Identifying Drug (Cocaine) Intake Events from Acute Physiological Response in the Presence of Free-living Physical Activity

Syed Monowar Hossain*, Amin Ahsan Ali*, Md. Mahbubur Rahman*, Emre Ertin†, David Epstein‡, Ashley Kennedy‡, Kenzie Preston‡, Annie Umbricht§, Yixin Chen¶ and Santosh Kumar*
*Dept. of Computer Science, University of Memphis, Email: {smhssain, aaali, mmrahman, santosh.kumar}@memphis.edu
†Dept of Electrical & Computer Engineering, The Ohio State University, Email: ertin.1@osu.edu
‡NIDA Intramural Research Program, Email: {depstein, kpreston}@intra.nida.nih.gov, ashley.kennedy@nih.gov
§Dept. of Psychiatry and Behavioral Sciences, Johns Hopkins University, Email: annieumbricht@jhu.edu
¶Dept of Computer Science and Engg., Washington University in St. Louise, Email: chen@cse.wustl.edu

***Abstract*— A variety of health and behavioral states can potentially be inferred from physiological measurements that can now be collected in the natural free-living environment. The major challenge, however, is to develop computational models for automated detection of health events that can work reliably in the natural field environment. In this paper, we develop a physiologically-informed model to automatically detect drug (cocaine) use events in the free-living environment of participants from their electrocardiogram (ECG) measurements. The key to reliably detecting drug use events in the field is to incorporate the knowledge of autonomic nervous system (ANS) behavior in the model development so as to decompose the activation effect of cocaine from the natural recovery behavior of the parasympathetic nervous system (after an episode of physical activity). We collect 89 days of data from 9 active drug users in two residential lab environments and 922 days of data from 42 active drug users in the field environment, for a total of 11,283 hours. We develop a model that tracks the natural recovery by the parasympathetic nervous system and then estimates the dampening caused to the recovery by the activation of the sympathetic nervous system due to cocaine. We develop efficient methods to screen and clean the ECG time series data and extract candidate windows to assess for potential drug use. We then apply our model on the recovery segments from these windows. Our model achieves 100% true positive rate while keeping the false positive rate to 0.87/day over (9+ hours/day of) lab data and to 1.13/day over (11+ hours/day of) field data.**

Keywords*—Drug Event Detection, Wearable Sensors, Electrocardiogram.

I. Introduction

Advances in mobile sensing are enabling a new vision of healthcare (called mobile health or *mHealth*), where consumers can monitor, manage, and improve health and well-being as they go about their daily lives [1]. Wearable, inexpensive sensors allow capture of health relevant data, such as measurements of heart, respiration, physical activity, location, etc. in the natural environment [2], [3]. These wearable sensors have become reliable instruments to collect sensor measurements in the field environment, and hence research has shifted from developing wireless sensor platforms to the processing of sensor data to infer health related events such as stress [4], smoking [5], drug use [6], and identify their antecedents and precipitants (i.e., high risk situations). Automated detection of these precipitants on a mobile phone can be used to trigger just-in-time intervention or treatment. For example, momentary exposure to greater physical disorder, social disorder, and drug activity in a neighborhood (as indicated by the NIfETy score [7]) and experiencing craving or stress could constitute a high-risk situation worthy of real-time intervention for a drug user wanting to quit. The first step in building such mHealth control systems for improving human health usually is to collect sensor data in the natural environment and locate the adverse health events in the time-series of sensor data so as to discover the causal role of various contexts in precipitating the adverse health event. Therefore, the development of reliable models for detecting adverse health events from mHealth sensor data in the natural field environment is a critical step in the development of just-in-time mHealth interventions.

Development of such models involve several challenges. First, appropriate sensor(s) are needed that can be used conveniently in the field settings for long enough duration to capture the health events of interest. The sensor should be robust for reliable data collection in the free-living natural environment. In addition, the sensor should be sensitive and specific enough to exhibit a detectable response to the health event of interest. Second, appropriate data collection experiment needs to be designed and conducted to collect the sensor data with appropriate labeling of the times when the health event of interest occurs (to train the model). Third, a robust computational model needs to be developed that is able to detect the event of interest from sensor data in the free-living natural environment.

There is increasing interest in the problem of automatically detecting drug use due to its high societal impact. Illicit drug use, affecting over 150 million people, is a major cause of mortality from fatal overdose and dependence, HIV, Hepatitis B, and Hepatitis C. Other adverse health effects include mental disorder, traffic accidents, suicides, and violence [8].

Drug intake events (e.g., cocaine use), known to acutely excite the autonomic nervous system, can potentially be detected from its response on wearable electrocardiogram (ECG) sensors. Heart Rate (HR) increase of more than 30% (for 16 mg of cocaine) [9], [10] have been observed in lab studies.

After (intravenous) cocaine administration, it usually takes 20-60 minutes for the heart rate to recover. Hence, the effect of cocaine use on ECG is quite acute and potentially distinct.

To the best of our knowledge, [6] is the first work to present a computational model for automated detection of cocaine use from ECG. They collect data from six participants in lab setting during cocaine administrations of 8 mg, 16 mg and 32 mg dosage levels. The authors develop a model to classify ECG cycles into drug and baseline classes. Although this method achieves high accuracy (average AUC > 0.9) on clean lab data, there are several challenges in using such a model for detecting cocaine use events in the field setting. For example, as we show in Section V-A, it is not trained to distinguish ECG cycles during drug use from that during physical activity. Second, it is not trained to be invariant to dosage amount and to the modality of administration. Third, since the model classifies each ECG cycle, it is not clear how to use this as a building block to develop a model for detecting the entire drug use event and to distinguish it from physical activity events. Although accelerometry measurements can be used to detect the occurrence of physical activity, drug use events in the field are usually concurrent with physical activity (i.e., subjects are not stationary following cocaine use), and hence acceleromtery alone can't distinguish the two events.

In this paper, we develop a physiologically-informed model to automatically detect drug events from their acute physiological response, in the presence of various confounders inherent in the free-living lifestyle. The key to reliably detecting drug use event is to incorporate the knowledge of autonomic nervous system (ANS) behavior in the model development. We decompose the activation effect of cocaine from the natural recovery behavior of the parasympathetic nervous system (PNS) that can be observed upon conclusion of a physical activity episode.

We first designed and conducted three user studies with active drug users — two in the lab and one in the field. In each study, participants wore the AutoSense sensor suite [2] that included an ECG sensor and accelerometers. The lab studies were conducted in residential facilities with 9 drug users (across 89 days). It included free-living lifestyle together with sessions of repeated cocaine administration of various doses under medical supervision. For the field study, 42 drug users wore the sensors for 4 weeks in the field so as to maximize the chances of capturing real-life cocaine use events.

We then develop efficient methods to screen and clean sensor data to handle noise and drift. Next, we develop a data preprocessing stage to identify and locate windows in ECG time series that exhibit physiological response of sufficient magnitude that may result from cocaine use. Activity-free recovery segments from these windows are assessed to determine whether this window is a result of cocaine use. For this purpose, we develop a dynamical system model of the parasympathetic nervous system (PNS) behavior from the heart rate recovery observed upon conclusion of cocaine-free physical activity episodes. In the case of cocaine use, the PNS recovery is dampened due to excitation of the sympathetic nervous system (SNS) from cocaine. The strength of SNS excitation weakens with metabolism of cocaine. Using lab data from cocaine administrations, we develop dynamical system models of both SNS activation and its weakening due to cocaine metabolism, in order to model the cocaine-dampened PNS recovery. We refer to these models collectively as our Autonomous Nervous System (ANS) model. Our ANS model classifies a window into cocaine class if the recovery portion of the window matches that of a cocaine-dampened PNS recovery, and otherwise, if it better matches natural PNS recovery.

We evaluate the performance of our ANS model on both lab and field data. We present some operating points from the ROC curve (see Figure 8). For 100% true positive rate, the model keeps the false positive rate to 0.87/day in the lab, even though 10+ candidate windows (due to significant physical activity) are found during 9+ hours/day of sensor wearing in the lab. On 922 days of the field data, we find 27 episodes of cocaine use with good quality sensor data from 13 participants. For these participants, there are 79 confirmed non-cocaine days as established from urine assessment. On these days, there are 1,171 major activity episodes. The model keeps the false alarm rate to 1.13/day, even though 11+ candidate windows are found during 11+ hours/day of sensor wearing in the field.

