Detection and Analysis of Interrupted Behaviors by Public Policy Interventions during COVID-19

Guimin Dong
University of Virginia

Lihua Cai
University of Virginia

Shashwat Kumar
University of Virginia

Debajyoti Datta
University of Virginia

Detection and Analysis of Interrupted Behaviors by Public Policy Interventions during COVID-19

Guimin Dong, Lihua Cai, Shashwat Kumar, Debajyoti Datta, Laura E. Barnes and Mehdi Boukhechba
Engineering Systems and Environment
University of Virginia
151 Engineers Way, Charlottesville, VA, 22901, USA

***Abstract*—In most countries around the world, various public policies and guidelines, such as social distancing and stay-at-home orders, have been put in place to slow down the spreading of COVID-19. Relying on traditional surveys to assess policy impacts on community level behavior changes may lead to biased results, and limit fine-grained understanding of human behavior dynamics over time. We propose to leverage mobile sensing to capture people's behavior footprints amid the COVID-19 pandemic, and understand their collective behavior changes with respect to existing policies. Specifically, we propose to extract a rich set of behavioral markers from raw mobile sensing data, including mobility, social interactions, physical activities, and health states, and apply them in a generalized behavior change analysis framework to measure and detect community level behavior changes in an epidemic context. We present how to combine change point detection algorithm and interrupted time series analysis to automatically detect three different measurements of behavior changes (e.g., level, trend, and variance changes), and provide insights supported by statistical inference. A case study using a dataset that we collected from a large mobile sensing study conducted in the United States is shown to demonstrate the proposed framework and method.**

***Index Terms*—Behavior Change, change point detection, COVID-19, interrupted time series analysis, mobile sensing.**

I. Introduction

Non-pharmaceutical interventions (NPIs) such as case isolation, household quarantine, school and workplace closure, and restrictions on travel have been implemented by various state and federal governments [1] to slow down the transmission of COVID-19, thereby reducing peak healthcare demand while protecting those most at risk from infection [1]. However, as the number of newly confirmed cases continues to rise or stays high (as shown in Figure 1), it is crucial to understand communities' responses to the different public regulations and policies, and assess their effectiveness using data-driven approaches [2], [3].

Several studies have been conducted to understand people's reactions (e.g., attitudes, and knowledge on infection risk) to different public policies and scientific communications, as well as informing adjustments in the already adopted NPIs [4], [5]. These subjective data collected by sample surveys are critical to inform policy decisions, but fail to provide fine-grained understanding of people's daily behavior changes under the employed policies. Mobile sensing provides a complementary avenue for passively tracking people's fine-grained activities by leveraging smartphone embedded sensors (e.g., accelerometer, GPS, and Bluetooth) [6], [7]. Meanwhile, passive data can be augmented by application logs to track people's app usages, and by mobile ecological momentary assessments (EMAs) to conveniently and accurately capture people's subjective data such as daily symptomatic reports and moods. Combining all these data sources, we are capable of characterizing community behavior responses to COVID-19 and the implemented mitigation strategies more objectively, leading to a better understanding of policy effectiveness and potential adjustments.

In order to measure people's behaviors, we propose a set of behavior markers using raw sensor data from a large mobile sensing study conducted in the U.S. These behavior markers are categorized into mobility, app usage, face-to-face social interaction, entropy of physical activities, and symptomatic self-reports. To detect and measure behavior changes, we propose to combine changing point detection (CPD) algorithm [8] with interruption time series analysis (ITSA) [9] to automatically uncover the most significant change points (e.g., specific dates) as well as the different types of changes (e.g., level, variance, and trend changes) associated with these change points in each behavior marker. We then place all critical time points in the context of various public policies and important COVID-19 related events to interpret their implications on people's behavior change. The contributions of our current work can be summarized below:

- We apply mobile sensing to track community level behavior changes to understand potential impacts from public interventions.
- We propose a set of novel community level behavior markers that characterize people's mobility, activities, social interactions, and health states.
- We propose a generalized behavior change analysis framework that enables statistical understanding of behavior changes over time in a pandemic context such as COVID-19. Specifically, we combine CPD algorithm and ITSA to automatically detect change points offline, and measure three different types of change associated with the detected change points.
- We place our discovered change points and their associated changes in the context of important public policies,

[1] https://www.nga.org/coronavirus/

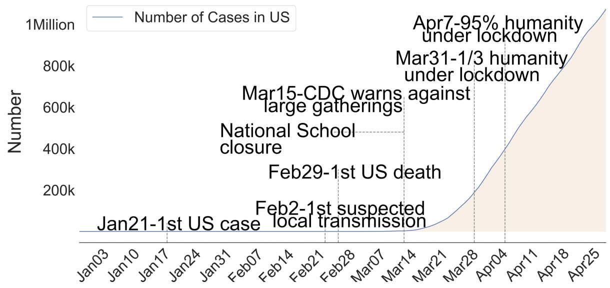

Fig. 1. US cases and public policies/guidelines from january to April, from Johns Hopkins Coronavirus Resource Center

guidelines, and COVID-19 related events to interpret their potential impacts on people's behavior changes, and discuss the associated policy implications to combat COVID-19.

The remainder of this work is presented as follows. In Section II, we summarize recent work on COVID-19, and existing methods for behavior change measurement and detection. In Section III, we define the proposed community level behavior markers, outline our generalized behavior change detection framework, and present the technical details on CPD and ITSA. Section IV introduces our mobile sensing data and summarizes our analysis plan, while Section V presents the results from our case study. We end the current work with a discussion in Section VI and conclusion marks in Section VII.

II. Related Work

Since the first COVID-19 outbreak in late December and early January in Wuhan China, many researches have been studying how to most effectively battle community transmission in COVID-19. A group of researchers from Imperial College London simulated the progression of the COVID-19 pandemic in two fundamental strategies – mitigation and suppression – based on corresponding configurations in different NPI combinations (e.g., home isolation of suspect cases, and family quarantine) and transmission rates, and reached the conclusion that only suppression can avoid overwhelming health care systems and hundreds of thousands of deaths in both the U.S. and U.K. [1]. Dehning et al. inferred the spreading rate of COVID-19 using the Susceptible-Infected-Recovered (SIR) model with three manually identified change points based on Germany public measures in fighting the pandemic [10]; while Deb et al. applied time series techniques to estimate the reproduction number of COVID-19 [11]. Researchers from [12] and [13] surveyed existing literature in behavioral science and psychology to pool together theories that can be leveraged to inform design of public interventions with respect to social distancing, scientific communications, and mental health coping during pandemic, as well as health behaviors such as hand-washing and circulation of misinformation.

The above works can help inform time critical decisions on public policies, but does not collect data directly from individuals, and thus can not be used to assess the direct impact of the deployed policies on human behaviors. Another group of works addressed this limitation by collecting survey data from representative samples of the general public to assess knowledge and perceptions of COVID-19 in different populations, regions, and phases of the pandemic, on topics such as disease transmission and common symptoms, falsehoods and discrimination [4]; COVID-19 severity, social responsibility values, social trust, self-interest, and their correlations on various behaviors including news monitoring, social distancing, disinfecting behaviors, and hoarding [5]; risk perception and protective behaviors in early March in the U.S. [14]; partisanship, health behavior, and policy attitudes [15]; effectiveness of public health messaging and mental health during quarantine [16]; social distancing compliance in Italy [17]; community responses including risk perception, information exposure, and preventive measures in Hongkong [18] and U.K. [19]; trust and spreading of misinformation in social media [20].

Instead of cross-sectional surveys, several other works applied mobile sensing to understand community level behaviors during COVID-19 pandemic. In [21]–[23], location visit data collected by mobile phones were leveraged to understand county-level average daily visits to non-essential businesses and travel distance. These works discovered that partisanship plays a critical role in people's responses to COVID-19 pandemic, and recommended the need of honest, clear, and consistent communications by political leaders and the media. Compared to our current work, we propose a rich set of community level behavior markers, including mobility, social interactions, app usages, physical activities, and symptomatic self-reports, to measure people's behavior changes with respect to public policies and guidelines. We study the temporal dynamics of changes in behaviors under both state and federal level COVID-19 intervention policies instead of aggregated measures conditioned on partisanship information. Many mobile sensing studies attempted to understand behavior dynamics with respect to different health outcomes such as depression [24], [25], sleep quality, physical activeness, social interactions [26], and schizophrenia [27], [28]. However, the majority of these works focused on predicting the targeted health outcomes using extracted features in supervised learning approaches. Our work builds on top of existing works by focusing on leveraging mobile sensing data to understand how communities respond to public policies that are put in place to combat the COVID-19 pandemic.

