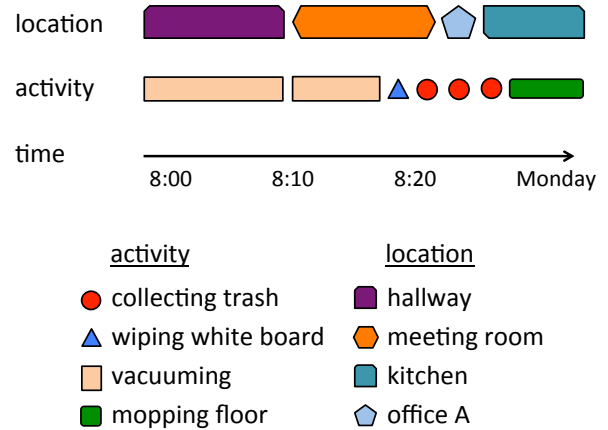


Fig. 2. Timeline of janitor’s workday. Time, location and activity information provides us with more insight into the janitor’s work patterns.

Given the above context information, our goal is to automatically extract work knowledge. We represent the knowledge as a set of rules (shown in Figure 3). Based on the context we can discover rules such as a simple time-based rule **R1**: “if it is 8am on Monday, a janitor will be vacuuming the floor” or a time-location-based rule **R2**: “if the janitor is in the kitchen in the morning, he is mopping the floor”. Additionally, rule-based representation can express exceptions such as **R3**: “if the meeting room is occupied, the janitor should move on to the office A and clean the meeting room later”.

Knowledge Base

Rule ID	Condition	Action
R1	<ul style="list-style-type: none"> 8:00 am Monday 	vacuuming □
R2	<ul style="list-style-type: none"> 8:30am kitchen □ 	mopping floor ■
R3	<ul style="list-style-type: none"> meeting room □ room is occupied 	office A □

Fig. 3. The knowledge base representing the extracted rules of users behavior patterns.

B. Utility of Extracted Knowledge

The extracted context information and knowledge can be used in the following way:

- 1) **Detailed Analysis:** The inferred location and activity information can be used for a detailed analysis of a worker’s workflow. Through the inference we can answer questions such as when and how many times a worker takes a brake, what is the busiest time of his day or what is his physical condition after performing a certain task.
- 2) **Summarization and Prediction:** Using mobile sensing a large amount of fine-grained data can be collected. This data can be used to extract patterns, which capture the worker’s natural behavior. Thus, we can summarize

the worker’s physical activities into a knowledge model, which can be used to predicting the worker’s future behavior.

- 3) **Process Optimization:** The extracted rules can be used by operational researchers to optimize work processes to increase workers’ productivity [14]. Moreover, tools such as expert systems [7] can leverage the extracted rules to automate subprocesses to reduce the required human effort.

C. Mobile Sensing and Its Limitations

Modern smartphones are equipped with a large range of sensors, which can be used for inferring information needed for the knowledge extraction. Accelerometer, magnetometer and gyroscope can be used for detecting motion and recognizing user’s physical activities [28]. GPS and Wi-Fi can be used for estimating user’s location outdoors and indoors [21]. Phone microphones and cameras can be used to detect the context information such whether a user in a presentation or in a kitchen talking to a colleague.

On the other hand mobile sensing has also some limitations. The sensing quality often depends on the position users carry their phones (e.g., in a pocket, inside a purse etc.) and on the quality of mobile sensors, which are typically low-cost and less precise compared with high-end sensors. Moreover, for professions with a limited amount of mobility (e.g., administrative staff working with desktop computers), accelerometer data would not necessarily provide enough information for the knowledge extraction. We expect the approach proposed in this work to be used in combination with existing approaches (e.g., computer logging) to deliver a more complete picture of the professionals’ work patterns. Mobile sensing can also be extended by incorporating additional sensors embedded in wearable devices such as wrist watches [19] or glasses[26] to capture other aspects of an professionals’ working experience.

IV. UNSUPERVISED KNOWLEDGE MINING FROM PHYSICAL ACTIVITIES

In the previous section, we discussed knowledge rules that can be extracted through mobile sensing in a real-world scenario. In this section, we will describe the process of knowledge mining, which consists of three stages: feature extraction from sensor readings, context-inference and knowledge rules extraction. The objective is to mine knowledge without human annotations by analyzing prominent behavior patterns.

A. Knowledge Extraction as a Natural Language Problem

Figure 4 shows the architecture of the proposed knowledge mining system. Comparing our system with a traditional corpus-based knowledge extraction system [17], both can be divided into 3 main components: raw data processing, information inference and knowledge extraction component. On the lowest level, our system captures raw data such as time, accelerometer, magnetometer and Wi-Fi readings. This data is then used to infer context information such as location and activity of an individual. By fusing the inferred information

from multiple sensor modalities, we extract knowledge using rule-based mining techniques.

