Detection of Cigarette Smoking Episodes in the Natural Environment

Amin A. Ali
Department of Computer Science
University of Memphis
aaali@memphis.edu

Abstract

One cannot over emphasize the harmful effects of smoking which has been proven to be the leading cause of mortality in the United States. Though there exists extensive research on developing intervention to aid smoking smoking cessation, in the absence of an automated smoking detection method, such programs have poor success rates. Automated detection of smoking episodes in the natural environment can revolutionize smoking research and lead to effective intervention.

In this dissertation, we present the outlines of a novel system to automatically detect smoking episodes from respiration measurements. Preliminary investigation of the feasibility of using respiration measurement provides ample promise for the development of a reliable and robust detector. We have developed an SVM classifier that is able to detect puffs with 87% accuracy. However, there are several challenges that still needs to be addressed, such as, accounting for between and inter subject differences in puff taking behavior, reducing the false positive rate by accounting for artifacts (e.g., bodily movement, conversation, laughter etc.), and detection of smoking episodes in the presence of misclassification. In order to meet these challenges, first, we have designed a new user study in order to collect more smoking data in the field with addition modalities such as accelerometer, gyroscope to detect the motion of the hand holding the cigarettes. Second, to locate puffs in the field data, we intend to use smoking topography devices. Third, as in a smoking session there are repetitive puffs in somewhat regular intervals, HMM models can be useful to smooth out misclassification and to identify smoking episodes using puffing patterns (e.g., inter puff durations).

1 Introduction

It is a well known fact 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. These diseases lead to early deaths and it has been shown that use of tobacco accounts for nearly one of every five deaths in the United States [1]. In addition to substantial progress in tobacco control, and health education programs, there has been much research in development of interventions to aid smoking cessation. However, most smoking cessation programs achieve low success rate (i.e., less than 10%) and one of the main reasons behind is that they do not have mechanisms to intervene at the right moment.

In smoking research one of the important goals is to identify the antecedents and precipitants (i.e., high risk situations, such as stress) of smoking lapses. Scientific user studies resort to observing and recording user’s context while smoking occurs. Thus it is imperative that these studies employ some method for detecting smoking episodes. However, most of the studies on smoking behavior mainly employ self-reporting methods for this purpose. These self-reporting methods range from basic pen-paper methods and retrospective recalls, to electronic diary keeping and ecological momentary assessments(EMA) [10]. These methods, in addition to putting extra burden on users, depend on user compliance, introduce subjective biases etc. that limit their utility. Although other methods such as use of CO monitors (e.g., piCO+, Micro+ [2], CReSS Pocket [3], and RespiTrace® [5]) have been reported in tobacco research literature, as discussed in section 2 they either require manual intervention from users/observers, or not unobtrusive enough to be used in the user’s daily life.

In this dissertation, we introduce a novel method for automatic detection of smoking episodes in the natural environment that is also operator independent. The system, uses machine learning models that use respiration measurements, collected using Autosense® [6]. The expected contribution of this work are, first, identification of robust respiration features that facilitate the identification of smoking puff respiration cycles. The features must be robust enough to be resilient to noise in the respiration measurement that is observed in the data collected in the field. Second, building a model that classifies the respiration cycles to puff and non-puff classes. Third, in the presence of misclassification, iden-
tifying the smoking episodes (i.e., the start and end of smoking events) reliably. Fourth, if EMA could be augmented with automated detection of smoking, we could reduce participant burden and be more confident in the completeness of our smoking records. Automatically detected smoking could also trigger a prompt so that participants could provide additional information about the episode—or could trigger an Ecological Momentary Intervention (EMI). In our preliminary investigation we built a smoking puff detector which achieves 87% accuracy. However, with a high rate of false positives and false negatives this classifier is not reliable enough to detect smoking episodes in the field. Moreover, there needs to be an in-depth investigation of between and within person differences in puffing behavior. To meet these challenges we have designed a new user study where we will recruit daily smokers who will wear Autosense along with a wrist band containing an accelerometer and gyroscope in order to capture the movement of the hand that holds the cigarette. Also we intend to use of CReSS device to mark the timing of each puff. These will allow identifying the puffs in the field more reliably and in turn helping to identify robust features for classification. However, in order to smooth out the misclassification and to identify smoking episodes we believe we have to look at higher level features such as number of puffs in a window, inter puff duration etc. and employ a Hidden Markov Machine like model to capture the puff patterns within a smoking session. In addition to collecting physiological data, we also collect real time self-reported craving, mood, activities and cue exposure using EMAs. This provides the opportunity to design effective interventions that utilizes the relationship between smoking events and the above mentioned contexts.