In [29], the authors applied test-retest correlation and piecewise linear regression to understand the stability and trend changes of physical and social activities, and a technique called latent growth curve (LCG) to model the mean-level and individual-level changes of physical and social activities over the 10-week academic term. However, their approaches require manual specification of change points. And in their case, the change point was chosen to be midterm. In comparison, we propose to leverage the change point detection (CPD) algorithm with interrupted time series analysis (ITSA) to automatically detect significant change points with three different types of changes (e.g., level, trend, and variance changes). In Section III, we will provide more details about our proposed community-level behavior markers, as well as the CPD algorithm and ITSA.

III. Methodology

A. A Generalized Behavior Change Detection Framework

We propose a generalized behavior change detection framework that can guide study of community level behavior change in the context of public crises or events such as COVID-19. This framework consists of five components as shown in Figure 2: **1) Form Research Questions**: First, we need to form research questions as it determines what sensor data to collect using mobile sensing. In the current work, our research question is to understand changes in people's behavior in response to public interventions put in place by governments to combat COVID-19 pandemic. **2) Design Behavior Markers**: We design digital behavior markers that can characterize people's behaviors. In this work, twelve different behavior markers are defined to describe people's mobility, app usages, social interactions, physical activities, and health states (e.g., self-reported symptomatic severity). These behavior markers can be extracted from raw sensing streams generated by smartphone embedded hardware sensors such as GPS, Bluetooth, and accelerometer, and software sensors such as app usage logs, and mobile EMAs. **3) Detect Change Points**: We want to detect at which points of each behavior marker time series significant changes may have occurred. In the current work, we apply an algorithm called Prophet [30], and extend it to automatically select the locally optimal number of change points according to Occam's razor or the law of parsimony. **4) Test Significance**: We apply interrupted time series analysis (ITSA) algorithm [31] to infer the significance of the changes before and after each detected change point. **5) Change Analysis**: Lastly, we measure changes for each significant change point in three different ways including level, trend, and variance changes. Specifically, level changes measure shifts in the underlying fixed levels of the behavior model; trend changes measure shifts in the rate of changes over time; and variance changes measure difference in the fluctuations between two consecutive time series segments.

B. Behavioral Markers

In this section we provide the mathematical formulations of the extracted behavior markers using raw mobile sensing data.

1) Mobility: Mobility-based behavioral markers represent people's collective activeness and dynamics with respect to travels and visits to different places. Place visit trajectory are generated based on GPS data using DBSCAN algorithm [32]. **Travel Distance (TD)** measures people's average daily total travel distance. Let $\{X_{1,i}, X_{2,i}, \dots, X_{N_{i,t},i}\}$ denote the sequence of anonymized GPS coordinate pairs, $N_{i,t}$ indicates the number of records for participant i on day t, and $i = \{1, 2, \dots, M_t\}$, where M_t indicates the number of participants on day t. Then TD can be formulated as: $TD(t) = \frac{1}{M_t}\sum_{i=1}^{M_t}\sum_{j=1}^{N_{i,t}-1} D(X_{j,i}, X_{j+1,i})$, where $D(\bullet)$ is the Haversine distance in kilometer given two pairs of latitude and longitude coordinates.

Entropy of Visits (EVP) measures the average daily randomness of place visits. Daily entropy of visits can be defined as: $EVP(t) = -\frac{1}{M_t}\sum_{i=1}^{M_t}\sum_{j=1}^{|C_{i,t}|} P_{i,t}(c_j) log P_{i,t}(c_j)$, where, for participant i on day t, $C_{i,t}$ is the set of unique places being visited, $c_j \in C_{i,t}$, $|C_{i,t}|$ is the number of unique places, and $P_{i,t}(c_j)$ is the proportion of times place c_j is being visited.

Unique Places (UP) denotes the average number of unique places visited by participants on a given day. $UP(t)$ is defined as: $UP(t) = \frac{1}{M_t}\sum_{i=1}^{M_t}|C_{i,t}|$, where $C_{i,t}$ is the same as in EVP.

Number of Visits (NV) shows average daily visits to different places. Given a daily sequence of place visit trajectory for participant i as $\{l_{1,i}, l_{2,i}, \dots, l_{N_{i,t},i}\}$, $NV(t)$ is defined as: $NV(t) = \frac{1}{M_t}\sum_{i}^{M_t} N_{i,t}$.

Radius of Gyration (RG) shows the range of travels and is defined as: $RG(t) = \frac{1}{M_t}\sum_{i=1}^{M_t}\sqrt{\frac{\sum_{j=1}^{N_{i,t}} D(X_{j,i},\mu_i)^2}{N_{i,t}}}$, where $\mu_i = \frac{\sum_{j=1}^{N_{i,t}} X_{j,i}}{N_{i,t}}$ [33].

Time Spent Out of Home (TSOH) represents the average time people spend outside their homes daily. Home labels are inferred using the place that each participant spent most time at between 10:00pm and 9:00am on the next day, which was previously adopted in [34]. For participant i and day t, the daily place visit trajectory $l_{1,i}, l_{2,i}, \dots, l_{N_{i,t},i}$ can be partitioned into home $H_{i,t}$ and non-home $NH_{i,t}$ subsets. Let $q_{j,i}$ denote the time duration of non-home visits for $j \in NH_{i,t}$. $TSOH(t)$ can be defined as: $TSOH(t) = \frac{1}{M_t}\sum_{i=1}^{M_t}\sum_{j=1}^{NH_{i,t}} q_{j,i}$.

2) App Usage: In this section, we define several app usage behavior markers to understand the changes in virtual activities, including financial, social, and entertainment app usages. We assigned each app to the 3 categories by using Play Store Scraper API [35]. We denote the duration of sequence of app usage as $f_{1,i}, f_{2,i}, \dots, f_{N_{i,t},i}$ for participant i on day t.

Financial app usage (FAU) measures the average cumulative time people spend on finance apps on a given day. $FAU(t)$ can be defined as: $FAU(t) = \frac{1}{M_t}\sum_{i=1}^{M_t}\sum_{j}^{N_{i,t}} f_{j,i} \times \mathbb{1}(app_{j,i} = Finance)$.

Social App Usage (SAU) measures the average cumulative time people spend on social apps on a given day and can be defined as: $SAU(t) = \frac{1}{M_t}\sum_{i=1}^{M_t}\sum_{j}^{N_{i,t}} f_{j,i} \times \mathbb{1}(app_{j,i} = Social)$.

Entertainment app usage (EAU) measures the cumulative time people spend on entertainment apps on a given day, where $EAU(t) = \frac{1}{M_t}\sum_{i=1}^{M_t}\sum_{j}^{N_{i,t}} f_{j,i} \times \mathbb{1}(app_{j,i} = Entertain)$.

3) Proximity of Face-to-face Social Interaction: We use the number of unique Bluetooth devices discovered between 9am and 10pm for each participant to approximate face-to-face social interactions outside home [36], [37]. Let us denote