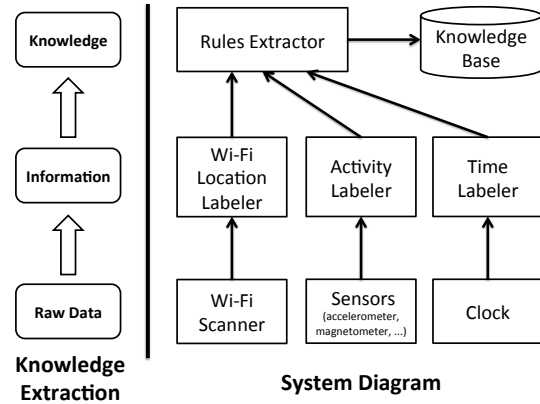


Fig. 4. System components for acquisition of data, information and knowledge in each of the knowledge extraction steps.

One of the key ideas of the proposed approach is to frame the sensor-based knowledge extraction problem as a NLP problem. As demonstrated by [4], there are many similarities between natural language and human activity:

- **Semantic compositionality:** The meaning of a user’s activity is composed from the meanings of its sub-activities.
- **Grammatical structure:** There are underlying hierarchical orders in user’s activities. For instance, in a high-level activity “Walking up the stairs”, a “left step” typically precedes a “right step” which might be followed by a “turn 90 degrees” and then again the reoccurring pattern of “left step” and “right step”.
- **Sequential representation:** When we speak or write, we do not produce a syntactic tree or semantic form to convey the meaning. Rather, what we produce is a unstructured sequence of words. Similarly, when a user performs an activity, the outcome as observed by mobile sensors is a time-series of sensor readings without any structure.

The motivation of framing the knowledge extraction problem as a NLP problem is to reuse the existing NLP techniques such as the Latent Dirichlet Allocation (LDA) [2], [11] for topic modeling and Association Rules Extraction [10], [13], [6], [8] for knowledge extraction from text. In this work we represent the stream of multi-dimensional raw sensor data of a user’s physical activity as a sequence of one-dimensional symbols (so called **labels** or “words”) and apply text-based knowledge extraction algorithms to discover the implicit knowledge.

In the following, we will introduce algorithms for inferring context labels from raw sensor readings. These labels are then used to extract rules, which capture the generalized work patterns.

B. Wi-Fi Location Labeling

Although GPS has been a dominant source for outdoor positioning, it is unreliable in indoor environments. In this sec-

tion, we describe an unsupervised indoor location recognition method leveraging Wi-Fi signatures. The proposed approach is not limited to Wi-Fi and can be extended to other location sources such as Bluetooth and RFIDs if available.

Many approaches have been developed that use Wi-Fi signals of surrounding access points (APs) for positioning purposes [15]. These approaches are built on the observation that there is a high correlation between observed received signal strength indicator (RSSI) and a mobile device’s physical location. At each physical location a mobile device can observe a vector of RSSIs, each RSSI corresponding to one nearby AP. A RSSI vector is considered to be a location indicator in the latent Wi-Fi space. Mapping a Wi-Fi latent space (e.g., a latent space location L_1) to a physical space location (e.g., “office 203” or “kitchen”) requires a time-consuming labeling effort known as Wi-Fi fingerprinting [21].

For knowledge extraction, we argue that we do not need to know the physical location label for each Wi-Fi latent location. As long as the Wi-Fi latent locations are consistent, it is useful for knowledge extraction such that we can tell the subject always performs an activity at location L_1 , even when not knowing the exact physical location of L_i .

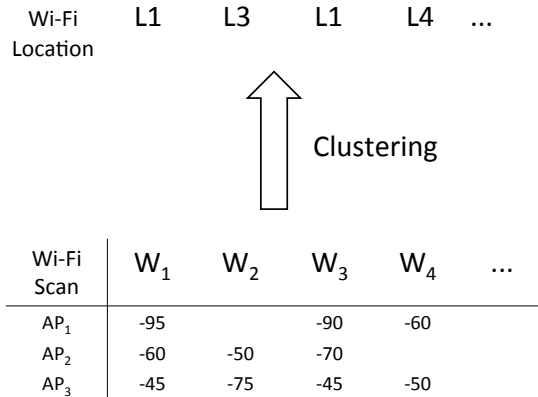


Fig. 5. Representing a sequence of n-dimensional Wi-Fi scans as a sequence of Wi-Fi latent locations labels.

Figure 5 shows the process of extracting location information of a user. First, a mobile device scans for the surrounding Wi-Fi APs and at each time t outputs an observed APs’ RSSI vector W_t . We cluster all W_t vectors into different groups based on the pair-wise similarities [21]. Each resulting cluster L_i contains similar RSSI vectors, which typically come from physical locations not too far away from each other. Figure 6 shows an example of clusters L_i mapped to the physical space. Finally, each Wi-Fi scan W_t is represented through its corresponding cluster L_i . This basically corresponds to mapping an n-dimensional vector into a one-dimensional label. The sequences of labels are used for further analysis as described in the following sections.

C. Activity Labeling

In order to infer the subject’s activity label, we use a two-stage unsupervised recognition approach shown in Figure 7.

