Analyzing the Restoring Impact of Uplift Events on Overwhelmed Teenagers via Microblogs

Qi Li, Yuanyuan Xue, Liang Zhao and Ling Feng *Senior Member, IEEE*

Abstract—As an important concept in psychological theory, restoring is an essential process in human's stress coping system. Timely and efficient restoring of stress could help teenagers get out of overwhelmed status as soon as possible. Previous research has explored the possibility of detecting teenagers' stress series and mining the impact of stressor events from social media. On the contrary, the research on auto-analyzing the restoring ability of uplift events still calls for more exploration, due to the uncertainty and complexity of various restoring situations. In this paper, we give a deep inside into the stress easing function of uplift events on the real data set of 124 high school students. A two-sample based statistical model is conducted to analyze the stressful behavioral correlations when uplift events happened to overwhelmed students from multiple perspectives. Experimental results show that our method could measure the restoring impact of school scheduled uplift events with 69.52% accuracy, and integrating such impact of uplift events helps reduce the stress prediction errors efficiently (with MSE, RMSE, MAPE and MAD reduced by 49.34%, 28.81%, 26.80% and 42.11%, respectively). Our exploration provides guidance for school and parents that which kind of uplift events could help relieve students' overwhelmed stress in both stress prevention and stress early stopping situations.

Index Terms—uplift event, stressor event, teenagers, microblogs

I. INTRODUCTION

Life is always full of ups and downs. According to the transactional model of stress [1], our stress mainly comes from daily hassles. The cumulative stress caused by the small and frequent stressful life events could drain people's inner resources, leading to psychological maladjustment, ranging from depression to suicidal behaviours [2]. As a global public health concern, suicide caused by psychological stress has become the second leading cause of death among young adults in college (Centers for Disease Control and Prevention, 2012).

On the other hand, positive life events (called *uplifts* in psychological theory) such as satisfying social interactions, excellent academic performance and pleasant entertainment activities are conceptualized in psychological literature as exerting a protective effect on emotional distress [3] [4] [5]. Compared with adults, young people exhibit more exposure to uplift events, as well as hassles, due to the immature inner status and lack of experience [6]. Researchers indicate that positive events mitigate the relation between negative events

[1]Dept. of Computer Science and Technology, Centre for Computational Mental Healthcare Research, Institute of Data Science, Tsinghua University, Beijing, China;
[2]Institute of Social Psychology, Xi'an Jiaotong University, Xi'an, China.
E-mail: {liqi13,xue-yy12}@mails.tsinghua.edu.cn,{fengling}@tsinghua.edu.cn, {zhaoliang0415}@xjtu.edu.cn,

and maladjustment in samples of adolescents experiencing family transitions [7]. The written expression of positive feelings has also be shown to prompt increased cognitive re-organization among an undergraduate student group [8].

Positive uplifts can not only help reinforce adolescents' sense of well-being, and help restore the capacity for dealing with stress, but also have been linked to medical benefits, such as improving mood, serum cortisol levels, and lower levels of inflammation and hypercoagulability [9]. Through examining the relationship between self-reported positive life events and blood pressure (BP) in 69 sixth graders, researchers found that increased perceptions of positive life events might act as a buffer to elevated BP in adolescents [10].

The protective effect of uplift events is hypothesized to operate in two ways: directly and indirectly by 'buffering' [4]. In the direct way, the more positive uplift events people experienced, the less distress they experience. While in the indirectly way, positive life events play its role by buffering the effects of negative events on distress. A pioneer experiment conducted by Reich and Zautra provided enlightening evidence for us [11]. In this experiment, sampled college students who reported initial negative events were encouraged to engage in either two or twelve pleasant activities during one-month, and compared with students in the controlled group experiencing no pleasant activities. Results indicated that participants in the two experimental groups reported greater quality of life compared with controlled students, and participants who engaged in twelve uplift events exhibited lower stress compared with whom engaging two or none uplifts, implicating the protective effect of uplift events on adolescents.

Previous exploration for the protective effect of uplift events on adolescents are mostly conducted in psychological area, relying on traditional manpower-driven investigation and questionnaire. The pioneer psychological researches provide us valuable implications and hypothesis, while limited by labor cost, data scale and single questionnaire based method. With the high development of social networks, today adolescents tend to express themselves and communicate with outside world through posting microblogs, at anytime and anywhere. The self-motivated expressions could deliver much information about their inner thoughts and life styles. In recent years, some research on psychological stress analysis based on social network has emerged, from basically detecting stress intensity from microblog content [12], [13], predicting future stress level in time series [14]–[17], to extracting stressor events and stressful intervals [18]. These researches explored applying psychological theories into social network based stress mining, offering effective tools for adolescent stress sensing. Neverthe-

less, few work takes an insight into the restoring function of uplift events, which plays an important role opposite to stress, as the essential way for adolescent psychological stress easing.

In this paper, we aim to continually mine the restoring impact of uplift events leveraging abundant data source from microblogs, to further provide guidance for school and parents that when and which kind of uplift events could help relieve students' overwhelmed stress in both stress prevention and stress early stopping situations. To model such a practical application problem, several challenges exist.

- *How to extract uplift events from microblogs and identify corresponding impact interval*? The impact of uplift events is highlighted when the teen is under stress, with various relative temporal order. Extracting such scenarios from teen's messy microblogs is the first and basic challenge for further analysis.
- *How to qualitatively and quantitatively measure the restoring impact conducted by uplift events*? There are multiple clues related to teens' behaviours from microblogs, i.e., depressive linguistic content, abnormal posting behaviours. The teen might act differently under similar stressful situations when the uplift event happens or not. It is challenging to find such hidden correlation between uplift events and teen's behavioural characters.

Moreover, for different types of uplift events, the restoring impact might be different. And for each individual, the protective and buffering effect for stress might also varies according to the personality. All these questions guide us to solve the problem step by step.

In this paper, we first conduct a case study on real data set to observe the posting behaviours and contents of stressful teens under the influence of uplift events. We conduct the case study on the real data set of 124 high school students associated with the school's scheduled uplift and stressor event list. Several observations are conducted to guide the next step research. Next, we extract uplift events and the corresponding impacted interval from microblogs. We define and extract structural uplift events from posts using linguistic parser model based on six-dimensional uplift scale and LIWC lexicons. Independent stressful intervals (SI) and stressful intervals impacted by uplifts (U-SI) are extracted considering temporal orders. To quantify the restoring impact of uplift events, we describe a teen's stressful behaviours in three groups of measures (stress intensity, posting behaviour, linguistic), and model the impact of uplift events as the statistical difference between the sets of SI and U-SI in two aspects: the two-sample based method is employed for variation detection, and the t-test correlation is conducted to judge the monotonous correlation.

The rest of the paper is organized as follows. We introduce related works in section II, and conduct the data observation in section III. The preliminaries and problem formulation are presented in section IV. We conduct the procedure for extracting uplift events and identifying the impact interval in section V, and introduce the detailed method for analyzing the restoring impact of uplift events in section VI. We present the experimental results in section VII, and discuss the future work in section VIII.

II. RELATED WORK

A. Protective function of uplift events

Many psychological researchers have focused on the restorative function of positive events and emotions with respect to physiological, psychological, and social coping resources. Folkman *et al.* [19] identified three classes of coping mechanisms that are associated with positive emotion during chronic stress: positive reappraisal, problem-focused coping, and the creation of positive events. The author also considered the possible roles of positive emotions in the stress process, and incorporated positive emotion into a revision of stress and coping theory in the work [20]. They conducted a longitudinal study of the care giving partners of men with AIDS and described coping processes that were associated with positive psychological states in the context of intense distress. Cohen *et al.* [4] hypothesized that protective effect of uplift events operates in both directly (i.e., more positive uplift events people experienced, the less distress they experience) and indirectly ways by 'buffering'.

Chang *et al.* [21] investigated the protective effect of positive events in a sample of 327 adults, and found that the positive association between loneliness and psychological maladjustment was found to be weaker for those who experienced a high number of positive life events, as opposed to those who experienced a low number of positive life events. This is assistant with the conclusion made by Kleiman *et al.* [22] that positive events act as protective factors against suicide individually and synergistically when they co-occur, by buffering the link between important individual differences risk variables and maladjustment. Through exploring naturally occurring daily stressors, Ong *et al.* [23] found that over time, the experience of positive emotions functions to assist high-resilient individuals to recover effectively from daily stress. In the survey made by Santos *et al.* [24], strategies of positive psychology are checked as potentially tools for the prophylaxis and treatment of depression, helping to reduce symptoms and for prevention of relapses. Through a three-week longitudinal study, Bono *et al.* [25] examined the correlation between employee stress and health and positive life events, and concluded that naturally occurring positive events are correlated with decreased stress and improved health.

B. Measuring the Impact of Uplift Events

To measure the impact of uplift events, Doyle *et al.* [26] conducted *Hassles and Uplifts Scales*, and concluded that the assessment of daily hassles and uplifts might be a better approach to the prediction of adaptational outcomes than the usual life events approach. Silva *et al.* [27] presented the *Hassles & Uplifts Scale* to assess the reaction to minor every-day events in order to detect subtle mood swings and predict psychological symptoms. To measure negative interpretations of positive social events, Alden *et al.* [28] proposed the interpretation of positive events scale (*IPES*), and analyzed the relationship between social interaction anxiety and the tendency to interpret positive social events in a threat-maintaining manner. Mcmillen *et al.* [29] proposed the *Perceived Benefit Scales* as the new measures of self-reported positive life changes after

traumatic stressors, including lifestyle changes, material gain, increases in selfefficacy, family closeness, community closeness, faith in people, compassion, and spirituality. Specific for college students, Jun-Sheng *et al.* [30] investigated in 282 college students using the *Adolescent Self-Rating Life Events Checklist*, and found that the training of positive coping style is of great benefit to improve the mental health of students.