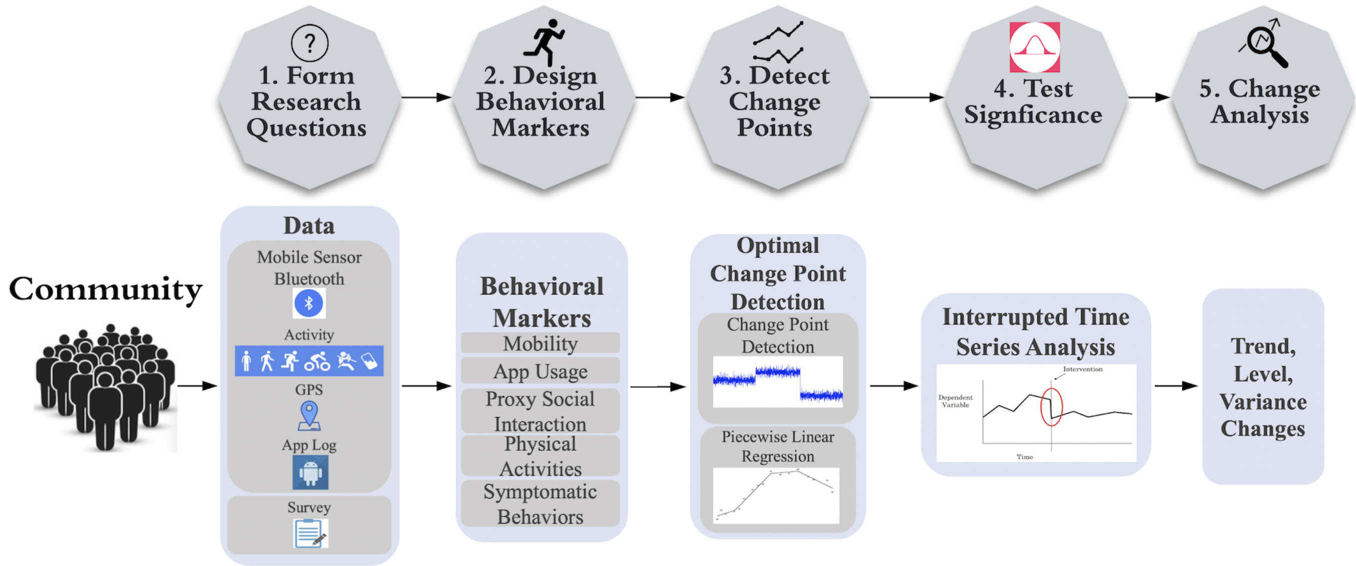

Fig. 2. A general framework of behavior change detection and analysis in public crisis and events such as COVID-19.

$e_{i,t}$ as the number of unique devices for participant i on day t.

Number of Bluetooth Encounters (NBE) measures the average daily detected unique Bluetooth devices between 9am and 10pm on a given day. $NBE(t)$ can be defined as: $NBE(t) = \frac{1}{M_t}\sum_{i=1}^{M_t} e_{i,t}$.

4) Entropy of Physical Activities: We define entropy of physical activities to measure people's daily activeness. These activity labels, including inVehicle, onBicycle, onFoot, Running, Still, Tilting, Unknown, and Walking, are generated by either Google Activity Recognition API [38] for Android devices, or CMMotion Activity Recognition API for iOS devices [39]. Let us denote $\mathbf{L}$ as the set of unique activity labels, and $l_j \in \mathbf{L}$ as an activity label.

Entropy of Activities (EA) measures the average daily randomness of physical activities and can be defined as: $EA(t) = -\frac{1}{M_t}\sum_{i=1}^{M_t}\sum_{j=1}^{|\mathbf{L}|} P_{i,t}(l_j) log P_{i,t}(l_j)$.

5) Symptomatic Self-reports: We also collected self-reported symptoms and their severity level daily on a 0-6 scale where 0 means no symptom and 6 means that the corresponding symptom is very severe. The symptoms include drowsiness, runny nose, headache, fever, fatigue, difficult breathing, diarrhea, cough, and chest pain. Let us denote $\mathbf{S}$ as the set of symptom labels, and $s_{j,i}$ denote the severity score of symptom $j \in \mathbf{S}$ from participant i.

Symptoms score (SS) measures the average daily symptomatic severity scores. $SS(t)$ can be defined as:$SS(t) = \frac{1}{M_t}\sum_{i=1}^{M_t}\sum_{j\in S} s_{j,i}$.

C. Change Point Detection

Change point detection (CPD) aims to identify change points in time series data. The formulation of CPD can be described as given a non-stationary random process $\mathbf{S} = \{s_1, s_2, \ldots, s_T\}$, where $s_t \in \mathbb{R}$ for $t \in \{1, 2, \ldots, T\}$. Given a piecewise stationary time series $\mathbf{S}$, there exists some unknown time $u_1, ..., u_K$ with either known or unknown K, such that the states of the time series $\mathbf{S}$ transfer from one state to another. The CPD problem then reduces to find the change time index $k \in \{1, 2, \ldots, K\}$ [8].

CPD problems can be solved by supervised or unsupervised learning approaches [40]. As annotations of change points is not available in this work, only unsupervised learning methods are considered. Since people's behaviors have both circadian and social cycles (e.g. daily and weekly routines), we adopt a linear decomposition method called Structural Time Series Model (STSM) by using Prophet [30], which decomposes time series data into three additive components including trend, seasonality, and holidays using Formula 1 given below.

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, \tag{1}$$

where $g(t)$ models non-periodic changes in trend, $s(t)$ captures seasonal changes, and $h(t)$ encodes holidays. The error term ϵ_t contains the unexplained part of the additive model and is assumed normally distributed. In $g(t)$, change points are explicitly defined. Suppose there exist S change points at s_j where $j \in \{1, ..., S\}$. The base growth rate k is assumed to be a constant. The adjusted rate associated with change point s_j is denoted as δ_j, and $\boldsymbol{\delta} \in \mathbb{R}^S$ is a vector contains all the adjusted rates. Then the rate at each time t can be expressed as $k + \boldsymbol{a(t)}^T\boldsymbol{\delta}$, where $\boldsymbol{a(t)} \in \{0,1\}^S$ is a change points indicator vector. Finally, the linear trend function can be defined as

$$g(t) = (k + \boldsymbol{a(t)}^T\boldsymbol{\delta})t + (m + \boldsymbol{a(t)}^T\boldsymbol{\gamma}), \tag{2}$$

where m is the offset parameter and $\boldsymbol{\gamma}$ is the adjustment of the rate to guarantee the continuity of the function. To fit the models, Stan's L-BFGS [41] was implemented in prophet to estimate the parameters with pre-defined prior probability. Thus with pre-specified number of change points, change points can be automatically detected.

To find the best number of change points to fit the time series data, we apply an automatic CPD algorithm using piecewise linear regressions (PLR) to select the local optimal number of change points, as shown in Algorithm 1 adapted from this work [42]. The general idea of Algorithm 1 is that fitting PLR segmented by change points are expected to improve the fitness over one simple linear regression to the time series data. If the fitted PLR with more change points does not decrease $RMSE$ beyond the chosen threshold ϵ, we stop adding change points to the regression. Here we empirically set the threshold $\epsilon = \frac{behavior\ marker\ scale}{1000}$, and the smaller the threshold ϵ is, the more sensitive PLR will be with respect to $RMSE$ (i.e., more change points could be detected). This process is analogous to early stopping in training machine learning models, that is terminating the training process when validation error starts to increase to prevent overfitting.

Algorithm 1: Change Points Detection with Piecewise Linear Regression Fit

```
Result: Best piecewise-linearly fit change points to the
        time series
Initialize a threshold ε and the maximal possible
  number of change points C;
Fit a simple linear regression by setting the time as
  explanatory variable and time series data as response
  variable;
Calculate root mean square error (RMSE) between
  fitted value and original time series data and denote
  as RMSE_0;
for iteration i = 1,2,..., C do
   Get i number of change points from Prophet; Fit
    PLR segmented by the change points; Calculate
    RMSE_i; if RMSE_{i-1} - RMSE_i > ε then
      Continue;
   else
      Return the change points selected at iteration
        i-1
   end
end
```

D. Interrupted Time Series Analysis

Interrupted time series analysis (ITSA) is a powerful quasi experimental design for evaluating the effectiveness of medical or political interventions by testing whether there are significant differences in trend and/or level changes in time series before and after the specified intervention times [43] with Equation 3 below:

$$Y_t = \beta_0 + \beta_1 T + \beta_2 X_t + \beta_3 T X_t, \tag{3}$$

where T is the number of selected time units (e.g. day, month, or year); X_t is a dummy variable taking value 0 or 1 indicating the pre and post intervention period respectively; TX_t is the time after intervention, and Y_t is the outcome at time t. In Equation 3, β_0 demonstrates the baseline level at $T = 0$, β_1 represents the slop of the pre-intervention trend, β_2 expresses the difference of y_t before and after the intervention time, and β_3 represents the changes of the time series trend in the post-intervention time period, and the slop of post-intervention trend can be obtained by $\beta_1 + \beta_3$.