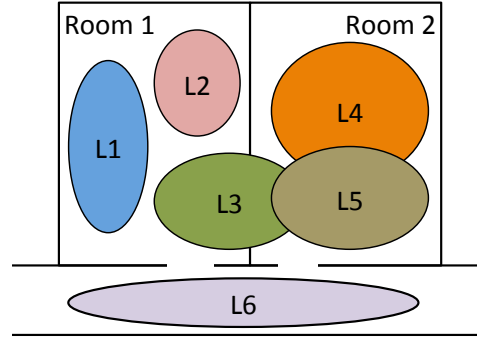


Fig. 6. Wi-Fi virtual location clusters L_i represented in a physical space.

This approach allows the system to recognize similar activities based on corresponding sensor reading patterns. The main idea is to leverage the structure and patterns of activities without having to know their semantic meaning. In the first stage of the activity recognition process we use raw sensor readings to infer motion labels. In the second stage, the inferred labels are used to identify high-level activities of a user. In the following, these two stages are described in detail.

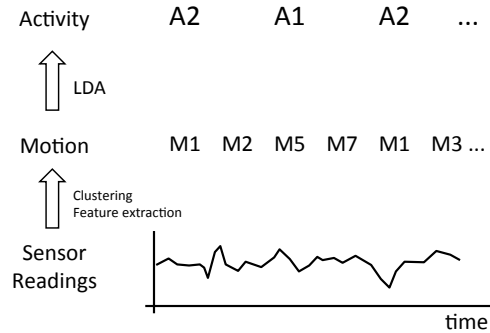


Fig. 7. Two-stage approach for activity inference. First, motion labels are identified through feature extraction and clustering. Then, by using LDA we infer activity labels and their activation scores.

1) *Motion Inference*: A “motion” is an atomic component of an activity. It is also known as a “primitive activity”. For example, a motion label can represent “moving forward” or “moving backward”. High-level activities are composed of low-level motions. For example, “mopping floor” is composed of “moving from one side to another with a constant back and forth movement of the hand” for a certain time period.

Acceleration: Given the 3-dimensional accelerometer data, we first compute the magnitude of the acceleration. Acceleration magnitude is independent from the orientation of the phone. This desired characteristic allows us to recognize similar activities independently from how the janitor stows the phone in a pocket. We apply a sliding window over the acceleration magnitude time-series and calculate mean and variance. These features allow us to differentiate between tasks of different motion intensities such as vacuuming compared to sitting on a chair and relaxing.

Orientation Change: The change of orientation is a good indicator for many activity types. For example, when the

janitor is mopping the floor, his body rotates from one side to another. The phone can detect significant changes in the orientation during this activity. When the janitor is vacuuming the carpet, his body leans mainly forwards and backwards. Thus, the rotation component of vacuuming is significantly lower than of the mopping activity. Similar to the acceleration features, we calculate the mean and variance of the magnitude of orientation changes over a sliding window.

For each window we extract 4 features: mean and variance of acceleration magnitude and mean and variance of the orientation change magnitude. Similarly to Wi-Fi scan, we quantize the feature space by clustering the 4-element vectors. Finally, we use the clustering results to represent each feature element as a motion label as shown in Figure 7.

2) *Activity Inference*: An activity can be represented by a sequence of relevant motions just as a topic can be described by a set of keywords. In this work, we use Latent Dirichlet Allocation (LDA) [2] to infer the underlying topic/activity distribution of a sequence of motion labels.

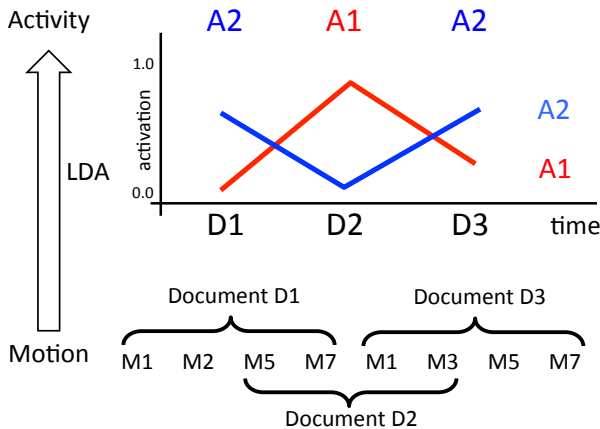


Fig. 8. Using LDA to first infer the activity activations of each topic.

LDA is commonly used in NLP to infer the latent topics of documents [2]. Given a set of documents and expected number of topics, LDA infers the words-in-topics distributions and topics-in-documents distributions. Words-in-topic distribution is a probability distribution of generating a certain word given the topic. Topics-in-document distribution is a probability distribution of a document being labeled with certain topics. Thus, from the result of LDA, we can infer a set of topics for a given document. Furthermore, each topic is assigned a score (so called activation score), which indicates a dominance of a topic in the document.

In this work, we consider activities as topics and subsequences of motion labels as documents. Thus, each subsequence of motion labels is assigned a distribution of activities with different activation scores. Figure 8 shows the process of using LDA for activity labeling. Given a sequence of motion labels we first extract documents. A document corresponds to a subsequence of motions labels over a time window. Empirically, we use sliding windows of 10 seconds wide. Two adjacent sliding windows overlap with each other for 50% of

the width in order to avoid losing information through the window transitions [4].

