Predicting “About-to-Eat” Moments for Just-in-Time Eating Intervention

Tauhidur Rahman
Information Science
Cornell University
Ithaca, New York
tr266@cornell.edu

Mary Czerwinski
VIBE Group
Microsoft Research
Redmond, Washington
marycz@microsoft.com

Ran Gilad-Bachrach
VIBE Group
Microsoft Research
Redmond, Washington
rang@microsoft.com

Paul Johns
VIBE Group
Microsoft Research
Redmond, Washington
paul.johns@microsoft.com

ABSTRACT

Various wearable sensors capturing body vibration, jaw movement, hand gesture, etc., have shown promise in detecting when one is currently eating. However, based on existing literature and user surveys conducted in this study, we argue that a Just-in-Time eating intervention, triggered upon detecting a current eating event, is sub-optimal. An eating intervention triggered at “About-to-Eat” moments could provide users with a further opportunity to adopt a better and healthier eating behavior. In this work, we present a wearable sensing framework that predicts “About-to-Eat” moments and the “Time until the Next Eating Event”. The wearable sensing framework consists of an array of sensors that capture physical activity, location, heart rate, electrodermal activity, skin temperature and caloric expenditure. Using signal processing and machine learning on this raw multimodal sensor stream, we train an “About-to-Eat” moment classifier that reaches an average recall of 77%. The “Time until the Next Eating Event” regression model attains a correlation coefficient of 0.49. Personalization further increases the performance of both of the models to an average recall of 85% and correlation coefficient of 0.65. The contributions of this paper include user surveys related to this problem, the design of a system to predict about to eat moments and a regression model used to train multimodal sensory data in real time for potential eating interventions for the user.

Keywords

Eating Habit Modeling, Next Eating Event Prediction, Just-in-Time Eating Intervention

1. INTRODUCTION

Irregular eating habits and disproportionate or inadequate dietary behaviors may increase the likelihood of severe health issues, including obesity. According to the World Health Organization (WHO), more than 1.9 billion adults 18 years and older were overweight in 2014 [9]. In the United States, two out of every three adults is considered overweight or obese, which cost about 147 billion US dollars in the year 2000 [2]. Obesity is a leading cause of preventable death second only to smoking, resulting in 2.5 million deaths per year [3]. In order to tackle this obesity epidemic, we need to look at its root cause, which is the energy imbalance between physical activity and eating. Although ubiquitous and wearable technologies (e.g., fitbit [7]) have already been proven successful at physical activity estimation, tracking of eating events is yet to be popular among most users. Manual self-reporting of various eating events with the help of a mobile device is the most common food journaling tool (Fitbit [7], MyFitnessPal [4], Caroll et al. [16]). However, the required high level of engagement very quickly leads the users towards fatigue and to lower level of compliance as the novelty fades away [15, 19, 17]. Many recent studies have proposed semi-automatic and fully automatic food journaling tools. These tools employ different sensor systems (e.g., camera, accelerometer, microphone) and human-in-the-loop techniques (e.g., Amazon Mechanical Turk, contextual recall interface) in order to address the problem with manual self-reporting of eating event. We present a systematic review of different eating event detection/tracking solutions in the Related Work section.

Although a passive current eating event detector/tracker could be very useful for building self-monitoring or self-reflective interfaces, for a Just-in-Time eating intervention, detecting the current eating event itself is insufficient. What we really need is to be able to detect moments when we are “about to eat” so that we can trigger healthy eating interventions just prior to the actual eating events. The intuition behind this claim is that, in order to change the course of action towards better and healthier eating behavior, Just-in-Time interventions might better influence users if they are triggered just prior to an actual eating event. Wansink et al. showed in a recent study [44] that adults consume about 92% of what we serve ourselves on our dinner plates, irrespective of our perceived self-control, emotional state or other external variables. As a result once the process of eating has begun, it is more difficult to alter or stop this process in the interest of health. There is almost no prior work that addressed this challenge of detecting “About-to-Eat” moments. Whether or not it is possible to tell ahead of time that one is going to have an eating event in the next N minutes using multimodal sensor data is the primary research question in this paper.

DH ’16, April 11-13, 2016, Montréal, QC, Canada

DOI: http://dx.doi.org/10.1145/2896338.2896359

Technique	Study	Sensor	Information
Manual Self-Report	Fitbit [7], MyFitnessPal [4], Caroll et al. [16]	-	-
Contextual Recall	Arab et al. [14], Reddy et al. [36]	Camera	Photos of Food
Crowd-sourcing	Noronha et al. [32], Thomaz et al. [41]	Camera	Photos of Food
Passive Sensing	Stellar and Shrager [39]	Strain Gauge	Tongue Pressure and Flexing during Chewing
Passive Sensing	Yatani and Truong [46], Amft et al. [12, 11, 13] Shuzo et al. [38]	Microphone	Body Sound
Passive Sensing	Amft et al. [10], Dong et al. [20] Thomaz et al. [40]	Inertial Sensors	Arm Movement
Passive Sensing	Kalantarian et al. [25], Rahman et al. [35] Sazonov et al. [37]	Piezo-electric Sensor	Swallow and Mastication
Passive Sensing	Kong and Tan [26], Zhu et al.[47]	Camera	Food Amount and Type
Passive Sensing	Liu et al. [27]	Microphone and Camera	Chewing, Photo of Food
Passive Sensing	Thomaz et al. [42]	Microphone	Ambient Sound
Instrumented Utensil	Kadomura et al. [24]	Bio-impedance Color Inertial Sensor	Food Chemistry Food Color Arm Movement
Scanning Shopping Receipts	Mankoff et al. [28]	OCR Scanner	Name of Food items Nutritional Content

Table 1: Overview of the eating event detection/food journaling related literature

In this work, we present a wearable sensing framework that predicts "About-to-Eat" moments and the "Time until the Next Eating Event". We use an array of sensors that are currently not available in a single device. Therefore, for the sake of the experiment, we use a Microsoft Band [8], an Affectiva Q sensor [1], a wearable microphone, and an Android smartphone application. Using Microsoft Band [8], we passively and continuously captured users' physical movement (raw accelerometer, gyroscope, step count, speed), caloric expenditure, heart rate, skin temperature, etc. The Affectiva Q sensor [1] also has a form factor of a wrist-band and measures electrodermal activity, which is a good indicator of psychological arousal. We used a wearable microphone (similar to BodyBeat [35]) that continuously monitors chewing and swallowing sounds and detects current eating events. Lastly, a smartphone was used for continuous and passive capture of GPS location and for recording self-reports before and after every eating event. Using all these physical and physiological variables, we extracted window-level features, selected relevant feature subsets and trained machine learning models that predict the "Time until the Next Eating Event" and detect "About-to-Eat" moments.

