mPuff: Automated Detection of Cigarette Smoking Puffs from Respiration Measurements

Amin Ahsan Ali, Syed Monowar Hossain, Karen Hovsepian, Md. Mahbubur Rahman,
Kurt Plarre, Santosh Kumar
Dept. of Computer Science, University of Memphis, TN 38152, USA
{aaali, smhssain, kshvspan, mmrahman, kplarre, skumar4}@memphis.edu

ABSTRACT

Smoking has been conclusively proved to be the leading cause of mortality that accounts for one in five deaths in the United States. Extensive research is conducted on developing effective smoking cessation programs. Most smoking cessation programs achieve low success rate because they are unable to intervene at the right moment. Identification of high-risk situations that may lead an abstinent smoker to relapse involve discovering the associations among various contexts that precede a smoking session or a smoking lapse. In the absence of an automated method, detection of smoking events still relies on subject self-report that is prone to failure to report and involves subject burden. Automated detection of smoking events in the natural environment can revolutionize smoking research and lead to effective intervention.

In this paper, we present *mPuff*, a novel system to automatically detect smoking puffs from respiration measurements, using which a model can be developed to automatically detect entire smoking episodes in the field. We introduce several new features from respiration that can help classify individual respiration cycles into smoking puffs or non-puffs. We then propose supervised and semi-supervised support vector models to detect smoking puffs. We train our models on data collected from 10 daily smokers and find that smoking puffs can be detected with an accuracy of 91% within a smoking session. We then consider respiration measurements during confounding events such as stress, speaking, and walking, and show that our model can still identify smoking puffs with an accuracy of 86.7%. The smoking detector presented here opens the opportunity to develop effective interventions that can be delivered on a mobile phone when and where smoking urges may occur, thereby improving the abysmal low rate of success in smoking cessation.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences

General Terms

Design, Experimentation, Measurement, Human Factors

Keywords

Smoking Detection, Wearable Sensors, Respiration

1. INTRODUCTION

Since the first US Surgeon General's report in 1964 there has been overwhelming and conclusive evidence that use of tobacco, especially in the form of cigarette smoking, causes cancer in different organs throughout the body, leads to cardiovascular and respiratory diseases, and harms reproduction [3]. Smoking induced diseases account for nearly one of every five deaths in the United States [24]. Smokers die 13-14 years younger and cost $193 billion annually [1]. In addition, almost 60% of children are still exposed to secondhand smoking which is also a known human carcinogen. To reduce such harmful effects of smoking, there needs to be substantial progress in tobacco control, health education programs, and development of interventions to aid smoking cessation.

Given the adverse impact of smoking on human health, significant research is conducted on development of smoking interventions. Eight (out of 27) divisions at National Institutes of Health (NIH) award research grants for smoking cessation. Of these, National Cancer Institute (NCI) alone awards $100+ million annually in smoking research. Despite extensive efforts, smoking continues to be prevalent. Seventy percent of adult smokers want to quit completely, while 40% try to quit each year - but most quit efforts end in relapse [7]. Each day about 3,000 people become new daily smokers[5].

Most smoking cessation programs achieve low success rate (i.e., less than 10%) because they are unable to intervene at the right moment. Smokers who are trying to quit need to avoid high-risk situations and if they get into a high-risk situation, need an intervention to break their urge. Given the ubiquity of smart phones, such a smoking cessation assistant app can be developed for smart phones that intervenes if a quitter is found to be in a high-risk situation. The challenge, however, is to automatically identify when the quitter is in a high-risk situation. Considerable amount of research work is focused on identifying the factors (called
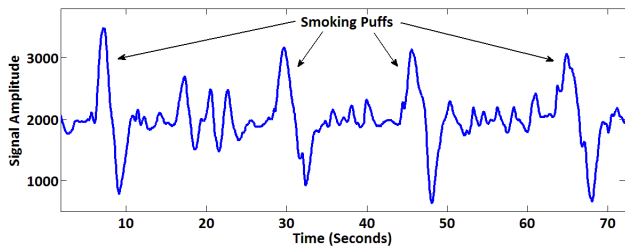
Figure 1: Respiration signals captured during a typical smoking episode.

antecedents) that lead to high-risk situations (and, eventually relapse) to design effective interventions [30, 11, 23] .

Identifying antecedents and precipitants of smoking lapse (i.e., an acute condition such as stress that causes a lapse) requires conducting scientific user studies in the natural environment so as to observe the psychological, social, and environmental factors that may be associated with smoking instances [19, 15, 18, 21, 27, 32, 31]. This is done by observing and recording the user's context when smoking occurs for daily smokers or when a smoking lapse occurs in those trying to quit [28]. These studies must have some mechanism of detecting when smoking occurs, so that physical, physiological, psychological, behavioral, social, and environmental contexts before, during, and after a smoking session can be identified. Most current studies on smoking behavior rely on various self-reporting techniques, where subjects are asked to self-report each smoking episode. These methods range from basic pen-paper methods and retrospective recalls, to electronic diary keeping and ecological momentary assessments(EMA) [29, 28, 20, 9]. These methods, in addition to imposing a burden on the study participants, have the limitation of introducing biases when recalling events, forgetting to report, among several others.

Technological methods to detect smoking episodes include carbon monoxide (CO) monitors such as piCO$^+$, CReSS Pocket[8], Micro$^+$[4], RespiTrace$^®$[13], and image processing [34]. As discussed in more detail in Section 2, each of these methods require manual intervention by subjects or an observer. To the best of our knowledge, there does not exist a method to automatically detect smoking events in real-time, in the natural environment, that is operator independent.

In this paper, we take the first step towards automatically detecting smoking in the natural environment by developing *mPuff*, a model to automatically detect smoking puffs from respiration measurements. mPuff uses respiration measurements collected from a respiration band that the user wears underneath their clothing. Detection of smoking puffs from respiration is feasible because they are associated with deep inhalation and deep exhalation (see Figure 1 for an example). It should be noted that a puff lasts only for one respiration cycle. Thus in order to detect puffs we need to find appropriate features that can help discriminate a smoking puff not only from usual respiration cycles, but also from those respiration cycles that may represent speaking, stress, or physical activity such as walking.

A smoking puff tends to lengthen the duration of a respiration cycle relative to its neighboring cycles. It also amplifies the degree of inhalation and exhalation (i.e., both

directions on y-axis) as compared to the usual level of peak and valley, as well as that compared to the neighboring cycles. We build on these insights to identify 12 new features in respiration measurements that together with 5 previously known features are used in mPuff to detect smoking puffs. A majority of these features are not person specific as they measure relative changes of some basic characteristics of the respiration cycles. For those features that depend on absolute values such as inspiration duration, we normalize it by computing the z-score of the feature values using the person-specific mean and standard deviation, thus accounting for the between-person differences. Therefore, mPuff self-calibrates to each person and does not need to be trained on a person prior to its usage.