The set of created documents is used as an LDA input. LDA outputs a set of activity activation scores for each document. This can be interpreted as a user performing multiple activities at the same time. For example, while a janitor vacuums a room, he also moves chairs around. The example in Figure 8 shows two activities A1 and A2. In document 1, the activity A2 has significantly higher activation than the activity A1. In this work, a document is finally labeled with an activity with the highest activation score (e.g., document 1 is labeled as A2). However, labeling a document through a set of soft labels is also possible and will be explored in our future work.

D. Rule-based Knowledge Representation

We showed how context information can be inferred from raw sensor readings. This context information is further used for the knowledge extraction. Our system models knowledge as a set of IF-THEN rules, each composed of a condition and its corresponding action:

IF (condition) **THEN** (action)

Both the condition and action are composed of one or multiple statements, which are concatenated through an AND-operator. A statement is an element of time, location or activity. A rule can have a following format:

IF (*time is Monday 8am*) **AND** (*person is in kitchen*)
THEN (*person is mopping a floor*)

The condition contains two statements, one in time-statement and one location-statement. The action contains an activity-statement.

E. Knowledge Rules Learning

We frame the knowledge discovery problem as a rule extraction problem. Thus, we can utilize a large body of work in the area of rule learning [6], [13], [8]. Table II shows an example input for the learning algorithms. In addition to the location and activity information, we also use the quantized time information such as day of the week, hour and minute in the learning process. Proposed approach is not limited to this set of features and can be extended by other sensor modalities.

Id	Day	Hour	Minute	Location	Activity
1	d2	h11	m00	L3	A1
2	d2	h11	m00	L6	A4
3	d2	h11	m20	L4	A5
4	d2	h11	m20	L5	A6
5	d2	h11	m20	L3	A3
6	d2	h11	m40	L4	A5

TABLE II

FEATURES USED FOR THE EXTRACTING KNOWLEDGE RULES.

Extracted rules are in the following format:

R1: **IF** d2 **AND** L2 **THEN** A3
R2: **IF** h8 **AND** A1 **THEN** L1

In addition to the actual rules, rule extraction algorithms output meta information such as **confidence score** and **frequency**

count for each rule. Since for each condition multiple actions can be applied, the confidence score indicates the dominance of the action for the given condition. The frequency count of a rule pattern in the dataset indicates how often the worker follows these rules during his workday. Both the confidence scores and frequency counts provide valuable information about the work pattern. For example, frequent rules with high confidence scores are associated with tasks that the worker performs everyday at a specific time and location.

Due to the unsupervised nature of our proposed approach, the semantic meaning of each rule is unknown. However, the proposed approach can be combined with traditional supervised activity recognition techniques [1] to infer the meaning of rules of interest. We first use our unsupervised approach to extract the knowledge rules. The information such as confidence, frequency or time is used to identify rules of interest (e.g., the operational researchers are interested in all frequent rules appearing on Monday mornings). Then we ask the janitor to manually annotate the conditions and actions of these rules once they occur again while the janitor performs his everyday tasks. The design of the discussed query system is, however, not in the scope of this paper.

V. EXPERIMENTS

We present multiple experiments to demonstrate how extracted context information and knowledge rules can be used to analyze worker’s behavior patterns. As described in Section III-A we recruited a janitor working at our university for this study. The janitor is asked to carry a mobile phone in the pocket while performing his everyday tasks, which consist of cleaning and maintenance more than 30 rooms in a building. We collect data during his work hours for a period of 1 month.

A. Wi-Fi Location Inference

In the first experiment, we show results of the proposed Wi-Fi location labeling approach. We cluster the Wi-Fi scans using k-means clustering with k empirically set to 50. Thus, we have 50 different location/cluster labels L_1, \dots, L_{50} . Then we represent each Wi-Fi scan with a corresponding cluster label. Figure 9 shows the janitor’s Wi-Fi location labels over one workday (bottom part of the figure). Each color indicates a unique Wi-Fi location of the janitor at a certain time of a day. Through this visualization we can observe time periods when the janitor has a high level of mobility, i.e., frequent location changes (e.g., around 10:00am he collects trash from all the rooms). We can also observe time periods when the janitor is stationary (e.g., around 10:10am he has a coffee break in the kitchen).

The upper part of Figure 9 shows the variance of the acceleration magnitude. This measure indicates the level of motion intensity. The higher the value, the more intense is the activity. When variance is low, the janitor is either sitting (e.g., taking a break) or performing an activity without making a significant body movement (e.g., washing dishes, talking to a manager, etc.).