2. RELATED WORK

2.1 Eating Event Tracking

Table 1 outlines the related works in the area of eating tracking and food journaling. The traditional way of doing food journaling is through various self-reporting tools such as paper diaries or mobile devices. Many wearable fitness trackers (e.g., fitbit [7]) and a few commercial smartphone applications (e.g., myfitnesspal [4]) provide their users with a food database to help them with the manual food logging. However, this method is still very time consuming and typically requires high level of engagement, which can quickly lead to lower level of compliance [15, 19]. In a very recent study on food journaling, Cordeiro et al. found that users' forgetfulness, difficulty to log and privacy concerns in social settings are some of the main cause of a low level of compliance [18]. As one misses food journal entries, slowly the food journal loses its value and credibility, which eventually leads them to abandoning the tool.

Researchers have been trying to develop automatic and passive sensing mechanisms to detect eating events in order to address these problems, but most solutions have included manual self-report. Some researchers have resorted to human computation or crowd-sourcing to tackle this problem given its difficulty. [32, 41]. Users have taken photos of their food using some wearable camera or smartphone's built-in camera, after which the photos are processed and uploaded to online crowdsourcing frameworks to get labels and nutritional information associated with the food. However, many users are resistant to the idea of sharing continuous unfiltered pictures with crowd workers. Some studies automated the process of food labeling using computer vision algorithms to recognize food [26, 47]. Food identification from pictures is a very difficult task for both humans and computer vision algorithms, due to various confounding factors including lighting conditions, quality of photos and type of food. In 1985, Stellar and Shrager used an oral strain gauge that measured tongue pressure and flexing during chewing to monitor eating events throughout the day [39]. More recently, Amft et al. [12, 11, 13] and Yatani and Truong [46] used wearable microphones to detect different eating events by recognizing the sounds that are generated in the process of mastication, or swallowing. Similarly, some other studies explored the applicability of body-worn inertial sensors to detect characteristic arm movements during eating [10, 20, 40]. Some studies used multiple sensor streams instead of relying on one and demonstrated an increase of performance of a current eating detector [27]. Although none of them have really taken off, the design and implementation of these ubiquitous wearable technologies have inspired us for this study. Some studies instrumented utensils related to eating in order to capture eating events. For example, Kadomura et al. instrumented a fork using bio-impedance, color and inertial sensor and detected eating related activities [24]. Although instrumentation of such utensils may work very well for constrained situations, this technique's major shortcoming comes from the fact that we typically don't carry a particular eating utensil and use it repeatedly. Mankoff et al. studied some indirect measures of eating behavior by processing and analyzing the name of the food items in shopping receipts [28].

Many of the techniques outlined in Table 1 can track current eating events in real time, which can be useful for building a self-reflective tool. We argue in this paper that, from an active eating intervention perspective, just current eating event tracking is not enough. Predictability of the "About-to-Eat" event is very important for Just-in-Time interventions, which we discuss next.

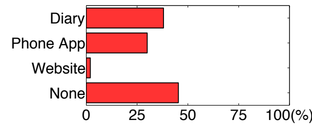
2.2 Just-in-Time Intervention

Many prior research efforts in a variety of health domains also found that timing of when to intervene is a crucial part of the general framework of Behavioral Intervention Technologies (BITs) [29]. Just-in-Time interventions are crafted to provide support at the opportune moment so that positive behaviors can be enlisted [30]. Many Just-in-Time interventions are triggered upon detecting certain events or conditions which are deemed as the optimal moment for intervention. Such optimal moments are often associated with high risk or vulnerability coupled with an ineffective coping response, which may easily lead someone towards decreased self-efficacy and possibly to relapse [45]. Many researchers working in the area of alcohol, drug problems, smoking addiction and stress [21, 45] used high risk moments as opportune moments for triggering Just-in-Time interventions as patient gets the opportunity to cope, divert or circumvent the course of life which constitutes the negative health outcome. For example, a Just-in-Time intervention for a recovering alcoholic patient might be a warning from the mobile device carried by the patient when it detects that she is approaching a liquor store. Similar Just-in-Time interventions are also explored in the context of smoking cessation where interventions are triggered when the participants have a high urge to smoke and just prior to a smoking event. In another recent work, Pina et al. [33] studied just-in-time stress coping interventions for parents of children with ADHD and found that interventions prompted just before a full escalation of stress were more useful as they were especially receptive to an intervention strategy at that time. Now, in the context of the Just-in-Time eating intervention the question is, "What is the optimal moment to nudge or intervene in order to change the course of one's action towards better and healthier eating behavior?" In a very recent study, Wansink et al. [44] showed that we (adults) are highly likely to eat close to the entirety of what we serve ourselves irrespective of our perceived self-control, emotional state or other external variables. Therefore, we argue that just-in-time interventions for eating disorders should be delivered before the eating event. Although prior work demonstrated ubiquitous and wearable technologies for automatic and passive eating event detection in real time, no work, to our knowledge, proposed similar technology for "About-to-Eat" event prediction.

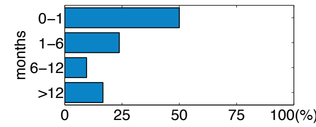
3. EATING TRACKING AND INTERVENTION SURVEY

We conducted an eating tracking and intervention survey to learn about the pitfalls of using existing food journaling tools and to sketch the design considerations for a Just-in-Time eating intervention. This survey not only informed us about the gap in existing eating tracking practices but also helped us to realize the importance of the "About-to-Eat" moments prediction for designing an effective Just-in-Time eating intervention. We surveyed 75 participants (30 female, 45 male), who were recruited from a research division of a technology company in the United States via its' local mailing list. The age of the participants varied from 17 to 56 with a median age of 34.5. The height ranged from 60 to 75 inches with a median height of 69 inches. The weight ranged from 102 lbs to 280 lbs with a median weight of 156 pounds. The participant pool was also diverse in terms of their Body Mass Index (BMI) scores. The percentage of participants falling under the categories of underweight, normal, overweight and obese are respectively 1%, 63.2%, 25% and 10.8%. Figure 1 shows the survey results.

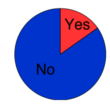
All of the survey participants were first asked about the types of eating tracking or food journaling tools they had ever used. While responding to this question, 34 out of 75 respondents shared that they had never used any such tools. Other responses included using a food diary (27 respondents), smartphone applications (21 respondents), or websites (2 respondents) for food tracking. Participants



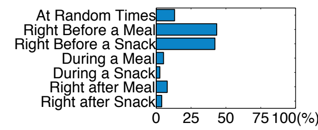
(a) Which of the above food journaling tools have you ever used?



(b) If you have ever used any of the food journaling tools (in question a), how long did you use it?



(c) If you have ever used any of the food journaling tools (in question a), do you still use it now?



(d) If you had a smartphone application that could assist you in healthy eating decisions, when would you want the app to intervene?

Figure 1: Eating tracking and intervention survey results.

who had ever used any food tracking tools were then also asked, "How long did you use the tool?". About 48.9% of the respondents told us that they had used the tools for less than a month, while only 24.4%, 9.7% and 17.0% of the respondents have used the tool respectively for 1-6 months, 6-12 months and more than a year. Responding to: "Do you still use the tool now?" an overwhelming majority (65 out of 75 respondents) shared that they do not use the eating tracking tools anymore. These statistics clearly show that most of the existing food tracking tools were used for only a brief period of time and then the lack of novelty and high level of engagement have led the majority (86.7%) of the users to quit. A recent study on food journaling conducted by Cordeiro et al. [18] echo the same findings of low levels of compliance with food journaling tools.