C. Analyzing adolescent stress from social media

With the high development of social network, researchers tend to digging user' psychological status from the self-expressed public data source. Billions of people record their life, share multi-media content, and communicate with friends through such platforms, e.g., Tencent Microblog, Twitter, Facebook and so on. Inspired by rich microblogging content, Xue *et al.* [12], [13] proposed to detect adolescent stress from single microblog utilizing machine learning methods by extracting stressful topic words, abnormal posting time, and interactions with friends. Lin *et al.* [31] construct a deep neural network to combine the high-dimensional picture semantic information into stress detecting. Based on the stress detecting result, Li *et al.* [14] [16] [15] adopted a series of multi-variant time series prediction techniques (i.e., Candlestick Charts, fuzzy Candlestick line and SVARIMA model) to predict the future stress trend and wave. Taking the linguistic information into consideration, Li *et al.* [17] employed a NARX neural network to predict a teen's future stress level referred to the impact of co-experiencing stressor events of similar companions. All above pioneer work focused on the generation and development of teens' stress, providing solid basic techniques for broader stress-motivated research from social networks.

To find the source of teens' stress, previous work [18] developed a frame work to extract stressor events from microblogging content and filter out stressful intervals based on teens' stressful posting rate. Based on such research background, this paper starts from a completely new perspective, and focuses on the buffering effect of positive events on restoring stress. Thus we push forward the study from how to find stress to the next more meaningful stage: how to deal with stress.

D. Correlation analysis for multivariate time series

Basic correlation analysis methods on time series focused on univariate data have been well studied. As the most widely adopted method, the Pearson correlation analysis [32] measures the linear correlation between two variables X and Y. One inevitable defect is that Pearson correlation is too sensitive to outlier values. To overcome such drawback, Spearman Rank correlation [33] and Kendall Rank correlation [34] are proposed based on Pearson correlation. While Pearson correlation estimates linear relationships, Spearman correlation estimates monotonic relationships (whether linear or not), and are calculated as the Pearson correlation between the rank values of two variables. The Kendall correlation mainly assesses the similarity of the orderings of the data when ranked by each of the quantities. The above correlation methods are usually used to estimate relationship between

TABLE I: Examples of uplift events expressed and extracted from teens' microblogs.

I am really looking forward to the spring outing on Sunday now. (Doer:*I*, Act:*looking forward*, Object:*spring outing*)
My holiday is finally coming [smile]. (Doer:*My holiday*, Act:*coming*, Object:*[smile]*)
First place in my lovely math exam!!! In memory of it. Object:*first place, math, exam, memory*)
You are always here for me like sunshine. (Doer:*You*, Object:*sunshine*)
Thanks all my dear friends to take the party for me. Happiest birthday! (Doer:*friends*, Act:*thanks*, Object:*party, birthday*)
Be yourself. Trust yourself and follow your heart. (Doer:*yourself*, Act:*trust*, Object:*heart*)
Feel proud of our play in the Games. Our class is always the family!!! (Doer:*Our*, Object:*class, family*)
A good film always makes bring comfort and happiness to me. (Doer:*me*, Act:*bring*, Object:*comfort, happiness*)
I know my mom is the one who support me forever, no matter when and where. (Doer:*mom*, Act:*support*)

TABLE II: Examples of stressor events expressed and extracted from teens' microblogs.

| |
|---|
| I don't know how long can I bear the nag. (Doer:*I*, Act:*bear*, Object:*nag*) |
| Parents like to judge everything around me with their emotion. (Doer:*parents*, Act:*judge*, Object:*everything*) |
| Hope that my uncle could revive earlier. (Doer:*my uncle*, Act:*revive*) |
| Every one betrayed me. (Doer:*every one*, Act:*betray*, Object:*me*) |
| I'm too weak to handle such a fierce competition. (Doer:*I*, Act:*too weak to handle*, Object:*competition*) |
| I just felt hurt, depressed, self-abased and sad. (Doer:*I*, Act:*feel hurt, depressed, self-abased and sad*) |
| My holiday is filled with all kinds of homework. (Doer:*My holiday*, Act:*fill with*, Object:*homework*) |
| Unescapably, it's time to go back to school. (Act:*go back*, Object:*school*) |
| When can you be aware of my heart-broken feeling again and again? (Doer:*you*, Act:*be aware of*, Object:*heart-broken feeling*) |

single-dimensional variables, and cannot be adopted directly in our microblog content based scenario.

For multivariate time series analysis, two-sample based methods are widely adopted. Such kind of methods are deduced to check whether two samples come from the same underlying distribution, which is assumed to be statistically unknown. Correspondingly, various kernel [35] and distance-based methods [36] (e.g., the nearest neighbor based method two-sample method) are proposed. Scholkopf *et al.* [35] proposed to transform the distance between two variables and nearest neighbors into a reproducing kernel Hilbert space (RKHS), and solve the problem using Maximum Mean Discrepancy. In work [36], Schilling *et al.* adopted the r-nearest neighbor based method to partition two set of event driven time series data. The global proportion of the right divided neighbors are calculated to estimate whether there exists statistically difference between the two sets. We use the r-nearest neighbor based two-sample method in our problem, thus to measure the distance and correlation between two multi-dimension variables.

TABLE III: Examples of school scheduled uplift events and stressor events.

Type	Date	Work Content	Grade
stressor event	2014/4/16	*first day of mid-term exam*	grade1,2
uplift event	2014/11/5	*campus art festival*	grade1,2,3

Fig. 1: Examples of school related stressor events, uplift events and a student's stress fluctuation
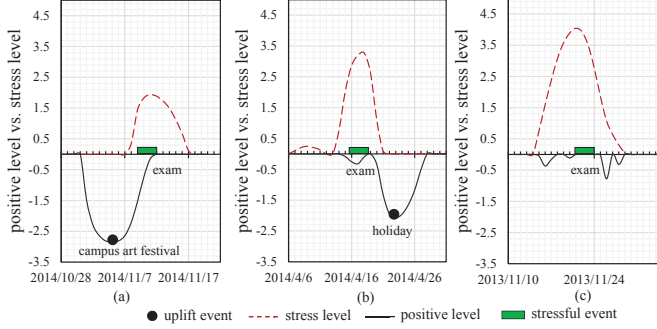
III. DATA OBSERVATION

We built our dataset based on two sources: 1) the microblogs of students coming from Taicang High School, collected from January 1st, 2012 to February 1st, 2015; and 2) list of scheduled school events, with exact start and end time. We filtered out 124 active students according to their posting frequency from over 500 students, and collected their microblogs throughout the whole high school career. Totally 29,232 microblogs are collected in this research, where 236 microblogs per student on average, 1,387 microblogs maximally and 104 posts minimally.

Uplift events and stressor events. The list of weekly scheduled school events (from February 1st, 2012 to August 1st 2017) are collected from the school's official website [1], with detailed event description and grade involved in the event. There are 122 stressor events and 75 uplift events in total. Here we give the examples of scheduled uplift and stressor events in high school life, as shown in Table III. There are 2-3 stressor events and 1-2 uplift event scheduled per month.

Stress detected from microblogs. Since our target is to observe the restoring impact of uplift events for teenagers under stress. Based on previous research [12], we detected the stress level (ranging from 0 to 5) for each post; and for each student, we aggregated the stress during each day by calculating the average stress of all posts. The positive level (0-5) of each post is identified based on the frequency of positive words (see Section 5 for details). Figure 1 shows three examples of a student's stress fluctuation during three mid-term exams, where the uplift event *campus art festival* was scheduled ahead of the first exam, the uplift event *holiday* happened after the second exam, and no scheduled uplift event was found nearby the third exam. The current student exhibited

[1]http://stg.tcedu.com.cn/col/col82722/index.html

differently in above three situations, with the stress lasting for different length and with different intensity.

To further observe the influence of uplift events for students facing stressor events, we statistic all the stressful intervals [18] detected surround the scheduled examinations over the 124 students during their high school career. For each student, we divide all his/her stressful intervals into two sets: 1) stressful intervals under the influence of neighbouring uplift events (e.g., *Halloween activity*), and 2) independent stressful intervals. Figure 2 shows five measures of each student during the above two conditions: the *accumulated stress*, the *average stress* (per day), the *length of stressful intervals*, the *frequency of academic topic words*, and the *ratio of academic stress among all types of stress*. For each measure, we calculate the average value over all eligible slides for each student.

Findings. Comparing each measure in scheduled exam slides under the two situations: 1) existing neighbouring uplift events or 2) no neighbouring scheduled uplift events, we find that students during exams with neighbouring uplift events exhibit less average stress intensity (both on accumulated stress and average stress), and the length of stress slides are relatively shorter. Further, we statistic the frequency of academic related topic words for each exam slide (as listed in Table IV), and look into the ratio of academic stress among all five types of stress. Results in Figure 2 shows that most students talked less about the upcoming or just-finished exams when uplift events happened nearby, with lower frequency and lower ratio. The stress intensity and type distribution detected from each student's microblogs varies due to personal life experience, posting habits and express styles. The statistic result shows clues about the stress-relieving ability of scheduled uplift events, and thus helps shape our problem as how to quantify the influence of uplift events, thus to provide further guidance for planning campus activities to help relive high school students' psychological stress effectively.