Contributions: This paper makes several key contributions. First, it establishes feasibility of collecting good quality physiological sensor data in the field environment from illicit drug users during real-life drug use events. These data provide a first look at physiological response to drug use episodes of up to 600 mg. Second, this work shows that it is indeed possible to develop a model for detection of health events from physiological sensor data collected in the lab settings that generalize to the free-living natural environment, providing a feasibility result for the critical question of lab-to-field generalizability. This is achieved by developing explainable models that model physiological mechanisms and hence enable decomposing the effect of physical activity on the physiological response. Third, automated detection of drug use events in real-time (in a scientific study) can be used to solicit self-reports on the mobile phone to obtain more information surrounding the context of drug use. Fourth, accurate localization of drug use events in the time series of sensor measurements can enable discovery of antecedents and precipitants that can be detected from sensors in the mobile phone alone (e.g., GPS, accelerometry). Such discoveries can then be used to construct novel just-in-time interventions to predict drug use events and break the urge. Even self-monitoring of these contexts can enhance a patients' awareness of vulnerable contexts and help them avoid such high risk situations [11], [12], [13].

Finally, the model of parasympathetic nervous system (PNS) can be of independent interest in assessing the fitness of individuals. Recovery from physical activity (i.e., the health of PNS) is traditionally used for estimating cardiovascular health in clinical settings [14], [15]. The ANS model developed in this work can be used to obtain stable estimates of PNS recovery in the natural environment. This new measure has a potential to be of similar utility as heart rate variability (HRV) in biofeedback applications for self-improvement of physiological health.

Real-life Usage of the Model: The work presented here is not intended to be used directly by drug users seeking help with cocaine abstinence. Rather, this model is intended to be used in scientific studies of drug use. Using the model, we are able to pinpoint where the cocaine event occurred in time, which cannot be reliably inferred from self-report. Locating cocaine use events in the time series can help identify predictors from

other sensor modalities (e.g., location, activity, stress, etc.). During real-life usage to help cocaine abstinence, the system does not detect cocaine use; rather, it predicts potential for lapse and prompts a user to break his/her urge when the predictors are detected. In this phase, the users are not wearing the physiological sensors. They only carry their smartphones.

We further point out that collecting physiological sensor data during real-life cocaine use episodes is an inherently challenging task due to the context and situation associated with illicit drug use. It took nearly 922 person days (10,449 hours) of sensor wearing in the field to find 27 cocaine use episodes with good quality ECG data. Fortunately, development of a model is a one-time activity. Once a model is published and predictors of cocaine use are discovered and reported (subject of ongoing work), the community does not need to go through this tedious and resource intensive process again. They can use our models and use the predictors we find in our data set in developing cocaine abstinence interventions.

Organization: Section II describes some related works and summarizes the key challenges in developing a reliable model for detecting cocaine use in the field. Data collection procedure and statistics of data collected is described in Section III. Data processing steps and model development are described in Section IV. Section V presents evaluation of our ANS model. Section VI concludes the paper and discusses future research.

II. Related Works and Key Challenges

There exists a substantial body of literature on the measurement of physiological responses to various drugs, such as cocaine, in the laboratory environment, where different doses of cocaine were administered to volunteers and physiological measurements (e.g., heart rate, blood pressure) recorded [9], [16], [10], [17], [18], [19]. These studies focussed on examining the influence of different routes of administration (e.g. smoked, intervenous and intranasal) on pharmacokinetic parameters and drug-induced behavioral and physiological effects of cocaine. Table 1 summarizes results from these works on the effect of cocaine on heart rate. We observe that the response time and duration of effect depends highly on the route of administration [20]. In case of smoked or intravenous administration, the onset of action is almost immediate and within 1-5 minutes the effect reaches its peak. In all cases, there is an increase in heart rate (or decrease in RR interval - the interval from one R-peak to the next in the ECG waveform) and blood pressure and both increase with dosage. An increase of 32% and 34% respectively for intervenous cocaine doses of 16 mg and 32 mg was reported in [9], [10], [21].

TABLE I. Response time and effect duration of cocaine intake on heart rate.

Route of Administration	Onset of Action	Peak Effect	Duration of Action
Smoking	3-5 sec	1-3 mins	5-15 mins
Intravenous	10-60 sec	3-5 min	20-60 min
Intranasal or mucusal	1-5 min	15-20 min	60-90 min
Gastrointestinal	$\leq$ 20 min	$\leq$ 90 min	$\leq$ 180 min

For characterizing physiological response to drug use, [22] used non-linear regression to model the heart rate during cocaine and placebo administration sessions. The placebo session consisted of a bolus intervenous injection and a 4-hour continuous infusion of a 0.2N saline solution. It used 40 mg, 60 mg and 80 mg cocaine intervenous doses followed by a 0.2N saline or cocaine infusion that lasted for 4 hours. Heart rate was modeled as the sum of baseline heart rate, an exponentially decaying model of conditioned response effect, and a non-linear drug effect model. The conditioned response effect was used to model the response that was observed a couple of minutes before the injections which also continued a couple of minutes after the injection, even during the placebo sessions. This model, however, cannot be readily used to detect drug use in the field as it does not distinguish from similar effects that are observed during other confounding events in daily life (e.g., physical activity).

In [12], [13], the authors describe preliminary investigation of a proposed project called "iHeal'', that can potentially detect subject's craving of drugs in their natural environment from sensor data on electrodermal activity, body motion, skin temperature, and, optionally, heart rate. Results of this project, however, are yet to be reported.

To the best of our knowledge, [6] presents the first work on an ECG morphological feature based classifier for detecting cocaine use. As discussed in Section 1, the model presented in [6] for classifying each ECG cycle into cocaine use or baseline does not lead to a model for detecting Cocaine use events in the field setting that involves free-living activities.

Key Challenges: In conclusion, detection of cocaine use from physiological measurements collected in the field setting is challenging. We now summarize some of these technical and experimental challenges. First, physiological measurements such as ECG are subject to several sources of noises and quality issues, especially when worn in the mobile environment. They include incorrect placement and poor attachment of electrodes. Second, reliably collecting wearable ECG measurements from illicit drug users during active drug use in their natural environment is challenging due to the scenario and context of its usage. From our multi-year effort to collect ECG data in this population, we observe that in many situations the participant simply chooses not to wear the sensors during drug intake, even though they wear the sensors daily for 11+ hours. This may be due to safety concerns associated with wearing devices in crime-prone neighborhoods, where they may be suspected to be wired by the police. Third, it is very difficult to recruit participants who are active drug users; therefore, in most published scientific studies, the number of participants is usually in single digits. Fourth, obtaining ground truth, i.e., self-reports from participants as they take drug is also difficult, partly due to the same reasons as above. Though urine assessment can indicate drug intake over the previous few days, they do not provide the exact timing of the drug use event. Fifth, dosage amount and the method of administration (in case of cocaine — smoking, intravenous, intranasal or gastrointestinal) has significant effect on the HR response. However, we cannot obtain training data in a lab setting by administering drugs that represent very high dosage levels observed in the field (up to 600 mg). Sixth, as data is collected in unconstrained environments, there are usually confounding factors that can have similar physiological response. For cocaine use detection, the common confounding factors are activity, caffeine intake, and the intake of other drugs. Since we can't ask participants to walk during drug administration in lab, we can't collect lab data that represents drug use mixed with physical activity.

Yet, the model developed must be able to distinguish drug use events from physical activity even when they co-occur. The model we present is the first work to handle all of the above issues and detect drug use events that occur in the field setting.

III. Data Collection — Lab and Field

We designed and conducted an in-residence user study with 3 cocaine dependent volunteers at Johns Hopkins University Medical School (termed "JHU Lab Study"). We conducted a second in-residence study with 6 cocaine using volunteers at National Institute on Drug Abuse Intramural Research Program (NIDA IRP) (termed "NIDA Lab Study"). We also conducted a field study with 42 active poly-drug users (for 4 weeks of sensor wearing per user) at NIDA IRP (termed "NIDA Field Study") to collect sensor data in the free-living environment. All studies were conducted upon approval from the Institutional Review Board (IRB) of the respective institutions.

Sensor Suite: We used a wearable wireless sensor suite called AutoSense [2]. AutoSense uses a flexible band worn about the chest to capture respiration data via inductive plethysmography (called RIP). The chest sensor unit also contains two-lead electrocardiograph (ECG), 3-axis accelerometer, temperature sensors (ambient and skin), and galvanic skin response (GSR). The sampling rates for the chest band were 21.33 Hz for RIP, 64 Hz for ECG, 10.67 Hz for each of the three axes of accelerometers and GSR, and 1 Hz for the two temperature sensors and the battery level. These samples were transmitted wirelessly using ANT radio [23] to a Sony Ericsson Xperia X8 smart phone at the rate of 28 packets/second, each of which was 8 bytes long, containing 5 samples. The sensors last more that 10 days on a 750 mAh battery. We used the FieldStream mobile phone software [2] for logging the data. On a full charge, the phone battery lasts over 11 hours.

In the following, we describe the study protocols and the measurements obtained from the three studies.