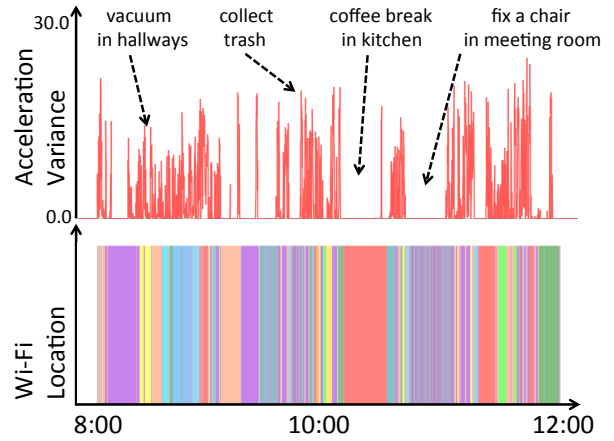


Fig. 9. The lower chart shows Wi-Fi locations of the janitor during his workday. Each color indicates a unique Wi-Fi location. The upper part of the figure shows the corresponding motion intensity of the janitor’s activities.

By combining the observations of both mobility and motion patterns, we get a better understanding of the janitor’s workflow. Table III shows example tasks for given mobility and motion intensity levels.

Mobility	Motion intensity	Example tasks
Low	Low	Relaxing in kitchen, fixing a chair
Low	High	Cleaning the bathroom
High	High	Collecting trash

TABLE III

EXAMPLE TASKS FOR A GIVEN MOBILITY AND MOTION INTENSITY LEVEL.

Figure 10 shows inferred Wi-Fi locations of the janitor over two consecutive workdays. Even though the exact physical location of the Wi-Fi location is unknown, we can observe similar sequential patterns over these two days.

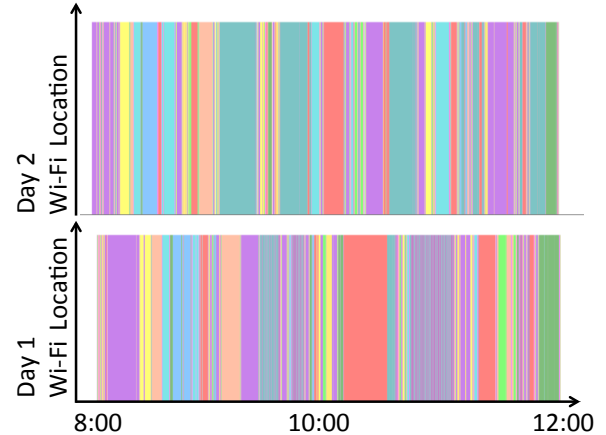


Fig. 10. Similar patterns of janitor’s Wi-Fi locations over 2 days. Best view in colors.

From the Wi-Fi location labels we can identify janitor’s location distributions during workdays. Figure 11 shows a histogram of occurrences of certain Wi-Fi locations. From the figure we can observe that a janitor usually visits the location

L19 early mornings and the location L15 rather in the middle of his workday.

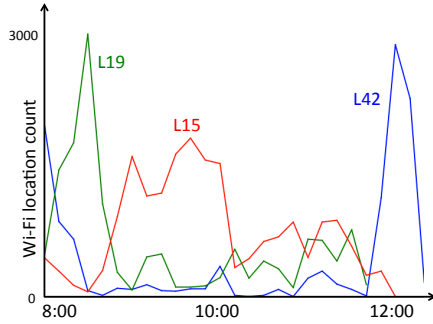


Fig. 11. Duration of visits at three locations for different time of the day. This shows that the janitor visits L19 usually at 8:20am, L15 around 9:30am and L42 at the beginning and ending of his work.

B. Activity Labeling

In the following, we show the results of using LDA for activity labeling. In this experiment, each LDA document contains 100 words (motion labels), which corresponds to 10 seconds of sensor data. For the initial configuration we set the number of LDA topics (i.e., the number of inferred activity types) to 10.

Figure 12 shows two activities which can be interpreted as "cleaning" and "maintenance". Each activity is assigned an activation score, which indicates how dominant is a certain activity at specific time. From the figure we can observe that the cleaning activity is highly dominant at the beginning of the day. This can be explained by the fact that the janitor starts with cleaning activities in the morning before students come to the campus. The maintenance activity such as "fixing chairs" is typically left to the later part of the workday.

By using LDA we can differentiate between phases when the phone is not carried by a janitor (e.g., the phone lies on the table) and when a janitor is sitting at one location. Even though raw accelerometer data appear to be similar (as shown on the top of Figure 12), the extracted features help to differentiate these two activities.

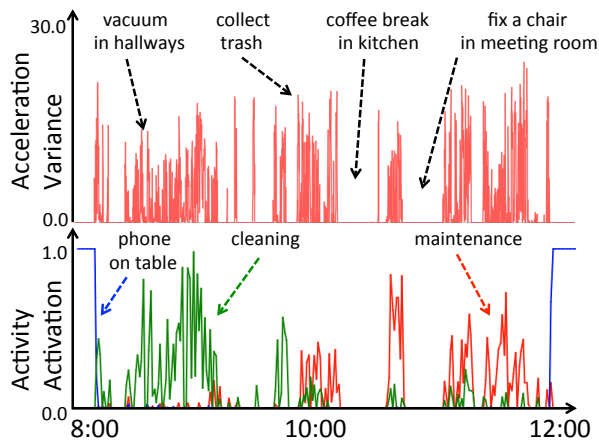


Fig. 12. Activity's activation scores indicate what time of the day the janitor performs which type of activities.