In order to develop the mPuff model from the above mentioned features, we use support vector machine (SVM) that is trained over the respiration features. Given that various smoking researchers may need different sensitivities to false positive and false negative rate, the model we use can be customized for a target false positive or false negative rate. Given imperfection in automated models, the smoking research studies may continue collecting self-reports. These self reports, however, may not always be located before the start of a smoking session. They in some cases may be located during or after a smoking session. We propose a semi-supervised support vector model that improves the accuracy of detecting smoking puffs by making use of the self-report markings. Our model can potentially be used as a building block to develop a full-fledged smoking detector that can identify those smoking episodes that may not have received a self-report marking.

To train mPuff, we collected respiration data during smoking from 10 volunteer daily smokers. During the collection of labeled data, we carefully marked each puff in a smoking session. To ensure generalizability of our model, we also used data sets from major confounding events, e.g., physical activity, conversation, and stress, that may cause similar patterns in respiration as smoking puffs. All these data sets constitute the training and testing data sets of the supervised SVM. These data sets are supplemented by collection of respiration data from 4 volunteers (out of the original 10 volunteers) who wore the sensors for 7 continuous days in their natural environment and self-marked smoking episodes. Together with the labeled data, this data set is used for the training of the semi-supervised model. For additional generalizability of our models, we divided the data set into training and test set, where 10-fold cross-validation was conducted on the training set, but the test set was reserved purely for testing.

Results: We find that smoking puffs can be detected with an accuracy of 91% within a smoking session. When applying the model to confounding events, we obtain an accuracy of 84.5% for the supervised SVM model, which improves to 86.7% by using a semi-supervised model that is able to use a much larger data set from the field. We also find that the accuracy of the classifier increases by more than 10% by using the newly proposed features. We applied mPuff to our data set to observe patterns of smoking behavior. We find that the average duration of a smoking session is 6.62 minutes, a smoking session contains an average of 12 puffs, among several other interesting statistics.

Future Applications: Our model opens the opportunity for automated detection of smoking episodes in the natural

environment. Since respiration measurements can be used to detect stress [25] as well, which has been found to be a leading predictor of smoking relapse, smoking research can potentially be revolutionized. It has been found that stress levels of abstinent smokers who relapsed rises hours before a lapse [30]. Now, it can be found out what happens in the minutes preceding a smoking lapse. Since several other contexts such as location, commuting, physical activity, and social interactions can also be detected on a smart phone, rich contextual analysis can be conducted to find true predictors of smoking lapses. Such analysis can then be used to design effective interventions which can be delivered on a mobile phone, when and where smoking urges may occur.

Organization of the paper: Section 3 describes the AutoSense sensor suite and the data sets we use for the development and evaluation of the mPuff puff detector. In section 4, we present the features used to train mPuff. We also describe the supervised and semi-supervised SVM models we use for puff and non-puff detection. Section 5 describes the performance of the classifiers and the effect of new features on improving the accuracy of classification. Section 6 describes the smoking topology statistics that we obtain by applying mPuff to our data sets. Section 7 presents some challenges in developing a model for detecting an entire smoking episode. It describes potential directions for research that leverage the patterns of conversation in the natural environment. We also propose a method for computing the confidence in detecting a smoking episode. Finally, section 8 concludes the paper.

2. RELATED WORKS

We discuss related works in two categories. We discuss technological methods to detect smoking and their shortcomings in Section 2.1. In Section 2.2, we discuss existing methods to analyze respiration measurements to make inferences of human states and point out why these methods are insufficient to detect smoking puffs from respiration.

2.1 Technological Methods to Detect Smoking

Technological methods to detect smoking episodes include carbon monoxide (CO) monitors such as piCO$^+$, CReSS Pocket [8], Micro$^+$ [4], RespiTrace$^®$ [13], and image processing [34]. piCO$^+$ and Micro$^+$ are handheld devices designed for use as motivational aid in smoking cessation programs. They display the amount of smoke inhalation and carbon monoxide levels in a single breath exhaled, measured through a mouthpiece attached to the devices. They also calculate and display the percentage of carboxyhaemoglobin in the blood, thereby providing visual motivation for the smoker to stop smoking. These device are, however, not intended to be used for automatically detecting smoking in an operator independent fashion.

CReSS Pocket/CReSS Micro [8] is a portable device that can be used to acquire the smoking behavioral information in the smoker's natural environment over weeks as they store the data on the device's memory. The subject is asked to insert a cigarette into a holder of CReSS and smoke through a mouthpiece attached to the device. The device then is able to compute several measures of smoking behavior including puffs per cigarette, puff volume, and puff duration and also the timestamps of cigarette insertion and removal. All this data can be downloaded later to a computer. Although CReSS has been used in some studies outside of the laboratory settings, it has been mostly in studies by tobacco companies to establish brand differences [22, 17, 33] by observing the smoking pattern and the degree of tobacco intake. For example, it was observed that with light cigarettes, smokers take more frequent puffs in order to inhale the same amount of tobacco as in a heavy cigarette [17], negating the purpose of making lighter cigarettes. CReSS requires subject's compliance since each time they smoke, they need to smoke through CreSS. Furthermore, it may be embarrassing for the subjects to use it in the natural environment, since the device on their mouth will be visible to others in the vicinity.

Respitrace$^®$ is a newer device that uses a respiration sensor, such as RespiBand Plus, that measures the chest's expansion as the wearer breathes in and out. The timing of each puff is marked manually by an observer who presses a push button switch when the subject places the cigarette on lips. Authors in [13] make use of these measurements to analyze post-puff breathing patterns in smoking. The use of Respitrace$^®$ has been restricted to lab settings to study smoking patterns since it requires manual marking of each puff.

If the place of smoking is under the coverage of a video camera, then movement of hands and presence of cigarette in the mouth can be detected by image processing to automatically detect smoking [34]. Use of this method, however, requires installment of video cameras in all locations where a subject may smoke. Alternatively, the subject may be asked to have a portable video camera (e.g., on a smart phone) pointed to them before they smoke, which again requires the involvement of subject each time they smoke.

In summary, each of the above technological methods require subject compliance and hence are not suited for widespread usage in smoking research. Therefore, smoking researchers continue to rely on self-report method today. The prospect of having a method such as mPuff that enables the development of models that can automatically detect all smoking episodes in the field from an unobtrusively wearable respiration sensor measurement excites the entire smoking research community given its potential to revolutionize smoking research.

2.2 Analytical Methods to Infer Human States from Respiration

There have been several recent works on detecting various human states from respiration measurements. In [25], respiration measurements are processed to infer physiological and perceived stress. Various features such as inhalation duration, exhalation duration, minute ventilation, and respiration rate are computed from each minute of respiration measurements. Machine learning models are then trained over these features to infer whether the subject is stressed. In [26], features used in [25] are supplemented with some new features such as B-Duration, computed from 30 seconds of respiration measurements to detect if the subject is speaking, listening, or quiet. These states are then composed together in a Hidden Markov Model (HMM) to identify conversation episodes.

While some features identified in the above works can help in detecting smoking puffs, the features and models used in these and other works are not directly applicable to detecting puffs. First, robust statistics such as mean and median of various features are used in the detection of stress

and conversation since there are several respiration cycles in 30 seconds or 1 minute. For smoking, each puff needs to be identified reliably. Second, there is a pattern of transitions among the speaking, listening, and quiet states that can be leveraged in an HMM to detect conversation episodes, such patterns are not observable in a smoking session. Third, the timing or number of respiration cycles between successive smoking puffs can vary widely among different subjects, and among different smoking session for the same subject, for instance, when smoking in a group or when smoking alone during work hours.