TABLE IV: Examples of academic related topic words.

exam, fail, review, score, grade, test paper, rank, pass, math, chemistry homework, recite, regress, fall behind, tension, stressed out, physics, nervous, mistake, answer, question, puzzle, difficult, lesson, careless,

IV. PRELIMINARIES

Based on the observation and psychological theory, we conduct our research under the assumption that uplift events can ease teenagers' stress, namely the positive restoring impact of uplift events. While stressor events stimulate human's stress, uplift events bring positive influence and stronger restoring ability to stressed people in various situations with multi-types [3] [4] [5]. Taking the three stress curves in Figure 1 for example, comparing the stress curves *a)*, *b)* with *c)*, when an uplift event (*campus art festival, holiday* here) happens, the overall stress intensity during the stressful period is reduced. An uplift event might happen before a teen's stress caused by scheduled stressor events (*example a*), conducting lasting easing impact; Meanwhile, an uplift event might also happen during (*example b*) or at the end of the stressful period, which might promote the teen out of current stressful status more

Fig. 2: Compare students' stress during exam intervals in two situations: 1) affected by neighboring uplift events (U-SI), 2) no uplift events occurred nearby (SI)
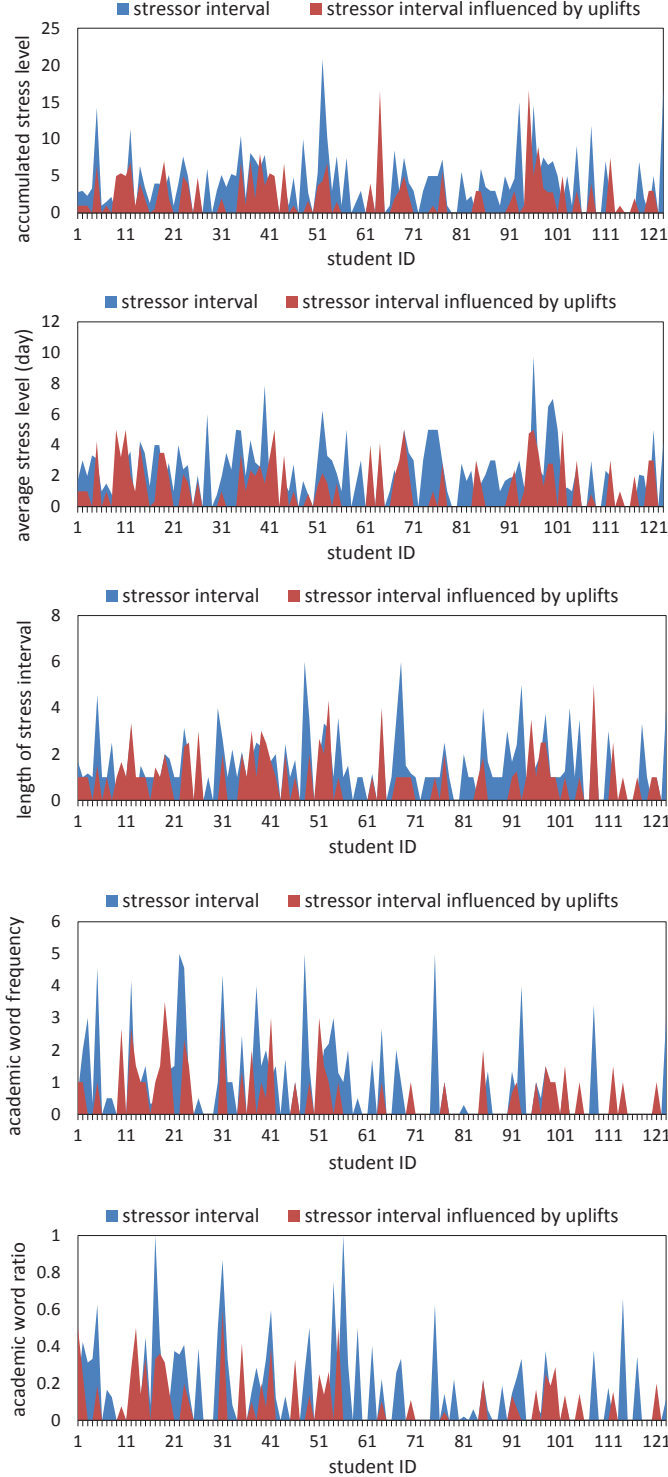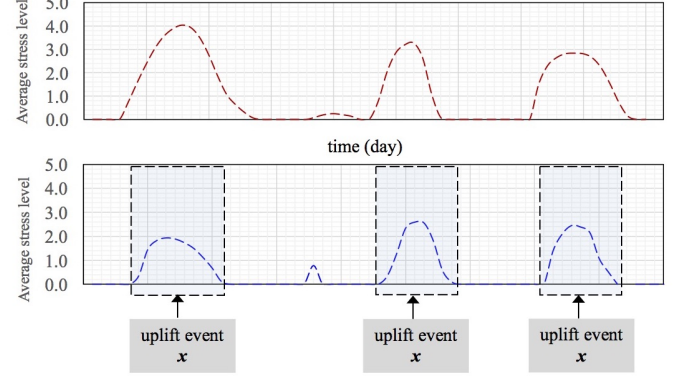

Fig. 3: Illustration of SI and U-SI stressful intervals.

quickly. To study the restoring impact of an uplift event, we structure its impact from three aspects:

- **Impact interval of uplifts**. To study the impact of uplift events from microblogs, two fundamental factors are identifying the exact time when the uplift event happens, and the corresponding stressful interval it impacts. The temporal order between uplift events and the teen's stress series varies in different situations, and its a challenge to match the uplift event to the right stressful interval it actually impacts.

- **Restoring patterns of uplifts**. As the restoring impact of uplift events relieves the teen's stress and exhibits in multiple aspects (e.g., the changes in posting behavior, linguistic expression, and stress intensity from microblogs), it's meaningful to extract the stress-related restoring patterns and describe the restoring impact of uplift events structurally.

- **Quantified impact of uplifts**. Different types of uplift events might conduct restoring impact with different intensity. In this paper, the ultimate problem we target to solve is how to quantify the restoring impact both qualitatively and quantitatively from teenagers microblogs.

Given an uplift event with specific type, we consider its restoring impact by comparing the teen's behavioral measures under two situations. As illustrated in Figure 3, in the original situation (i.e., sub-series A), the teen's stress is caused by a stressor event, lasting for a period, and no other intervention (namely, uplift event) occurs. We call the set of such stressful intervals as **SI**. In the other comparative situation (i.e., sub-series B), the teen's stressful interval caused by the same type of stressor is impacted by an uplift event with type x, we call the set of such stressful intervals as **U-SI**. Thus the difference under the two situations SI and U-SI could be seen as the restoring impact conducted by the uplift event of type x.

Next, we give the formal definition for uplift events and stressor events from the perspective of linguistic structure.

Definition 1: **Uplift event**. Let $u = [type, \{role, act, descriptions\}]$ be an uplift event, where the element *role* is the subject who performs the *act*, and *descriptions* are the key words related to u. According to psychological scales [30],

[37], teenagers' uplift stressors mainly focus on six aspects, as $\mathbb{U}$ = { *entertainment', 'school life', 'family life', 'pear relation', 'self-cognition', 'romantic'*}, $\forall u$, $u._{type} \in \mathbb{U}$.

Definition 2: **Stressor event**. Similar to stressor event, let $e = [type, \{role, act, descriptions\}]$ be a stressor event. According to psychological questionnaires [26], [38]–[40], we classify stressor events into five types, as $\mathbb{S}$ = { *'school life', 'family life', 'pear relation', 'self-cognition', 'romantic'*}, $\forall e$, $e._{type} \in \mathbb{S}$.

The examples of teens' microblogs describing uplift events and stressor events are listed in Table I and Table II. For the post '*I have so much homework today!!!*', its elements are *role* = 'I', *act*='have', *descriptions* = 'homework', and the *type* = 'school life'. For the post '*Expecting Tomorrow' Adult Ceremony[Smile][Smile]* ', we translate it into *act* = 'expecting', *description* = 'Adult Ceremony', and the *type* = 'self-cognition'.

Problem: For an uplift event u with type $U^{'}$, a stressor event e with type $S^{'}$, let $F:(u, U^{'}, e, S^{'}) \rightarrow A$ (A is a multidimensional vector) be the restoring influence of uplift event u conducted on the stress caused by stressor event e. We aim to quantify such influence A from multiple views.

V. IDENTIFY UPLIFTS AND IMPACT INTERVAL

In this section, we first introduce the procedure to extract uplift events and stressful intervals from teens' microblogs. The uplift events are extracted from microblogs applying LTP natural language processing segmentation and parser models [41]. Stressful intervals are identified using probability based statistical method according to the teen's stressful posting frequency. We judge whether each stressful interval is correlated with neighboring uplift events, thus to classify all stressful intervals into two sets: SI and U-SI.

A. Uplift Events

Lexicon. We construct our lexicon for six-dimensional uplift events from two sources. The basic positive affect words are selected from the psychological lexicon SC-LIWC (e.g., *expectation*, *joy*, *love* and *surprise*) [42]. Then we build six uplift event related lexicons by expanding the basic positive words from the data set of teens' microblogs, and divide all candidate words into six dimensions corresponding to six types of uplift events, containing 452 phrases in *entertainment*, 184 phrases in *family life*, 91 phrases in *friends*, 138 phrases in *romantic*, 299 phrases in *self-recognition* and 273 phrases in *school life*, with totally 2,606 words, as shown in Table V. Additionally, we label *role* words (i.e., *teacher*, *mother*, *I, we*) in the uplift lexicon.

Parser relationship. For each post, after word segmentation, we parser current sentence to find its linguistic structure, and then match the main linguistic components with uplift event related lexicons in each dimension. The parser model in Chinese natural language processing platform [41], [43] is adopted in this part, which identifies the central verb of current sentence first, namely the *act*, and constructs the relationship between the central verb and corresponding *role* and *objects* components. By searching these main elements in uplift event

TABLE V: Examples of topic words for uplift events.