We asked our survey participants a few more questions related to interventions. In order to learn about users' intervention timing preferences, the participants were asked "If you had a smartphone application that could assist you in healthy eating decisions, when would you want the app to intervene?". The most common response (33 out of 75) was that they wanted the application to intervene right before a meal or a snack. Thus about 43% of the respondents preferred to be intervened during the "About-to-Eat" moments to be nudged towards healthy eating decision making. This result backs up the fundamental premise of this paper: the usefulness of "About-to-Eat" moment detection. It clearly shows that there is an agreement between our assumption about the usefulness of the "About-to-Eat" moment detection and the preference of users. Some qualitative, open-ended responses included: "If the app could intervene before (or at the start) of a snack to remind me of my calorie intake for the day and make me more conscious of my choices that would be ideal", "Maybe at restaurants, to help me decide what to order.", "Before meal preparation", etc. Both the quantitative and qualitative feedback indicated that most of the participants wanted the interventions to be triggered right before an eating event (either a meal or a snack) at home or at a restaurant, preferably before serving or ordering. One participant nicely summarized it all by stating: "a little before I get hungry, just in time for me to make a good decision and then have time to take action on it".

In order to identify potential smartphone-based eating intervention, our survey also asked "What would you want a healthy eating smartphone app to do for you?". 42 out of 75 respondents wanted to use a smartphone application calorie calculator for the food that they are about to eat. 20 respondents wanted to receive a reminder to eat a balanced meal. 36 survey participants wanted to receive a reminder of calorie allowance. 37 people wanted to see a visualization of what has been eaten so far. A few participants (5) wanted a breakdown of calories from protein, carbohydrates, fat and even sodium content. One participant reported, "For me it's less about calories and more about varied intake - my favorite apps were ones that helped me keep track of food groups and nutrients. It was helpful to realize I'd hardly eaten any veggies, or I was over on fats". Another interesting suggestion was related to drinking water. It said "Remind me to drink 8 glasses of water, and even track hydration vs dehydration."

4. DATA COLLECTION

"About-to-Eat" moments are temporal episodes that precede an eating event (e.g., breakfast, lunch, dinner or snack). In this study we wanted to explore if we can predict "About-to-Eat" moments with the help of a set of common ubiquitous and wearable technologies. In order to achieve this goal we recruited 8 participants (3 female, 5 male) for this study from a research division of a technology company. The participants' age ranged from 26 to 54 years. After we received our participants consent, they were asked to fill in a pre-experimental survey in order to obtain background information about our participants. They were then introduced to the four ubiquitous and wearable technologies and were asked to use these technologies for 5 days.

4.1 Ubiquitous and Wearable Technologies

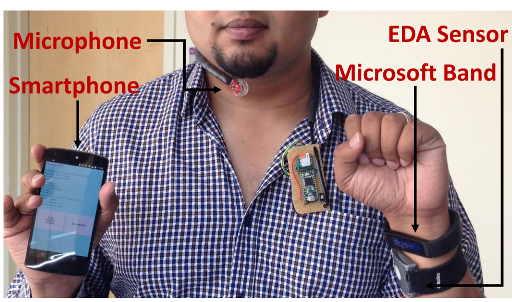


Figure 2: The ubiquitous and wearable technologies used for our data collection is outlined here.

Prior work on binge eating behavior and emotional eating has demonstrated that several physiological variables including heart rate, electrodermal activity (EDA), finger pulse amplitude and heart rate variability show different trends during food exposure period [43, 31]. Informed by these prior works, we incorporated some of these relevant sensor streams like heart rate and EDA in our work. In addition to that, eating is a very habitual daily activity. Knowing location (GPS) traces, the timing and duration of various physical activity traces (step count, calorie expenditure, speed traces), when users do particular types of gestures/movements (via accelerometer and gyroscope) could all help in predicting "About-to-Eat" moments. In order to capture all these relevant physical and physiological health variables, we used four wearable technologies including a Microsoft Band [8], an Affectiva Q sensor [1], a wearable microphone and an Android smartphone. Figure 2 shows all the four technologies on a user. Table 2 lists all the raw sensor streams and their preprocessing steps.

The Wearable Microphone was used to detect current eating events by detecting characteristic eating sounds like mastication and swallowing. The microphone was directly attached to the skin around Laryngopharynx area in the neck. The wearable microphone was made using an off-the-shelf electret microphone [6], Teensy ARM microcontroller unit [5] and a SD card interface. Although detecting current eating events were not the focus of this paper, we used it to explore how this microphone-based eating moment detection works in the wild. In addition to that, a current eating event detector can obviate the need of manual logging for training or personalizing "About-to-Eat" moment prediction model. The Android Smartphone Application was used for three reasons. Firstly, the smartphone application continuously and passively records GPS and network location in terms of latitude and longitude at every minute. Secondly, the smartphone application receives Microsoft Band's sensor data over Bluetooth and stores it in its external memory. We have used Microsoft Band's software development kit [8] for developing this part of the application. Thirdly, it allows the participants to manually self-report the start and end of a particular eating event, affect, stress level, craving, hunger, satiation. In the next subsection, we discuss the user interface of the android smartphone application.

4.2 Self-Report User Interface

The Android smartphone application contains a self-report user interface, which is shown in Figure 3. The primary goal of the self-report user interface was to enable participants to log the starting time and ending time of different eating events. Right before and after an actual eating event, our participants can declare ground truth start time and end time of an eating event simply by tapping on a button (shown in Figure 3b). For the sake of consistency of our data collection, we explicitly advised our participants to tap on the "Start of Eating Event" button exactly before having the first bite (neither while deciding which food to order/cook nor while waiting for the food to be cooked/prepared in a home or restaurant setting). Similarly, the end of the eating event was defined by the moments after the last bite or gulp of the food and can be self-reported by tapping on the "End of Eating Event" button. In addition to this, the UI elicits other contextual and perceptual information around the eating event. When one taps on the "Start of Eating Event" button in Figure 3b, the user is also asked about their current emotional state using the Photographic Affect Meter (PAM in Figure 3c) [34], meal type (Breakfast, Brunch, Lunch, Dinner or Snack), intensity of desire/craving and hunger (in a numeric scale from 1 to 7) right before an eating event. Following an actual eating event, the participant can also log their emotional state, amount of food consumed, satiation and healthiness (all using a numeric scale from 1 to 7) of the food.