Fourth, the accuracy of detecting a smoking puff may need to be customizable for various use cases. A study on observing smoking behavior may want good accuracy for detecting an entire smoking session, whereas another study on abstinent smokers may want to detect individual puffs, at the cost of a higher false positive rate, since even a single puff can lead to a full relapse. The first puff in such scenarios constitutes the moment of lapse (also called first lapse) and is the main event which is used in the analysis for identifying antecedents and precipitants of smoking relapse. It is critical to be able to obtain the timing of the first lapse, and the entire subject data may need to be removed if the first lapse is not detected [28]. False positives may be acceptable, especially, if the analysis is to be done post-facto, where the goal is to be able to pinpoint the timing of relapse. Each puff may be presented to the subject to identify the one that may correspond to the actual first lapse. None of the above works present a way to customize the model to a given rate of false positive or false negative.

Fifth, none of the above works use a semi-supervised model to use the data collected from the field to improve the accuracy of the model. In a smoking research study, subjects are usually asked to mark each smoking episode in the field. These marks, however, may be before, during, or after the smoking episode. Some smoking episodes may have no marks at all. Hence, the marks provide a label for smoking episodes, but these labels are a noisy source. We develop a semi-supervised approach to make use of these noisy labels to improve the accuracy of our model. In summary, the requirements for the development of mPuff are more stringent than other works on detecting psychological and behavioral states from respiration measurements and hence require a new approach to model development.

3. DATA COLLECTION

In this section, we describe the AutoSense sensor suite we used to capture respiration measurements and the data collection procedure for collecting respiration measurements for developing, training, and testing the mPuff model.

3.1 The AutoSense Sensor Suite

We use the Autosense sensor suite [16] that includes a Respiratory Inductive Plethysmograph (RIP) band to measure relative lung volume and breathing rate. AutoSense also includes ECG, galvanic skin response and 3-axis accelerometer sensors, but only respiration measurements are used in this work. The sampling rate of RIP is 21.3 samples/sec. RIP uses a conductive thread that is sewn in a zigzag fashion to the elastic band. An alternating current source is applied to the resulting loop of wire, which, in turn, generates a magnetic field that opposes the current whose strength is proportional to the area enclosed by the wire according to

Figure 2: Respiratory Inductive Plethysmograph band (in blue color) is worn around the chest area and the wearable AutoSense sensor unit clips to the belt. A 3-axis Acceloerometer, ECG, and Galvanic Skin Response sensors are also included in the same sensor unit. Two coins (a quarter and a penny) are also shown in the picture to indicate the form factor.

Lenz's law. The ratio of the magnetic flux to the current is called self-inductance. Therefore, changes to the chest circumference can be measured by measuring the changes to the self inductance of the band. The inductance measurement purely depends on the geometry of the band and is not related to the tension in the band. As a result, the measurement is not prone to the trapping of the band and associated artifacts due to changes in tension. The sensors transmit data to an Android mobile phone in real-time over a low-power wireless link. We use the FieldStream mobile phone software available in [2]. Using the FieldStream software, we obtained the raw data files collected in the phone.

3.2 Data Collection for Model Development

To develop a model for smoking detection, we collected data from 10 volunteer participants over 13 individual smoking sessions. Each participant was a daily smoker. They wore the RIP sensor of the Autosense suite in their natural environment and were accompanied by an observer. The observer marked each puff the participant took by pressing a button on the mobile phone that also received the respiration measurements via wireless channel. The timing of the button press was saved. Marking the puff times on the same phone that received the measurements reduced the time lapse between the markings and sample timestamps. In order to get a more precise marking, the data from smoking session was visualized with the markings. The markings of puffs were then adjusted to match each puff, which is visually distinctive due to deep inhalation and exhalation associated with a puff.

Out of the 10 participants, 4 participants wore the sensor suite for 7 days in their natural environment during their awake hours. They were asked to self-report the each time they smoked a cigarette. We modified the interface of data collection software (FieldStream framework) on the phone to facilitate self reporting. Though the participants were asked to mark the smoking sessions as they light up a cigarette, we do not expect that the self reported times exactly corresponds to the beginning of the sessions; they can be any-

where (before, during, or after) in the vicinity of the smoking episode. Visual inspection confirmed this hypothesis. In total, we have 136 self-reports of smoking from these 4 participants.

3.3 Data Sets for Model Evaluation

We expect a smoking session to be confused with acute stress, conversation, and physical activity, since they all affect respiration measurements in a similar pattern as smoking. We call these events confounding events. In order to evaluate the model's performance on different confounding events, we use the data collected in our previous user studies [25, 26] that also used the AutoSense sensor suite. The first data set is from a study on 21 participants who were exposed to three real-life stressors (e.g. public speaking, mental arithmetic, and cold pressor tasks) in a lab setting [25]; successive stress periods were separated by rest periods. The second data set consists of conversation episodes from 12 participants. Total amount of data collected for this set is around 46 hours [26]. In both of these studies, the participants were suggested to behave naturally and not instructed in anyway that would confine their movement or posture. In addition to these two data sets, we also collected data from 5 participants that captured different levels of physical activities ranging from running and walking to sitting quietly. This data set consists of 1 hour worth of data.

4. PUFF IDENTIFICATION

In this section, we describe the development of the mPuff model that classifies each respiration cycle into *smoking puffs* and *non-puffs*. We first run a peak-valley detection algorithm to find the peaks and valleys in each respiration cycle. We use the peak-valley detection method proposed in [26]. Once peaks and valleys are located, features of interest can be computed to use in a classification algorithm. In Section 4.1, we describe the features we identify for use in puff detection. We then describe (in Section 4.2) the development of a supervised classifier that uses labeled data to detect smoking puffs. This model can be configured for prescribed target false positive or false negative rate. We next describe (in Section 4.3) a semi-supervised model that uses self-report labels obtained from field data to improve the accuracy of puff detection.

4.1 Features

Other works that use respiration measurements to infer human states such as [25, 26] make inferences on time windows that are 30 seconds or longer. Hence, these works are able to compute statistics over multiple respiration cycles making them robust to noise and outliers. In contrast, in this work, we need to make accurate decision at the level of an individual cycle since a smoking puff lasts only one respiration cycle. This makes the task of identifying appropriate features more challenging.

We identify 17 distinct features that are computed from the respiration signal. We identify 5 features from existing work and propose 12 new features. The new features are selected through visual inspection of data collected during smoking and other non-smoking episodes such as physical activity, stress and conversation. Computation of the features involves the identification of the respiration cycles, which are composed of an inhalation and an exhalation period. We now define all features in the following and illus-
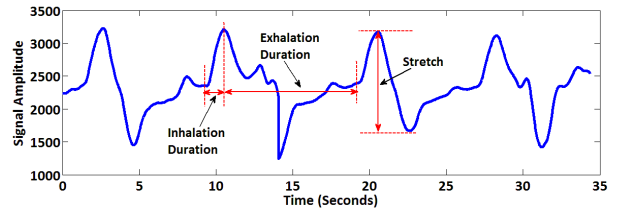

**Figure 3: Illustration of three features extracted from respiration cycles.**

trate three of them in Figure 3.