Dimension	Example words	Total
entertainment	hike, travel, celebrate, dance, swimming, ticket, shopping, air ticket, theatre, party, Karaoke, self-driving tour, game, idol, concert, movie,	452
school life	reward, come on, progress, scholarship, admission, winner, diligent, first place, superior, hardworking, full mark, praise, goal, courage,	273
romantic	beloved, favor, guard, anniversary, concern, tender, deep feeling, care, true love, promise, cherish, baby, kiss, embrace, dating, reluctant,	138
pear relation	listener, company, pour out, make friends with, friendship, intimate, partner, team-mate,	91
self-cognition	realize, achieve, applause, fight, exceed, faith, confidence, belief, positive, active, purposeful,	299
family life	harmony, filial, reunite, expecting, responsible, longevity, affable, amiability, family, duty,	184

related lexicons, we identify the existence and type of any uplift event. Due to the sparsity of posts, the *act* might be empty. The *descriptions* are collected by searching all nouns, adjectives and adverbs in current post. In such way, we extract structured uplift events from teens' microblogs.

B. Impact Interval

When an uplift event happens, it can conduct positive impact on stressed teens, which may exhibit in multi-perspectives, including the length of corresponding stressful interval, the stress intensity and stressful behaviors during the interval. Basically, in this part, we identify stressful intervals from timeline thus to support further quantifying the influence of an uplift event. Splitting interval is a common time series problem, and various solutions could be referred. Here we identify the teen's stressful intervals in three steps. In the first step, we extract uplift events, stressor events and filter out candidate intervals after a smoothing process. Then, applying the Poisson based statistical method proposed in [18], we judge whether each candidate interval is a confidential stressful interval. Finally, we divide the stressful intervals into two sets: the SI set and the U-SI set, according to its temporal order with neighboring uplift events.

Smooth stress series. Since a teen's stress series $S = \{s_1, s_2, \cdots, s_n\}$ detected from microblogs (aggrated by day) are discrete points, we adopt the loess (local regression using weighted linear least squares and a 2nd degree polynomial) method [44] to highlight characteristics of the stress curve. In loess method, we need to set the parameter *span*, which represents the percentage of the selected data points in the whole data set (ranging from 0 to 1), and determines the degree of smoothing. We discuss the influence of different *span* settings on the result of interval segmentation in the experiment section.

Select candidate intervals. We filter out candidate stressful intervals on the smoothed stress series $S^{'} = \{s^{'}_1, s^{'}_2, \cdots, s^{'}_n\}$. Let the sub-series $w_{<a,b>} = [s^{'}_a, \cdots, s^{'}_b]$ as a *wave*, where $s^{'}_v = vally(w_{<a,b>})$ is the minimum stress value, $s^{'}_p = peak(w_{<a,b>})$ is the maximal stress value during $\{s^{'}_a, \cdots, s^{'}_b\}$, and $s^{'}_a \leq s^{'}_{a+1} \leq \cdots \leq s^{'}_p \leq s^{'}_{p+1} \leq \cdots \leq s^{'}_b$.

A candidate interval $I = <w_1, \cdots, w_i, \cdots, w_m>$ is identified with following rules:

- $s_1' = 0$, $s_m' = 0$. $\forall s_j' \in \{s_2', \cdots, s_{m-1}'\}$, $s_j' > 0$.
- Let wave w_i be the biggest wave in current candidate interval, with $peak(w_i) = \omega$, $\forall$ wave $w_j \in I$, $peak(w_j) <= peak(w_i)$.
- For waves w_k before the interval biggest wave w_i , i.e., $\forall w_k \in <w_1, \cdots, w_{i-1}>$, $peak(w_{k+1}) >= peak(w_k)$, $vally(w_{k+1}) >= peak(w_k)$.
- For waves w_k behind the interval biggest wave w_i, i.e., $w_k \in <w_i, \cdots, w_m>$, $peak(w_{k+1}) <= peak(w_k)$, $vally(w_{k+1}) <= peak(w_k)$.

Identify stressful intervals. For each candidate interval, a Poisson based probability model [18] is adopted to measure how confidently the current interval is a stressful interval. Here a teen's stressful posting rate under stress (λ_1) and normal conditions (λ_0) are modeled as two independent poisson process:

$$Pr[N = n|\lambda_i] = \frac{e^{-\lambda_i T}(\lambda_i T)^n}{n!} \qquad (1)$$

where $i \in \{0, 1\}$, $n = 0, 1, \cdots, \infty$. We expect that $\lambda_1 > \lambda_0$, and measure the probability as $P(\lambda_1 > \lambda_0|N_1, T_1, N_0, T_0)$, where N_1, N_0 are the number of stressful posts, and T_1, T_0 are time duration corresponding to λ_1 and λ_0. Without loss of generality, we assume a Jeffreys non-informative prior on λ_1 and λ_0, and infer the posterior distribution $P(\lambda_1|N_1)$ and $P(\lambda_0|N_0)$ according to Bayes Rule. Thus for current interval I_1 and historical normal interval I_0, the quantified probability $\beta = P(\lambda_1 > \lambda_0|I_1, I_0) \in (0, 1)$ indicates the confidence whether I_1 is a stressful interval.

Intervals impacted by uplift events. In this part, we filter out two sets of stressful intervals: stressful intervals without the impact of uplift events (SI), and stressful intervals under the impact of uplift events (U-SI). For a detected stressful interval $I = <t_1, \cdots, t_n>$, we consider the temporal order between I and any detected uplift event u happened at time point t_u:

- If the uplift event u happens during the stressful interval, i.e., $t_u \in [t_1, t_n]$, the uplift interval I is judged as $I \in SI$.
- For the uplift event happening nearby a stressful interval, we also consider the probability that it conducts impact on the teen's stressful interval. Here the gap between t_u and I is limited to ξ, i.e., if $t_u \in [t_1 - \xi, t_1) \cup (t_n, t_n + \xi]$, then $I \in SI$.

If a stressful interval satisfies none of the above conditions, we classify it into the U-SI set.

VI. IMPACT OF UPLIFT EVENTS

To quantify the restoring impact of uplift event, in this section, we propose to model the impact as the teen's behavioral differences in two cases: 1) stressful intervals unaffected by uplift events (SI), and 2) stressful intervals impacted by uplift events (U-SI). Multiple stress and positive emotion related measures are proposed to describe the correlation between SI and U-SI, and we quantify such differences as correlations using a two-sample based statistical method.

A. Restoring Patterns and Behavioral Measures

To extract the restoring patterns of each type of uplift events, we describe a teen's positive and stressful behavioral measures in SI and U-SI sets from three aspects: posting behavior, stress intensity, and linguistic expressions.

Posting behavior. Stress could lead to a teen's abnormal posting behaviors, reflecting the teen's changes in social engagement activity. For each stressful interval, we consider three measures of posting behaviors in each time unit (day), and present each measure as a corresponding series. The first measure is *posting frequency*, representing the total number of posts per day. Research in [18] indicates that overwhelmed teens usually tend to post more to express their stress for releasing and seeking comfort from friends. Further, the second measure *stressful posting frequency* per day is based on previous stress detection result and highlights the stressful posts among all posts. Similarly, the third measure is the *positive posting frequency*, indicating the number of positive posts per day. The forth measure *original frequency* is the number of original posts, which filters out re-tweet and shared posts. Compared to forwarded posts, original posts indicate higher probability that teens are talking about themselves. Thus for each day in current interval, the teen's posting behavior is represented as a four-dimension vector.

Stress intensity. We describe the global stress intensity during a stressful interval through four measures: *sequential stress level, length, RMS,* and *peak*. Basically, *stress level* per day constructs a sequential measure during a stressful interval, recording stress values and fluctuation on each time point. The *length* measures the lasting time of current stressful interval. As uplift events might conduct impact on overwhelmed teens, and postpone the beginning or promote the end of the stressful interval, we take the *length* as a factor representing the interval stress intensity. To quantify the intensity of fluctuations for stress values, we adopt the *RMS* (root mean square) of stress values through the interval as the third measure. In addition, the *peak* stress value is also a measure to show the maximal stress value in current interval.

Linguistic expressions. We extract the teen's positive and stressful expressions from the content of posts in SI and U-SI sets, respectively. The first linguistic measure is the frequency of *positive word*, which represents the positive emotion in current interval. The second measure is the frequency of *uplift event topic words* in six dimensions, reflecting the existence of uplift events. Another important factor is wether existing *self-mentioned words* (i.e., 'I','we','my'). Self-mentioned words show high probability that the current stressor event and stressful emotion is related to the author, rather than the opinion about a public event or life events about others.

Except uplift-related linguistic descriptions, we also take stressful linguistic characters as measures, which is opposite with positive measures, while also offers information from the complementary perspective. The first stressful linguistic measure is the frequency of *stressor event topic words* in five dimensions, which represents how many times the teen mentioned a stressor event, indicating the degree of attention for each type of stressor event. The frequency of *pressure*

words is the second stressful linguistic measure, reflecting the degree of general stress emotion during the interval. We adopt this measure specifically because in some cases teens post very short tweets with only stressful emotional words, where topic-related words are omitted.

Next, based on the posting behavior, stress intensity and linguistic measures from both the stressful and positive views, we quantify the difference between SI and U-SI sets, thus to measure the impact of uplift events.