E. Behavior Change Measures

In this section, we define three different measures of changes associated with each learned significant change point, including level, trend, and variance changes. ITSA can measure both level and trend changes, and provide tests for their statistical significance for each change point. In addition, we also compare variance difference between two time series segments around each change point, and use the Levene's test [44] to provide statistical significance. **Level change**, as represented by β_2, is defined as the difference between the original value on the intervention time point and the value extrapolated by the pre-intervention fitted linear regression to the intervention time point. **Trend change**, as represented by β_3, is defined as the difference between the slop of the pre-intervention fitted linear regression and the slop of the post-intervention fitted linear regression [9]. The p-values of the estimated coefficients β_2 and β_3 in Equation 3 indicate whether there are significant level and trend change respectively. The estimated value of β_2 and its sign measure the magnitude and direction of level change respectively. The estimated value of β_3 and its sign measure the magnitude and direction of the trend change, respectively. **Variance change** is estimated using empirical estimate of variance in each sample segment. The change points that are tested with any significant trend, level, or variance change are denoted as **significant change point (SCP)**.

IV. Study Design

The data presented in this work is part of a multicohort study leveraging smartphones for earlier diagnosis of illness. The study was approved by the Western Institutional Review Board (IRB) with board protocol number 20770 and the date of approval on 08/24/2017. To date, over 2700 participants from 24 states have been recruited to participate for up to 1-year in the study. From this data we extracted a total of 598 participants that have data from Jan 1st, 2020 to the end of this study (April 30th, 2020) to explore behavioral patterns amid the COVID-19 pandemic. The mean age of the participants is $40.98 (sd = 14.63)$. 66.36% of the participants are female, with 67.2% White, 18.3% Black, 7.3% Asian, 3.6% Multiple Races , 3.2% Hispanic , and 0.4% others. More details about the larger study can be provided by the corresponding author. Among the 598 participants, 50% are full time workers, 18% are full time students, 14% are working part time, 8% have retired, 6% are care takers, and 4% are

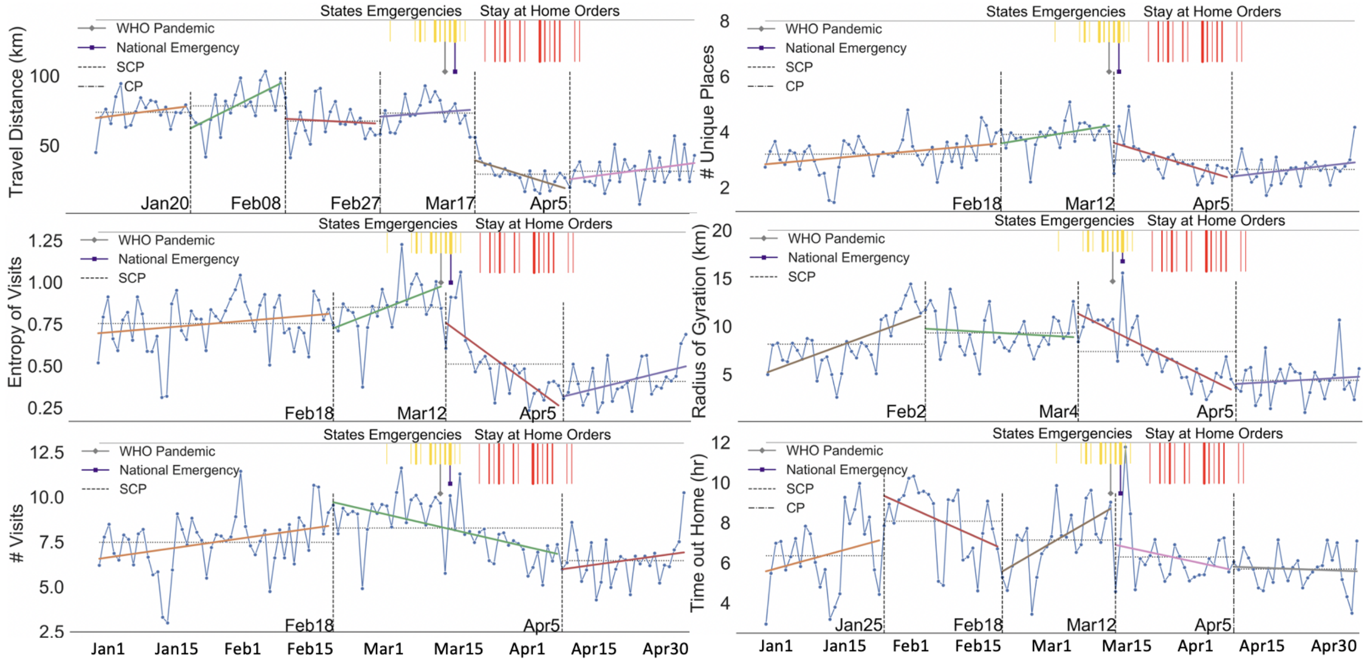

Fig. 3. Mobility behavior markers. Vertical yellow lines dropping down from the top indicate dates of declaration of state emergencies, and vertical red lines dropping down from the top indicate dates of issuance of stay-at-home orders from various states in the data. Dashed horizontal lines within each segment indicate the average in each segmented time series. The solid lines are the fitted piecewise linear regression line using ITSA.

temporarily unemployed. Participants received up to $300 for participating in the study.

Participants were asked to install and run a mobile app called ReadiSens for up to 4 months. ReadiSens is a cross-platform app built on top of Sensus [45] to passively collect GPS location data every 30 mins, Bluetooth Encounters data every 15 mins in Android and every 30 mins in iOS, activity data when activities change in Android and every 2 hours in iOS. App usage logs were captured by ReadiSens for Android devices only. EMAs were delivered at 8pm everyday to collect self-reported influenza-like symptoms including: drowsiness, runny nose, headache, fever, fatigue, difficult breathing, diarrhea, cough, and chest pain. Collected data were periodically uploaded to Amazon Web Services (AWS) Simple Storage Service (S3). All data were encrypted and anonymized to protect participants' data security and privacy. Any identifying information was omitted before being stored on AWS. GPS data were anonymized by omitting the integer parts of their longitudes and latitudes. Bluetooth's name and MAC addresses were automatically hashed on Readisense.

We also augment the mobile sensing data with manually curated time points, when the federal and state governments declared national and state emergencies and issued stay at home orders. All the states that the participants came from declared state emergency between Feb 29th and Mar 15th, and issued stay-at-home order between Mar 19th and Apr 7th. To analyze community level behavior changes, we aggregated individual mobile sensing data based on the defined community level behavior markers in Section III-B. Each behavior marker time series was run through the proposed automatic CPD to obtain a set of change points that give the best overall model fit in piecewise linear regressions using the Prophet algorithm [30]. These sets of change points are then fed into the ITSA method presented in Section III-D to test their significance with three different change measures including the level, trend, and variance changes associated wth each change point. A change point in a given behavior marker is said to be significant if at least one of these three change measures is significant, and is denoted SCP for convenience. We use a 95% significance level for all significance tests.

V. Results

A. *Mobility Changes*

Table I summarizes all significant change points from each behavior marker. As shown in Figure 3, travel distance started with a fitted line that is flat up until Feb 25th, which is detected as an insignificant CP when the two segments associated with Feb 25th are compared against each other. Then Mar 6th, Mar 15th and Mar 24th were detected and tested as SCPs, of which only trend changes were significant at a 95% confidence level (as shown in Table I. Specifically, the trend first decreased by 3.93 on Mar 6th, then continued to decrease further by 2.84 on Mar 15th, but reverted after Mar 24th with a relatively modest positive overall slope given by estimate of $\beta1 + \beta3$. We believe these results were correlated with the declarations of state emergencies, national emergency, and WHO declaring global pandemic before Mar 15th, sustained through some early stay-at-home orders from various states before Mar 24th.