Existing Features. We first describe five features that have previously been proposed for identifying stress and conversation events from respiration [25, 26]. **Inhalation Duration** corresponds to the time elapsed from a valley of a respiration cycle, to the subsequent peak. The amplitude difference in signal values between these points is the maximum expansion of the chest during a respiration cycle (see Figure 3). **Exhalation Duration** corresponds to the time duration between a peak and the subsequent valley in a respiration cycle. **IE Ratio** is defined as the ratio of inhalation duration to the exhalation duration in a respiration cycle. **Respiration Duration** is the sum of inhalation and exhalation duration. **Stretch** is defined to be the difference between the maximum (legitimate) amplitude, and the minimum (legitimate) amplitude the signal attains within a respiration cycle (see Figure 3). These features have been shown to be effective in identification of stress and conversation from respiration [25, 26]. As we show in Section 5, using these features provides an accuracy of 73.55%, which improves to 86.7% once the new features described below are added.

New Features. Figure 4 shows the respiration patterns during smoking and three confounding events (stress, conversation and running). We observe that features such as stretch and expiration duration are distinct during a respiration cycle containing a puff as compared to respiration cycles in speaking, stress, or activity, and hence have discriminatory power. We further observe that the relative change in stretch and exhalation duration from one respiration cycle to the neighboring cycles are higher when there is a puff involved during smoking. On the other hand, we do not see such magnitude of change during running or conversation events. This is because it is quite unusual to take two consecutive puffs without taking any breaths in between. We define the first difference feature and ratio features in order to capture these information concerning relative change.

We also observe that during a puff, the respiration signals stretch in both upward (called *Upper Stretch*) and downward directions (called *Lower Stretch*), extending the peak amplitude and reducing the valley amplitude respectively, as compared to usual respiration cycles. This suggests that the measurement of relative change in the upper stretch and lower stretch as compared to the running mean of the valley in respiration cycles, can further improve the accuracy of identifying smoking puffs. We now describe the 12 new features.

- **Forward and Backward First Differences** of a fea-

(a) Smoking

(b) Stress

(c) Running

(d) Conversation
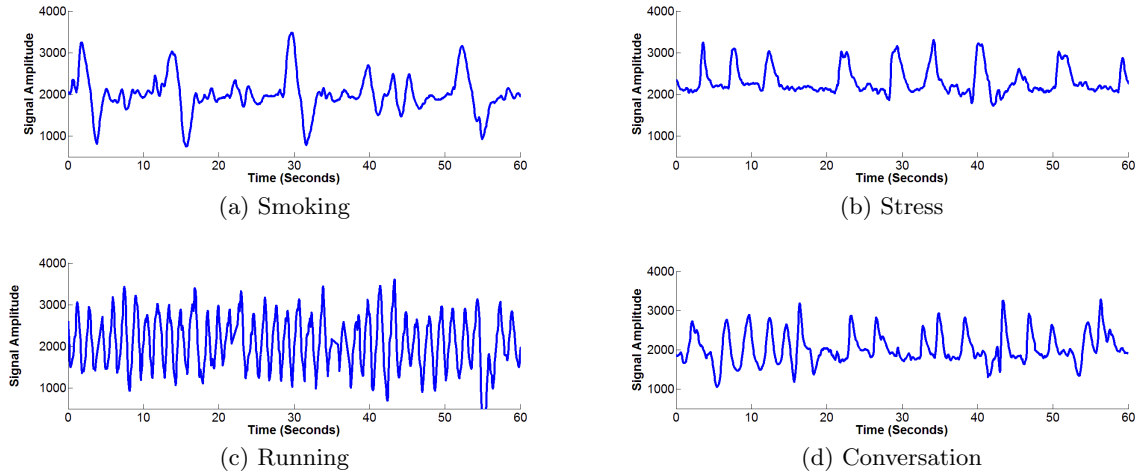
Figure 4: The four figures above show the respiration signal during smoking and three confounding events. We observe that the stretch of a respiration cycle is higher during running and puffs, as compared to speaking and stress. We further observe that unlike during running events, during smoking sessions, the non-puff cycles around the puff do not have as high of a stretch. This simple visual inspection suggests the use of change in stretch relative to its neighboring cycles in discriminating puffs from conversation, stress, and running events. Similar observations can be made for other new features such as relative change in exhalation duration, and upper and lower stretch as described in Section 4.1.

ture is derived by computing the first order differences of the feature values from their previous and next feature values respectively. We compute these first order differences for inhalation, exhalation and respiration durations and stretch and use them as features. Altogether, this procedure creates a total of eight new features.

- **Stretch Ratio** of a particular cycle c, is defined as the ratio of its stretch to the average stretch value in a window of five cycles, with the window centered on Cycle c. When computing the average, we exclude Cycle c. We use a window of five cycles because we never see successive puffs occurring in a window of five cycles, if the window is centered on a puff cycle.

- **Exhalation Ratio** of a particular cycle is similarly computed from the average exhalation duration in a window of five cycles.

- **Upper and Lower Stretch** values are the two features computed from the stretch of each cycle, by splitting it into two parts. The upper stretch magnitude is computed by taking the difference of peak amplitude and running mean value of the valley amplitudes of signal cycles ($ValleyMean$). Similarly, the lower stretch magnitude is computed by taking the absolute difference of minimum amplitude in a respiration cycle and $ValleyMean$. During the computation of the running mean, $ValleyMean$, any valley amplitude two standard deviations away from the current mean value is discarded in the computation.

In order to visualize the effectiveness of the features, we use the **Fisher's Linear Discriminant method**[10] to project the 17-dimensional data vectors to a single dimension using $y = \mathbf{w}^T x$. The idea is to adjust the components

of $\mathbf{w}$ in such a way that the projection maximizes the class separation. The discriminant method maximizes the difference between the projected class means while minimizing the projected class scatter. This method can actually be used as a classifier, albeit a weak one, as much of the information inherent in the data gets lost in the projection. It, however, provides an easy way to visualize the separation of the classes and separation of the classes in the one dimension does hint to the fact that the features may be promising for classification in the higher dimensional space. Figure 5 shows the projected data points and it can be readily observed that there exists good enough separation of the puff and non-puff classes.

4.2 Classifier: Supervised Learning Model

The supervised classifier we use to detect individual puffs is learned using a supervised learning algorithm. The standard classification supervised learning framework is formulated as follows. Given an example set of input observations $X = \mathbf{x}_i \in \mathbb{R}^n | i = 1 \ldots N$, e.g., matrix of n RIP features by N inhalation cycles, with corresponding class labels $Y = y_i \in +1, -1 | i = 1 \ldots N$, e.g., puffs and not puffs, the training algorithm learns a classification function $f_\alpha(\mathbf{x})$, which estimates the class label $\hat{y}$ for a given unlabeled/novel input observation $\mathbf{x}$. The function f is parameterized by the parameter vector α.

