

# Analyzing and Identifying Teens Stressful Periods and Stressor Events from a Microblog

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

**Abstract**—Increased health problems among adolescents caused by psychological stress have aroused worldwide attention. Long-standing stress without targeted assistance and guidance negatively impacts the healthy growth of adolescents, threatening the future development of our society. So far, research focused on detecting adolescent psychological stress revealed from each individual post on microblogs. However, beyond stressful moments, identifying teens stressful periods and stressor events that trigger each stressful period is more desirable to understand the stress from appearance to essence. In this paper, we define the problem of identifying teens stressful periods and stressor events from the open social media microblog. Starting from a case study of adolescents' posting behaviors during stressful school events, we build a poisson-based probability model for the correlation between stressor events and stressful posting behaviors through a series of posts on Tencent Weibo (referred to as the microblog throughout the paper). With the model, we discover teen's maximal stressful periods and further extract details of possible stressor events that cause the stressful periods. We generalize and present the extracted stressor events in a hierarchy based on common stress dimensions and event types. Taking 122 scheduled stressful study-related events in a high school as the ground truth, we test the approach on 124 students' posts from January 1, 2012 to February 1, 2015, and obtain some promising experimental results: (stressful periods: *recall* 0.761, *precision* 0.737, and  $F_1$ -*measure* 0.734) and (top-3 stressor events: *recall* 0.763, *precision* 0.756, and  $F_1$ -*measure* 0.759). The most prominent stressor events extracted are in the *self-cognition* domain, followed by the *school life* domain. This conforms to the adolescent psychological investigation result that problems in school life usually accompanied with teens inner cognition problems. Compared with the state-of-art top-1 personal life event detection approach, our stressor event detection method is 13.72% higher in *precision*, 19.18% higher in *recall*, and 16.50% higher in  $F_1$ -*measure*, demonstrating the effectiveness of our proposed framework.

**Index Terms**—Adolescent, microblog, stressful period, stressor event.



## 1 INTRODUCTION

**Motivation.** Stress is an important health issue. 77% of adults reported regularly experiencing physical symptoms caused by stress, and an estimated 75% to 90% of visits to primary care physicians were stress-related [1]. Teens are more stressed out than adults. According to an annual survey published by the American Psychological Association [2], teens routinely said that their school-year stress levels were far higher than they thought was healthy and their average reported stress exceeded that of adults. Around 30% of teens reported feeling sad or depressed because of stress and 31% felt overwhelmed. Another 36% said that stress makes them tired and 23% said they had skipped meals because of it. On average, teens reported their stress level was 5.8 on 10-point scale, compared with 5.1 for adults. Stress seems to be getting worse for some teens, according to the survey. About 31% of kids said their stress level had increased in the past year, twice as many as those who said it went down, and 34% said they expected their stress level would rise in the coming year. There is no doubt that with the fast speed of society and high expectation from adults and themselves, teens will become more stressful than ever before.

Stress is the way human body responds to challenges. The causes of stress are known as *stressors* and there are literally hundreds of different types of stressors. Any event that one finds

threatening and difficult to cope with can be a potential cause of stress. Typical stressor events that may cause youth stress include schoolwork and career decisions, physical and cognitive changes of puberty, family and peer conflicts, crammed schedules, dating and friendships, and so on. The ways in which teens cope with these stressor events can have significant short- and long-term consequences on their physical and emotional health. Difficulties in handling stress can lead to mental health problems, such as depression, anxiety disorders, and even suicide [3]. Due to the immaturity and lack of life experiences, teenagers may not always take effective measures to deal with the stress. Thus, timely discovery and effective intervention prevention of adolescents mental disease is necessary and urgent.

**Existing solutions.** Signs of teens stress could show up in their behaviors, emotions, and bodies, and can thus be analyzed and discovered through various psychological signals (e.g. heart rate variability HRV [4], electrocardiogram ECG [5], galvanic skin response GSR [6], blood pressures, and electromyogram [7]) and physical signals (e.g. gestures, voice, eye gaze, facial expressions, pupil dilation and blink rates [8], [9]). Recent studies demonstrated the feasibility of stress detection through the open social media microblogs (e.g. Twitter, Tencent Weibo, and Sina Weibo) [10]–[15]. Through teen's posting content (linguistic text, emoticons, and repetitive punctuation), posting time, and social interaction with friends (being liked/reposted/cared and comment-response acts) underneath each post, teens stress category (academy, inter-personnel, self-cognition, or affection) and stress level could be detected.