B. Quantify the Correlation

In our problem, there are two sets of stressful intervals to compare: the SI set and the U-SI set, containing stressful intervals unaffected by uplift events and stressful intervals impacted by uplift events, respectively. The basic elements in each set are stressful intervals, i.e., the sequential stress values in time line, which are modeled as multi-dimensional points according to the three groups of measures in section VI-A. Thus we formulate this comparison problem as finding the correlation between the two sets of multi-dimension points. Specifically, we adopt the multivariate two-sample hypothesis testing method [45], [46] to model such correlation. In this two-sample hypothesis test problem, the basic idea is judging whether the multi-dimension points (i.e., stressful intervals) in set SI and set U-SI are under different statistical distribution. Assuming the data points in SI and U-SI are randomly sampled from distribution $F^{(1)}$ and $F^{(2)}$, respectively, then the hypothesis is denoted as:

$$H_0 : F^{(1)} = F^{(2)} \quad versus \quad H_1 : F^{(1)} \neq F^{(2)}. \quad (2)$$

Under such hypothesis, H_0 indicates points in SI and U-SI are under similar distribution, while H_1 means points in SI and U-SI are under statistically different distributions, namely uplift events have conducted obvious restoring impact on current stressed teen. Next, we handle this two-sample hypothesis test problem based on both positive and stressful behavioral measures (i.e., *posting behavior*, *stress intensity* and *linguisitc expressions*), thus to quantify the restoring patterns of uplift events from multi perspectives.

As a classic statistical topic, various algorithms have been proposed to solve the two-sample hypothesis testing problem. Since each point in the two sets (SI and U-SI) is depicted in multi-dimensions, here we take the KNN (k nearest neighbors) [36] based method to judge the existence of significant difference between SI and U-SI. For simplify, we use the symbol A_1 to represent set SI, and A_2 represent set U-SI, namely A_1 and A_2 are two sets composed of stressful intervals. In the KNN algorithm, for each point ℓ_x in the two sets A_1 and A_2, we expect its nearest neighbors (*the most similar points*) belonging to the same set of ℓ_x, which indicates the difference between the points in the two cases.

Following section VI-A, for each teen, three groups of behavioral measures are considered: *posting behavior*, *stress intensity* and *linguistic expressions*, indicated as $< D_p, D_s, D_l >$, respectively. To measure the correlation for each group of positive and stressful behavioral measures, the Euclidean distance is adopted to calculate the distance of structured points in A_1 and A_2.

For each point $\ell x \in A = A_1 \bigcup A_2$, let $NN_r(\ell_x, A)$ be the function to find the $r-$th nearest neighbor of ℓ_x. Specifically, according to the three group of measures, three sub-functions of $NN_r(.)$ are defined as $PNN_r(.)$, $SNN_r(.)$ and $LNN_r(.)$, corresponding to the teen's posting behaviors, stress intensity and linguistic expressions in each stressful interval, respectively.

For point ℓ_x with posting behavior matrix $\boldsymbol{D_p^x}$, stress intensity matrix $\boldsymbol{D_s^x}$, and linguistic expression matrix $\boldsymbol{D_l^x}$, the r-th nearest neighbor of ℓ_x in each measure is denoted as:

$$PNN_r(\ell_x, A) = \{y | min\{||\mathbf{D}_p^x - \mathbf{D}_p^y||_2\}, y \in (A/\ell_x)\}$$
$$SNN_r(\ell_x, A) = \{z | min\{||\mathbf{D}_s^x - \mathbf{D}_s^z||_2\}, z \in (A/\ell_x)\}$$
$$LNN_r(\ell_x, A) = \{w | min\{||\mathbf{D}_l^x - \mathbf{D}_l^w||_2\}, w \in (A/\ell_x)\}$$
$$(3)$$

The r-th nearest neighbor considering all three groups of measures is denoted as:

$$NN_r(\ell_x, A) = \{v | min\{a \times ||\mathbf{D}_p^x - \mathbf{D}_p^v||_2 + \quad (4)$$
$$b \times ||\mathbf{D}_s^x - \mathbf{D}_s^v||_2 + c \times ||\mathbf{D}_l^x - \mathbf{D}_l^v||_2\}, v \in (A/\ell_x)\} \quad (5)$$

In this study, we set $a = b = c = 1/3$. Next, let $I_r(\ell_x, A1, A2)$ be the function denoting whether the r-th nearest neighbor is in the same set with ℓ_x:

$$I_r(\ell_x, A_1, A_2) = \begin{cases} 1, & if \ell_x \in A_i \&\& NN_r(\ell_x, A) \in A_i, \\ 0, & otherwise \end{cases}$$
$$(6)$$

Let $T_{r,n}$ denote the proportion that pairs containing two points from the same set among all pairs formed by $\ell_x \in A$ and its k nearest neighbors:

$$T_{k,n} = \frac{1}{n \times k} \sum_{i=1}^{n} \sum_{j=1}^{k} I_j(x, A_1, A_2) \quad (7)$$

The value of $T_{k,n}$ shows how differently the points in the two testing sets (SI and U-SI) perform in three groups of measures. If the value of $T_{r,n}$ is close to 1, it can be shown that the two underlying distributions $F^{(1)}$ and $F^{(2)}$ for SI and U-SI are significantly different, indicating current uplift events conduct obvious restoring impact on the teens' stress series. Let $\lambda_1 = |A_1|$ and $\lambda_2 = |A_2|$, the statistic value Z is denoted as:

$$Z = (nr)^{1/2}(T_{r,n} - \mu_r)/\sigma_r \quad (8)$$
$$\mu_r = (\lambda_1)^2 + (\lambda_2)^2 \quad (9)$$
$$\sigma_r^2 = \lambda_1 \lambda_2 + 4\lambda_1^2 \lambda_2^2 \quad (10)$$

where μ_r is the expectation and σ_r^2 is the variance of Z. Based on hypothesis test theory [46], when the size of the testing set (λ_1 and λ_2) are large enough, Z obeys a standard Gaussian distribution.

Thus we judge whether the uplift events have conducted significant restoring impact on the teen's stress series as follows: if $f(SI, USI) = (nr)^{1/2}(T_{r,n} - \mu_r)/\mu_r^2 > \alpha$ ($\alpha = 1.96$ for $P = 0.025$), then the hypothesis H_1 is true.

C. Temporal Order

To measure the intensity of stress changes in the two sets (SI and U-SI) of intervals, for each stressful interval, we further

quantify its stress intensity by comparing with the front and rear adjacent intervals, respectively. For a stressful interval $I = <t_i, t_{i+1}, \cdots, t_j>$, let $I^{front} = <t_m, \cdots, t_{i-1}>$ be the adjacent interval before I, and $I^{rear} = <t_{j+1}, \cdots, t_n>$ be the rear adjacent interval of I. The length of I^{front} and I^{rear} are set to $|I|$. For the set of stressful intervals SI composed of $<I_1, I_2, \cdots, I_N>$, the corresponding sets of adjacent front and rear intervals are denoted as SI^{front} and SI^{rear}. Similarly, for the set of stressful intervals $U - SI = <UI_1, UI_2, \cdots, UI_M>$ impacted by uplift events, the corresponding sets of adjacent front and rear intervals are denoted as USI^{front} and USI^{rear}. We compare the intensity of stress changes in following four situations, where $g(.)$ is the function comparing two sets.

① $g(SI, SI^{front})$ returns if intensive change happens when stressful intervals begin.

② $g(SI, SI^{rear})$ returns if the teen's stress change intensively after the stressful intervals end.

③ $g(USI, USI^{front})$ returns if intensive change happens when stressful intervals affected by uplift events appears.

④ $g(USI, USI^{rear})$ returns if stress change intensively after stressful intervals affected by uplift events end.

In our problem, taking the comparison between SI and SI^{rear} for example, the basic computation element $I_k \in SI \cup SI^{rear}$ in both sets is a multi-dimension interval. Here we adopt the t-test method as the intensity computation function $g(.)$. The t-test algorithm measures if intensive positive or negative monotonous correlation exists between two sample sets. The function $g(.) = t_{score} \in (-1,1)$ is represented as:

$$g(SI, SI^{rear}) = \frac{\mu_{SI} - \mu_{SI^{rear}}}{\sqrt{\frac{(n_1-1)\sigma_{SI}^2 + (n_2-1)\sigma_{SI^{rear}}^2}{n_1+n_2-2}\left(\frac{1}{n_1} - \frac{1}{n_2}\right)}} \quad (11)$$

where μ_{SI} and $\mu_{SI^{rear}}$ are the mean stress values of intervals in sets SI and SI^{rear}, and σ_{SI} and $\sigma_{SI^{rear}}$ are the variance stress values of intervals in sets SI and SI^{rear}, respectively. If $g(SI, SI^{rear}) > \alpha$, stress intensity in SI^{rear} show significant decrease compared with SI (monotonic negative effect). If $g(SI^{front}, SI) < -\alpha$, stress intensity in SI show significant increase compared with SI^{front} (monotonic positive effect). Here we adopt $\alpha = 1.96$, $P = 0.025$. We conduct comparison for above four situations, to observe whether the occurrence of uplift events relieve the monotonic negative effect of $g(SI, SI^{rear})$ and the monotonic positive effect of $g(SI^{front}, SI)$.

D. Overall Algorithm

The overall pipeline for identifying the restoring impact of uplift events is presented in algorithm 1. For an uplift event u with type U', a stressor event e with type S', the overall algorithm is represented as $F : (u, U', e, S') \to A$. To quantify the restoring impact of uplift events, we first extract uplift events and stressful intervals from the teen's microblogs. All stressful intervals are classified into two sets: the set of stressful intervals affected by uplift events (SI), and the set of stressful intervals impacted by uplift events (U-SI). To

---

Algorithm 1: Identify the restoring impact of uplift events.