Our puff from non-puff classification function is learned using the industry-standard Support Vector Machines (SVM) algorithm, which has been demonstrated to be highly competitive on a great number of problems and tasks, owing to its combination of high learning capacity, i.e., ability to learn highly complex classification functions, with a mathematically rigorous handling of the overfitting/training error trade-off via regularization in the space of kernel functions. Formally, the SVM classification function is defined, using the dual formulation, as
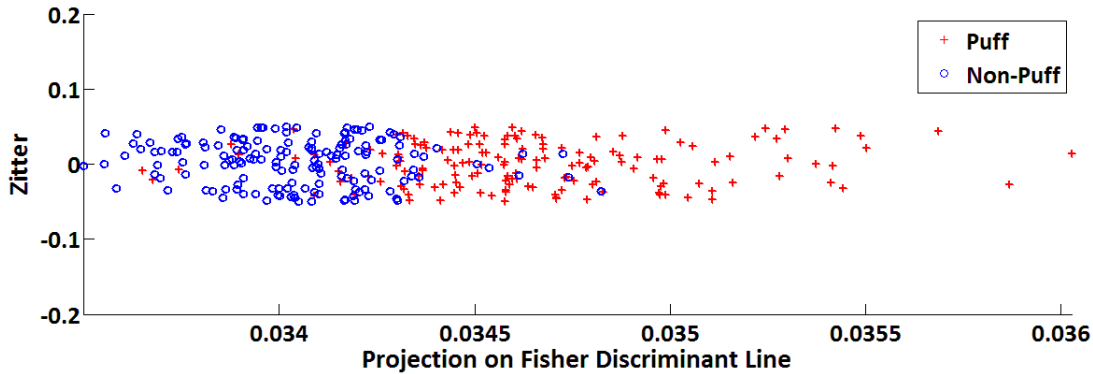
Figure 5: Projection of data on the Fisher's Linear Discriminant Line. The y-axis is used to spread the data points on the line for visual aid. The x-axis presents the value of the projection.

$$f_\alpha(\mathbf{x}) = \text{sign}\left(g_\alpha(\mathbf{x})\right).$$
$$g_\alpha(\mathbf{x}) = \sum_{i \in X_{\text{SV}}} \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b. \tag{1}$$

In the above formulation, α_i's and bias constant b are the parameters learned in the course of training, and the set X_{SV} contains the training observations, called support vectors, that define the boundary separating the two classes. The function $K(;)$ is the special kernel function that allows SVM to learn highly complex functions f, corresponding to a highly non-linear separation boundaries. This so-called kernel-trick makes it possible to implicitly transform all observations to a space of much higher dimensionality, called kernel feature space, where difficult problems are simplified. Formally, the kernel function $K(\mathbf{u}, \mathbf{v})$ corresponds to a dot product between the original vectors $\mathbf{u}$ and $\mathbf{v}$ in the kernel feature space, making the explicit transformation unnecessary. One powerful class of kernels is the Radial Basis Functions class: $K(\mathbf{u}, \mathbf{v}) = e^{\gamma \|\mathbf{u} - \mathbf{v}\|^2}$, leading to classification functions capable of classifying very complex datasets.

The training formulation is based on regularized empirical risk minimization, whereby the algorithm minimizes the error on the training observations, while minimizing the L2-norm of the function f, which works to minimize its complexity. The formal definition of the SVM primal learning problem is

$$\underset{\mathbf{w},\mathbf{b}}{\text{argmin}} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_i^N \max(1 - y_i(\mathbf{w}^T \mathbf{x}_i + b), 0), \tag{2}$$

where the primal variables/parameters w and b represent the linear decision boundary. Note that minimizing the L2-norm of f is equivalent to maximizing The more useful formulation, however, is the Wolfe dual, which also specifies the implicit transformation of the problem into the kernel feature space.

$$\underset{\alpha}{\text{argmax}} \sum_i^N \alpha_i - \frac{1}{2}\sum_{i,j}^N \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j)$$
$$0 \le \alpha_i \le C, \tag{3}$$
$$\sum_{i=1}^n \alpha_i y_i = 0.$$

The algorithm requires several user-defined parameters. The two main ones are: 1) the C constant, which directly penalizes the error on the train set, as per equation (2), by which it indirectly manages the trade-off between overfitting and training error, and 2) the choice of kernel function, along with any constants in it. We use the RBF kernel, and vary the γ hyper-parameter, in conjunction with the C constant, in order to attain the best performing function f. As per standard practice, we defined a set of candidate C's and a set of candidate γ's, and tried all combinations of values from these sets. For both C and γ, the candidate values ranged from 2^{-10} to 2^{10}, increasing in steps of $2^{0.5}$.

We follow the standard cross-validation approach to evaluate and fine-tune the learned model, with the number of cross-validation partitions equal to 10. We use the Area Under ROC Curve (AUC) metric to assess the performance of the model, which is a valid metric to use with cross-validation. The choice of AUC, as opposed to the accuracy (number of correctly classified observations), is preferred in our case because we have highly imbalanced class sizes. After performing cross-validation on the data set of puffs and non-puffs, we found the optimal C to be 4, and the optimal γ to be 5.65685. These values could then be used to train a single model for classifying puffs and non-puffs. In the next section, we present the performance characteristics of this fine-tuned model on the training set (evaluated via cross-validation), as well as on a separately withheld test set.

4.2.1 Optimal classification threshold selection

After we have fine-tuned the hyper-parameters C and RBF γ by finding the highest cross-validated AUC, we can introduce additional bias for one of the two classes in the problem, and fine-tune it either to improve accuracy (especially in cases of imbalanced class sizes) or to reflect our

greater preference for minimizing the false positive rate or the false negative rate. We can modify the original formulation (1) by adding an additional bias constant λ to function $g_\alpha(\mathbf{x}) = \sum_{i \in X_{\text{sv}}} \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b + \lambda$. Intuitively, a positive λ pushes the separating boundary closer to the negative class, in which case it introduces bias for the positive class. Similarly, a negative λ biases against the positive class.

In our problem, we assumed equivalent preference for low false positive rate and low false negative rate. Therefore, we fine-tuned λ to maximize the accuracy. We found that this optimal λ is 0.103129, which gives preference to the positive class. This means that while the trained classifier is able to rank the cycles well, owing to high AUC, it is somewhat skewed in the direction of the positive class, making us misclassify some negative (non-puffs) observations, but achieving a higher recall rate. Figure 6 contains the ROC curve plots for the fine-tuned model on the training set (6(a)), as well as on the withheld test set (6(b)). The filled square and filled circle on each plot correspond to the default and optimal choices for the additional bias, λ. Note that the optimal choice in both cases leads to higher accuracy (ACC in the plot) by means of higher recall.

4.3 Using Unlabeled Data: Semi-Supervised Model

In this section, we describe a specific semi-supervised model that makes use of unlabeled data obtained from the field data collected from the four participants (see section 3.2). Semi-supervised learning has been shown to improve the classification accuracy when there is a scarcity of labeled data. We use the Semi-supervised Support Vector Machine (S3VM) model [35] that extends the basic supervised SVM to incorporate unlabeled data.