Behaviors	Behavioral Markers	SCP	Level Change		Trend Change		Variance Change	Adjusted R^2
Mobility	Travel Distance	2020-01-20	$+-$	**-17.32 (0.02)**	$++$	1.33 (0.06)	127.85 (0.11)	0.25
		2020-02-08	$+-$	**-25.19 (0.00)**	$+-$	**-1.97 (0.01)**	-72.38 (0.48)	0.28
		2020-03-17	$+-$	**-35.13 (0.00)**	$+-$	**-1.37 (0.01)**	-27.28 (0.29)	0.85
		2020-04-05	$++$	6.86 (0.25)	$-+$	**1.38 (0.00)**	36.98 (0.13)	0.11
	Entropy of Visits	2020-02-18	$+-$	-0.02 (0.71)	$++$	**0.01 (0.01)**	0.00 (0.97)	0.13
		2020-03-12	$+-$	**-0.19 (0.02)**	$+-$	**-0.03 (0.00)**	0.01 (0.36)	0.67
		2020-04-05	$++$	**0.04 (0.52)**	$-+$	**0.02 (0.00)**	-0.02 (0.15)	0.47
	Number of Visits	2020-02-18	$++$	**1.38 (0.01)**	$+-$	**-0.10 (0.00)**	-0.13 (0.65)	0.23
		2020-04-05	$+-$	-0.89 (0.16)	$-+$	**0.10 (0.00)**	-0.83 (0.11)	0.55
	Unique Places	2020-03-12	$+-$	**-0.56 (0.05)**	$+-$	**-0.08 (0.00)**	0.11 (0.54)	0.53
		2020-04-05	$++$	0.00 (0.98)	$++$	**0.07 (0.00)**	-0.14 (0.51)	0.28
	Radius of Gyration	2020-02-02	$+-$	-1.28 (0.24)	$+-$	**-0.21 (0.00)**	-3.06 0.29)	0.81
		2020-03-04	$++$	**2.69 (0.01)**	$--$	**-0.22 (0.00)**	2.47 (0.99)	0.46
		2020-04-05	$++$	0.53 (0.61)	$-+$	**0.28 (0.00)**	4.66 (0.09)	0.56
	Time Spent Out Home	2020-01-25	$++$	**2.32 (0.02)**	$+-$	-0.17 (0.17)	-0.52 (0.67)	0.25
		2020-02-18	$+-$	-1.38 (0.10)	$-+$	**0.25 (0.00)**	-2.66 (0.99)	0.29
		2020-03-12	$+-$	**-1.75(0.03)**	$+-$	**-0.19 (0.02)**	-0.48 (0.50)	0.23
App Usage	Social App Usage	2020-03-12	$+-$	**-1.50(0.03)**	$-+$	**0.15(0.00)**	**2.75(0.00)**	0.24
		2020-04-05	$+-$	**-1.72 (0.01)**	$+-$	**-0.17 (0.00)**	0.19 (0.52)	0.25
	Ent App Usage	2020-02-18	$+-$	-0.02(0.78)	$-+$	**0.01(0.01)**	0.07(0.12)	0.03
		2020-04-05	$+-$	**-0.47 (0.00)**	$++$	**0.04 (0.00)**	**0.11 (0.00)**	0.39
SocialInt	#Bluetooth Encounter	2020-03-04	$++$	-0.12 (0.99)	$+-$	**-5.87 (0.00)**	1404.41 (0.86)	0.33
		2020-04-05	$++$	10.52 (0.62)	$-+$	**5.12 (0.00)**	**-3790.80 (0.00)**	0.44
Physical Activities	Entropy of Activity	2020-01-25	$++$	**0.11 (0.01)**	$+-$	**-0.01 (0.00)**	-0.003 (0.33)	0.48
		2020-02-18	$+-$	0.007 (0.86)	$-+$	**0.01(0.02)**	0.002 0.44)	0.13
		2020-03-12	$+-$	**-0.17(0.00)**	$+-$	**-0.003 (0.45)**	0.004 (0.06)	0.37
Symptom reports	Symptoms Score	2020-02-08	$++$	**1.34 (0.02)**	$++$	**0.14 (0.01)**	-0.01 (0.77)	0.71
		2020-02-27	$+-$	**-2.66 (0.00)**	$++$	**0.21 (0.00)**	5.68 (0.07)	0.65
		2020-03-17	$+-$	**-4.05 (0.00)**	$+-$	**-0.39 (0.00)**	**-1.17 (0.00)**	0.62
		2020-04-05	$++$	0.05 (0.92)	$-+$	**0.10 (0.04)**	-4.55 (0.47)	0.12

TABLE I

Significant change points (SCP) with their level, trend, and variance changes. The values in parenthesis are p-values. Bold cells indicate significance at 95% level. The Adjusted R^2 values are the fitness metrics in the segmented linear regression using ITSA. The left and right signs under Level Change column indicate the positive or negative value of β_0 and β_2, respectively; and the left and right signs under Trend Change column indicate the positive or negative value of β_1 and β_3, respectively. All values are cut off at two decimal digits. Bold values denote significant results ($p \leq 0.05$)

After Mar 24th, we observed from the data that people started to become more active despite more states issuing stay-at-home orders, but at a very low growth rate. This can be seen from the flat slope in the segment after Mar 24th for TD in Figure 3. This slow increasing of TD could be explained by the already low-level in TD before Mar 24th, and may indicate that voluntary stay-at-home orders over time may loose their effectiveness if no sufficient external interventions (e.g., public communications to reinforce the importance of stay-at-home orders) were exerted on people. However, we need to keep in mind the possibility that in less population dense areas, residents could engage in physical activities such as taking a walk outside while being able to maintain social distancing with others. The overall fit of the segments associated with each change point was quite reasonable with adjusted-R^2 all being around or above 0.5 level.

For Entropy of Visits (EV), as shown in Figure 3, there are two insignificant change points including Feb 6th and Feb 29th, and four significant change points including Jan 13th, Jan 25th, Mar 12th, and Mar 24. From Table I, all four SCPs of EV are significant in trend changes, with the exception of Mar 24th, which was also significant in variance change. From Figure 3, we observe that the SCPs are disrupted by the two CPs (Feb 06th and 29th), with relative small differences in mean values in each segments created by the change points up until Mar 12th. Regardless of being significant or not, the trend changes before Mar 12 appeared to be ups and downs with relative small changes in magnitude (e.g., 0.03 on Jan 13th and 25th). Since the first state emergency was declared on Feb 29th, followed by declarations from a few other states until Mar 12th, the trend changes didn't have significant drops until Mar 12th, indicating some lagged effects on EV from public policies. After Mar 12th, people started to pay visits to a less diverse types of places, showing the potential impacts of WHO declaring global pandemic, national emergency, and the public interventions had on people's visits to non-essential businesses and venues. The overall fit of the segments associated with the last two change points was very good with adjusted-R^2 being around 0.8.

For Number of Visits (NV), as shown by Figure 3, both detected change points, Feb 12th and Mar 24th, were tested as SCPs. From Table I, Feb 12th was significant in both trend and variance change, while Mar 24th was significant in both level and trend changes. From Figure 3, there is a moderate increase in trend after Feb 12th, but the overall average NV would have been very closed before and after Feb 12th if not for the upsurge in NV between Mar 1st and around Mar 15th. This surge in NV could be explained by the early batch of state emergency declarations. On March 24th, as the first few stay-at-home orders issued in several states in our data, NV sharply dropped in both trend and level changes, indicating a potential significant impacts generated by the policies. The overall fit of the PLRs associated with the first SCP (Feb 12th) was low due to large variance in the segments. The opposite result was recorded for the second SCP (Mar 24th).

For Unique Places (UP), as shown by Figure 3, Jan 15th, Jan 29th, Mar 10th and Mar 24th were the detected and tested as SCPs, while Feb 12th and 25th are tested not significant. From Table I, all four SCPs except Jan 29th were significant with trend change, while Jan 29th was significant with level change, and Mar 24th was also significant with variance change. From Figure 3, we observed a similar pattern to that of EV, with the four SCPs being disrupted by two CPs. The mean UPs and trends in each segment appeared to be ups and downs regardless of significance levels up until Mar 10th. On Mar 10th and 24th, the trend maintained decreasing with a more significant drop in mean UV after Mar 10th, indicating a potential significant impacts from the early declarations of state emergencies, which preceded the following national emergency and WHO declaring global pandemic. As more states declare state emergencies and issued stay-at-home orders, the impacts were maintained throughout the end of the data period on April 15th. The overall fit was good for the last two SCPs in March.