5. PREDICTING "ABOUT-TO-EAT" MOMENTS

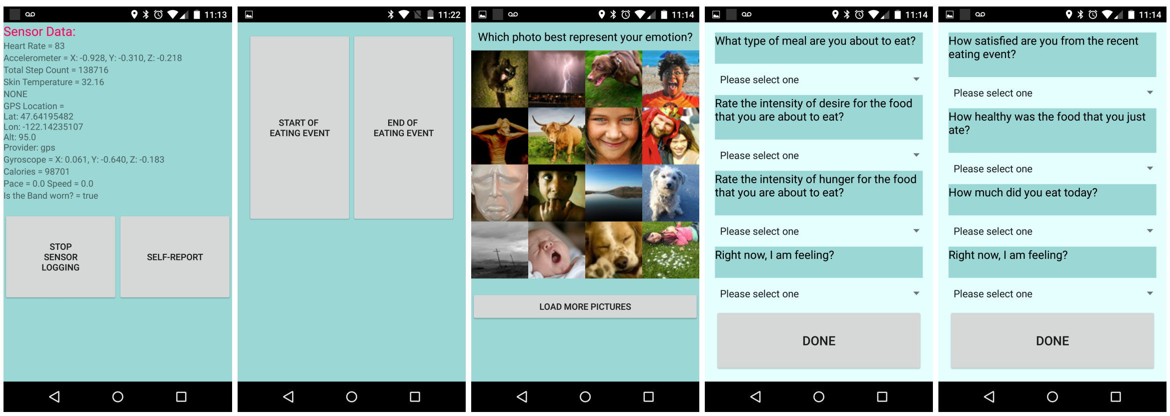
Modeling "About-to-Eat" moments consists of different steps including data cleaning and preprocessing, feature extraction, feature selection and machine learning.

5.1 Preprocessing

The primary purpose of the data preprocessing step is to make sure that the raw sensor data streams are clean and are off high fidelity. All the sensor streams captured by the ubiquitous computing frameworks (i.e., Microsoft Band, Affective Q sensor and Android application) are time-stamped with their own real-time clock. At first, the sensor streams are time aligned. The Microsoft Band's sampling rate varies a bit, so The sensor streams captured by the Microsoft Band are resampled to a fixed frequency. Microsoft Band

Sensor	Stream	Preprocessing
Microsoft Band	Heart Rate	Resampling
	Skin Temperature	Resampling
	Accelerometer	Norm of 3 dimensional linear acceleration
	Gyroscope	Norm of 3 dimensional angular acceleration
	Step Count	Resampling, Differentiation to estimate instantaneous value
	Calorie	Resampling, Differentiation to estimate instantaneous value
	Speed	Resampling
Affectiva Q Sensor	Electrodermal Activity	Remove mean, Lowpass filter with cutoff at 0.05Hz is applied to estimate EDA Tonic signal, Bandpass filter with cutoffs at 0.05Hz and 1Hz is applied to estimate EDA Phasic signal
Wearable Microphone	Body Sound	Chewing and Swallowing Recognition
Smartphone Application	GPS and Network Location	Extracted Latitude and Longitude
	Self-Report	Log start and end time of eating, Affect, Stress, Craving, Hunger, Satiation

Table 2: The list of ubiquitous and wearable technologies used for data collection



(a) (b) (c) (d) (e)

Figure 3: (a) Homepage of the Android application (b) Tap at the Self-report button in the homepage prompts user to select between "Start of Eating Event" or "End of Eating Event" button. (c) After tapping on either "Start of Eating Event" or "End of eating event" button, the user is at first asked to record their current emotional state using Photographic Affect Meter (PAM). (d) User is then asked about their type of meal, craving, hunger and stress using this questionnaire at the start of an eating event. (e) User inputs their satisfaction, healthiness, amount of consumption and stress at the end of an eating event.

returns cumulative values for both the calorie expenditure and the step count. To get the instantaneous values for both of these sensor streams, we use differentiation on the interpolated signal. The raw Electrodermal Activity (EDA) signal can be decomposed into two parts: the long term and instantaneous response of physiological arousal (called tonic and phasic respectively). The long term slow changing part of the raw EDA signal is the EDA tonic signal, where the faster changing part is considered as the EDA phasic signal. We applied a Butterworth low pass filter with a cutoff frequency of 0.05 Hz to estimate the EDA tonic signal and used a band pass filter with cutoff frequencies at 0.05 and 1 Hz to estimate the EDA phasic signal. Table 2 lists all the preprocessing steps applied to the raw sensor streams.

5.2 Feature Extraction

The first step of feature extraction starts with windowing the processed sensor time series and here we have consider two types of windowing parameters: the feature extraction window size and feature extraction window shift size. While the feature extraction window size determines the duration of processed sensor time series data in a particular window, the shift determines the amount of time shift in two adjacent windows. A small feature extraction window size could capture instantaneous characteristics or properties in the sensor time series, while features extracted with a coarse feature extraction window size could provide information about long term trend. As we wanted to determine an optimal feature extraction window size, we used different window sizes from 5 to 120 minutes. The optimal window length is determined empirically based on the performance of our prediction model. Irrespective of the differences among different feature extraction window sizes, we have used a constant window shift of one minute. It means that n^{th} window is shifted by one minute in comparison to $(n-1)^{th}$ window. Also note that the window shift size determines the resolution of the prediction/inference, as the machine learning models outputs a label for each window.

In order to extract features, a set of statistical functions was applied on each window. Table 3 lists all of the statistical functions that were applied for feature extraction to capture the different aspects of the windowed sensor streams. In addition to all of these sensor streams based window-level features, we extracted two additional types of features corresponding to each window. Firstly, we used current time in minutes since the start of the day as a feature. The second feature type was previous eating event-based feature. It includes the time since the last eating event in minutes and the number of previous eating events since the beginning of the day. In order to detect eating events (not to predict them) the wearable microphone along with a mastication and swallowing sound detection algorithm can be used. We followed the mastication and swallowing sound detection algorithm presented in a recent study [35]. In total, we extracted 158 window-level features.

Type	Statistical Functions	Acronym
Extremes	Minimum	min
	Maximum	max
Average	Mean	mean
	Root Mean Square	RMS
Quartiles	1st, 2nd and 3rd Quartile	qrtl25, qrtl50 qrtl75
Dispersion	Standard Deviation	std
	Interquartile Range	iqrl
Peaks	Number of peaks	numOfPeaks
	Mean Distance of Peaks	meanDistPeaks
	Mean Amplitude of Peaks	meanAmpPeaks
Rate of Change	Mean Crossing Rate	mcr
Shape	Linear Regression Slope	slope

Table 3: The list of statistical functions applied to the windowed sensor data for extracting window-level features

5.3 Feature Selection

In our work, we used the Correlation-based Feature Selection (CFS) criteria [23] along with the best first search to select the subset of features. The CFS algorithm evaluates the goodness of features based on two criteria: firstly, all the features in the feature subset are highly indicative of the target class, and secondly, the features in the feature subset are highly uncorrelated with each other, thus each feature can provide complementary information. Figure 4 shows the contribution of different feature groups for predicting "About-to-Eat" moments. How much a drop in performance a classifier concedes if a person-independent "About-to-Eat" moment classifier is trained without a particular group of features, is the metric we used to estimate the contribution of that particular feature group. As a performance metric we used the F measure.