The basic intuition behind S3VMs is that if we have unlabeled data together with labeled data the decision boundary obtained by the learning algorithm should be such that it separates the labeled data with a maximal margin, while simultaneously maximizing its distance to unlabeled examples. The second part in the model formulation is motivated by the notion that the model should have as little ambiguity in classifying the unlabeled examples as possible, even if there's not assurance that these classifications are correct (given that there are no labels for them). This intuition is incorporated by modifying the objective function of the basic SVM, and is given by

$$\operatorname*{argmin}_{\mathbf{w}, \mathbf{b}} \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_{i}^{N} \max(1 - y_i(\mathbf{w}^T \mathbf{x}_i + b), 0)$$
$$+ C^* \sum_{j}^{M} \max(1 - |\mathbf{w}^T \mathbf{z}_j + b|, 0),$$

where, $\{\mathbf{z}_j \in \mathbb{R}^n | j = 1 \ldots M\}$ is a set of unlabeled input data. We essentially add penalty in the objective function for the unlabeled data points that are too close to the decision boundary, specifically for which $-1 < |\mathbf{w}^T \mathbf{z}_j + b| < 1$, thereby, forcing the decision boundary to go through a low density area of the unlabeled instances. The S3VM experiments were conducted with the SVMlin toolbox [6].

The challenge in developing the semi-supervised learning model is to identify a feature that can connect the self-report to the smoking puffs, knowing that the self-report can be before, during, or after a smoking episode. The new feature we identify is the time distance of the respiration cycle in consideration to the closest self-report timestamp in the field data. As there should exist a marking before, during, or after every smoking session reported, the time distance from each cycle to the nearest self-report time should help the learning algorithm. Note that we do not have actual self-report time for the labeled data, but we can reasonably assume the existence of a hypothetical accurate self-report at the beginning of the carefully labeled smoking sessions (see section 3.2).

5. TRAINING & EVALUATION OF PUFF DETECTION MODELS

In this section, we present the performance of the classifiers for detecting smoking puffs. The training data set for the supervised classifier is comprised of the instances of puffs cycles and instances of non-puff cycles from the smoking sessions. The other sources of non-puff instances are data from the stress, conversation, and physical activity data sets. These data sets do not include any smoking events. Inclusion of these data sets are required in order to create a robust classifier that should generalize to the natural field environment better than a classifier that uses only the puff and non-puff cycles from the smoking session. Moreover, as stress, conversation and physical activity have been shown to be inferrable from respiration, they form the set of plausible confounding factors in smoking detection. In total, we have 161 puff instances and the same number of non-puff instances. The non-puff instances with equal proportion come from smoking sessions and the 3 confounding data sets.
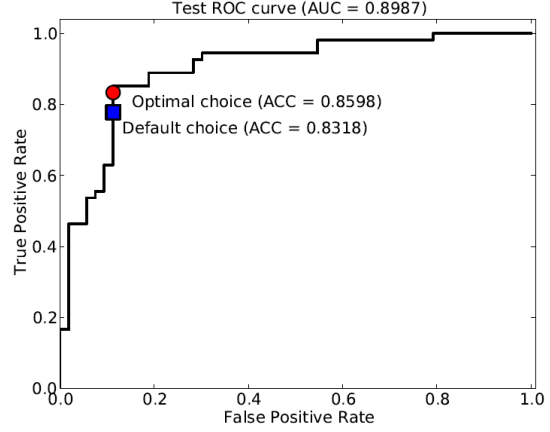
The training data set for the semi-supervised model set includes the same labeled data set as described above and a large amount of unlabeled data taken from the field data sets obtained from 4 participants. However, not all data from the field data is included as it amounts to 28 hours worth of data. Such a large amount of data proved to be infeasible to run on the SVMlin tool. Experimenting with different amounts of data, we ended up including 10 times the amount of cycles in the labeled data. We ensured that data was included both from the neighborhood of self-report times (± 5 minutes) and far away of from any self reports, thereby ensuring that the unlabeled data included both puffs and non-puffs.

Training and Testing Performance: Tables 1 and 2 present the performance of the classifiers on labeled data set. Also for greater generalizability, we split our labeled data set into training and test sets — one subset contains 66% of the whole data set and other subset contains the remaining 34%. Table 1 shows the performance of the classifiers on whole data set with 10-fold cross validation. As mentioned above for supervised classification, we use Support Vector Machine (SVM). We then split the labeled dataset into training and testing sets. For this case, the classifier performance on the test data is presented in Table 2. Data used in the supervised classifier are carefully labeled. But, when we add the noisy labeled data set from the field, we use the S3VM classifier which is a semi-supervised support vector machine. For the first experiment, when we use all the labeled data, SVM provides 84.5% accuracy; S3VM is able to improve this accuracy to 86.7%. In the second experiment, when we apply these classifier models on the testing data set, we observe the performances of the classifiers are also similar
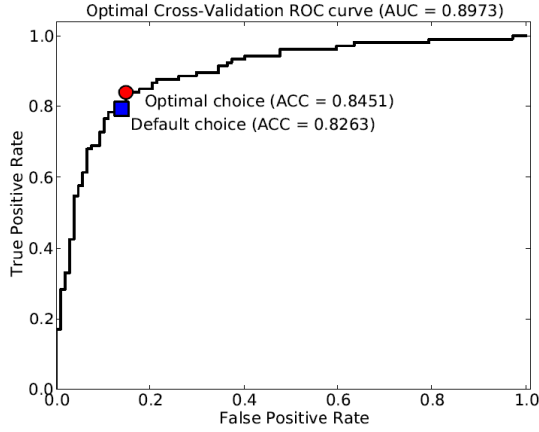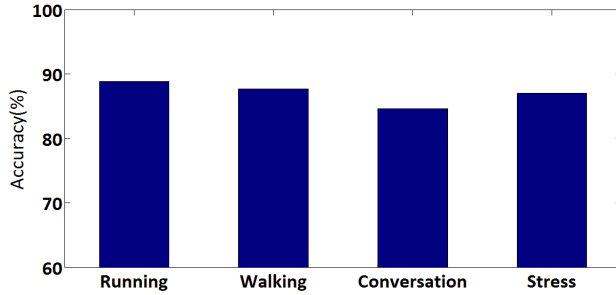
| (a) Cross-validation ROC curve | (b) Test ROC curve |

Figure 6: ROC curves for the fine-tuned model, with $C, \gamma, \lambda = (4, 5.65685, 0.103129)$, **plotted for, (a), the training set (via cross-validation) and, (b), the test set.**

Figure 7: Accuracy of the classifier on the datasets for four confounding events.

to the training accuracies (as shown in Table 2). In both of these experiments, we get high precision, high recall or false positive rate, and high true positive rate. Moreover, if the training set consists of only the non-puffs taken from the smoking session the testing accuracy is 91.43% for the SVM classifier. This indicates that the classifier is quite efficient in detection of puffs and non-puffs in the absence of confounding factors.

Figure 7 presents the accuracy incurred when the puff detection model is run on the stress, conversation, and physical activity datasets. We note that that these datasets do not include any smoking sessions and therefore respiration cycles not correctly identified implies it was detected as a puff. We observe that conversation is the most confounding event for the puff detection model. One reason is that during speaking, we tend to take deep breaths at the beginning of the speech and that is sometimes detected as puffs. To remedy this problem, in the next section, we propose a technique that can be applied on the data in the field when the conversation and smoking detection models are both available on the phone.