C. Accuracy of Extracted Rules

From the raw data, we infer context information such as location and activity, which are used for the rule extraction. In this experiment, we focus on rules, which have activities as the rule's actions, i.e., rules that describe what the janitor should do under certain conditions. We evaluate how well do the extracted rules generalize over an unseen dataset by using 10-fold cross-validation. In each cross-validation run, we use 90% of the data to build the rule set. This rule set is then used to predict actions of the remaining 10% of the unseen dataset.

1) *Accuracy of rule extraction algorithms:* For the initial experiment we set the number of location labels to 50 and the number of activity labels (LDA topics) to 10. We use the following rules extraction algorithms for the evaluation: Ripper [6], Decision Table [13], and PART [8],

Table IV shows the accuracy of the 10-fold cross validation. As the baseline we use a simple algorithm, which computes the most frequent activity label in the dataset and always uses this label for prediction. For the initial configuration and for many other tested configurations, PART consistently outperforms other rules extraction algorithms. Therefore, for the simplicity we show only the results of the PART algorithm in following experiments.

Algorithm	Accuracy
Baseline	38.2%
Ripper	53.9%
DecisionTable	63.3%
PART	71.3%

TABLE IV
ACCURACY OF THE 10-CV WHEN APPLYING THE ASSOCIATION RULES EXTRACTION ALGORITHMS ON THE DATASET.

For the janitor's dataset collected during 1 month we derive 67333 feature rows (as shown in Table II). Using the PART algorithm 1301 rules are extracted. The histogram of the rules' confidence is shown in Figure 13. Rules capture the generalized behavior pattern of the janitor. Rules with high confidence represent situations, when given a condition there is only one possible action (e.g. if it is 10:10am, the janitor will always be sitting in the kitchen).

Figure 14 shows the histogram of the rules' frequencies. We can observe that there many rules with low frequency, i.e., these rules capture the situations that rarely happen. Table V shows the rules, which have a confidence higher than 0.95 and occur more than 50 times. The condition of the extracted rules do not have to contain all the attributes. For example, the first rule does not have the day attribute. This rule can be interpreted as: "If it is 10:20am and the janitor is at location L4, he will be performing the activity A7, independently of which day of the week it is."

2) *Prediction accuracy vs. number of activities:* In this experiment, we analyze the relationship between the prediction accuracy and the number of activities. As shown in Figure 15, the prediction accuracy increases when the number of activities decreases. Intuitively, if we want LDA to extract only 2 types of activities, the results of LDA corresponds to 2 high-level

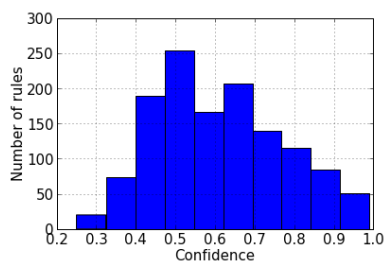


Fig. 13. Confidence of the rules extracted by using PART.

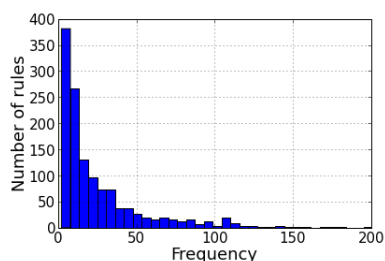


Fig. 14. Frequency of the rules indicate how often they were observed in the dataset.

activities. Thus, instead considering low-level activities such as “mopping the floor” or “vacuuming”, the inferred activity corresponds to high-level “cleaning”. For this reason, the lower is the number of activities, the more general are the rule’s actions and therefore the easier it is to predict actions given the conditions.

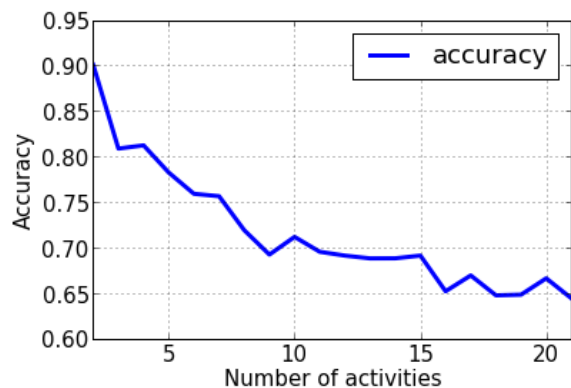


Fig. 15. It is easier to predict a janitor’s action with the learned rules if there is a smaller number of activities allowed in the knowledge base.

3) *Varying the document length in LDA*: So far we used a document containing motion labels captured during a window of 10 seconds. The document with this setting contains on average 100 words. Figure 16 shows the results when we increase the length of the document. By increasing the length we have a higher chance to capture long-duration activities. This would allow us to capture activities of an office worker, who spends the most of the time sitting at a desk and performing the same task. However, the duration of a janitor’s activity is relatively short. For example, when he mops a floor

Day	Hour	Minute	Location	Activity
	h10	m20	L4	A7
d6	h10	m20	L19	A3
d6	h9	m50	L31	A2
d5	h8		L42	A3
d2	h10	m20		A1
d3	h15		L4	A5

TABLE V

HIGH CONFIDENCE RULES, WHICH OCCUR FREQUENTLY IN THE DATASET.