For Radius of Gyration (RG), as shown by Figure 3, Mar 12th and 24th were detected and tested as SCP, while Feb 29th was tested not significant. From Table I, both Feb 25th and Mar 24th were significant in trend changes, with Mar 24th also being significant in variance change. From Figure 3, we observed significant amount of variances before Feb 29th. However, the mean RG was similar before and after Feb 29th. On Mar 12th, amid the early declared state emergencies, and the WHO global pandemic, RG started a significant decreasing trend until Mar 24th that leads to lower mean RG in the last two segments. After Mar 24th, the trend reverted to positive from negative but was nearly flat. A similar implication to TD can be drawn here, as people gradually became more active after getting accustomed to all the public interventions in the COVID-19 pandemic. To maintain the low level of geographical range in activities, more interventions need to be considered. The overall fit was fairly high with adjusted-R^2 values being around 0.8 for both SCPs.

For Time spent out of home (TSOH), as shown by Figure 3, the four detected SCPs were Jan 29th, Feb 25th, Mar 10th, and Mar 24th, while Jan 15th and Feb 12th were not significant. From Table I, Jan 29th was only significant in the level change, Feb 25th was only significant in trend change, Mar 10th was significant both in trend and variance changes, while Mar 24th was significant in level, trend, and variance changes. From Figure 3, we observed ups and downs in both level and trend, and significant amount of fluctuations in TSOH up until Mar 10th, which was preceded by the first batch of state emergency declarations from various states in our data. After Mar 10th, with the WHO global pandemic and the national emergency declarations mixed with more state emergencies and some stay-at-home orders, people spent less and less time outside their home, leading to family quarantine. The overall fit was fairly high for the two SCPs in March with adjusted-R^2 values being above 0.7.

Overall, people's mobility behaviors were significantly impacted by COVID-19 events and public policies. For the majority of the proposed community level mobility behavior markers, we observed small but significant changes before March 10th or 12th, which were preceded by the early state emergencies, and around both the WHO global pandemic and the U.S. national emergency declarations. Significant change points detected during this long time span from Jan 1st to Mar 10th or 12th could be considered natural fluctuations in human activities as the COVID-19 developed around the globe, but not yet causing havoc in the western world. However, all six behavior markers consistently indicated large significant changes after Mar 10th or 24th when the COVID-19 pandemic became increasingly severe, and the U.S. became the new epidemic center.

B. App Usage Changes

For app usages, financial app had two detected CPs on Feb 18th and Apr 5th. However, both of them were not significant. Social app had two SCPs on Mar 12th and Apr 5th. Between Mar 12th and Apr 5th, participants on average spent an increasing amount of time using social apps. Entertainment app had two SCPs on Feb 18th and Apr 5th. After Apr 5th, participants on average sharply increased their entertainment app usage potentially as a consequence of the stay at home orders. Figure 4 visualizes the app usage behavior markers with the same background information as in Figure 3. The figure for social app usage highlights increasing usage of social apps between Mar 12th and Apr 5th, and the figure for entertainment app usage shows significant surge in entertainment app usage after Apr 5th. One possible interpretation is that between Mar 12th and Apr 5th, participants relied on social media to consume information regarding COVID-19; while after Apr 5th, as more participants worked from home, they spent more time with entertainment apps.

C. Social Interaction Changes Using Bluetooth Encounters

We use number of unique devices detected from Bluetooth encounters to approximate in-person social interactions. There were two SCPs on Mar 4th and Apr 5th for number of Bluetooth encounters. Between Mar 4th and Apr 5th, participants on average had decreasing daily in-person social interactions.

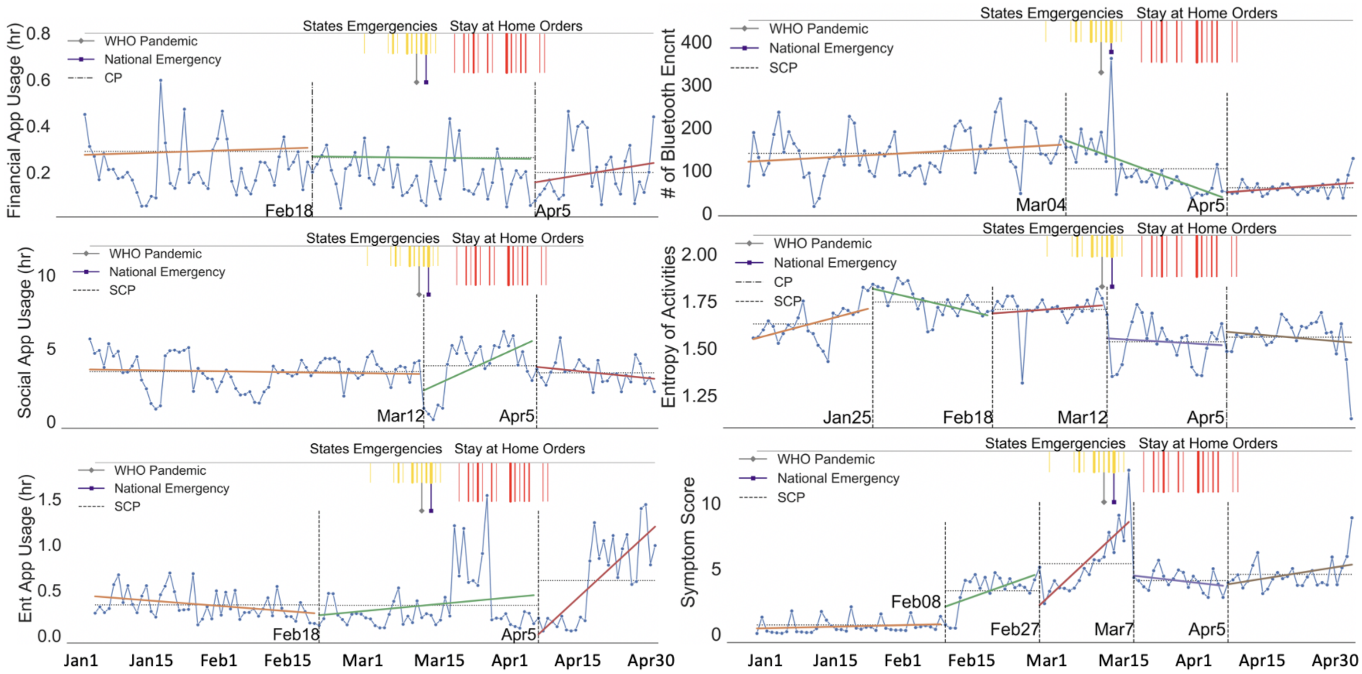

Fig. 4. Behavior markers of app usages, physical activities, Bluetooth encounters, symptomatic self-reports. Vertical yellow lines dropping down from the top indicate dates of declaration of state emergencies, and vertical red lines dropping down from the top indicate dates of issuance of stay-at-home orders from various states in the data. Dashed horizontal lines within each segment indicate the average in each segmented time series. The solid lines are the fitted piecewise linear regression line using ITSA.

After Apr 5th, in-person social interaction maintained at a stable and relatively low level. Figure 4 visualizes these behaviors, which could be explained by the issuance of state emergencies and stay at home orders by most states in the U.S.

D. Physical activity Changes

For physical activity changes, entropy of activities had three detected SCPs on Jan 25th, Feb 18th and Mar 12th, and one CP on Apr 5th. Figure 4 visualizes this behavior marker. The significant level drop and decreasing trends after Mar 12th indicated that participants engaged in less diverse physical activities such that the distribution of time spent in each activity category by the participants was relatively stable and dominated by the 'Still' category.

E. Symptomatic Score Changes

For self-reported symptoms, symptoms score had four SCPs on Feb 8th, Feb 27th, Mar 17th, and Apr 5th. From Jan 1st to Mar 7th, the symptoms score continued to increase with positive trends. After Mar 7th, this trend was interrupted as most states declared state emergencies. Subsequently, the symptoms score dropped dramatically and maintained a relatively stable level. Figure 4 visualizes this pattern.

F. Implications for Policies to Combat COVID-19

Figure 5 summarizes all the detected change points for each behavior marker. Unique dates that are detected as significant change points by two or more behavior markers including Jan 29th (number of unique places, time spent outside home, and usage of entertainment apps), Feb 12th (number of visits, usage of financial apps, and usage of entertainment apps), Feb 25th (radius of gyration, time spent outside home, and usage of entertainment apps), Mar 10th (number of unique places, and time spent outside home), and Mar 24th, which were detected by 10 out of the 12 behavior markers, except number of bluetooth encounters and symptomatic scores. Notice that there are only two unique dates detected in February (Feb 12th and 25th). These two dates were sandwiched by the event of 1st suspected community transmission on Feb 2nd, and 1st U.S. death on Feb 29th (see Figure 1). On Feb 29th, Washington state became the first state to announce state emergency in the U.S.. Also, seven unique dates were detected as significant in March (Mar 3rd, 6th, 10th, 12th, 14th, 15th, and 24th). As can be seen from Figure 1, on Mar 15th, the CDC warned against large gatherings, and schools were closed nation wide, and on Mar 31st, one third of population in the world were home quarantined. These two events were amid the declaration of state emergencies throughout the first half of March, and the subsequent stay-at-home orders by the majority of states in the U.S.. Under this backdrop, the entire situation became quite fluid and could cause diverse responses and changes in people's behaviors, which could be highly dependent on the severity of COVID-19 infections near where they resided.