Input: SI: Set of stressful intervals caused by S';
          U-SI: Set of stressful intervals affected by U';
Output: Restoring impact of uplift U' on stressor S': A

1 **Initialize:** $H_1, H^{front}, H^{rear} = false$;
2 **if** $f(SI, USI) > \alpha$ **then**
3 $\quad \lfloor H_1 = ture$;
4 **if** $g(SI, SI^{rear}) > \alpha$ && $g(SI, SI^{rear}) > g(USI, USI^{rear})$ **then**
5 $\quad \lfloor H^{rear} = true$;
6 **if** $g(SI^{front}, SI) < -\alpha$ && $g(SI, SI^{front}) < g(USI, USI^{front})$ **then**
7 $\quad \lfloor H^{front} = true$;
8 **return** $A = <H_1, H^{front}, H^{rear}>$;

---

judge if SI are statistically different with U-SI, next, the two-sample hypothesis testing method is conducted on the two sets with multi positive and stressful measures (posting behavior, stress intensity and linguistic expressions). To further judge the monotonous restoring intensity of each type of uplift events, the final step comes to comparing SI and U-SI with adjacent intervals, respect to temporal order.

VII. EXPERIMENTS

In this section, we first identify uplift events and corresponding restoring performance from microblogs, and compare the results with scheduled positive events collected from the school's official website. Further, we integrate the impact of uplift events into traditional stress prediction in time line, and verify whether the restoring patterns of each type of uplift events could help improve the prediction performance, thus to show the effectiveness of our method for quantifying the impact of uplift events, as well as the easing function of uplift events during the process of dealing with stress.

A. Experimental setup

1) Data set: We take the same data set in previous case study (section III]), which contains 29,232 microblogs of 124 students from Taicang High School of Jiangsu Province from Tencent Microblog platform[2], posted from 2012/1/1 to 2015/2/1. To protect the privacy, all usernames are anonymized in the experiment. We collected the school's weekly plans published on its official website. Among the 273 school scheduled events filtered out, 122 events are study-related stressor events, and 75 are uplift events. Based on the scheduled time of stressor and uplift events, we identified 518 scheduled academic related stressful intervals (SI) and 259 academic stressful intervals impacted by four typical scheduled uplift events (U-SI) (in Table VII) from the students' microblogs. Further experiments are conducted on the two sets to verify the impact of uplift events from multi perspectives.

[2]http://t.qq.com/

TABLE VI: Monotonous stress intensity changes in U-SI and SI intervals compared with adjacent intervals.

	School life		Romantic		Peer relationship		Self-cognition		Family life		All types	
	U-SI	SI	U-SI	SI	U-SI	SI	U-SI	SI	U-SI	SI	U-SI	SI
# Interval	365	514	536	587	128	391	564	609	321	481	1,914	2,582
Front $\rightarrow$ I	0.7260	0.7879	0.6903	0.7751	0.7422	0.8159	0.7004	0.7767	0.6791	0.7796	0.7017	0.7851
I $\rightarrow$ rear	0.7589	0.7840	0.7463	0.7905	0.7813	0.8261	0.7500	0.7915	0.7414	0.7942	0.7513	0.7955

2) Baseline methods: We adopt the commonly used Pearson correlation algorithms to compare with the two sample statistical method in this study. As a widely adopted measure of the linear correlation between two variables, the Pearson correlation method computes a value in the range $(-1, 1)$, where 1 denotes total positive linear correlation, 0 denotes no linear correlation, and -1 is total negative linear correlation. The Pearson correlation between two variables X and Y is presented as:

$$\rho_{X,Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \qquad (12)$$

where μ_X and μ_Y is the mean value of X and Y, respectively; $E(.)$ is the expectation function; and $E[(X - \mu_X)(Y - \mu_Y)]$ is the covariance of the two variances X and Y.

In our two sample statistical procedure, to calculate the distance between two n dimension points X and Y, we adopt the L_2 measures, denoted as $L_2 = \sqrt{\sum_{i \in [1,n]}(X_i - Y_i)^2}$ (the Euclidean metric).

3) Metrics: To check the performance of uplift event extraction and the validation of our assumption, we first compare the experimental results with the school's scheduled positive events. Further, to measure the effectiveness of our method for quantifying the restoring impact of uplift events, we integrate the impact of uplift events into traditional stress series prediction problem, and verify whether the restoring pattern of uplift events could help improve the prediction performance. Here we choose the SVARIMA (Seasonal Autoregressive Integrated Moving Average) algorithm [47], which is proved to be suitable for teens' linear stress prediction problem [14], due to the seasonality and non-stationarity of teens' stress series. The basic stress prediction is conducted using SVARIMA approach, in the set of stressful intervals impacted by uplift events (U-SI). Since stressor events cause the fluctuation of stress series from normal states, to eliminate the interference, we simply consider the prediction problem in those stressful intervals rather than randomly picked out stress series. Further, the impact of uplift events are utilized as adjust values to modify the stress prediction result.

Four metrics are adopted to measure the stress forecasting problem, where *MSE*, *RMSE* and *MAD* measure absolute errors and *MAPE* measures relative errors. For all real stress value $\overline{s_i}$ and predicted stress value s_i in predicting sequence $< s_1, \cdots, s_n >$, $MSE = \frac{1}{n}\sum_{i \in [1,n]}(s_i - \overline{s_i})^2$, $RMSE = \frac{1}{n}\sqrt{\sum_{i \in [1,n]}(s_i - \overline{s_i})^2}$, $MAD = \frac{1}{n}\sum_{i \in [1,n]}|s_i - \overline{s_i}|$, and $MAPE = \frac{1}{n}\sum_{i \in [1,n]}|s_i - \overline{s_i}|/s_i$.

TABLE VII: Quantify the impact of scheduled uplift school events using KTS and baseline method.

	Practical activity	*Holiday*	*New year party*	*Sports meeting*	*All*
Size of U-SI	219	339	235	226	1,019
Pearson	54.52%	78.39%	63.39%	58.74%	69.52%
KTS[1]	55.65%	70.97%	56.45%	54.84%	65.32%

[1]KTS denotes the knn-based two sample method adopted in this research.

Fig. 4: Correlation towards each types of stressor events
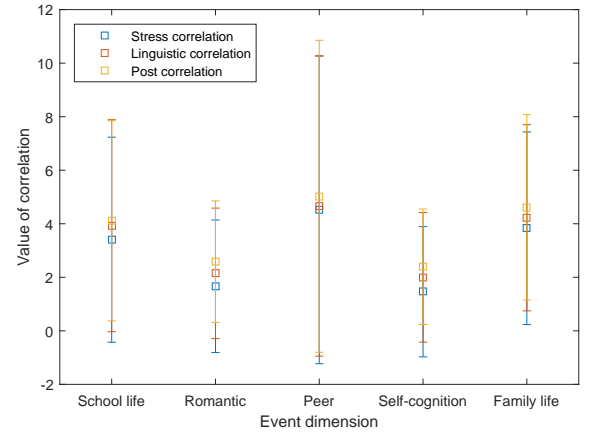
B. Impact and monotonous stress trends of uplift events

Basically, we focused on four kinds of scheduled positive events: *practical activity*, *holiday*, *new year party* and *sports meeting*. For each of the four scheduled positive events, we quantify the restoring impact and temporal order based on corresponding SI and U-SI interval sets of the 124 students. Table VII shows the experimental results, where 54.52%, 78.39%, 63.39%, 58.74% significant restoring impact are detected for the four specific scheduled positive events, respectively, with the total accuracy to 69.52%. For comparison, our knn-based two sample method (denoted as *KTS*) outperforms the baseline method with the best improvement in *new year party* to 10.94%, and total improvement to 6%. The correlation of uplift events for *linguistic expression*, *stress intensity* and *post behaviors* towards five types of stressor events are shown in Figure 4, among which the uplift events conduct most intensive restoring impact in 'school life' and 'peer relationship' dimensions.

Further more, to verify the monotonous stress changes when an uplift event impacts a stressful interval, we collected 1,914 stressful intervals in U-SI, and 2,582 stressful intervals

TABLE VIII: Compare the stress forecast performance under three restoring patterns of uplift events.

	None				Uplift (L)				Uplift (S)				Uplift (P)			
	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE	MAD
School life	0.0856	0.2926	0.4852	0.1146	0.0259	0.1609	0.2991	0.0923	0.0297	0.1723	0.3135	0.0899	0.0223	0.1493	0.3438	0.0931
Romantic	0.0703	0.2651	0.3555	0.1083	0.0291	0.1706	0.2832	0.0919	0.0379	0.1947	0.2941	0.1026	0.0332	0.0835	0.2746	0.1240
Peer relationship	0.2800	0.5292	0.3256	0.1697	0.3140	0.5604	0.3626	0.1202	0.2972	0.5452	0.3060	0.1298	0.2557	0.1472	0.3481	0.1458
Self-cognition	0.0445	0.2110	0.3066	0.1895	0.0345	0.1857	0.2721	0.1653	0.0366	0.1913	0.2557	0.0754	0.0245	0.0862	0.2863	0.1447
Family life	0.1602	0.4002	0.3291	0.1587	0.0889	0.2982	0.2891	0.0944	0.0378	0.1944	0.2952	0.0842	0.1827	0.0979	0.3148	0.1131
All	0.1281	0.3579	0.3604	0.1482	0.0985	0.3138	0.3012	0.1128	0.0878	0.2964	0.2929	0.0964	0.1037	0.1128	0.3135	0.1241

|	Uplift (L&S)				Uplift (L&P)				Uplift (S&P)				Uplift (L&S&P)			
	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE	MAD	MSE	RMSE	MAPE	MAD
School life	0.0283	0.1682	0.2934	0.0824	0.0261	0.1616	0.2770	0.0768	0.0342	0.1849	0.2629	0.0590	0.0132	0.1149	0.2364	0.0717
Romantic	0.0219	0.1480	0.2532	0.0839	0.0180	0.1342	0.2644	0.0952	0.0176	0.1327	0.2549	0.0823	0.0251	0.1584	0.2507	0.0891
Peer relationship	0.2361	0.4859	0.3182	0.1300	0.2349	0.4847	0.3283	0.1189	0.2351	0.4849	0.3558	0.1297	0.2341	0.4838	0.3096	0.1093
Self-cognition	0.0329	0.1814	0.2942	0.0946	0.0262	0.1619	0.2791	0.0858	0.0245	0.1565	0.2740	0.0945	0.0144	0.1200	0.2580	0.0739
Family life	0.1489	0.3859	0.2750	0.1244	0.0395	0.1987	0.2853	0.0939	0.0484	0.2200	0.2946	0.0992	0.0378	0.1944	0.2645	0.0848
All	0.0936	0.3060	0.2868	0.1031	0.0689	0.2626	0.2868	0.0941	0.0720	0.2683	0.2884	0.0929	0.0649	0.2548	0.2638	0.0858