THE FIRST RULE CAN BE INTERPRETED AS: “If it is 10:20am and the janitor is at location L4, he will be performing the activity A7, independently of which day of the week it is.”.

he often stops for a certain time period to let people walk by. Due to the frequent changes of the activity, using documents with long length results in lower prediction accuracy.

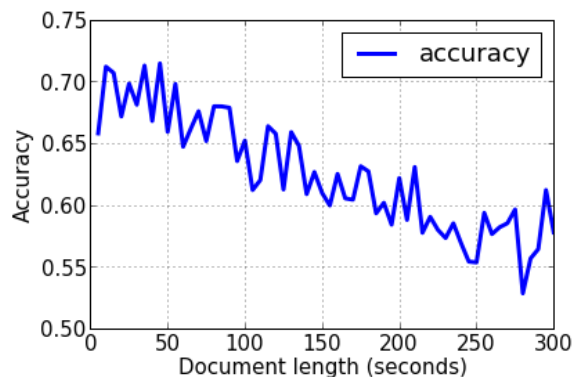


Fig. 16. By increasing the document length, we can capture long lasting activities. However, janitor’s activities have typically a short time span. Thus, by increasing the document length, we achieve lower prediction accuracy when applied on the janitor’s dataset.

4) *Extending LDA from words to n -grams*: The traditional LDA model is based on the concept of bag-of-words, which does not take into the consideration the order of words in a document. This has the advantage of being able to capture more flexible structures compared to n -gram models [20] or grammar induction techniques [23]. On the other hand, the traditional LDA approach does not capture the local dependencies/orders of words, which can be crucial to certain topics. Thus, we extend the standard LDA by using n -grams instead of words, i.e., we use “bag-of- n -grams” instead of “bag-of-words” to infer the latent topics of documents.

Figure 17 shows the prediction accuracy of the rule sets by setting n to various values. The highest accuracy of the prediction is achieved by setting n to 4. This is similar to the findings of n -gram language modeling in statistical natural language processing research: higher order n -grams models can improve the performance, however, going beyond 4-gram or 5-gram rarely improves the results and sometimes hurts the system performance due to data sparseness [5].

VI. CONCLUSION AND FUTURE WORK

In this work, we showed the feasibility of using smart phones to observe an employee’s mobility and physical activities. We proposed an unsupervised approach for work knowledge extraction by framing the problem as a natural

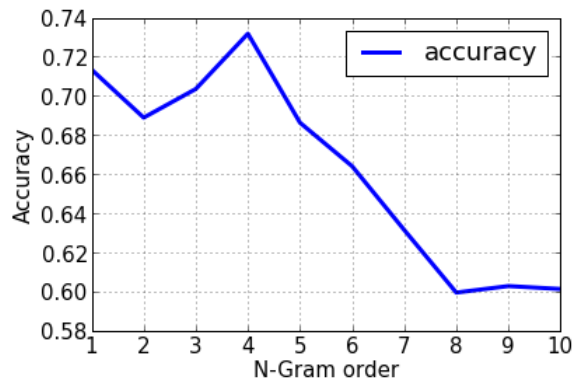


Fig. 17. Using n-gram up to $n = 4$ allows us to capture the useful information about the order of words / motions labels in a document. However, n-gram with higher order causes sparsity issues resulting in the decrease of the prediction accuracy.

language processing problem. Through the experiments we showed that the extracted knowledge rules capture the general patterns of an worker's behavior.

Our current system is limited by the sensing capability of the mobile devices. In the future work, we will explore other sensing modalities such as using smart watches or digital glasses to capture various aspects of an worker's experience. Moreover, we will extend our approach by building a system, which leverages the extracted knowledge rules to automatically identify the activities of interest and ask a worker to label them.

VII. ACKNOWLEDGEMENT

This work is supported in part by the National Science Foundation through the Smart and Connected Health program under the award IIS1344768.