If we view the aforementioned approaches as *symptom-based point-wise* exploration towards teens stress detection, this study

- Dept. of Computer Science and Technology, Centre for Computational Mental Healthcare Research, Institute of Data Science, Tsinghua University, China.  
E-mail: {liqi13,xue-yy12,jing-zhao11}@mails.tsinghua.edu.cn, {jjia,fengling}@tsinghua.edu.cn

approaches the problem from the *cause* of the stress, and aims at a *period-wise* detection solution under the assumption that teen's stress is caused by one or more stressor events and will last for a period of time. Short-lasting acute stress is not considered in the current study, as long-lasting stress has a more serious impact on teens physical and psychological health than acute stress due to sustained high levels of chemicals released in the response. From this perspective, stressful periods are more worth paying attention to than stressful points.

Another difference from the previous work is that, despite the unique advantage of microblogs over the sensory approaches in offering linguistic textual information for analysis, the aforementioned approach stopped at detecting the general stress category, and none went further to extract specific events that may cause the stress. From example, it is *quarreling with friends* that leads to teen's *inter-personnel* stress. Knowing stressor event details is helpful in understanding and guiding stressful teens towards healthy ways of thinking and reacting.

**Our work.** To this end, this paper pushes forward the problem of teens psychological stress detection from emotional symptom-oriented level to event-oriented level. By investigating the correlation between stressor events and stress reflected through a series of posts on Tencent Weibo (<http://t.qq.com>) [16], we discover teen's stressful periods, and further extract details of stressor events that trigger the stress. Exploiting the correlation between stressor events and stress for stressful period detection is also advantageous in handling data uncertainty when a teen's post is missing and false alarms of intermittent stressful points.

However, detection of stressful periods and stressor events from the microblog is not a trivial task. Three fundamental questions need to be addressed: (1) *what is the latent correlation between stressor events and teens stressful reflection on the microblog?* (2) *what is counted as a teen's stressful period?* and (3) *what are the criteria for teen's stressor events extraction?*

In this paper, we stand upon and extend the previous stressful posts detection result. Motivated by the empirical findings from a user study at a high school that *the rate of teens posting stressful posts is higher during stressor-event periods than non-stressor-event periods*, we propose to model the correlation between stressor events and teens stressful periods through a statistic measure, which is the Bayesian posterior probability of teens stressful posting rates during event and non-event periods. As this probability only depends on stressful post rates rather than on specific stressful post time points during event and non-event periods, the model is attractive for statistic inference in large-scale systems in terms of both measurement and computation. With the model, we can analyze and identify maximal stressful periods, from which a set of stressor events are extracted.

In the study, we follow the categorization of teens stressor events, containing daily hassles and life events, that worldwide teens likely encounter from five *dimensions*, namely, *school life*, *family life*, *peer relation*, *romantic relation*, and *self-cognition* [17]–[20]. Each dimension contains a set of items, representing *event types*, as listed in Table 1. A stressor event is presented in the form of [Dimension, Event-type, Event-instance (description, doer, act, object, time, location)], influence), where Dimension and Event-type are from Table 1, and Event-instance refers to the time-specific event occurrence with the involved role, act, object, time, location, linguistic description, and its influence measurement upon stress.

Due to teens causal and informal expression on the microblog, it may be hard to identify and distinguish concrete event instances with complete occurring time and roles information. Therefore, we generalize and present the extracted stressor events based on common dimensions and event types. Considering that a stressor event tends to stimulate the increase of a teen's stress level, we rank the stressor events at the level of instance, type, and dimension based on their negative stress impact that, the higher the stress impact is, the greater teen's stress level is increased, the longer the teen's stress will sustain, and the more the teen may post about. The experimental study in a high school with 124 participants shows that our method achieves promising performance in detecting stressful periods (*recall* 0.761, *precision* 0.737, and  $F_1$ -*measure* 0.734) and stressor events *recall* 0.763, *precision* 0.756, and  $F_1$ -*measure* 0.759). We compare the performance of our stressor event detection approach with the state-of-art top-1 personal life event detection approach, achieving 13.72% higher in *precision*, 19.18% higher in *recall*, and 16.50% higher in  $F_1$ -*measure*.