On March 24th, the world continued to witness the in-

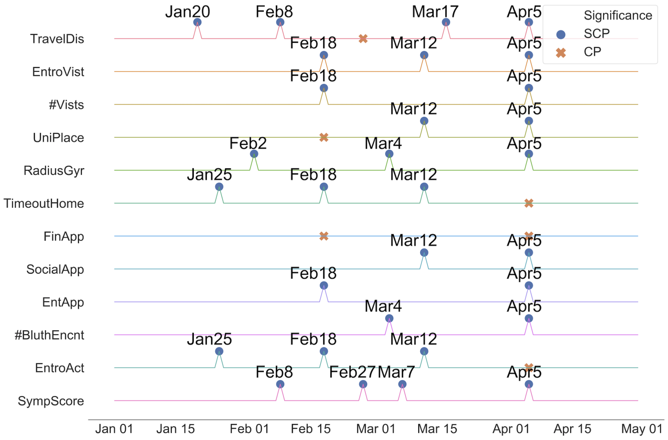

Fig. 5. Change points (CPs) and significant change points (SCPs) for all behavioral markers between Jan 1st and Apr 30th.

creasingly severe damages caused by COVID-19 pandemic. Japan had postponed the 2020 Tokyo Olympics to 2021; India locked down 1.3 billion people in the country; new cases grew exponentially in Italy, Spain, Germany, and Iran [2]; while in the U.S., the total number of confirmed cases reached beyond 50,000 from all 50 states, and it was the deadliest day with under 200 deaths in a day. By then, nearly 50% of the U.S. population had been ordered to stay home [3]. It was consistently identified as significant change point on Mar 24th among our proposed behavior markers, with the majority of them started off a downward trend, or stayed at very low level until Apr 15th, which was the last date of our data.

Our proposed framework with the adopted change point detection and interrupted time series analysis algorithms was able to reveal all the behavior changes in the various dates summarized in Figure 5. Aiming to understand policies effectiveness, our framework can be adopted by local communities with larger representative samples to monitor community level behavior changes over the course of various public policies. In events such as COVID-19, in order to control community level virus transmission, a system of interdependent and complex policies have to be carried out in a coordinated manner. It is immensely challenging to assess policy effectiveness over time to provide timely feedback to policy makers. We believe the current behavior monitoring framework using mobile sensing can help achieve this goal. There are two aspects to it: 1) it detects when significant changes have occurred with respect to public interventions, thus facilitate understanding of potential policy effectiveness in people's behavior changes; 2) it informs current behavior states of communities (e.g., after Mar 24th, various behavior markers had positive trends), upon which policy makers can re-evaluate the needs of making new policies or adding reinforcements to existing ones.

[2]https://www.spglobal.com/en/research-insights/articles/covid-19-daily-update-march-24-2020

[3]https://www.cnn.com/world/live-news/coronavirus-outbreak-03-24-20-intl-hnk/index.html

VI. Discussion

Assessing community level behavior changes with respect to public interventions such as stay-at-home order and work/school closures can reveal policy effectiveness, and provide feedback for policy adjustments. The advancements of sensing technologies and the ubiquity of personal smartphones make possible behavior monitoring using mobile sensing. We propose a general framework for behavior change detection and monitoring using mobile sensing, and find that our approach is able to identify dates, on which community level behavior changes correlated with important COVID-19 events and public policies, indicating potential policy impacts on people's behaviors. We discuss limitations in our current work, and lay out some future directions below.

A. Limitations

When perform ITSA for detected change points in behavior marker time series, we assumed there exists an external intervention on each change point, and that the assumed intervention between two time series segments could only affect the immediate post-intervention time series segment. However, in practice, a sequence of interventions may have concurrent and cumulative impacts on people's behaviors. This would make interpretations of policy effectiveness difficult if not at all impossible.

In addition, there exists privacy concerns when we apply mobile sensing to collect personal data. Raw sensing data can be utilized in ways that lead to privacy breaching. In order for users to be willing to install and use these mobile apps, transparency about data usage and sufficient data protection measures are required to convince them.

Last but not least, the current CPD+ITSA approach is offline, meaning that analyses will be conducted in retrospective. Online CPD algorithms may be desirable to measure changes as data are made available. Our current approach also treated each behavior marker in isolation, which ignores their mutual correlations. Instead of using univariate time series approach, multivariate approach can potentially provide us different insights about people's behavior changes.

B. Future Direction

We plan to investigate new behavior markers that can represent different aspects of human behaviors in the future. We will also investigate multivariate time series approaches for behavior change detection and measurement. A popular linear approach is ARIMA. Other nonlinear time series approaches for behavior change detection such as dynamic time warping, and sequential pattern mining can be explored using our proposed framework. Furthermore, ITSA with multiple interventions is also worth investigating to examine the potential impacts of sequence of interventions (e.g., public policies) on behavior changes. Additionally, to address privacy concerning issues, we can apply advanced data anonymization and differential privacy techniques, such that maximize the usage of the privacy protected data with the lowest risk of exposing personal information. Lastly, we want to explore online CPD algorithms for more timely behavior change detection and monitoring, and providing feedback to inform policy maker in improving existing policies, or making new ones.

VII. Conclusion

In the current work, we propose a behavioral change monitoring framework using mobile sensing to help evaluate public policy effectiveness amid the COVID-19 pandemic. Under our framework, we present analyses on a unique mobile sensing dataset from a large mobile sensing study in the U.S. that overlapped with COVID-19 pandemic, using changing point detection and interrupted time series analysis algorithms. The results revealed a set of significant change points in a dozen of behavior markers on mobility, app usages, in-person social interactions, physical activities, and self-reported symptomatic severity scores. We measure behavior changes using three metrics including level, trend, and variance changes, and interpret the insights learned from the analyses in the backdrop of COVID-19 developments and public policies at various levels in the U.S.. Our work shows that mobile sensing is a promising technology to monitor community behaviors, understand their dynamics, and detect any significant changes with respect to external interventions.

Acknowledgement

This work was supported by the DARPA Warfighter Analytics using Smartphones for Health (WASH) program.

References

[1] N. Ferguson, D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg *et al.*, "Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand," 2020.

[2] A. Madan, M. Cebrian, D. Lazer, and A. Pentland, "Social sensing for epidemiological behavior change," in *Proceedings of the 12th ACM international conference on Ubiquitous computing*, 2010, pp. 291–300.

[3] A. Madan, M. Cebrian, S. Moturu, K. Farrahi *et al.*, "Sensing the" health state" of a community," *IEEE Pervasive Computing*, vol. 11, no. 4, pp. 36–45, 2011.

[4] P. Geldsetzer, "Knowledge and perceptions of covid-19 among the general public in the united states and the united kingdom: A cross-sectional online survey," *Annals of internal medicine*, 2020.

[5] B. Oosterhoff *et al.*, "Psychological correlates of news monitoring, social distancing, disinfecting, and hoarding behaviors among us adolescents during the covid-19 pandemic," 2020.

[6] E. Shmueli, V. K. Singh, B. Lepri, and A. Pentland, "Sensing, understanding, and shaping social behavior," *IEEE Transactions on Computational Social Systems*, vol. 1, no. 1, pp. 22–34, 2014.

[7] G. Dong, M. Boukhechba, K. M. Shaffer, L. M. Ritterband, D. G. Gioeli, M. J. Reilley, T. M. Le, P. R. Kunk, T. W. Bauer, and P. I. Chow, "Using graph representation learning to predict salivary cortisol levels in pancreatic cancer patients," *Journal of Healthcare Informatics Research*, pp. 1–19, 2021.