A. In-Residence Study Protocols

In the "JHU Lab Study," three non-treatment seeking cocaine dependent volunteers (37-41 years old, 2 males) enrolled in a behavioral pharmacology residential study at Johns Hopkins University. They wore the AutoSense for at least 8 hours daily on the weeks when cocaine self-administration sessions were scheduled. During a safety session, participants self-administered intravenous cocaine doses of placebo (1 mg), 10 mg, 20 mg and 40 mg (every 30 minutes) via a Patient-Controlled Analgesia (PCA) pump. Safety data were collected only for two participants. During study weeks 1, 3 and 5, participants went through a series of cocaine self-administration sessions. Dose Response session on Monday consisted of three doses 45 minutes apart — placebo (1mg), 20 mg and 40 mg. On Tuesday, Wednesday and Thursday, a *sample* of cocaine dose was self-administered in the morning (double-blind randomized, out of placebo, 20 mg or 40 mg). After 2 hours, they were offered 7 *choices* for either the same morning drug dose or decreasing amounts of money in a sample/choice session. In summary, each participant went through 13 days of cocaine administration (1 safety session, and 3 study weeks of 4 days each). During rest of the awake hours of the day when wearing AutoSense, the participants self reported smoking and craving events as well as some events that may trigger craving (e.g., watching TV, watching movies, or playing video games).

In the ongoing "NIDA Lab Study," healthy cocaine users are admitted to a secure residential research unit at NIDA IRP. They undergo baseline assessments on Day -1, receive training on Day 0, and receive single doses of intravenous cocaine (25 mg) on Days 1, 5 and 10. On Days 1, 5, and 10, dried blood spot specimens are collected up to 3 times daily over 1.5 hours. Single oral doses of acetazolamide (15 mg) are given on Days 2-5 and quinine (80 mg) on Days 7-10. Blood, oral fluid, and breath specimens are collected for up to 71 hours, 70 hours, and 22 hours, respectively, after drug administration on Days 1, 4, 5, 9 and 10. Participants wear AutoSense on Days 1, 3, 4, 5, 8, 9 and 10 for up to 12 hours each day. Six participants have completed this lab study.

Data obtained from the in-residence studies are used to develop the model for detecting cocaine use since it consists of clean and carefully labeled activity-free cocaine use events, as well as various confounding events such as physical activity, games, smoking, etc., that are free from cocaine use.

B. Field Study Procedure

Methadone-maintained poly-drug (Cocaine, Heroin, etc.) users (different from those in the lab study) were recruited for NIDA field study. Participants wore AutoSense for four one-week periods, in their natural free-living environment. Participants were asked to self-report drug craving and use events in the field by pressing a button on a study smart phone (different from the one that collected ECG sensor data). Whenever they self-reported a drug-use event, they were asked to provide additional information on the timing of the drug use by choosing one of the options provided on the mobile device. The options were: (1) Less than 5 minutes ago, (2) 5-15 minutes ago, (3) 15-30 minutes ago, and (4) More than 30 minutes ago (in which case they were asked to input an estimate of the time of drug use). Urine samples were collected three times weekly (Monday, Wednesday, and Friday) during weeks when participants wore the sensors. Forty two participants have participated in the field study. Data collected in this study are used to understand the challenges in detecting drug use in the field environment and to validate our model for detecting cocaine use in the field setting.

C. Data Collected

In the "JHU Lab Study," the first participant wore the sensors only during weeks 3 and 5. Twelve days of data is collected from the first participant, 24 days from the second participant, and 22 days from the third participant, making for a total of 58 days. Of these 8, 13, and 13 days respectively are from cocaine administration sessions. An average of 9.55 hours of data was collected per day, for a total of 554 hours of good quality usable data.

In developing the model, we did not use the data corresponding to the 10 mg dose as we do not observe significant physiological response for this dosage level. From the remaining 36 sessions, sensor data is lost for 9 episodes (5 episodes from 20 mg and 4 episodes from 40 mg), due to ECG electrode detachment and/or displacement, leaving 27 episodes

(13 episodes from 20 mg and 14 episodes from 40 mg) for use in model development and evaluation.

From the 3 participants who went through the choice session, only 1 chose cocaine, providing 18 instances of active cocaine administration during the choice sessions. We do not use the choice session instances in modeling since cocaine injections were administered only 15 minutes apart, not allowing enough time for the physiology (i.e., heart rate) to recover before re-administration. Further, the effects of multiple doses accumulate, making it quite different from a single cocaine dose response.

In the ongoing "NIDA Lab Study," 31 days of data have thus far been collected, of which 14 days had cocaine sessions. Each cocaine day had only one cocaine session of 25 mg. A total of 280 hours (or, 9 hours/day) of good quality ECG data have been collected from these 31 days. Two of these participants are still in the protocol.

In the on-going "NIDA Field Study," we have thus far collected 922 person days of field data from 42 participants (4 participants did not complete). A total of 10,449 hours of ECG data has been collected (11.33 hours/day, on average). On those days, participants self-reported 211 instances of illicit drug (142 for cocaine) use events. Each self-report has the time of drug use, type of drug (e.g., Cocaine, Heroin, Methamphetamine, opiates, THC, Benzodiazepine), quantity, and how the drug was administered (e.g., smoking, sniffing/snorting, oral, or intravenous). Sometimes, the participants reported the use of multiple types of drug at the same time.

Among the 42 participants, 20 participants actually reported cocaine use. Among the 142 reported cocaine uses, 3 were for 50 mg, 86 for 100 mg, 1 for 150 mg, 34 for 200 mg, 8 for 300 mg, 9 for 400 mg and 1 for 600 mg. For modality of use, smoking was most popular (52 out of 142), followed by intravenous (48), sniffing (39), and oral (3).

We observe several issues with data collection in field. In several instances, data is unavailable due to not wearing the sensor or of not usable quality due to improper or loose contact of ECG electrodes, electrode detachment, loosening of electrical connectors, drying out of gel, noise from physical movement, etc. We adopt a method proposed in [24] for determining acceptability of ECG signals. In addition, a human expert visualized the signals and corrected any error in the labeling produced by the automated algorithm though they are very few in number. Since the participants did not report the exact time of the drug use, we measure the availability and quality of our sensor data for each person day to verify the usability of the collected data. Also, the self-report may not always be accurate. Hence, we visually inspect the ECG, and accelerometer signals together with the drug use self-reports to ascertain whether data in the vicinity of self-report indeed exhibits a drug use response.

Out of the 142 self-reports of cocaine use (from 103 person days), sensors were not worn on 17 days (for 22 reports). Of the remaining, 25 episodes are reported after the participant took off the sensor at night. Of the remaining 95, sensor was taken off during drug use and then put back on in 3 cases. Accelerometry sensor was not working in 6 cases, ECG sensor data was missing around the report in 6 cases, and ECG data quality was unacceptable in 53 cases. Hence, we are left with 27 instances with good quality ECG data. Among the 27 instances of cocaine uses, 14 were for 100 mg, 8 for 200 mg, and 2 for 300 mg and the remaining 3 instance were for 400 mg. These 27 instances come from 320 person days of good quality data, from 13 different participants (out of the 20 who reported cocaine use), where physiological response in the vicinity of self report of drug use can be observed. On these 320 days, we have a total of 3,631 hours of data.

Urine Reports: Urine samples were collected three times weekly during weeks when participants wore AutoSense. We classify each day using the result from the urine test via the following simple rule. If cocaine is detected from the urine sample collected, one can infer that in the last 4 days from that day, there must be at least one cocaine use. We label all these days as potential cocaine use day. On the other hand, if the urine report is negative, one can assume that the last 24 hours is a cocaine-free day. Also, if a participant has two or more negative urine reports in a row, with no self-report of cocaine use, we can safely infer that they didn't use cocaine at all during those days. However, if the participant self reports cocaine use on a particular day, we mark that day as a potential cocaine day, even if the urine report doesn't reflect it. Using these rules, we identified 385 potential cocaine use days out of the 922 days of data collection.

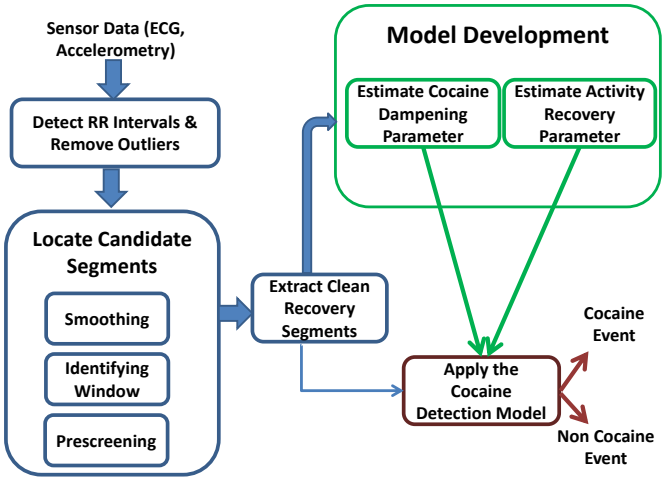

Fig. 1. Sensor data are first processed to find the RR interval timeseries and activity. Next, the start and end of cocaine and activity windows are identified and the recovery portions of the cocaine and activity responses are extracted. Two model parameters are learned during the traning phase and the tranined models are applied on the recovery response portion of the candidate windows extracted from the test data set.