[1] Three restoring pattern measures: 'L' represents *linguistic expression*, 'S' represents *stress intensity*, and 'P' represents *posting behavior*.

impacted by uplift events in SI. For each stressful interval in SI and U-SI, we quantify its stress intensity by comparing with the front and rear adjacent intervals, respectively. Here four situations are considered and compared according to the temporal order in Section VI-C, as shown in Table VI, where the *ratio of intervals* detected with monotonous increase from the *front interval* to *stressful interval* (denoted as *front→ I*), and monotonous decrease from the *stressful interval* to the *rear interval* (denoted as *I → rear*) are listed. Under the impact of uplift events, both the ratio of intensive stress increase in *front→ I* and the ratio of intensive stress decrease in *I → rear* are decreased, showing the effectiveness of the two sample method for quantifying the impact of uplift events, and the rationality of the assumption that uplift events could help ease stress of overwhelmed teens.

C. Integrate the impact of uplift events into stress forecasting

Next, we integrate the impact of uplift events into stress prediction. The experimental set contains 1,914 stressful intervals under the impact of uplift events (U-SI). The number of each type of uplift events and stressful intervals are lists in Table VI. As shown in Table VIII, the original prediction result using only SVARIMA method achieves 0.1281 MSE, 0.3579 RMSE, 0.3604 MAPE and 0.1482 MAD ($L = 7$, $\alpha = 0.5$). Then we integrate the impact of each type of uplift events into stress prediction. Specifically, for uplifts with obvious restoring impact (under the L&S&P pattern), the average stress level during historical restoring intervals are integrated to modify the result, with adjusting the parameter α (details see VII-E). After the modification, the prediction performance achieves 0.0649 MSE, 0.2548 RMSE, 0.2638 MAPE and 0.0858 MAD, reducing the prediction errors efficiently (with MSE, RMSE, MAPE and MAD reduced by 49.34%, 28.81%, 26.80% and 42.11%, respectively). Further, we present the prediction result under the impact of uplift events under different lengths of prediction windows, ranging from 1 to 10 days, as shown in 6. With the window length increasing, the prediction error shows decreasing trend in all metrics. The reason is that longer prediction window takes more previous predicted results, and the error accumulates with more predicted values taken into the

Fig. 5: Stress forecast performance under the L&S&P pattern of uplift events.
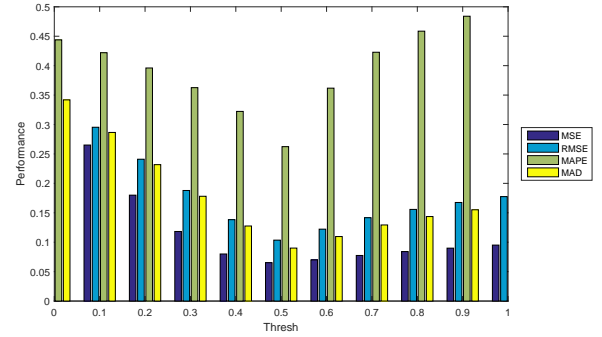
next step prediction. Among the five dimensions of events, the prediction for school life stress achieves the best performance. On one side, more uplift events and stressors about school life events are detected from teens microblogs, providing sufficient data in prediction. On the other side, stress coming from school life is the most common stress in the student group, with relative stable periodicity and high frequency.

D. Contribution of each restoring measure

We conduct experiments with different restoring patterns included respectively to show its contribution to the impact of uplift events during prediction. Four groups of situations are considered here, as shown in Table VIII, considering 1) all the stress intensity, linguistic expression and post behavior measures (the L&S&P pattern), 2) any two of the three measures included (the L|S, L&P, and S&P patterns), 3) only one of the three measures included (the L, S, or P patterns), and 4) none measure included. We integrate the impact of uplift events under the four situations into stress prediction using the parameter α, as overlapping $\alpha \times S_{historical}$, where $S_{historical}$ is the average stress level in historical restoring intervals. The detailed adjust process of α is presenting in section VII-E. Here we present the prediction result when

Fig. 6: Teens' stress forecast performance under different lengths of predicting windows.
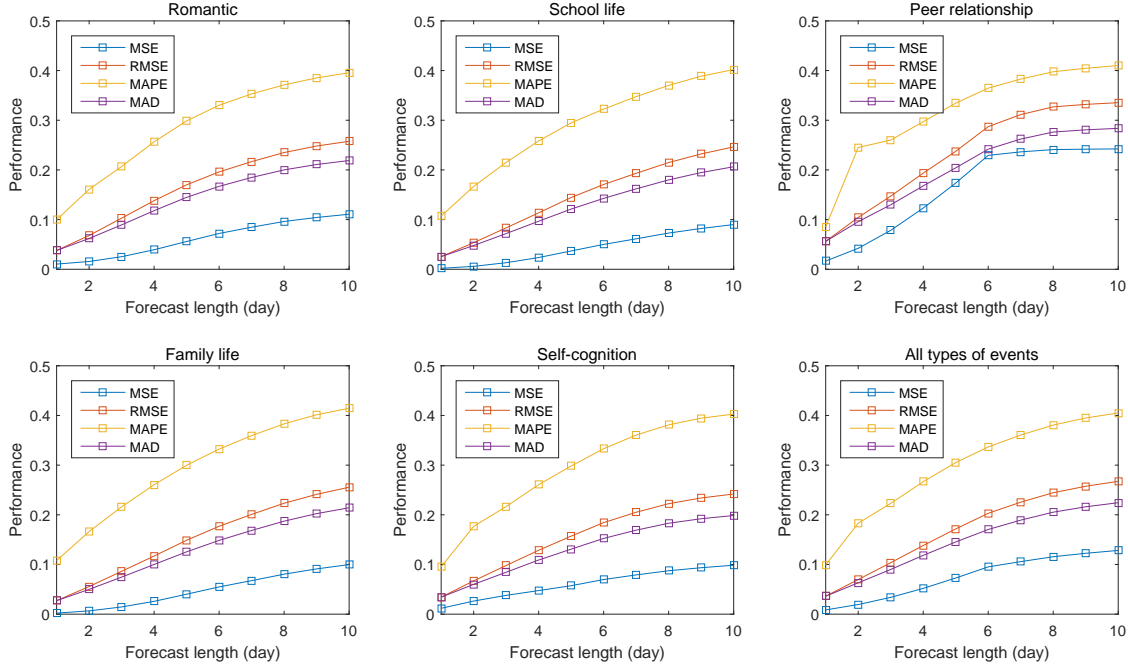
$\alpha = 0.5$ in each dimension of stress respectively. Results show that the correlation in the L&S&P pattern outperforms other patterns (0.0649 MSE, 0.2548 RMSE, 0.2638 MAPE and 0.0858 MAD), showing the effectiveness of considering all the three correlations.

E. Parameter settings

The parameter α is adjusted when integrate the impact of uplift events into stress prediction. For each of the four groups of restoring patterns, we adjust α in the effect of $\alpha \times L$. We calculate the corresponding prediction result for each teen respectively, and show the result of the whole testing group using the averaging performance. Figure 5 shows the changing trend under the L&S&P pattern. The prediction error decreases first and then increases, and the best performance is achieved when α is nearby 0.52, with 0.0649 MSE, 0.2548 RMSE, 0.2638 MAPE and 0.0858 MAD as the average performance of the whole experimental data set. Multiple methods for integrating the impact of uplift event into stress prediction could be adopted. In this paper we adopt the simple one to verify the effectiveness of our model in quantifying the impact of uplift events, and the setting of parameter α could be changed due to different individuals and data sets.

VIII. CONCLUSION

In this paper, we give a deep inside into the stress easing function of uplift events on the real data set of 124 high school students. A two-sample based statistical model is conducted to analyze the stressful behavioral correlations when uplift events happened to overwhelmed students from multiple perspectives.

Experimental results show that our method could measure the restoring impact of school scheduled uplift events efficiently, and integrating the impact of uplift events helps reduce the stress prediction errors. Our research provides guidance for school and parents that which kind of uplift events could help relieve students' overwhelmed stress in both stress prevention and stress early stopping situations. Our future work will focus on digging the overlap impact of multiple uplift events in more complex situations, as well as the frequent appearing patterns of different types of uplift events and stressor events, thus to provide more accurate analysis and restoring guidance for individual teenagers.