REFERENCES

- [1] L. Bao and S. S. Intille. Activity recognition from user-annotated acceleration data. In *Pervasive Computing*, pages 1–17. Springer, 2004.
- [2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022, Mar. 2003.
- [3] O. Brdiczka and V. Bellotti. Identifying routine and telltale activity patterns in knowledge work. In *Semantic Computing (ICSC), 2011 Fifth IEEE International Conference on*, pages 95–101. IEEE, 2011.
- [4] P.-W. Chen, S. K. Chennuru, S. Buthpitiya, and Y. Zhang. A language-based approach to indexing heterogeneous multimedia lifelog. In W. Gao, C.-H. Lee, J. Yang, X. Chen, M. Eskenazi, and Z. Zhang, editors, *Proceedings of the 12th International Conference on Multimodal Interfaces / 7. International Workshop on Machine Learning for Multimodal Interaction, ICMI-MLMI 2010, Beijing, China, November 8-12, 2010*, page 26. ACM, 2010.
- [5] S. F. Chen and J. Goodman. An empirical study of smoothing techniques for language modeling. In A. Joshi and M. Palmer, editors, *Proceedings of the Thirty-Fourth Annual Meeting of the Association for Computational Linguistics*, pages 310–318, San Francisco, 1996. Morgan Kaufmann Publishers.
- [6] W. W. Cohen. Fast effective rule induction. In *Proceedings of the Twelfth International Conference on Machine Learning*, pages 115–123. Morgan Kaufmann, 1995.
- [7] J. Diederich, I. Ruhmann, and M. May. Kriton: a knowledge-acquisition tool for expert systems. *International Journal of Man-Machine Studies*, 26(1):29–40, 1987.
- [8] E. Frank and I. H. Witten. Generating accurate rule sets without global optimization. pages 144–151. Morgan Kaufmann, 1998.
- [9] H. Hashimoto, I. Yoshida, Y. Teramoto, H. Tabata, and C. Han. Extraction of tacit knowledge as expert engineer's skill based on mixed human sensing. In *RO-MAN, 2011 IEEE*, pages 413–418, 2011.
- [10] R. C. Holte. Very simple classification rules perform well on most commonly used datasets. In *Machine Learning*, pages 63–91, 1993.
- [11] T. Huynh, M. Fritz, and B. Schiele. Discovery of activity patterns using topic models. In *Proceedings of the 10th international conference on Ubiquitous computing*, UbiComp '08, pages 10–19, New York, NY, USA, 2008. ACM.
- [12] T. W. Jackson, S. Tedmori, C. J. Hinde, and A. I. Bani-Hani. The boundaries of natural language processing techniques in extracting knowledge from emails. *Journal of Emerging Technologies in Web Intelligence*, 4(2):119–127, 2012.
- [13] R. Kohavi. The power of decision tables. In *Proceedings of the European Conference on Machine Learning*, pages 174–189. Springer Verlag, 1995.
- [14] A. Kusiak. Process planning: a knowledge-based and optimization perspective. *Robotics and Automation, IEEE Transactions on*, 7(3):257–266, 1991.
- [15] H. Lin, Y. Zhang, M. Griss, and I. Landa. Wasp: an enhanced indoor locationing algorithm for a congested wi-fi environment. In *Proceedings of the 2nd international conference on Mobile entity localization and tracking in GPS-less environments*, MELT'09, pages 183–196, Berlin, Heidelberg, 2009. Springer-Verlag.
- [16] A. Majchrzak, C. Wagner, and D. Yates. Corporate wiki users: results of a survey. In *Proceedings of the 2006 international symposium on Wikis*, pages 99–104. ACM, 2006.
- [17] W. B. J. Martz and M. M. Shepherd. Testing for the transfer of tacit knowledge: making a case for implicit learning. *Decision Sciences Journal of Innovative Education*, 1(1):41–56+, 2003.
- [18] S. McDonald. Studying actions in context: a qualitative shadowing method for organizational research. *Qualitative research*, 5(4):455–473, 2005.
- [19] E. Morganti, L. Angelini, A. Adami, D. Lalanne, L. Lorenzelli, and E. Mugellini. A smart watch with embedded sensors to recognize objects, grasps and forearm gestures. *Procedia Engineering*, 41:1169–1175, 2012.
- [20] L. Nguyen, H.-T. Cheng, P. Wu, S. Buthpitiya, and Y. Zhang. Pnlum: System for prediction of next location for users with mobility. In *Nokia Mobile Data Challenge 2012 Workshop. p. Dedicated challenge*, volume 2, 2012.
- [21] L. T. Nguyen and J. Zhang. Wi-fi fingerprinting through active learning using smartphones. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, UbiComp '13 Adjunct, pages 969–976, New York, NY, USA, 2013. ACM.
- [22] N. M. Oliver, B. Rosario, and A. P. Pentland. A bayesian computer vision system for modeling human interactions. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 22(8):831–843, 2000.
- [23] H.-K. Peng, P. Wu, J. Zhu, and J. Y. Zhang. Helix: Unsupervised grammar induction for structured activity recognition. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining, ICDM '11*, pages 1194–1199, Washington, DC, USA, 2011. IEEE Computer Society.
- [24] M. Polanyi. *The Tacit Dimension*. Terry lectures. Anchor Books, 1967.
- [25] J. Ryu, Y. Jung, K.-m. Kim, and S. H. Myaeng. Automatic extraction of human activity knowledge from Method-Describing web articles. *Proceedings of the 1st Workshop on Automated Knowledge Base Construction*, 2010.
- [26] T. Starner. Wearable computing: through the looking glass. In *Proceedings of the 17th annual international symposium on International symposium on wearable computers*, pages 125–126. ACM, 2013.
- [27] R. Sternberg and J. Horvath. *Tacit Knowledge in Professional Practice: Researcher and Practitioner Perspectives*. Taylor & Francis, 1999.
- [28] P. Wu, H.-K. Peng, J. Zhu, and Y. Zhang. Senscare: Semi-automatic activity summarization system for elderly care. In J. Zhang, J. Wilkiewicz, and A. Nahapetian, editors, *Mobile Computing, Applications, and Services*, volume 95 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 1–19. Springer Berlin Heidelberg, 2012.