IV. Data Processing and Modeling

In this section, we describe the data processing steps and model development. Figure 1 presents an overview of the data processing stages. First, we describe the processing of ECG data in order to obtain an RR interval time series. We also describe the activity detection method that we use to identify segments in the time series that may correspond to physical activity. We then describe the process we use to localize the drug and activity episodes in the RR interval time series. Next, we provide simple rules to screen out some of these segments or windows that obviously do not come from drug episodes.

Each of the windows contains two parts. First part corresponds to the excitation of the SNS system up to the point when heart rate reaches it's peak value. We call this the activation response. The second part corresponds to the portion of the window that represents the recovery of the physiology. We extract these two portions and construct models for the RR interval curve corresponding to recovery during cocaine-free activity episodes and activity-free cocaine episodes. Detection of cocaine use episodes makes use of these two models.

A. RR interval detection

We follow similar preprocessing and RR interval detection method as presented in [6]. We first adopt a method proposed in [24] for determining acceptability of ECG signals. We then apply the Tompkin's algorithm [25] to detect R-peaks. We remove outliers in the time series of RR intervals (time between two R-peaks) using the outlier removal algorithm presented in [26] and remove the DC offset from each interval to remove the baseline drift.

B. Activity detection

In order to determine the presence of physical activity from accelerometery, we use a simple threshold based activity detector using the 3-axis on-body accelerometer (placed on chest). Phone accelerometer data was not used because the phone may not be on the person and thus may miss some physical activity episodes. We make use of existing physical movement detection approach [27], [28], [29] and adapt it to fit our data. As the placement of the accelerometer and the participant population is different from that presented in prior works, we collected training data to determine an appropriate threshold for detecting activity.

We collected labeled data under walking and running (266 minutes), and stationary (183 minutes) states from seven participants while they wore AutoSense. After noise and drift removal, we extract the standard deviation of magnitude, from 10 second windows, which is independent of the orientation of the accelerometers. We scale this signal using the 99^{th} and 1^{st} percentile values of the signal using $m = (m-h)/(max-l)$, where m, h and l are the samples' 99^{th} percentile and 1^{st} percentile values. We find that a threshold of 0.35 is able to distinguish stationary from non-stationary states with an accuracy of 93% in 10-fold cross-validation.

C. Model Development

To develop a model, we first develop a method to smooth the RR interval time series, identify the windows that exhibit sufficient change so as to result from a potential drug use event, locate the start and end points of such windows, and then extract portions of this window from which the parameters of the parasympathetic nervous system (PNS) and the parameters for dampening introduced by Cocaine metabolism induced sympathetic nervous system (SNS) activation can be estimated. These steps are described in Section IV-C1. Section IV-C2 presents the rationale and the actual model for detecting drug use events in the selected candidate windows of RR intervals.

1) Candidate Window Selection & Preparation: Figure 2 shows examples of cocaine responses observed during lab administration of cocaine. We observe that the effect of cocaine can last a long time (more than 30 minutes) and the entire response window (which can vary in length as well as intensity of response) must be detected as one single event. Hence, the problem of detecting drug use is a time series pattern recognition problem. The challenge then is to locate candidate windows and identify its boundaries so as to assess it for drug use response. This challenge is compounded by dynamic variations in RR intervals and difficulty in locating the start of cocaine or activity response and the end of recovery.

This problem has similarity with the dynamic fluctuations in stock prices. Therefore, in order to identify the potential windows and to identify the start and end of the window that may indicate a cocaine use event, we use the Moving Average Convergence Divergence method (MACD) [30]. This method is widely used to compute an indicator that investors in stock markets use to identify the stock price rise and fall trends. The indicator is also used as an oscillator indicator that is used to identify when the market moves sideways, i.e., when the price oscillates within a narrow range. MACD thus is more stable than (price) trend following indicators. We make use of this MACD procedure to identify the windows that correspond to 'large' rise and fall trends of the RR intervals. MACD readily provides the start and end of the windows. The MACD procedure makes use of Exponential Moving Averages (EMA), which gives more weight (a constant in this case) to recent values compared to simple moving averages. MACD is computed as follows:

$$\text{MACD Line} = EMA_{slow} - EMA_{fast} \quad (1)$$

$$\text{Signal Line} = \text{EMA of MACD Line} \quad (2)$$

Each of the EMA's takes one parameter — the window size on which the EMA is to be computed. EMA_{slow} is computed on a longer window than that of EMA_{fast}. The crossover points of the MACD line and the Signal line corresponds to the fall and rise points. We learn the parameter (i.e., window size) from the data we collect in the "JHU Lab Study." We mark the start and end of the windows for cocaine data by visual inspection. For activity, we select start and end of the window with the help of accelerometry and visual inspection. In total, we mark 27 cocaine windows and 272 activity windows. The search space for the parameters of the slow and fast moving averages are set to [1, 180] minutes and [2, 90] minutes respectively. The search space for the parameter of EMA_{slow} is same as that of EMA_{fast}. We obtain the parameters that achieve the minimum error in finding the start and end points. The search algorithm chooses crossover points that are nearest to the start and end points marked for each set of parameter values. The error in this decision is computed as the sum of the distance from the crossover points chosen by the algorithm and the start and end points marked via visual inspection.

The input to the MACD process is smoothed RR interval time series that removes high-frequency variations in RR intervals. We use a simple moving average over preceding 10 minutes to smooth the signal. We also learn the optimal window size of this moving average window in this

process. Figure 2 shows the MACD, signal line as well as the crossover points selected in this process. The parameters learned for EMA_{slow}, EMA_{fast}, and EMA for Signal Line are 35 minutes, 4 minutes and 3.67 minutes respectively. These parameters are used to find the windows on the field data too.

Screening: Next, we use the following two rules to screen out some of the candidate windows before we apply our model. First, we compute average width and height of the windows from the lab data. If the candidate window c is too 'small' compared to the drug windows (i.e., if width or height $<$ (mean $-$ 3 $*$ standard deviation) of that of width and height of windows corresponding to drug response), it is discarded. In this case, the physiological response is insufficient to be that from drug use. Second, out of all the 10 second accelerometer measurement windows within the first 5 minutes from the start of a candidate window c, if a majority of them are detected as activity we discard the window c. In this case, the activation of heart rate is the result of physical activity.

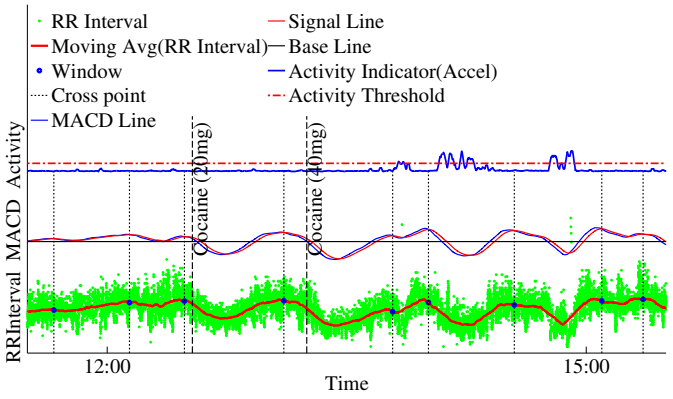

Fig. 2. Illustration of MACD based candidate window selection method. The crossover points, where the MACD and signal lines cross (marked using dotted horizontal lines), indicate the start and end points of the candidate windows.

Extracting recovery episodes: We extract both the recovery and activation part of the windows using the crossover points from the MACD process. For a particular window, the start and end of activation (A_{start} and A_{end} respectively) must occur after the start point of the window. They are defined as the last crossover point above the zero line and first crossover point below the zero line. On the other hand, the start and end of the recovery portion (R_{start} and R_{end} respectively) must occur after the activation part. They are defined as the last crossover point below zero line and first crossover above zero line. Finally, to model the recovery process, we extract the first subsequence of the recovery curve (defined by R_{start} and R_{end}) that is clean, i.e., not affected by activity.