2 Related Work

A sensor-based systems that automatically and continuously collects data and makes inferences from it without altering the smoking behavior, can circumvent methodological difficulties in smoking research. Additional features of such a system should focus on the comfortability of the use of the system and its unobtrusiveness as it is to be worn or carried by the user throughout the day. There are a couple of breath CO monitor devices, such as piCO+, Micro+ [2] and CReSS Pocket [3], that have been used in different smoking studies where the main goal has been to identify differences in smoking patterns, and cigarette types or brands [8, 7]. However, one has to actually exhale into the mouthpiece attached to these devices. This makes the use of such devices ill suited for actual smoking intervention or cessation programs. A promising alternative method is the use of a respiration band, such as RespiBand Plus, that measures the chest’s expansion as the wearer breathes in and out. Authors in [5] make use of these measurements to analyze smoking patterns. Use of a respiration band is relatively unobtrusive as it can be worn underneath the clothing and thus does not draw unwanted attention. However, the method used in [5] to actually mark the puffs in the respiration signal was semi-automatic. To the best of our knowledge, there has not been any attempts in real time detection of smoking, in the natural environment, that is also operator independent.

3 Preliminary Work

In this section we describe the smoking puff detection model. In this phase, data from ten volunteers were collected and each of the puffs were carefully marked. The participants wore the Autosense chest band that provides the respiration data by measuring the expansion and contraction of the chest via inductive plethysmography at the sampling rate of 21.33 Hz. The measurements are transferred over ANT radio to a mobile phone. In total we collected 161 puff instances. Four of these ten participants also provided data from their natural environment as they wore autosense for seven days and self-reported start of smoking episodes. In this case, the individual puffs were not marked.

Through visual inspection of the data we identified 17 features from respiration measurements that capture different morphological characteristics of respiration cycles. Out of these, five features were obtained from our prior work on respiration measurement [9] illustrated in Figure 1. The new features, derived from these basic features of respiration cycles over a window, tries to amplify the difference of a puff cycle from the neighboring non-puff cycles (see Figure 2, for details see [4]).

A SVM classifier was trained to classify the respiration cycles. Instances of non-puff class were collected from smoking sessions and data set of three different confounding factors (stress, conversation and physical activity). The accuracy for test data set, when data was split into training and test data sets, was 86% with 11% false positive rate. We also trained a semi-supervised classifier that made used of the unlabeled data, collected from the field, and it improved the accuracy of detecting puffs to 87.27% with 8% false positive rate.

4 Challenges

In this section we describe the challenges to be met in order to build a reliable smoking episode detector. First, although the current puff detection model achieves 87% classification accuracy, it can be readily seen that it is bound to produce a large number of false positive and false negatives when it is applied to participant’s field data. As there is, on average, 14 respiration cycles per minute, number of cycles amount to more than 20,000 cycles in the usual 14 hours of wearing time. Then the total number of misclassified cycles amount to approximately 3000 per day. Second, it can be observed form the data collected so far is that, there exists substantial between subject differences in the puffing behavior. The amount of smoke people tend to inhale during puffs vary quite a lot. Also the inter puff distance (in terms of time / number of respiration cycles) tend to vary in different circumstances such as when they are smoking alone, in

![Figure 1. Illustration of three features extracted from respiration cycles.](image-url)
a conversation with other people, or in motion (i.e., walking, driving etc.). Third, one needs to account for confounding factors such as bodily movement, physical activity, stress, conversation, as well as events such as laughter, sneezing etc that modifies the respiratory pattern. Fourth, we need to account for misclassifications when we want to identify the start and end of smoking episodes.

5 Detection in the Field

In this section we describe the approach we plan to follow to meet the challenges mentioned in section 4.

Data Collection: In order to build a robust classifier, first, we require a large number of examples of smoking from a number of people in different contexts.

With this objective of collecting data where all the location of smoking puffs are easily identifiable we designed another user study that uses Autosense along with the CReSS pocket smoking topology device. The participants will be asked to smoke using the CReSS device, which logs the timing of each puff. However, as this additional device is rather obtrusive and may cause embarrassment as well as increasing the burden of the participants, we do not envision it to be used in smoking cessation programs. Thus we plan to investigate other sensor modalities that can assist the puff detector without increasing the burden of the users too much.

Enhancement to the Sensor Suite: One modality that appears to be useful is an 3-d accelerometer. Smokers usually hold the cigarettes in their dominating hand and raises the hand to the mouth to take a puff. Thus a 3-d accelerometer on the wrist can be useful to detect this motion. We also plan to investigate the incremental advantage of adding a gyroscope with the accelerometer, in order to reliably detect that particular gesture.

Puff and Smoking Event detection: With the new set of data collected in the field we need to re-train the puff detector. We believe that ensuring low false positives is more crucial in different applications. Parameter search for the SVN model is to be set to achieve this. Next, we plan to train an Hidden Markov model to capture the pattern of occurrence of puffs within a smoking episode. The output of the SVN classifier are to be used as the observables. Use of HMM can smooth out the misclassified outputs and identify the start and end of smoking episodes.

6 Biography

Amin A Ali is a PhD candidate in the department of Computer Science at the University of Memphis, United States, under the supervision of Dr. Santash Kumar. He obtained his BSc and MSc in Computer Science and Engineering from University of Dhaka, Bangladesh. His expected graduation is Fall 2013.

7 References