To our knowledge, this is the first attempt in the literature that analyzes and identifies teens stressful periods and stressor events from the microblog. The poisson-based probability model for the correlation between stressor events and stressful posting behaviors, together with the event-oriented stress detection method proposed in this paper targets at solving the psychological health problem for adolescents from appearance to essence, thus enabling targeted guidance and support from teens parents, teachers, peers, and mental health care professionals in stress coping.

The reminder of the paper is organized as follows. We review some closely related work in Section 2. We empirically study teens abnormal posting behaviors incurred by stressor events in Section 3. A microblog-based framework for inferring maximal stressful periods and extracting stressor events is then presented in Section 4. We report our performance study in Section 5, and discuss implications of the study in Section 6. We conclude the paper in Section 7.

## 2 RELATED WORK

### 2.1 Stress Detection

#### 2.1.1 Stress Detection via Questionnaires

Cohen's Perceived Stress Scale (PSS-14) [21] is commonly used to measure human's stress level worldwide in psychology. The scores of PSS-14 indicate different stress levels (i.e., none, light, moderate, and strong). According to Ting and Huang [22], the PSS-14 score value 26 is the boundary value of PSS-14 questionnaire. Score falling into [27,75] indicates strong stress, score in [25,26) indicates moderate stress, score in [1,24) indicates light stress, and score 0 indicates no stress.

Pozos-Radillo *et al.* [23] examined the correlation and predictive value between the Academic Stress Inventory (ASI) and the Stress Symptom Inventory (SSI) among 527 university students. The findings showed that IEA situations corresponding to classroom intervention, mandatory work, and doing an exam predicted high-level chronic stress; being a female and 18, 23, and 25 years old were associated mostly to stress. It was concluded that accurate identification of stressors could help understand stress and its harmful effects on college students.

Michie [24] looked into pressures at work which could cause high and long lasting levels of stress, aiming to judge whether people were doing enough to prevent that harm. Based on structural equation modeling (SEM), Wohn and LaRose [25] distributed and

Dimension	Event-Type
School Life (25 types)	Having to study things you do not understand
	Feeling tired of study
	Keeping up with schoolwork
	Having too much homework
	Having exams or tests
	Having difficulty in some subjects
	Teachers expecting too much from you
	Getting along badly with teachers
	Lack of respect from teachers
	Being scolded by teachers
	Being punished corporally
	Getting up early in the morning to go to school
	Entering a higher school
	Going to school
	Failing to be admitted to a university
	Not getting enough time for leisure/fun
	Worrying about losing behind
	Wasting time
	Unsatisfying exam results
	Too many interruptions
	Failing to reach expectations
	Unsatisfying teaching quality
	Failing in awards competition
	Being suspended from school
	Changing class or school
Family Life (17 types)	Having bad relationship with family members
	Being beaten/scolded by parents
	Parents working away from home for a long time
	Parents hassling you about the way you look
	Household disasters
	Disagreements between you and parents
	Serious disease of family member
	Lack of understanding by your parents
	Death of family member
	Not being taken seriously by your parents
	Parents expecting too much from you
	Not having enough money to buy the things you want
	Family having financial difficulties
	Bad habits of parents
	Having to take on new financial responsibilities
	Parents' divorce
	Parents' quarreling
Peer Relation (6 types)	Being hassled for not fitting in with peers
	Trying to keep up with the Joneses
	Cannot bear the roommate's behaviors
	Rumor or satire from peers
	Misunderstood by others
	Fierce competition among peers
Self-Cognition (15 types)	Unsatisfied with appearance/weight
	Having to make decisions about future work or education
	Bad sleep conditions
	Having to take on new family responsibilities
	Change of daily routine
	Doubting the meaning of life
	Getting sick or hurt
	Depression
	Troubles of menstruation
	Regretting for past behaviors
	Internet addiction
	Losing face in public
	Worrying about future
	Being frightened/threatened unexpectedly
	Losing things or being stolen
Romantic Relation (6 types)	Secret adoration
	Sexual behavior
	Getting along with your boy/girl-friend
	Breaking up with your boy/girl-friend
	Too curious about sex
	Being ignored or rejected by the person you want to go out with

TABLE 1: Typical adolescent stressor event types

analyzed scales sampling 387 university students, to survey for life stress, learning self-efficacy, social self-efficacy, and smart phone addiction.