As can be seen in Figure 4, dropping all the feature groups individually causes a bit of performance drop except location features. When we dropped location features, we see an increase in the performance of our person-independent "About-to-Eat" moments detector. It basically means that the location-based features fail to capture any general trends about "About-to-Eat" moments and they introduce a lot of noise in the feature-space. As a result dropping the location-based features are more beneficial for training person-independent model. This is because geo-location during "About-to-Eat" moments are highly person specific traits (e.g., eating habit, occupational constraints), thus it cannot find any general trends that generalizes across individuals. The top contributing feature was step count followed by calorie expenditure. Step count at certain time from certain location (e.g., home or workplace) towards another location (e.g., restaurants and cafes) could give vital information about an upcoming eating event. Similarly, certain amount of calorie expenditure could also be an indirect indicator of hunger or craving and thus it could be informative about an "About-to-Eat" moments. Among the inertial sensors, the gyroscope features contributed more than accelerometer features, as it could capture the characteristic hand gesture from activities prior to an eating event like typing on a keyboard, opening door, walking etc. The current time also contributed a bit, as our eating is governed by a routine. But as different participants had slightly different routines, it did not turn out to be the top most feature for the generalized model. It is also interesting to note from Figure 4 that both EDA and heart rate contributed minimally towards the "About-to-Eat" moment detector. Notice that the feature contribution here is estimated with respect to a person-independent "About-to-Eat" classifier, so feature group (like location) that did not contribute for person-independent model, could contribute significantly for a person-dependent model. We discuss person-dependent features in the Personalized Model section below.

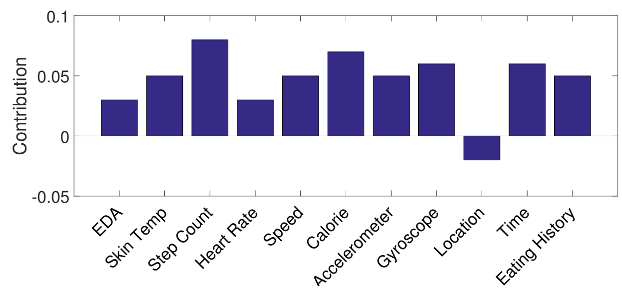


Figure 4: Contribution of different feature groups for training a Tree Bagger person-independent "About-to-Eat" event classifier

5.4 Training "About-to-Eat" Moment Detector

In order to train our "About-to-Eat" moment detector we, at first, trained an "About-to-Eat" moment classifier that differentiates between the "About-to-Eat" moments and the moments further in the past. Secondly, we also trained a regression model that predicts the "Time until the Next Eating Event".

5.4.1 "About-to-Eat" Moment Classifier

"About-to-Eat" moment is defined by a certain time period preceding the start of an eating event. Here this time period is called "About-to-Eat" definition window. In order to label different feature extraction windows into the two different classes for this classifier (which are "About-to-Eat" moment or "Non-About-to-Eat" moment), we used the end point of the feature extraction window. If the end point of the feature extraction window is within the definition window, we label the entire feature extraction window as "About-to-Eat" moment as the current feature extraction window already entered the "About-to-Eat" moment definition window. The goal of our classifier is to distinguish between these two classes. As our feature selection method suggested that location features introduced noise in the feature space, we did not use any location-based features for training our person-independent classifier. Table 4 presents the performance of different "About-to-Eat" event classifiers with and without feature selection from a Leave-One-Person-Out (LOPO) cross-validation experiment. From Linear Model to Tree-based and Support Vector Machine (SVM)-based classifiers, as we increase the complexity of the model, the performance of the classifier increased. It tells us that probably the non-linear, highly convoluted distribution of features in the feature space demands non-linear mapping from the features to the class. As can be seen in Table 4 Tree Bagger model (also known as Random Forest) outperforms all the rest of the models for both with and without feature selection. The Tree Bagger trained on a selected feature subset outperforms all the rest machine learning models and reaches a recall, precision and F score of respectively 0.77, 0.67 and 0.69. This result shows that a person-independent "About-to-Eat" detector, trained on a few behavioral information and with simple machine learning algorithm, can achieve a reasonable performance.

In order to investigate how the feature extraction window size affects the overall performance of our classifier and in order to select an optimal feature extraction window size, we extracted our features with different window sizes ranging from 5 minutes to 120 minutes by keeping all the other parameters constant and estimating the trained classifier's performance. Figure 5 shows how the performance of our classifier changes as the feature extraction window size changes. Both very small and coarse feature extraction

Model	Without Feat. Sel.			With Feat. Sel.		
	R	P	F	R	P	F
Linear Model	0.63	0.55	0.61	0.67	0.65	0.66
RepTree	0.61	0.49	0.60	0.65	0.62	0.64
SVM	0.63	0.60	0.62	0.67	0.62	0.66
Tree Bagger	0.74	0.59	0.64	**0.77**	**0.67**	**0.69**

Table 4: Performance of person-independent "About-to-Eat" event classifier trained with and without CFS feature selection in terms of Recall(R) and Precision (P) and F measure (F).

window sizes we have bigger disparity between precision and recall. For feature extraction window size of 60 to 80 minutes we have high, yet more consistent performance with respect to recall and precision. Very small feature extraction window sizes are susceptible to noise, while very coarse window size fails to capture the recent trends in the sensor time series which are essential for building our "About-to-Eat" classifier. Lastly, notice that feature extraction window size does not change the granularity or resolution of our "About-to-Eat" classifier's inference, as we kept feature extraction window shift constant (1 minute).

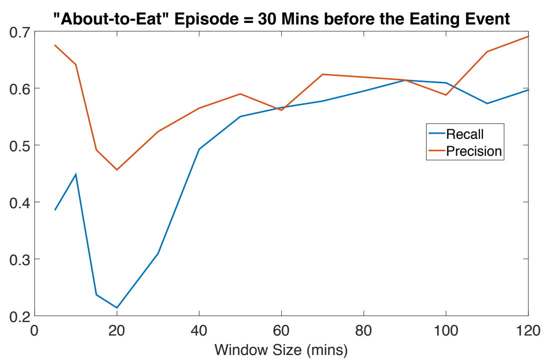


Figure 5: F measure of the "About-to-Eat" event classifier with Tree Bagger across different feature extraction window sizes

Similar to the feature extraction window size, the "About-to-Eat" definition window size plays a significant role. As we increase the size of the definition window size, we make the "About-to-Eat" moments less stringent. In order to further investigate any effect of different definition window sizes on the performance of our classifier, we evaluate our classifier for different definition window sizes. The definition window size will certainly have an effect on the performance of our classifier also changes, which can be seen in Figure 6. As we increase the definition window size, the performance of the "About-to-Eat" moment detection increases. In other words, if we increase the definition window size, the "About-to-Eat" moment detector has more opportunity to capture the subtle patterns in the feature space and predict the "About-to-Eat" moments correctly. However, practically too big of a definition window size is not very useful, as we want to trigger our intervention as close to the start of an eating event as possible.