2) Autonomous Nervous System (ANS) Model: To develop a model to detect drug use events, we consider an abstract model describing the interaction of the ANS with the cardiovascular system during physical activity and drug intake. ANS is a control system that regulates heart and respiration rate, cardiac output and constriction of dilation of blood vessels to meet the demands imposed on the body. The ANS is divided into two separate systems — PNS and SNS acting in concert. Roughly speaking, PNS is responsible in conservation of energy by reducing heart, respiration rate and

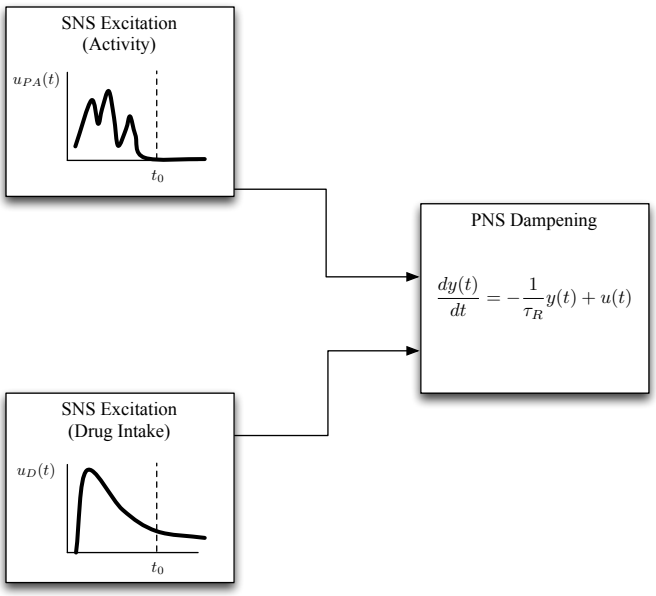

Fig. 3. ANS model of heart rate recovery.

blood pressure providing a dampening effect. In contrast, SNS provides activation by increasing heart rate, blood pressure and cardiac output to meet the demands of physical activity, fear, stress, etc. Generally speaking, SNS works at a faster time scale than PNS and can be excited by many inputs, not all directly observable. In this paper, we consider modeling heart rate recovery regulated by PNS with and without drug usage events. We learn a simplified abstract model with few parameters driven by physiology from the lab data to build a statistically optimal detector. The concise model avoids overfitting and provides parameters readily interpretable as time constants and dosage.

We would also like to note that the activation data portion of the response is not used in our modeling. There are several difficulties associated with using this portion of the data. First, during activation there is the combined effect of Vagal withdrawal and SNS activation as well as the metabolism of cocaine. Therefore, in order to model the activation curve, we need to estimate the parameters associated with Vagal withdrawal and SNS activation. However, the activation portion is quite short, not offering sufficient data for robust modeling.

We focus, therefore, on modeling the RR interval time series $r(t)$ (inverse of the instantaneous heart rate) during recovery periods. We assume that the RR interval time series $r(t)$ can be modeled as noisy samples of a signal $y(t)$ measured from the resting baseline RR interval B.

$$r(t) = B - y(t) + n(t) \tag{3}$$

We assume that the PNS can be modeled by a first-order differential equation:

$$\frac{dy(t)}{dt} = -\frac{1}{\tau_R}y(t) + u(t), \tag{4}$$

where $u(t)$ models the excitatory input from the SNS. In the absence of any excitatory input $u(t) = 0$. In such a case, the solution corresponds to an exponential decay model, with a

recovery time constant of τ_R. We model the natural recovery process as

$$y_R(t) = y(t_0)e^{-\frac{(t-t_0)}{\tau_R}}. \quad (5)$$

Such exponential decay models have been previously suggested as models of heart rate recovery from physical exercise [31]. Here, we consider two sources of excitatory input — physical activity and drug intake. Both can elevate the heart rate. We focus on recovery portions where the excitatory input due to activity is absent, so that we need to model only one source of activation. We note that although the activity is missing during our chosen recovery periods, the initial heart rate elevation could be due to both physical activity and drug intake, complicating the detection of drug usage and estimation of dosage. We model the excitation due to drug intake as an exponentially decaying process since its intensity decays due to cocaine metabolism[1] with time-constant τ_D:

$$u_D(t) = u(t_0)e^{-\frac{(t-t_0)}{\tau_D}}.$$

Our ANS model of recovery in the absence of physical activity is given in Figure 3. Inserting this input into the differential equation governing dynamics of the PNS, results in the following model of recovery curve under drug intake:

$$y_{DR}(t) = y(t_0)e^{-\frac{(t-t_0)}{\tau_R}} - \frac{u(t_0)}{K}e^{-\frac{(t-t_0)}{\tau_R}} + \frac{u(t_0)}{K}e^{-\frac{(t-t_0)}{\tau_D}}, \quad (6)$$

where the constant K is given by $K = (1/\tau_R - 1/\tau_D)$. This represents the rate of recovery by the PNS system when cocaine is still in the bloodstream (i.e., not fully metabolized).

We note that although we build upon established model for modeling the recovery behavior of the parasympathetic nervous system (PNS) from activity (i.e., in the absence of cocaine), but our proposed model introduces two new components to account for the dampening of PNS recovery due to cocaine stimulus and to tease it out from the effect of activity. Prior models of recovery from cocaine (e.g., [22]) consider only the composite recovery of heart rate for cocaine events and hence can't separate out the effects of physical activity. Our proposed ANS model is the first model that considers the dampening effect of cocaine and the decay of dampening due to cocaine metabolism by the body in PNS recovery, and accounts for the recovery from physical activity alone. The first component on the right hand side of (6) represents the natural recovery of the heart rate, while the 2nd and 3rd components represent the cocaine dampening and the decay of dampening due to metabolism of cocaine, respectively.

The detection of drug intake events can now be posed as a binary hypothesis testing problem of identifying the underlying curve from the two models in (5) and (6) from noise corrupted recovery segments of RR interval time series. We assume the noise on the time samples of $y(t)$ is identical and independently distributed with Gaussian distribution. The nuisance parameters of intercept $y(t_0)$ and initial drug intensity $u(t_0)$ are unknown. Therefore, we follow a generalized likelihood ratio test approach where maximum likelihood estimate of these parameters are used. The detector for drug intake events from RR interval samples $r(t)$ take the form of

$$\frac{\|B - r(t) - y_{RD}(t; \hat{y}(t_0), \hat{u}(t_0))\|^2}{\|B - r(t) - y_R(t; \hat{y}(t_0))\|^2} < \theta$$

where the nuisance parameters are found using least-square fit to the segment under test. The threshold θ has to be chosen to set an appropriate balance between probability of detection and probability of false alarm. For the detected drug episodes the parameter estimate $\hat{u}(t_0)$ is the intensity of remaining drug excitation at the beginning of the recovery segment. If reliable estimates of the drug intake time can be formed, this could be used to extrapolate $u(t)$ to drug-intake time to find an estimate of the drug dosage, which we leave for future work.

The ANS model parameters consist of resting baseline B, recovery time constant τ_R and drug metabolism time constant τ_D. In general, these parameters have to be learned for each individual separately. But, obtaining ground truth data on drug events for each individual is impractical. Therefore, we follow the following unsupervised procedure to construct semi-individual models of ANS activity. For each participant j, we extract segments for recovery from physical activity occurring naturally in their daily life and fit model in (5) and learn the PNS model parameters of the resting baseline B^j and recovery time constant τ_R^j for each subject.

For time-constant determination, the exponential moving averages (EMA) from section IV-C1 cannot be employed for the RR interval signal $y(t)$, since the long time averaging windows used in EMA's will cause an upward bias in the estimate of the recovery time constants. On the other hand, raw RR interval data collected in the field contains significant amount of outliers. These outliers are not well modeled with Gaussian statistics, since they have large amplitudes with a proportion larger than indicated by the tails of the normal distribution. As a result, least square estimates of τ_R and τ_D will be highly sensitive to outliers present in the unsmoothed RR interval data. To have robust estimates of these time constants we follow the robust regression framework introduce by Huber [32]. Specifically, we consider Huber's modified cost-function $\rho(e(t))$,

$$\rho(e) = \begin{cases} 0.5e^2 & \text{for } |e| \le k \\ k|e| - 0.5k^2 & \text{for } |e| > k \end{cases}, \quad (7)$$

where $e(t)$ is the error signal defined as the difference between the data $y(t)$ and the models in (5) and (6). The constant $k = 1.345\sigma$, with σ is the expected variance of RR interval data. The modified cost function eliminates large bias introduced by non-Gaussian outliers present in RR interval data. Then, for the nine participants with lab data on drug usage we fit the drug recovery model in conjunction with their PNS model of (B^j, τ_R^j) to get an average estimate of drug metabolism constant τ_D minutes (for each dose level) that we use uniformly for every participant in our testing.