REFERENCES

[1] F. P. D. Susan, "Stress: Appraisal and coping," pp. 1–460, 1984.
[2] M. K. Nock, G. Borges, E. J. Bromet, C. B. Cha, R. C. Kessler, and S. Lee, "Suicide and suicidal behavior." *Epidemiologic Reviews*, vol. 30, no. 1, pp. 133–154, 2008.
[3] L. H. Cohen, J. Mcgowan, S. Fooskas, and S. Rose, "Positive life events and social support and the relationship between life stress and psychological disorder." *American Journal of Community Psychology*, vol. 12, no. 5, pp. 567–87, 1984.
[4] S. Cohen and H. M. Hoberman, "Positive events and social supports as buffers of life change stress," *Journal of Applied Social Psychology*, vol. 13, no. 2, pp. 99–125, 2010.
[5] D. J. Needles and L. Y. Abramson, "Positive life events, attributional style, and hopefulness: Testing a model of recovery from depression." *Journal of Abnormal Psychology*, vol. 99, no. 2, p. 156, 1990.
[6] R. Vitelli, "Hassles, uplifts and growing older," https://www.psychologytoday.com/blog/media-spotlight/201406/hassles-uplifts-and-growing-older, jun 09, 2014.
[7] K. W. Doyle, S. A. Wolchik, S. R. Dawsonmcclure, and I. N. Sandler, "Positive events as a stress buffer for children and adolescents in families in transition." *Journal of Clinical Child and Adolescent Psychology*, vol. 32, no. 4, pp. 536–545, 2003.

[8] F. L. Coolidge, "A comparison of positive versus negative emotional expression in a written disclosure study among distressed students," *Journal of Aggression Maltreatment and Trauma*, vol. 18, no. 4, pp. 367–381, 2009.

[9] S. Jain, P. J. Mills, K. R. Von, S. Hong, and J. E. Dimsdale, "Effects of perceived stress and uplifts on inflammation and coagulability." *Psychophysiology*, vol. 44, no. 1, pp. 154–160, 2010.

[10] J. L. Caputo, D. L. Rudolph, and D. W. Morgan, "Influence of positive life events on blood pressure in adolescents," *Journal of Behavioral Medicine*, vol. 21, no. 2, pp. 115–129, 1998.

[11] G. Shahar and B. Priel, "Positive life events and adolescent emotional distress: In search of protective-interactive processes," *Journal of Social and Clinical Psychology*, vol. 21, no. 6, pp. 645–668, 2002.

[12] Y. Xue, Q. Li, L. Feng, G. Clifford, and D. Clifton, "Towards a micro-blog platform for sensing and easing adolescent psychological pressures," in *Proc. of Ubicomp.* poster, 2013.

[13] Y. Xue, Q. Li, L. Jin, L. Feng, D. A. Clifton, and G. D. Clifford, *Detecting Adolescent Psychological Pressures from Micro-Blog*, 2014.

[14] Y. Li, J. Huang, H. Wang, and L. Feng, "Predicting teenager's future stress level from micro-blog," in *IEEE International Symposium on Computer-Based Medical Systems*, 2015, pp. 208–213.

[15] Y. Li, Z. Feng, and L. Feng, "When a teen's stress level comes to the top/bottom: A fuzzy candlestick line based approach on micro-blog," in *Revised Selected Papers of the International Conference on Smart Health*, 2015, pp. 241–253.

[16] ——, "Using candlestick charts to predict adolescent stress trend on micro-blog ?" *Procedia Computer Science*, vol. 63, pp. 221–228, 2015.

[17] Q. Li, L. Zhao, Y. Xue, L. Jin, and L. Feng, "Exploring the impact of co-experiencing stressor events for teens stress forecasting," in *International Conference on Web Information Systems Engineering*, 2017, pp. 313–328.

[18] Q. Li, Y. Xue, L. Zhao, J. Jia, and L. Feng, "Analyzing and identifying teens stressful periods and stressor events from a microblog," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 5, pp. 1434–1448, 2017.

[19] S. Folkman and J. T. Moskowitz, "Stress, positive emotion, and coping," *Current Directions in Psychological Science*, vol. 9, no. 4, pp. 115–118, 2010.

[20] S. Folkman, "Positive psychological states and coping with severe stress," *Social Science and Medicine*, vol. 45, no. 8, pp. 1207–21, 1997.

[21] E. C. Chang, M. Muyan, and J. K. Hirsch, "Loneliness, positive life events, and psychological maladjustment: When good things happen, even lonely people feel better! ," *Personality and Individual Differences*, vol. 86, pp. 150–155, 2015.

[22] E. M. Kleiman, J. H. Riskind, and K. E. Schaefer, "Social support and positive events as suicide resiliency factors: Examination of synergistic buffering effects," *Archives of Suicide Research*, vol. 18, no. 2, pp. 144–155, 2014.

[23] A. D. Ong, C. S. Bergeman, T. L. Bisconti, and K. A. Wallace, "Psychological resilience, positive emotions, and successful adaptation to stress in later life." *Journal of Personality and Social Psychology*, vol. 91, no. 4, pp. 730–49, 2006.

[24] V. Santos, F. Paes, V. Pereira, O. Ariascarrin, A. C. Silva, M. G. Carta, A. E. Nardi, and S. Machado, "The role of positive emotion and contributions of positive psychology in depression treatment: systematic review." *Clinical Practice and Epidemiology in Mental Health Cp and Emh*, vol. 9, no. 1, p. 221, 2013.

[25] J. E. Bono, T. M. Glomb, W. Shen, E. Kim, and A. J. Koch, "Building positive resources: Effects of positive events and positive reflection on work stress and health," *Academy of Management Journal*, vol. 56, no. 6, pp. 1601–1627, 2013.

[26] A. D. Kanner, J. C. Coyne, C. Schaefer, and R. S. Lazarus, "Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events," *Journal of Behavioral Medicine*, vol. 4, no. 1, p. 1, 1981.

[27] M. T. A. Silva, E. A. Manriquesaade, L. G. Carvalhal, and M. Kameyama, "The hassles and uplifts scale," *Estud.psicol*, vol. 25, no. 1, pp. 91–100, 2008.

[28] L. E. Alden, C. T. Taylor, T. M. Mellings, and J. M. Laposa, "Social anxiety and the interpretation of positive social events." *Journal of Anxiety Disorders*, vol. 22, no. 4, pp. 577–90, 2008.

[29] J. C. Mcmillen and R. H. Fisher, "The perceived benefit scales: Measuring perceived positive life changes after negative events," *Social Work Research*, vol. 22, no. 3, pp. 173–187, 1998.

[30] H. U. Jun-Sheng, "Influence of life events and coping style on mental health in normal college students," *Chinese Journal of Clinical Psychology*, 2008.

[31] H. Lin, J. Jia, Q. Guo, Y. Xue, Q. Li, J. Huang, L. Cai, and L. Feng, "User-level psychological stress detection from social media using deep neural network," pp. 507–516, 2014.

[32] J. Cohen, J. Cohen, D. Cohen, J. Cohen, J. Cohen, J. Adad, J. Cohen, J. Cohen, J. Cohen, and J. Cohen, "Statistical power analysis for the behavioral science," *Technometrics*, vol. 84, no. 334, p. 19C74, 1988.

[33] C. Spearman, "The proof and measurement of association between two things," *The American journal of psychology*, vol. 100, no. 3-4, p. 441, 1987.

[34] A. I. Mcleod, "Kendall: Kendall rank correlation and mann-kendall trend test," 2011.

[35] B. Scholkopf, J. Platt, and T. Hofmann, "A kernel method for the two-sample-problem," vol. abs/0805.2368, no. 2007, pp. 513 – 520, 2006.

[36] M. Schilling, "Multivariate two-sample tests based on nearest neighbors," *Publications of the American Statistical Association*, vol. 81, no. 395, pp. 799–806, 1986.

[37] A. Kanner, J. Coyne, C. Schaefer, and R. Lazants, "Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events," *Journal of Behavioral Medicine*, vol. 4, pp. 1–39, 1981.

[38] H. U. Yan, F. B. Tao, and S. U. Pu-Yu, "Compilation and reliability and validity assessment of multidimensional life events rating questionnaire for middle school students," *Chinese Journal of School Health*, vol. February 31, no. 2, pp. 146–159, 2010.

[39] G. Jiang, "The development of the chinese adolescent life events checklist," *Chinese Journal of Clinical Psychology*, vol. 8, no. 1, pp. 10–14, 2000.

[40] L. Baoyong and H. Ying, "The development of the life stress rating scale for middle school students," *Chinese Mental Health Journal*, vol. 16, no. 5, pp. 313–316, 2002.

[41] M. Zhang, W. Che, G. Zhou, A. Aw, and et al., "Semantic role labeling using a grammar-driven convolution tree kernel," *Audio Speech and Language Processing IEEE Transactions*, vol. 16, no. 7, pp. 1315 – 1329, 2008.

[42] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: Liwc and computerized text analysis methods," *Proc. of JLSP*, vol. 29, no. 1, pp. 24–54.

[43] W. Che, Z. Li, and T. Liu, "Ltp: A chinese language technology platform," in *Proc. of ACL*, 2010, pp. 13–16.

[44] W. Cleveland and S. Devlin, "Locally weighted regression: An approach to regression analysis by local fitting," *Publications of the American Statistical Association*, vol. 83, no. 403, pp. 596–610, 1988.

[45] Q. Li, L. Zhao, Y. Xue, L. Jin, M. Alli, and L. Feng, "Correlating stressor events for social network based adolescent stress prediction," 2017.

[46] R. A. Johnson and D. W. Wichern, "Applied multivariate statistical analysis third ed," *Technometrics*, vol. 25, no. 4, pp. 385–386, 2012.

[47] B. Shumway and D. Stoffer, *Time Series Analysis and Its Applications*. Springer New York, 2006.