5.4.2 *"Time until the Next Eating Event" Regression Model*

In order to predict the "Time until the Next Eating Event", we trained several regression models including linear, RepTree, SMO and Tree Bagger regression model using Weka toolbox [22]. The preprocessing, feature extraction and feature selection stage remained the same. The "Time until the Next Eating Event" was estimated from the end point of every feature extraction window in minutes.

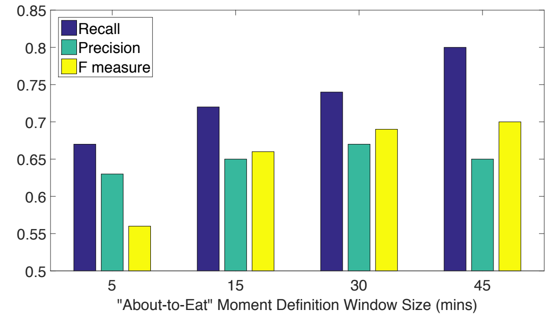


Figure 6: Performance of the "About-to-Eat" event classifier with Tree Bagger across different different "About-to-Eat" definition window sizes

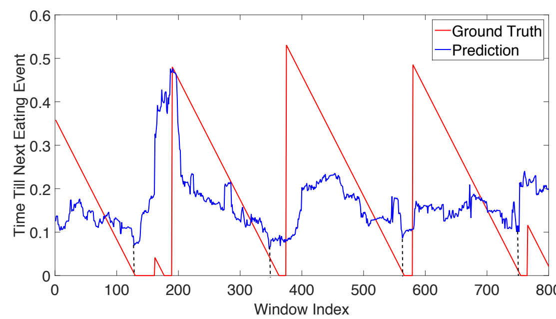


Figure 7: Shows how the predicted "Time until the Next Eating Event" matches with the reference value. The model is trained with Tree Bagger with CFS feature selection

If the time till the next eating event from the endpoint of any particular window (of feature extraction) was 5 hours or more, then we simply ignored those windows as those windows capture very different life events (e.g., sleeping) which are irrelevant to the problem at hand. Table 5 illustrates the results of different types of regression models with and without CFS feature selection. We estimated the performance of different regression models in terms of the Pearson correlation coefficient (ρ) and the mean absolute error (MAE). The best regression performance of 0.49 Pearson correlation coefficient and 0.18 of mean absolute error is achieved again with the Tree Bagger when it is trained with the selected feature subset. Figure 7 shows how our "Time until the Next Eating Event" regression model performs with respect to the reference. The reference "Time until the Next Eating Event" value is considered to be zero during the eating event. Notice that the lowest values of the predicted "Time until the Next Eating Event" graph lies right before the start of an eating event.

Model	Without Feat. Sel.		With Feat. Sel.	
	ρ	MAE	ρ	MAE
Linear Regression	0.10	0.47	0.19	0.26
RepTree	0.27	0.22	0.36	0.22
SMO Regression	0.28	0.25	0.33	0.23
Tree Bagger	0.40	0.20	**0.49**	**0.18**

Table 5: Performance of "Time until the Next Eating Event" regression model trained with and without CFS feature selection. The performance is measured in terms of Pearson Correlation Coefficient (ρ) and Mean Absolute Error (MAE) from a Leave-One-Person-Out (LOPO) cross-validation experiment.

In order to identify the optimal feature extraction window size, we extracted features with different window sizes and evaluated Tree Bagger model with feature selection. Figure 8 shows the performance of prediction model reaches its highest value when the window size is 100 minutes. Features extracted with any window size lower or greater than 100 minutes fails to capture the full dynamics of "About-to-Eat" moments in the sensor streams and thus fails to reach the best performance. This result echoes the same observation about the effect of feature extraction window size on the performance of the model that we presented in Figure 5.

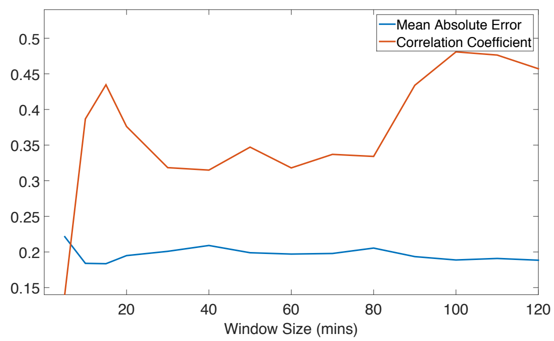


Figure 8: Correlation coefficient of the "Time until the Next Eating Event" Tree Bagger regression model across different feature extraction window sizes

5.4.3 *Personalized Model*

The performance of both the classification and regression could be further improved by learning person-dependent models with some labeled data from the target user, as the model could incorporate the idiosyncrasies (e.g., person-specific eating pattern, lifestyle) of the target user. To test our hypothesis, we trained person-dependent "About-to-Eat" moment classifier on the feature subset selected by the CFS feature selection and "Time until the Next Eating Event" regression model with Tree Bagger using 10-fold cross validation. Notice that in this experiment both train and test set are two mutually exclusive portion of the data that came from the same person. The person-dependent "About-to-Eat" moment classifier achieved a recall, precision and F-score of respectively 0.85, 0.82 and 0.84. Similarly, when we again trained a person-dependent "Time until the Next Eating Event" regression model, the performance reached a Pearson correlation coefficient of 0.65. Both person-dependent models clearly outperform the person-independent ones, which validates our hypothesis about personalization.

We further investigated on the features that are selected by the person-specific model. We found that different feature groups contribute differently for predicting "About-to-Eat" moments of different individuals. For example, participant *P* typically skips his breakfast and has his lunches at a particular neighborhood right next to his work place at a particular time of the day. He also has a quick snack in the afternoon right before he leaves to the gym to play table tennis with his friends. Later in the evening, he typically gets his dinner at home. Now, if we look at the top selected features for training a personalized model, we get location, time of the day, calorie and step count based features to be the top feature groups. Just these four groups of features captures his regular life style around different eating event. As a result the person-dependent model also could achieve a very high performance with an F-score of 0.90 (by the "About-to-Eat" moment classifier) and Pearson correlation coefficient of 0.83 (by "Time until the Next Eating Event" regression model). Now, if we look at participant *S* (who is a home maker and 28 years old), we get a completely different set of feature groups that have the most information about the person-dependent model. Unlike participant *P*, location turned out to be one of the least informative features for participant *S*, as she typically takes all of her meals at home. Inertial (accelerometer and gyroscope) sensor based features are selected as one of the most informative features as that could pick up various household chores. Heart rate and EDA-based features contribute more for training person-dependent model for participant *S* than they did for participant *P*.