V. Evaluation

We first present the performance of the ECG cycle based detector presented in [6] on our in-residence lab data. Next, we present estimates of parameters for our ANS model. To provide a feel for our model, we provide an illustration of our

[1]Drug metabolism is usually modeled as an exponentially decaying function since the rate of metabolism is proportional to its remaining concentration.

model by showing their fitting with data from a participant. Finally, we show the performance of our ANS model on the data collected in the in-residence studies and the field study.

A. Performance of Cycle Classifier

To assess the applicability of a model developed in [6] for detecting cocaine usage in free-living conditions, we implemented the classifier in [6] and applied it to the data collected from our in-residence study. As mentioned in Section I, this model classifies each ECG cycle into cocaine and and baseline classes. The in-residence study, in our case, did not include a baseline session, therefore the training data for non-cocaine class consists of data collected from 30 minutes to 2 minutes prior to the first cocaine administration of the day, since subjects are resting for this period. Data only for durations when there was no physical activity were labeled as non-cocaine. This is determined from the accelerometer data. If activity is detected for a particular activity detection window, we discard data up to 5 minutes after that window to make sure the data does not contain any activity. As for the cocaine class, data from the interval (valley -2 minutes) to (valley $+8$ minutes) from each cocaine response window are included. The cocaine response windows are identified by the Moving Average Convergence Divergence (MACD) method described in Section IV-C1. The valley of the window is defined to be the point when the RR interval reaches its minimum value (i.e., where HR peaks). Again, we only admit data that is not effected by activity. Total data cases for cocaine and non-cocaine class for participant #1 amounts to 2,080 cycles each. A support vector machine (SVM) classifier is trained and we obtain AUC = 0.91 for 10-fold cross-validation, which is similar to that reported in [6]. We then apply this classifier on the whole data set for that participant. To show the applicability of this model on non-baseline data we show the output of the classifier in Figure 4.

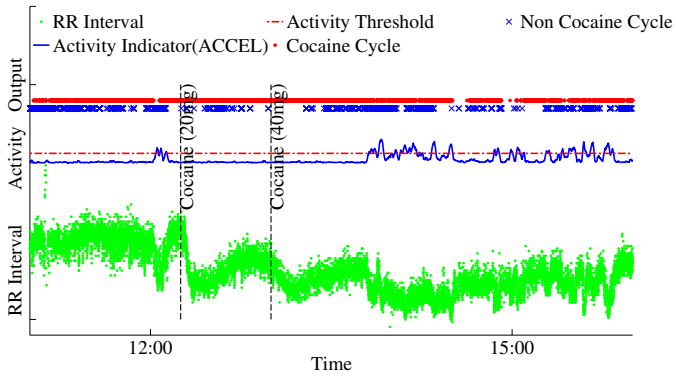

Fig. 4. The red dots on the top indicate all the cycles marked by the ECG cycle based classifier as cocaine class. The model produces too many false positive outputs rendering it unsuitable for the use in field setting.

We observe that although the classifier performs well for the 20 mg case, it's performance for both 40 mg and placebo (1 mg) cases are not as good. Also, we find that the classifier produces a large number of false positives when there is physical activity since it is not trained for this class. This indicates that not only RR intervals but also other features (hypothesized in [6] to be insensitive to physical activity) are confounded by activity. It remains open whether the model proposed in [6] can be improved by training it with free-living activity data. Another approach to improving it may be to combine this data-driven approach of classifying each ECG cycle with our model-based approach.

B. Evaluation of the ANS Model

We first estimate parameters of our ANS model and then evaluate its performance on the in-residence and field data.

Parameter Estimates: Our ANS model needs two parameters to be estimated — τ_R and τ_D. The $1/\tau_R$ and $1/\tau_D$ values represent the half-time (in minutes) to recover from cocaine-free physical activity and for cocaine to metabolize, respectively. τ_D is estimated from activity-free cocaine episodes, while τ_R is estimated from cocaine-free activity episodes. Table II presents the data obtained from the in-residence and the field studies. From both in-residence studies, we obtain a total of 41 cocaine episodes from 40 days of data. A total of 497 non-cocaine significant physical activity episodes are found on 49 non-cocaine days.

Since there are several episodes of cocaine-free recovery from activity for each participant, we obtain a person specific estimate of τ_R for each participant (in both lab and field) and use it in testing. Table III shows estimates for τ_R, which are also used in clinical settings for assessing cardiovascular fitness [14], [15]. We observe that it takes approximately 3 minutes on lab data (and 4 minutes on field data, due to higher intensity of activity episodes) for half-recovery of heart rate in the absence of cocaine.

To estimate τ_D, we need cocaine recovery curves that are free from activity. Since we have activity-free recovery from cocaine for only in-residence participants ($n = 9$), we obtain person-specific estimates for τ_D only for these participants. Table II shows estimates for τ_D for various dosage amounts. We observe that the half-time to cocaine metabolism (i.e., $1/\tau_D$) increases with increase in dosage. The median half-time for cocaine metabolism is 43 minutes.

When fitting the model to recovery curves of RR intervals, we use person-specific estimates of τ_R, but a person-independent estimate (i.e., median over 9 lab participants) of τ_D. Figures 5 and 6 show examples of fitting of the models to recovery of RR intervals from an activity episode and from a cocaine use episode. Figure 7 shows an example of fitting an activity recovery model and a cocaine recovery model onto an RR recovery curve from a cocaine use episode. We observe that the recovery rates of the two curves are quite distinguishable, with cocaine recovery model providing a better fit.

Performance on In-residence Data: Although we obtain person-specific estimates for τ_R, we need a person-independent estimate of τ_D for use on the field data. To determine the best estimate of τ_D, we obtain dose-specific estimates of τ_D (for 20 mg, 25 mg, and 40 mg) and dose-independent estimate of τ_D (by using all dosage). To determine their suitability, we test their performance on lab data. We test their performance on the dose-specific in-residence dataset (In-residence-Dose) (i.e., on all the cocaine of a particular dose and all non-cocaine episodes of the in-residence studies). This provides an indication of their expected performance if all ocaine episodes are of the same

amount. We also test their performance on all lab data (In-residence-All) to assess their invariance to dosage level.

We compute the number of false positives (i.e., the number of non-cocaine episodes identified as cocaine episodes) per day, without misidentifying any cocaine episodes (i.e., true positive rate of 100%). Results appear in Table IV. We observe that the 25 mg model provides the lowest false positive rate on the In-residence-Dose dataset. But, for the In-residence-All dataset, the lowest false positive/day occurs for τ_D obtained from the 40 mg cocaine data. We, therefore, use this time constant in all of our testing. The ROC for the 40 mg model with the data of all participants combined is presented in Figure 8 with a representative operating point shown.

TABLE II. DATA STATISTICS FOR IN-RESIDENCE AND FIELD STUDIES.

Study	Non-cocaine		Cocaine	
	# windows	# days	# windows	# days
In-residence Studies	497	49	41	40
Field Study	740	64	27	25

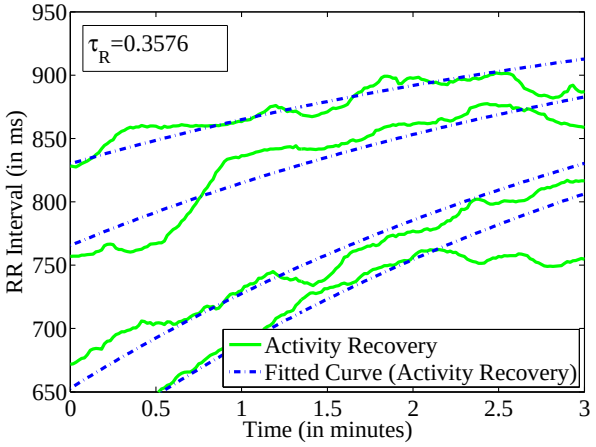

Fig. 5. Curve fitting for activity recovery. The four solid green curves represent four raw RR interval timeserieses of activity recovery of one participant. Fitting the recovery model (dotted blue curves) we find the estimate of τ_R. Based on this τ_R value we observe that it takes approximately 2.78 minutes for the heart rate to recover to half of the resting baseline for this participant.

Performance on Field data: As mentioned in Section III, there are 27 cocaine instances from 13 participants for which we have good quality data. These 27 instances come from 25 different days on which the participant both self reported cocaine use and are confirmed with positive urine reports. We also have 740 episodes of non-cocaine windows of significant physical activity that come from 64 confirmed non-cocaine days. We obtain person-specific estimates of τ_R. Table IV presents the results of applying the dose-specific and dose-independent models on the field data. An ROC curve for the 40 mg model is presented in Figure 8 with a representative operating point shown for the field.