Feature Analysis: The performance of the classifiers presented in the Tables 1 and 2 uses all the 17 features. Here, we present the effect of adding the 12 new features on the

performance of the classifier. We partition the set of all the new features into 3 sets, namely, the set of first differences, the set of the stretch and exhalation duration ratio, and the set of Upper and Lower stretches. All the accuracy values reported are for the SVM classifier. The basic set of features comprising only the existing 5 features produces an accuracy of 73.55% on the whole labeled data set. From Table 3, we observe that among the three new sets of features, adding the first differences to the basic set improves the accuracy most, reaching up to 81.1% . With this set of features, adding the stretch and exhalation duration ratios increases the performance of the classifier most. With these 2 new sets of features, we obtain an accuracy of 82.7%.

Table 3: Effect of new features on the classification accuracy. The classification accuracy obtained using only the existing features is 73.55%. They constitute the basic set of features. Three different sets of features are introduced in this paper. Let S_1 **= set of forward and backward first differences of basic features,** S_2 **= {Stretch ratio, Exhalation ratio} and** S_3 **= {Upper Stretch, Lower Stretch}. X denotes the inclusion of the set to the basic set of features.**

S_1	S_2	S_3	Accuracy(%)
X			**81.1**
	X		76.33
		X	77.86
X	X		**82.7**
X		X	81.32
	X	X	75.79

6. SMOKING TOPOGRAPHY STATISTICS

Table 4 summarizes mean and standard deviation of various measures of interest, e.g., duration of smoking session, puff duration, inter-puff interval, and frequency of puffs per smoking session. These measurements are calculated from

Table 1: Performance of classifiers on training data with 10-fold cross-validation.

Classifier	Performance Matrices			
	Accuracy(%)	Precision	TP rate	FP rate
SVM	84.5	0.85	0.84	0.15
S3VM	86.7	0.91	0.81	0.08

Table 2: Performance of classifiers on testing data when data is split into training and test sets.

Classifier	Performance Matrices			
	Accuracy(%)	Precision	TP rate	FP rate
SVM	85.98	0.88	0.83	0.11
S3VM	87.27	0.91	0.83	0.08

the data collected from the smoking sessions where puffs are marked carefully using a cell phone. As described in 3.2 there were 13 smoking session from 10 individuals in the data set.

Our observations are generally consistent with previous lab-based and field-based research on smoking topography and puff-analysis [17, 12, 14], thereby validating the use of respiration bands to collect respiration measurements of smoking in the field.

7. DISCUSSION

In this section, we consider two issues on detection of entire smoking episodes. We first discuss how the accuracy of detecting smoking episodes may be improved further by using insights from conversation patterns since conversation appears to be a frequent daily human activity that confounds detection of smoking. Next, we describe how the results of puff detection accuracy can be used with an extensive real-life data set to obtain confidence level in the detection of an entire smoking episode depending on the number of puffs contained in a smoking session.

Leveraging Patterns of Natural Conversations. Results on the confounding data sets indicate that during conversations, the puff detector may raise false positives. Upon running the conversation detector on smoking sessions, we find that the entire smoking session is detected as a long conversation. According to the findings in [26], most natural conversations are short, with an average duration of 3.8 minutes. Less than 20% of conversations are 6 minutes or longer. The average duration of a smoking session (M), on the other hand, is 6.62 minutes. We can use this difference in durations to eliminate a major source of false alarm in the detection of smoking events. One approach is to apply a conversation detector on the respiration measurements after applying the puff detector and then clear puff markings during those intervals that are detected as conversation, but have a duration less than $M - 2 * SD$ minutes, where SD is the standard deviation of a smoking session duration. This approach will be useful only in those scenarios where the objective is to detect entire smoking episodes, and not single puffs (such as in a smoking cessation study). We plan to evaluate such approaches on data sets collected from real-life to evaluate its impact on both false positive rate and false negative rate.

Confidence in Detecting an Entire Smoking Episode. Since a smoking session usually consists of multiple puffs, the confidence in detecting a smoking session should increase with increasing number of puffs contained in it. We can compute the confidence of detecting of a smoking session using the following procedure. Let X be a binary random variable, where $X = 1$ indicates that a puff detected by the smoking detector is in a smoking session, and $X = 0$, otherwise. Also, in a sequence of output produced by the puff detector, when applied on a times series of respiration measurements, let $Y(i)$ denote the number of puffs detected in N respiration cycles preceding and including the ith cycle. A good choice for N is the average number of respiration cycles in a smoking session, which from our data set is equal to 72. Then, $\Pr[X = 1|Y(i) = y]$ is the probability of a the ith cycle which is detected to be a puff being in a smoking session, given that in the last N cycles y puffs were observed. We can use the Bayes' rule to obtain this probability. More specifically,

$$\Pr[X = 1|Y = y] = \frac{\Pr[Y = y|X = 1]\Pr[X = 1]}{\Sigma_{i \in \{0,1\}}\Pr[Y = y|X = i]\Pr[X = i]}$$

This enables the computation of $\Pr[X = 1|Y = y]$ if we have estimates for $\Pr[Y = y|X = i]$ and $\Pr[X = i]$, for $i \in \{0, 1\}$. These probabilities represent the behavior of the smoking detector on data set obtained from real-life of various subjects, if the subjects self-mark each smoking session in their mobile environment while wearing a respiration sensor. Any puff within the neighborhood of a smoking self-report can be used to estimate $\Pr[Y = y|X = 1]$ and puff instances that are not in the neighborhood of a self-report can be used to estimate $\Pr[Y = y|X = 0]$. $\Pr[X = i]$ can be estimated using the true positive rate and false positive rates of the puff detector. This approach can be applied to real-life data sets to obtain the corresponding confidence levels in detecting entire smoking episodes when using our smoking detector.

8. CONCLUSION

Extensive research in smoking literature is aimed towards the development of efficient smoking interventions. However, the absence of a real time smoking event detector hampers the development of interventions that work well. Consequently, the success rate of smoking cessation programs is abysmally low. As a first step towards building a reliable smoking episode detector, in this work we presented a model to automatically detect smoking puffs in the natu-

Table 4: Statistics of smoking topography obtained from the labeled data set (average value and standard deviation).

Statistic	Avg. $\pm$ St.Dev.
Duration of smoking session (minutes)	6.62 ± 1.66
Puff duration (seconds)	1.09 ± 0.53
Inter-puff interval (seconds)	28.38 ± 14.57
Number of puff per smoking session	12.38 ± 0.92

ral environment from respiration measurement. We achieve 86.7% accuracy on the detection of puffs even when there exists potential confounding events in the collected data. For the purpose of building the model, we identified 12 new respiration features that are found to be effective compared to the use of only the existing respiration features available in the literature. We also presented a semisupervised moded that improves the accuracy of the model when we provide unlabeled data with self reports that is also collected in the natural environment from participants who are daily smokers. Since other contextual factors such as stress, location, social activity can also be detected on a mobile phone, our work opens the opportunity for effective smoking research by allowing the collection of different contextual factors, that are potential predictors of smoking, in the natural environment of smokers.