6. DISCUSSION

The best person-independent "About-to-Eat" moment classifier (Tree Bagger) reaches an F-score of 0.69. Similarly, when we trained a person-independent regression model that tries to predict the "Time until the Next Eating Event", we achieved a Pearson correlation coefficient of 0.49. The performance of the person-independent models clearly underscores the fact that the sensor streams captured by wearable wrist bands and a smartphone contain important information about "About-to-Eat" moments and the "Time until the Next Eating Event". As person-independent model captures generalized trends in the feature space (e.g., general eating schedule, physical variables around "About-to-Eat" moments), it fails to capture the person-specific traits or idiosyncrasies around one's eating behavior. In our analysis we demonstrated that personalized model could overcome this problem by learning person-specific trends in the feature space and increased the performance of "About-to-Eat" moment classifier to an F-score of 0.84 and the "Time until the Next Eating Event" regression model to a Pearson correlation coefficient of 0.65. Although in our study we trained our person-specific model by data from our target user in an offline manner, the multimodal sensor streams along with the online learning algorithm could learn a fairly accurate person-dependent "About-to-Eat" model on the fly if the system gets labels to train on. However, training a person-dependent model in this manner requires active user engagement and a lot of labeled data. One alternative to getting labeled data via self-report could be the wearable microphone that can tell current eating event. Thus users can use the wearable microphone along with the other wearable and smartphone technologies for a few days to generate enough training data for the person-dependent model to be built. The person-dependent feature selection analysis clearly shows that the features used by person-dependent model highly dependent on eating habit and life style. As the life style and eating habit remained constant over our rather small data collection period, our person-dependent model could find some internal trends in the feature space. However, if one's life style changes drastically due to a big life event, the person-dependent model needs to be again calibrated with the help of more labeled data.

At the end of our data collection through a semi-structured interview, we received qualitative feedback and comments from our participants on the wearability, form factor, privacy and social acceptability issues of the ubiquitous and wearable technologies. Although all these different socio-ethical aspects were not considered in this paper, here we included a limited discussion so that we can improve them in future iterations. Most of our participants were accustomed to smartphone and have been using it for telecommunication and entertainment purposes. Although a few of the participants requested a bigger font in our smartphone application, overall they found it to be sufficient. Many of our participants found Microsoft Band to be a very useful wearable fitness device. As a result the inclusion of these two technologies in our data collection framework was very well received. In fact many of our users expressed their interest to keep using our smartphone application. They found it useful as a food logger that helped them to be mindful about their eating event. However, all the participant found it difficult to use the wearable microphone that passively tracks different eating events by detecting mastication and swallowing sound. As one needs to wear the wearable microphone around the neck, it was also highly visible. As a result it was not very privacy sensitive. Many of our participants complained about the social acceptability, as they had to explain the purpose of the wearable microphone to their friends and family, which is undesirable and may contribute to leaving the technology all together. Our participants suggested that a form factor, which could be easily worn and hidden in clothes for a long time, could be a more meaningful form factor for the wearable microphone, which should be considered for future studies.

7. LIMITATIONS AND FUTURE WORK

Like most studies our study is not without limitations. In order to explore the feasibility of "About-to-Eat" moment detection, we collected a limited amount of (8) participant data, where each participant recorded data only for a few (5) days. Encouraged by the promising results of this feasibility study, we want to run a longitudinal study to collect a larger dataset with more participants during a longer time span. A larger dataset will enable us to test the generalizability of our "About-to-Eat" prediction models. In future, we want to learn personalized "About-to-Eat" moment detector online and assess its accuracy. Although we gathered ideas about potential Just-in-Time eating interventions triggered at "About-to-Eat" moments in this paper, in our current study these eating interventions are not triggered in an online manner. This feasibility study with a limited time span could also not estimate their efficacy of these intervention for behavior change. With our future longitudinal study we plan to trigger various Just-in-Time eating interventions as the personalized models detects "About-to-Eat" moment and assess the efficacy of such a system for behavior change.

There is a huge body of literature that shows the link among food, emotion and stress [16]. Affect also influences our perception of hunger and satisfaction. As our existing data collection scheme already collects information about different affective states (i.e., emotion and stress), in future we also would like to explore the relationship between "About-to-Eat" moments and affective states. Information about various affective states during "About-to-Eat" moments could also be useful select the most effective Just-in-Time eating intervention from a pool of interventions.

8. CONCLUSION

In this feasibility study, we explored if we could reliably detect "About-to-Eat" moments and predict the "Time until the Next Eating Event". We explored both person-dependent and person-independent models using a multimodal sensor dataset with four different ubiquitous and wearable technologies including a Microsoft Band, an Affectiva Q sensor, a wearable microphone and a smartphone. The wearable sensing framework captures physical activity, location, heart rate, electrodermal activity, skin temperature and caloric expenditure. Using signal processing and machine learning on this raw multimodal sensor stream, we trained a person-independent "About-to-Eat" moment classifier that reaches an average recall of 0.77. The person-independent "Time until the Next Eating Event" regression model attains a correlation coefficient of 0.49. By building person-dependent model, we can further boost the performance of both models.

9. ACKNOWLEDGMENTS

We thank all the members of Visualization and Interaction for Business and Entertainment (VIBE) research group and hardware laboratory at Microsoft Research for their invaluable suggestions, and feedback.

10. REFERENCES

[1] http://qsensor-support.affectiva.com. In *http://qsensor-support.affectiva.com*.
[2] https://en.wikipedia.org/wiki/obesityintheunitedstates. In *https://en.wikipedia.org/wiki/ObesityintheUnitedStates*.
[3] https://en.wikipedia.org/wiki/preventablecausesofdeath. In *https://en.wikipedia.org/wiki/Preventablecausesofdeath*.
[4] https://www.myfitnesspal.com. In *https://www.myfitnesspal.com/*.
[5] https://www.pjrc.com/store/teensy31.html. In *https://www.pjrc.com/store/teensy31.html*.
[6] https://www.sparkfun.com/products/9964. In *https://www.sparkfun.com/products/9964*.
[7] http://www.fitbit.com. In *http://www.fitbit.com*.
[8] http://www.microsoft.com/microsoft-band. In *http://www.microsoft.com/microsoft-band*.
[9] http://www.who.int/mediacentre/factsheets/fs311. In *http://www.who.int/mediacentre/factsheets/fs311*.
[10] O. Amft, H. Junker, and G. Troster. Detection of eating and drinking arm gestures using inertial body-worn sensors. In *Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*, pages 160–163, Oct 2005.
[11] O. Amft, M. Kusserow, and G. Troster. Bite weight prediction from acoustic recognition of chewing. *Biomedical Engineering, IEEE Transactions on*, 56(6):1663–1672, 2009.
[12] O. Amft, M. Staeger, and G. Troester. Analysis of chewing sounds for dietary monitoring. In *UbiComp '05*, pages 56–72, 2005.
[13] O. Amft and G. Troster. Methods for detection and classification of normal swallowing from muscle activation and sound. In *Pervasive Health Conference and Workshops, 2006*, pages 1–10, Nov 2006.
[14] L. Arab, D. Estrin, D. H. Kim, J. Burke, and J. Goldman. Feasibility testing of an automated image-capture method to aid dietary recall. *European journal of clinical nutrition*, 65(10):1156–1162, 2011.
[15] E. Barrett-Connor. Nutrition epidemiology: how do we know what they ate? *The American journal of clinical nutrition*, 54(1):182S–187S, 1991.
[16] E. A. Carroll, M. Czerwinkski, A. Roseway, A. Kapoor, P. Johns, K. Rowan, and m. c. schraefel. Food and Mood: Just-in-Time Support for Emotional Eating. In *Proceeding of IEEE Affective Computing and Intelligent Interaction (ACII) '13*, 2013.
[17] E. K. Choe, N. B. Lee, B. Lee, W. Pratt, and J. A. Kientz. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 1143–1152. ACM, 2014.
[18] F. Cordeiro, E. Bales, E. Cherry, and J. Fogarty. Rethinking the mobile food journal: Exploring opportunities for lightweight photo-based capture. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3207–3216. ACM, 2015.
[19] M. R. Craig, A. R. Kristal, C. L. Cheney, and A. L. Shattuck.