Figure 9 shows the application of the model on a sample of field data with cocaine use event reported. Our candidate window selection and screening method identifies several potential cocaine response windows. (These are representatives of the 740 non-cocaine windows.) In Figure 9, these windows are represented by dotted vertical lines marking the start and

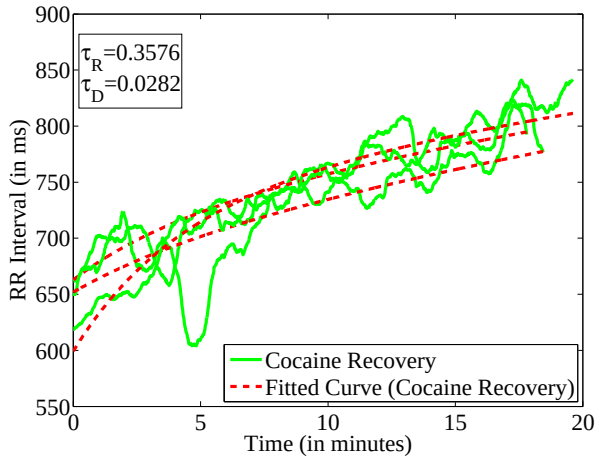

Fig. 6. Curve fitting for cocaine recovery. Using the activity recovery model we find the estimate of τ_R from the activity recovery data of this participant. τ_D is estimated from this participant's cocaine response data (solid green curves) by fitting the cocaine recovery model that uses the estimated value of $\tau_R = 0.3576$. In this case the estimated value for τ_D is 0.0282 which gives the half life of cocaine in the blood to be 35.46 minutes.

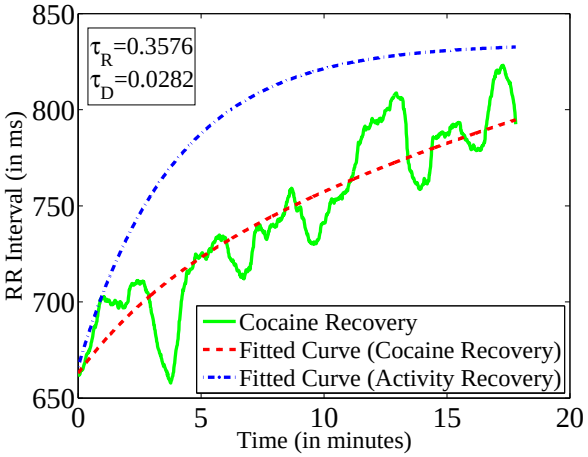

Fig. 7. Fitting of an example cocaine recovery curve to the models. The dotted blue and red curves represent the fitted activity recovery and cocaine recovery models respectively. It is clearly observed that the cocaine recovery model [represented by Equation-6] that takes into account both τ_R and τ_D performs better.

end times of each window. The output of the model for each of these windows is presented using blue or red markers over them. We find that the model detects two cocaine windows. The first window from the left marked with a red dot is the actual cocaine event (identified by the self report). The model identifies one more window as a cocaine response window. It can very well be the case that there was a second drug intake episode that the participant failed to report; there are several instances of repeated drug use episodes in the field.

More generally, from visual inspection of false positive cases, we observe that sometimes the candidate window selection method fails to identify the start of a recovery segment correctly, leading to a false positive. Hence, further reduction in false positives can be achieved by improving the window marking method.

TABLE III. MODEL PARAMETERS ($1/\tau_R$ AND $1/\tau_D$) IN MINUTES FOR DIFFERENT COCAINE DOSE AMOUNTS. (L) AND (F) REFERS TO THE $1/\tau_R$ VALUES COMPUTED FOR THE IN-RESIDENCE AND FIELD STUDIES RESPECTIVELY. $1/\tau_D$ VALUES FOR 20MG, 25MG, 40MG, AND COMBINED MODELS ARE PRESENTED IN THE COLUMNS LABELED $1/\tau_D(20)$, $1/\tau_D(25)$, $1/\tau_D(40)$ AND $1/\tau_D(C)$ RESPECTIVELY.

Measures	$1/\tau_R(L)$	$1/\tau_R(F)$	$1/\tau_D(20)$	$1/\tau_D(25)$	$1/\tau_D(40)$	$1/\tau_D(C)$
Range	2.41-6.75	2.00-9.82	27.55-75.75	35.09-81.97	40.16-61.35	35.09-81.97
Median	3.18	4.06	30.30	45.23	51.02	43.29

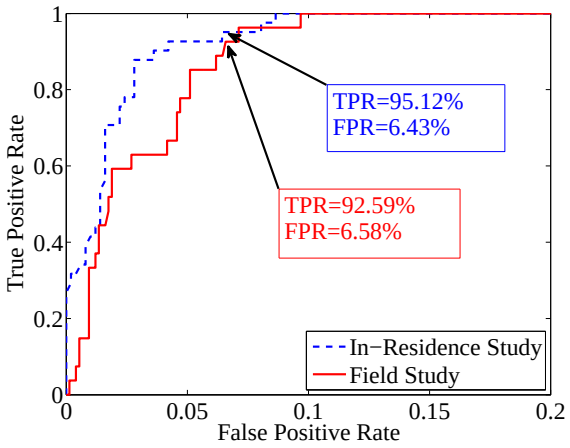

Fig. 8. ROC for detection of drug on the in-residence and field studies using the 40 mg model. The two points on the ROC curves presents two suitable operating points that have similar performances. If we allow for misclassification of drug events (TPR 95%) we can achieve false alarms rate of $< 7\%$.

TABLE IV. PERFORMANCE OF THE DETECTOR ON IN-RESIDENCE AND FIELD STUDY DATA FOR VARIOUS ESTIMATES OF τ_D. WE REPORT (AVG) NUMBER OF FALSE POSITIVE/DAY FOR A TRUE POSITIVE RATE OF 100%. L-ALL IS FOR THE DOSE-INDEPENDENT ESTIMATE OF τ_D.

Dataset	L-20	L-25	L-40	L-All
In-residence-All	0.91	0.91	0.87	0.91
In-residence-Dose	0.91	0.22	0.65	-
Field	1.13	1.14	1.13	1.14

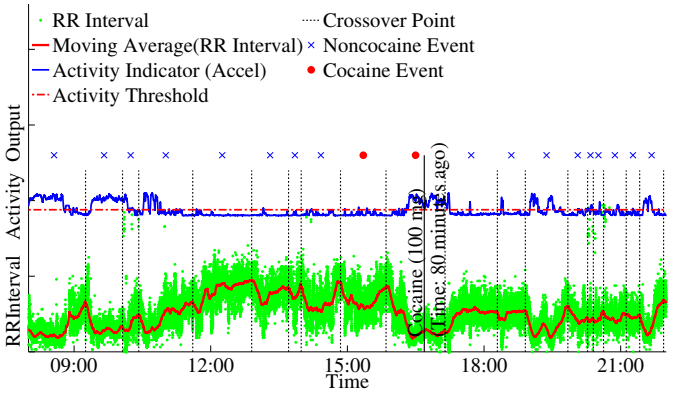

Fig. 9. Output of the model on the field data. Here we find that there is one self-report of cocaine use and the participant self reports that he/she had cocaine 80 minutes prior to the self-reporting time (marked by the vertical line). The candidate windows are marked by the dotted lines. The output of the model is presented using small red dots and blue crosses over each window representing cocaine and activity respectively. We see the model makes one error - it wrongly detects one non-cocaine window as a cocaine window.

VI. CONCLUSIONS AND FUTURE WORKS

Mobile health can help people improve their health by continuous monitoring of health and behavior on their mobile phone and by delivering timely interventions to motivate healthy lifestyle and abstinence from risky behaviors. Automated detection of various health states and behaviors from sensor data is key to realizing the mobile health vision. Such capabilities can then be used in scientific user studies to identify precipitating contexts that precede undesirable health states and risky behaviors. Automated detection of these predictors can be used to trigger the delivery of timely intervention on the mobile phone as preventative measures. Our work contributes to realization of this mobile health vision by showing that automated detection of drug use is feasible, opening the doors for development of just-in-time interventions.

The ANS model for detecting drug use can itself be improved for both accuracy and generality. First, we use a person-independent estimate of drug metabolism rate. This can be made more person-specific by incorporating demographic information such age, gender, weight, body mass index, etc. that are known to affect the metabolic rate. Second, the model can be improved further by using measurements from other physiological sensors such as respiration, galvanic skin response, electrodermal activity, etc. Third, the model can be expanded to include other illicit drugs. Psychostimulant drugs that may have similar response as Cocaine include Amphetamine and Methamphetamine, Methylphenidate (Ritalin), Methcathinone (an emerging drug under the name of Bath Salt), and MDMA (Ecstasy).

In addition to stimulating further research on mobile health, this work motivates three new sensor data processing issues for future research.