[8] C. Truong, L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods," *Signal Processing*, vol. 167, p. 107299, 2020.

[9] E. Kontopantelis, T. Doran, D. A. Springate, I. Buchan, and D. Reeves, "Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis," *bmj*, vol. 350, p. h2750, 2015.

[10] J. Dehning, J. Zierenberg, F. P. Spitzner, M. Wibral, J. P. Neto, M. Wilczek, and V. Priesemann, "Inferring change points in the spread of covid-19 reveals the effectiveness of interventions," *Science*, 2020.

[11] S. Deb and M. Majumdar, "A time series method to analyze incidence pattern and estimate reproduction number of covid-19," *arXiv preprint arXiv:2003.10655*, 2020.

[12] P. D. Lunn, C. A. Belton, C. Lavin, F. P. McGowan, S. Timmons, and D. A. Robertson, "Using behavioral science to help fight the coronavirus," *Journal of Behavioral Public Administration*, vol. 3, no. 1, 2020.

[13] J. J. Van Bavel, P. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. Crockett, A. Crum, K. Douglas, J. Druckman, J. Drury *et al.*, "Using social and behavioural science to support covid-19 pandemic response," 2020.

[14] T. Wise, T. D. Zbozinek, G. Michelini, C. C. Hagan *et al.*, "Changes in risk perception and protective behavior during the first week of the covid-19 pandemic in the united states," 2020.

[15] S. Kushner Gadarian, S. W. Goodman, and T. B. Pepinsky, "Partisanship, health behavior, and policy attitudes in the early stages of the covid-19 pandemic," *Health Behavior, and Policy Attitudes in the Early Stages of the COVID-19 Pandemic (March 27, 2020)*, 2020.

[16] S. Barari, S. Caria, A. Davola, P. Falco, T. Fetzer, S. Fiorin, L. Hensel, A. Ivchenko, J. Jachimowicz, G. King *et al.*, "Evaluating covid-19 public health messaging in italy: Self-reported compliance and growing mental health concerns," *medRxiv*, 2020.

[17] G. Briscese, N. Lacetera, M. Macis, and M. Tonin, "Compliance with covid-19 social-distancing measures in italy: The role of expectations and duration," National Bureau of Economic Research, Tech. Rep., 2020.

[18] K. O. Kwok, K. K. Li, H. H. Chan, Y. Y. Yi, A. Tang, W. I. Wei, and Y. S. Wong, "Community responses during the early phase of the covid-19 epidemic in hong kong: risk perception, information exposure and preventive measures," *medRxiv*, 2020.

[19] C. J. Atchison, L. Bowman, C. Vrinten, R. Redd, P. Pristera, J. W. Eaton, and H. Ward, "Perceptions and behavioural responses of the general public during the covid-19 pandemic: A cross-sectional survey of uk adults," *medRxiv*, 2020.

[20] G. Pennycook, J. McPhetres, Y. Zhang, and D. Rand, "Fighting covid-19 misinformation on social media: Experimental evidence for a scalable accuracy nudge intervention," 2020.

[21] J. M. Barrios and Y. V. Hochberg, "Risk perception through the lens of politics in the time of the covid-19 pandemic," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, no. 2020-32, 2020.

[22] H. Allcott, L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Y. Yang, "Polarization and public health: Partisan differences in social distancing during covid-19," *Available at SSRN 3570274*, 2020.

[23] G. Dong, L. Cai, D. Datta, S. Kumar, L. E. Barnes, and M. Boukhechba, "Influenza-like symptom recognition using mobile sensing and graph neural networks," in *Proceedings of the Conference on Health, Inference, and Learning*, 2021, pp. 291–300.

[24] F. Wahle, T. Kowatsch, E. Fleisch, M. Rufer, and S. Weidt, "Mobile sensing and support for people with depression: a pilot trial in the wild," *JMIR mHealth and uHealth*, vol. 4, no. 3, p. e111, 2016.

[25] G. Dong, M. Tang, L. Cai, L. E. Barnes, and M. Boukhechba, "Semi-supervised graph instance transformer for mental health inference."

[26] N. D. Lane, M. Lin, M. Mohammod, X. Yang, H. Lu, G. Cardone, S. Ali, A. Doryab, E. Berke, A. T. Campbell *et al.*, "Bewell: Sensing sleep, physical activities and social interactions to promote wellbeing," *Mobile Networks and Applications*, vol. 19, no. 3, pp. 345–359, 2014.

[27] R. Wang, M. S. Aung, S. Abdullah, R. Brian, A. T. Campbell, T. Choudhury, M. Hauser, J. Kane, M. Merrill, E. A. Scherer *et al.*, "Crosscheck: toward passive sensing and detection of mental health changes in people with schizophrenia," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 886–897.

[28] R. Wang, W. Wang, M. S. Aung, D. Ben-Zeev, R. Brian, A. T. Campbell, T. Choudhury, M. Hauser, J. Kane, E. A. Scherer *et al.*, "Predicting symptom trajectories of schizophrenia using mobile sensing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–24, 2017.

[29] G. M. Harari, S. D. Gosling, R. Wang, F. Chen, Z. Chen, and A. T. Campbell, "Patterns of behavior change in students over an academic term: A preliminary study of activity and sociability behaviors using smartphone sensing methods," *Computers in Human Behavior*, vol. 67, pp. 129–138, 2017.

[30] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.

[31] A. K. Wagner, S. B. Soumerai, F. Zhang, and D. Ross-Degnan, "Segmented regression analysis of interrupted time series studies in medication use research," *Journal of clinical pharmacy and therapeutics*, vol. 27, no. 4, pp. 299–309, 2002.
[32] A. Ram, J. Sunita, A. Jalal, and K. Manoj, "A density based algorithm for discovering density varied clusters in large spatial databases," *International Journal of Computer Applications*, vol. 3, 06 2010.
[33] S. Hoteit, S. Secci, S. Sobolevsky, C. Ratti, and G. Pujolle, "Estimating human trajectories and hotspots through mobile phone data," *Computer Networks*, vol. 64, pp. 296 – 307, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128614000656
[34] M. Boukhechba, P. Chow, K. Fua, B. A. Teachman, and L. E. Barnes, "Predicting social anxiety from global positioning system traces of college students: feasibility study," *JMIR mental health*, vol. 5, no. 3, p. e10101, 2018.
[35] D. Liu, "Play store scraper," 2019, https://pypi.org/project/play-scraper/.
[36] K. Katevas, K. Hänsel, R. Clegg, I. Leontiadis, H. Haddadi, and L. Tokarchuk, "Finding dory in the crowd: Detecting social interactions using multi-modal mobile sensing," in *Proceedings of the 1st Workshop on Machine Learning on Edge in Sensor Systems*, 2019, pp. 37–42.
[37] A. Matic, V. Osmani, A. Maxhuni, and O. Mayora, "Multi-modal mobile sensing of social interactions," in *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. IEEE, 2012, pp. 105–114.
[38] G. Developer, "Activity recognition api," 2020, https://developers.google.com/location-context/activity-recognition.
[39] A. Developer, "Cmmotionactivity," 2020, https://developer.apple.com/documentation/coremotion/cmmotionactivity.
[40] S. Aminikhanghahi and D. J. Cook, "A survey of methods for time series change point detection," *Knowledge and information systems*, vol. 51, no. 2, pp. 339–367, 2017.
[41] B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell, "Stan: A probabilistic programming language," *Journal of statistical software*, vol. 76, no. 1, 2017.
[42] J.-R. Yu, G.-H. Tzeng, and H.-L. Li, "General fuzzy piecewise regression analysis with automatic change-point detection," *Fuzzy sets and systems*, vol. 119, no. 2, pp. 247–257, 2001.
[43] J. L. Bernal, S. Cummins, and A. Gasparrini, "Interrupted time series regression for the evaluation of public health interventions: a tutorial," *International journal of epidemiology*, vol. 46, no. 1, pp. 348–355, 2017.
[44] J. L. Gastwirth, Y. R. Gel, and W. Miao, "The impact of levene's test of equality of variances on statistical theory and practice," *Statistical Science*, pp. 343–360, 2009.
[45] H. Xiong, Y. Huang, L. E. Barnes, and M. S. Gerber, "Sensus: a cross-platform, general-purpose system for mobile crowdsensing in human-subject studies," in *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, 2016, pp. 415–426.