The prevalence and impact of atypical days in 4-day food records. *Journal of the American Dietetic Association*, 100(4):421–427, 2000.

[20] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover. Detecting periods of eating during free-living by tracking wrist motion. *Biomedical and Health Informatics, IEEE Journal of*, 18(4):1253–1260, 2014.

[21] D. H. Gustafson, B. R. Shaw, A. Isham, T. Baker, M. G. Boyle, and M. Levy. Explicating an evidence-based, theoretically informed, mobile technology-based system to improve outcomes for people in recovery for alcohol dependence. *Substance use & misuse*, 46(1):96–111, 2011.

[22] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, 11(1), 2009.

[23] M. A. Hall. Correlation-based Feature Selection for Machine Learning. *PhD Thesis*, April 1999.

[24] A. Kadomura, C.-Y. Li, K. Tsukada, H.-H. Chu, and I. Siio. Persuasive technology to improve eating behavior using a sensor-embedded fork. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 319–329. ACM, 2014.

[25] H. Kalantarian, N. Alshurafa, and M. Sarrafzadeh. A wearable nutrition monitoring system. In *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*, pages 75–80, June 2014.

[26] F. Kong and J. Tan. Dietcam: Regular shape food recognition with a camera phone. In *Body Sensor Networks (BSN), 2011 International Conference on*, pages 127–132. IEEE, 2011.

[27] J. Liu, E. Johns, L. Atallah, C. Pettitt, B. Lo, G. Frost, and G.-Z. Yang. An intelligent food-intake monitoring system using wearable sensors. In *Wearable and Implantable Body Sensor Networks (BSN), 2012 Ninth International Conference on*, pages 154–160, May 2012.

[28] J. Mankoff, G. Hsieh, H. C. Hung, S. Lee, and E. Nitao. Using low-cost sensing to support nutritional awareness. In *UbiComp 2002: Ubiquitous Computing*, pages 371–378. Springer, 2002.

[29] D. C. Mohr, S. M. Schueller, E. Montague, M. N. Burns, and P. Rashidi. The behavioral intervention technology model: an integrated conceptual and technological framework for ehealth and mhealth interventions. *Journal of medical Internet research*, 16(6), 2014.

[30] I. Nahum-Shani, S. N. Smith, A. Tewari, K. Witkiewitz, L. M. Collins, B. Spring, and S. A. Murphy. Just-in-time adaptive interventions (jitais): An organizing framework for ongoing health behavior support. *Technical Report No. 14-126*, 2014.

[31] C. Nederkoorn, F. Smulders, R. Havermans, and A. Jansen. Exposure to binge food in bulimia nervosa: finger pulse amplitude as a potential measure of urge to eat and predictor of food intake. *Appetite*, 42(2):125–130, 2004.

[32] J. Noronha, E. Hysen, H. Zhang, and K. Z. Gajos. Platemate: crowdsourcing nutritional analysis from food photographs. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 1–12. ACM, 2011.

[33] L. Pina, K. Rowan, A. Roseway, P. Johns, G. R. Hayes, and M. Czerwinski. In situ cues for adhd parenting strategies using mobile technology. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, pages 17–24. ICST, 2014.

[34] J. P. Pollak, P. Adams, and G. Gay. Pam: a photographic affect meter for frequent, in situ measurement of affect. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 725–734. ACM, 2011.

[35] T. Rahman, A. T. Adams, M. Zhang, E. Cherry, B. Zhou, H. Peng, and T. Choudhury. Bodybeat: A mobile system for sensing non-speech body sounds. In *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '14, pages 2–13, 2014.

[36] S. Reddy, A. Parker, J. Hyman, J. Burke, D. Estrin, and M. Hansen. Image browsing, processing, and clustering for participatory sensing: Lessons from a dietsense prototype. In *Proceedings of the 4th Workshop on Embedded Networked Sensors*, EmNets '07, pages 13–17, 2007.

[37] E. Sazonov, S. Schuckers, P. Lopez-Meyer, O. Makeyev, N. Sazonova, E. L. Melanson, and M. Neuman. Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. *Physiological measurement*, 29(5):525, 2008.

[38] M. Shuzo, G. Lopez, T. Takashima, S. Komori, J.-J. Delaunay, I. Yamada, S. Tatsuta, and S. Yanagimoto. Discrimination of eating habits with a wearable bone conduction sound recorder system. In *IEEE Sensors '09*, pages 1666–1669, 2009.

[39] E. Stellar and E. E. Shrager. Chews and swallows and the microstructure of eating. *The American journal of clinical nutrition*, 42(5):973–982, 1985.

[40] E. Thomaz, I. Essa, and G. D. Abowd. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 1029–1040, 2015.

[41] E. Thomaz, A. Parnami, I. Essa, and G. D. Abowd. Feasibility of identifying eating moments from first-person images leveraging human computation. In *Proceedings of the 4th International SenseCam & Pervasive Imaging Conference*, pages 26–33. ACM, 2013.

[42] E. Thomaz, C. Zhang, I. Essa, and G. D. Abowd. Inferring meal eating activities in real world settings from ambient sounds: A feasibility study. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, pages 427–431. ACM, 2015.

[43] C. Vögele and I. Florin. Psychophysiological responses to food exposure: an experimental study in binge eaters. *Journal of Eating Disorders*, 21(2):147–157, 1997.

[44] B. Wansink and K. A. Johnson. The clean plate club: about 92% of self-served food is eaten. *International Journal of Obesity*, 39(2):371–374, 2015.

[45] K. Witkiewitz and G. A. Marlatt. Relapse prevention for alcohol and drug problems: that was zen, this is tao. *American Psychologist*, 59(4):224, 2004.

[46] K. Yatani and K. Truong. Bodyscope: A wearable acoustic sensor for activity recognition. UbiComp '12, pages 341–350, 2012.

[47] F. Zhu, M. Bosch, C. J. Boushey, and E. J. Delp. An image analysis system for dietary assessment and evaluation. In *ICIP*, pages 1853–1856, 2010.