- Generalizable approaches are needed for detecting events of interest from physiological time series data in the presence of numerous confounders encountered in the natural environment. Our work shows the feasibility of doing so for cocaine use that has a pronounced impact on physiology. But, new research is needed to develop explainable (i.e., non black box) methods for other events (e.g., stress, conversation, smoking, eating, etc.) that may not have as pronounced of an impact on physiology.
- Generalizable methods are needed to smooth noisy physiological time series data to see broad trends and to mark the boundary of events of interest in the time series. Our work illustrates an approach to identify the windows of interest in ECG time series for drug use response identification. But, further research is needed to investigate its applicability to other sensing modalities such as respiration, EEG, etc.
- We provide a generative model that tracks the recovery

portion of the heart rate response, when physical activity concludes. Such models are succinct (need estimation of only one parameter) and can be used to generate synthetic data for simulation. Again, more generalized generative models are needed that can model the activation portion of the response curve as well so that the entire response curve can be synthetically generated. In addition to generating synthetic data such models can help build more robust detectors by modeling the entire response curves.

ACKNOWLEDGMENT

We thank Daniel Agage, Michelle Jobs, and Matthew Tyburski at National Institute on Drug Abuse (NIDA), Porche Henry and Tiffany Duren from Johns Hopkins Medical School for their tremendous effort with data collection. We also thank Deepak Ganesan, Benjamin Marlin, Annamalai Natarajan, and Abhinav Parate from University of Massachusetts, Amherst for sharing their Matlab code for cleaning up ECG data and for identifying ECG cycles affected by cocaine response. This work was supported in part by the Intramural Research Program (IRP) of NIDA at National Institutes of Health, by NSF grants CNS-0910878 (funded under the American Recovery and Reinvestment Act of 2009 (Public Law 111-5)), CNS-1212901, CNS-1017701, CCF-1215302, IIS-1343896, IIS-1231754, and by NIH Grants U01DA023812, R01DA027065, and R01DA035502 from NIDA.

REFERENCES

[1] S. Kumar, W. Nilson, M. Pavel, and M. Srivastava, "Mobile health: Revolutionizing healthcare through trans-disciplinary research," *IEEE Computer*, vol. 46, no. 1, pp. 28–35, 2013.

[2] 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 SenSys*, 2011, pp. 274–287.

[3] "Zephyr Bioharness," http://www.zephyr-technology.com/bioharness-bt, Accessed: February 2014.

[4] 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 IPSN*, 2011, pp. 97–108.

[5] A. Ali, S. Hossain, K. Hovsepian, M. Rahman, K. Plarre, and S. Kumar, "mpuff: automated detection of cigarette smoking puffs from respiration measurements," in *ACM IPSN*, 2012, pp. 269–280.

[6] A. Natarajan, A. Parate, E. Gaiser, G. Angarita, R. Malison, B. Marlin, and D. Ganesan, "Detecting cocaine use with wearable electrocardiogram sensors," in *ACM Ubicomp*, 2013, pp. 123–132.

[7] C. Furr-Holden, K. Campbell, A. Milam, M. Smart, N. Ialongo, and P. Leaf, "Metric properties of the neighborhood inventory for environmental typology (nifety): An environmental assessment tool for measuring indicators of violence, alcohol, tobacco, and other drug exposures," *Evaluation review*, vol. 34, no. 3, pp. 159–184, 2010.

[8] L. Degenhardt and W. Hall, "Extent of illicit drug use and dependence, and their contribution to the global burden of disease," *The Lancet*, vol. 379, no. 9810, pp. 55–70, 2012.

[9] M. Fischman, C. Schuster, L. Resnekov, J. Shick, N. Krasnegor, W. Fennell, and D. Freedman, "Cardiovascular and subjective effects of intravenous cocaine administration in humans," *Archives of General Psychiatry*, vol. 33, no. 8, pp. 983–989, 1976.

[10] C. Muntaner, K. Kumor, C. Nagoshi, and J. Jaffe, "Intravenous cocaine infusions in humans: Dose responsivity and correlations of cardiovascular vs. subjective effects," *Pharmacology Biochemistry and Behavior*, vol. 34, no. 4, pp. 697–703, 1989.

[11] K. Johnson, A. Isham, D. Shah, and D. Gustafson, "Potential roles for new communication technologies in treatment of addiction," *Current psychiatry reports*, vol. 13, no. 5, pp. 390–397, 2011.

[12] E. Boyer, R. Fletcher, R. Fay, D. Smelson, D. Ziedonis, and R. Picard, "Preliminary efforts directed toward the detection of craving of illicit substances: the iheal project," *Journal of Medical Toxicology*, vol. 8, no. 1, pp. 1–5, 2012.

[13] E. Boyer, D. Smelson, R. Fletcher, D. Ziedonis, and R. Picard, "Wireless technologies, ubiquitous computing and mobile health: Application to drug abuse treatment and compliance with hiv therapies," *Journal of Medical Toxicology*, vol. 6, no. 2, pp. 212–216, 2010.

[14] C. Cole, E. Blackstone, F. Pashkow, C. Snader, and M. Lauer, "Heart-rate recovery immediately after exercise as a predictor of mortality," *New England Journal of Medicine*, vol. 341, no. 18, pp. 1351–1357, 1999.

[15] L. Cahalin, R. Arena, V. Labate, F. Bandera, C. Lavie, and M. Guazzi, "Heart rate recovery after the 6 min walk test rather than distance ambulated is a powerful prognostic indicator in heart failure with reduced and preserved ejection fraction: a comparison with cardiopulmonary exercise testing," *European journal of heart failure*, vol. 15, no. 5, pp. 519–527, 2013.

[16] R. Kloner, S. Hale, K. Alker, and S. Rezkalla, "The effects of acute and chronic cocaine use on the heart," *Circulation*, vol. 85, no. 2, pp. 407–419, 1992.

[17] S. Reed, M. Haney, S. Evans, N. Vadhan, E. Rubin, and R. Foltin, "Cardiovascular and subjective effects of repeated smoked cocaine administration in experienced cocaine users," *Drug and alcohol dependence*, vol. 102, no. 1, pp. 102–107, 2009.

[18] E. Cone, "Pharmacokinetics and pharmacodynamics of cocaine," *Journal of Analytical Toxicology*, vol. 19, no. 6, pp. 459–478, 1995.

[19] R. Foltin, A. Ward, M. Haney, C. Hart, and E. Collins, "The effects of escalating doses of smoked cocaine in humans," *Drug and alcohol dependence*, vol. 70, no. 2, pp. 149–157, 2003.

[20] M. Egred and G. Davis, "Cocaine and the heart," *Postgraduate medical journal*, vol. 81, no. 959, pp. 568–571, 2005.

[21] M. Perez-Reyes, S. Di Guiseppi, G. Ondrusek, A. Jeffcoat, and C. Cook, "Free-base cocaine smoking," *Clinical Pharmacology & Therapeutics*, vol. 32, no. 4, pp. 459–465, 1982.

[22] D. Noe and K. Kumor, "A pharmacokinetic–pharmacodynamic model of heart rate during cocaine administration in humans," *Clinical Pharmacology & Therapeutics*, vol. 49, no. 4, pp. 426–432, 1991.

[23] "ANT Radio," http://www.thisisant.com/, Accessed: February 2014.

[24] K. Plarre, A. Raij, S. Guha, M. al'Absi, E. Ertin, and S. Kumar, "Automated detection of sensor detachments for physiological sensing in the wild," in *Wireless Health 2010*. ACM, 2010, pp. 216–217.

[25] J. Pan and W. Tompkins, "A real-time qrs detection algorithm," *Biomedical Engineering, IEEE Transactions on*, vol. 32, no. 3, pp. 230–236, 1985.

[26] G. Berntson, K. Quigley, J. Jang, and S. Boysen, "An approach to artifact identification: application to heart period data," *Psychophysiology*, vol. 27, no. 5, pp. 586–598, 1990.

[27] S. Reddy, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, "Determining transportation mode on mobile phones," in *Wearable Computers, 2008. ISWC 2008. 12th IEEE International Symposium on*, 2008, pp. 25–28.

[28] L. Atallah, B. Lo, R. King, and G. Yang, "Sensor placement for activity detection using wearable accelerometers," in *Body Sensor Networks (BSN)*, 2010, pp. 24–29.

[29] J. Kwapisz, G. Weiss, and S. Moore, "Activity recognition using cell phone accelerometers," *ACM KDD*, vol. 12, no. 2, pp. 74–82, 2011.

[30] J. Murphy, *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. New York Institute of Finance, 1999.

[31] J. Freeman, F. Dewey, D. Hadley, J. Myers, and V. Froelicher, "Autonomic nervous system interaction with the cardiovascular system during exercise," *Progress in cardiovascular diseases*, vol. 48, no. 5, pp. 342–362, 2006.

[32] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.