9. ACKNOWLEDGMENTS

10. REFERENCES

[1] Smoking-attributable mortality, years of potential life lost, and productivity losses - United States, 2000 - 2004. *Morbidity and Mortality Weekly Report*, 57(45):1226–1228, 2008.

[2] Fieldstream. http://www.github.com/FieldStream/FieldStream, Accessed: October 2011.

[3] How tobacco smoke causes disease - the biology and behavioral basis for smoking-attributable disease - a report of the surgeon general. http://www.surgeongeneral.gov/library/ tobaccosmoke/report/index.html, Accessed: October 2011.

[4] Smokerlyzer. http://www.bedfont.com/smokerlyzer, Accessed: October 2011.

[5] Substance abuse and mental health administration: Results from the 2009 national survey on drug use and health: National findings. http://www.oas.samhsa. gov/NSDUH/2k9NSDUH/tabs/Cover.pdf, Accessed: October 2011.

[6] Svmlin. http://vikas.sindhwani.org/svmlin.html, Accessed: October 2011.

[7] Treating tobacco use and dependence: 2008 update - clinical practice guideline. http://www.surgeongeneral.gov/tobacco/ treating_tobacco_use08.pdf, Accessed: October 2011.

[8] CReSS. http://www.borgwaldt.de/cms/ borgwaldt-kc/produkte/rauchmaschinen/ geraete-zur-rauchtopographie/cress-pocket.html, Accessed: October 2011.

[9] H. Ashton, D. Watson, R. Marsh, and J. Sadler. Puffing frequency and nicotine intake in cigarette smokers. *The British Medical Journal*, pages 679–681, 1970.

[10] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[11] H. Brendryen, P. Kraft, and H. Schaalma. Looking inside the black box: Using intervention mapping to describe the development of the automated smoking cessation intervention'happy ending'. *The Journal of Smoking Cessation*, 5(1):29–56, 2010.

[12] R. Bridges, J. Combs, J. Humble, J. Turbek, S. Rehm, and N. Haley. Puffing topography as a determinant of smoke exposure. *Pharmacology Biochemistry and Behavior*, 37(1):29–39, 1990.

[13] F. Charles, G. Krautter, and D. Mariner. Post-puff respiration measures on smokers of different tar yield cigarettes. *Inhalation toxicology*, 21(8):712–718, 2009.

[14] C. Collins, D. Epstein, C. Parzynski, D. Zimmerman, E. Moolchan, and S. Heishman. Puffing behavior during the smoking of a single cigarette in tobacco-dependent adolescents. *Nicotine & tobacco research*, 12(2):164–167, 2010.

[15] K. Doherty, T. Kinnunen, F. Militello, and A. Garvey. Urges to smoke during the first month of abstinence: relationship to relapse and predictors. *Psychopharmacology*, 119(2):171–178, 1995.

[16] E. Ertin, N. Stohs, S. Kumar, A. Raij, M. al'Absi, T. Kwon, S. Mitra, S. Shah, and J. W. Jeong. Autosense: Unobtrusively wearable sensor suite for inferencing of onset, causality, and consequences of stress in the field. In *ACM Conference on Embedded Networked Sensor Systems (SenSys)*, pages 274–287, 2011.

[17] D. Hammond, G. Fong, K. Cummings, and A. Hyland. Smoking topography, brand switching, and nicotine delivery: results from an in vivo study. *Cancer Epidemiology Biomarkers & Prevention*, 14(6):1370–1375, 2005.

[18] J. Hughes and D. Hatsukami. Signs and symptoms of tobacco withdrawal. *Archives of General Psychiatry*, 43(3):289, 1986.

[19] N. Hymowitz, M. Sexton, J. Ockene, and G. Grandits.

Baseline factors associated with smoking cessation and relapse. *Preventive Medicine*, 20(5):590–601, 1991.

[20] D. Kalman. The subjective effects of nicotine: methodological issues, a review of experimental studies, and recommendations for future research. *Nicotine & Tobacco Research*, 4(1):25–70, 2002.

[21] J. Killen and S. Fortmann. Craving is associated with smoking relapse: Findings from three prospective studies. *Experimental and Clinical Psychopharmacology*, 5(2):137, 1997.

[22] S. Kolonen, J. Tuomisto, P. Puusteinen, and M. Airaksinen. Puffing behavior during the smoking of a single cigarette in a naturalistic environment. *Pharmacology, biochemistry and behavior*, 41(4):701–706, 1992.

[23] K. Matheny and K. Weatherman. Predictors of smoking cessation and maintenance. *Journal of Clinical Psychology*, 54(2):223–235, 1998.

[24] A. Mokdad, J. Marks, D. Stroup, and J. Gerberding. Actual causes of death in the united states, 2000. *JAMA: the journal of the American Medical Association*, 291(10):1238, 2004.

[25] K. Plarre, A. Raij, S. M. Hossain, A. A. Ali, M. Nakajima, M. al'Absi, T. Kamarck, S. Kumar, M. Scott, D. Siewiorek, A. Smailagic, and L. E. Wittmers. Continuous inference of psychological stress from sensory measurements collected in the natural environment. In *ACM Information Processing in Sensor Networks (IPSN)*, pages 97–108, 2011.

[26] M. M. Rahman, A. A. Ali, K. Plarre, M. al'Absi, E. Ertin, and S. Kumar. mconverse: Inferring conversation episodes from respiratory measurements collected in the field. In *ACM Wireless Health*, 2011.

[27] S. Shiffman. Reflections on smoking relapse research. *Drug and alcohol review*, 25(1):15–20, 2006.

[28] S. Shiffman, J. Paty, M. Gnys, J. Kassel, and M. Hickcox. First lapses to smoking: Within-subjects analysis of real-time reports. *Journal of Consulting and Clinical Psychology*, 64(2):366–379, 1996.

[29] S. Shiffman, D. Scharf, W. Shadel, C. Gwaltney, Q. Dang, S. Paton, and D. Clark. Analyzing milestones in smoking cessation: illustration in a nicotine patch trial in adult smokers. *Journal of consulting and clinical psychology*, 74(2):276–285, 2006.

[30] S. Shiffman and A. Waters. Negative affect and smoking lapses: A prospective analysis. *Journal of Consulting and Clinical Psychology*, 72(2):192–201, 2004.

[31] M. Stitzer and J. Gross. Smoking relapse: the role of pharmacological and behavioral factors. *Progress in clinical and biological research*, 261:163–184, 1988.

[32] G. Swan, M. Ward, and L. Jack. Abstinence effects as predictors of 28-day relapse in smokers* 1. *Addictive Behaviors*, 21(4):481–490, 1996.

[33] J. Veilleux, J. Kassel, A. Heinz, A. Braun, M. Wardle, J. Greenstein, D. Evatt, and M. Conrad. Predictors and sequelae of smoking topography over the course of a single cigarette in adolescent light smokers. *Journal of Adolescent Health*, 48(2):176–181, 2011.

[34] P. Wu, J. Hsieh, J. Cheng, S. Cheng, and S. Tseng. Human smoking event detection using visual interaction clues. In *International Conference on Pattern Recognition*, pages 4344–4347. IEEE, 2010.

[35] X. Zhu and A. Goldberg. *Introduction to Semi-Supervised Learning*. Morgan & Claypool Publishers, 2009.