### 2.1.2 Stress Detection via Physical and Physiological Signals

In recent years, sensor technology is widely used for human stress status detection and monitoring, based on the changing status of various physical and psychological signals. Typically used physical signals are gestures, voice, eye gaze, facial expressions, pupil dilation, and blink rates. In [8], smart phones were used as a voice sensor by collecting voice variations when people talking naturally, focusing on cognitive stress and stressor frequency estimation. Using a continuous sensing App on Android phones, Wang *et al.* [9] studied the impact of workload on stress, sleep, activity, mood, sociability, mental well-being, and academic performance for college students in a class. Kocielnik *et al.* [26] presented stress information derived from sensor measurements in the context of person's activities, aiming to provide the user with meaningful, useful, and actionable information. Bousefsaf *et al.* [4] detected mental stress using a low-cost webcam, which could recover the instantaneous heart rate signals from video frames of human faces conveniently.

Physiological signals popularly used for stress detection include heart rate variability (HRV), electroencephalogram (EEG), electrocardiogram (ECG), galvanic skin response (GSR), blood pressures, and electromyogram [27]–[29]. Sierra *et al.* [6] described the behavior of an individual under stressing stimuli in terms of HR and GSR, and implemented a fuzzy logic based stress detection system.

Combining observed driving behaviors, Rigas *et al.* [7] integrated multiple physiological signals (i.e., skin conductance, electromyogram, electrocardiogram, and respiration) to detect drivers' stress. Hamid *et al.* [30] proved the negative correlation between the ratio of EEG Power Spectrum and PSS, and suggested combining physiological signals and PSS for stress detecting. Sun *et al.* [5] collected accelerometer, ECG, and GSR signals from 20 volunteers doing activities (sitting, standing, walking), and differed the changing of physiological signals between physical activity and mental stress.

Xu *et al.* [31] proposed a cluster-based analysis method for personalized stress evaluation using physiological signals. It exploited the homogeneity of subjects in the change of their physiological features due to stress, and clustered subjects who shared similar patterns of stress response together. A cluster-wise stress evaluation was then performed using the general regression neural network to accommodate inter-subject differences.

### 2.1.3 Stress Detection on Social Networks

With social network becoming a new channel for sharing personal information, many researchers turn to detect stress by leveraging open self-expressed social media platforms (e.g., Twitter, Facebook, Tencent Weibo, and BBS). Compared to traditional sensors, the later are low-cost and full of "big data" [32], [33].

Shen *et al.* [34] constructed a two-stage supervised learning framework to identify potential depression candidates, based on the content and temporal features extracted from their write-ups on BBS. Moreno *et al.* [35] adopted negative binomial regression analysis to evaluate college students' Facebook disclosures which met DSM criteria for a depression symptom or a major depressive episode (MDE). Campisi *et al.* [36] studied the relation

between Facebook, stress, and incidence of URI (upper respiratory infection) in undergraduate college students, indicating that the impact of stress on the URI incidence rate increased with the size of the social network. Jennifer *et al.* [37] evaluated the disclosure of depression from college students' Facebook profiles, and proved that in-person communication from friends or trusted adults was the most preferred stress relieve method for depression detected from Facebook. Daniel *et al.* [38] proved that the medical students could benefit from a stress management intervention based exclusively on Facebook through a pilot study. Wohn and LaRose [25] studied the relationship between loneliness, varied dimensions of Facebook use, and college adjustment among first-year students, concluded that loneliness was a stronger indicator of college adjustment than any dimension of Facebook usage. As the first work for impression management on Facebook through emotional disclosure, Lin *et al.* [39] showed that individuals were more likely to express positive relative to negative emotions on Facebook than in real life. Wohn *et al.* [40] investigated the relationships between loneliness, anxiousness, alcohol, and marijuana use in the prediction of college students connections with others on Facebook as well as their emotional connectedness to Facebook. Xue *et al.* [10], [11] investigated a number of teens typical posting behaviors that might reveal adolescent stress, and applied five classifiers to teens stress detection. Based on the detection result, three timely intervention steps (encouraging teens to read or do something, notifying guardians at the worst case) were designed to help pressurized teenagers cope with their stress [12]. Lin *et al.* [13], [14] trained a deep sparse neural network to detect psychological stress from cross-media microblogs.

## 2.2 Stressful Events Detection

### 2.2.1 Stressful Events Detection via Psychological Questionnaires

Traditionally, detection of stressful events is based on psychological questionnaires for different groups of people. Specific to adolescents, several acknowledged questionnaires and investigations are done [17], [23], [41]–[44]. Bobo *et al.* [41] developed an Adolescent Hassles Inventory (AHI) upon the Hassles Scale [17]. Using a stratified random sampling method, Wang and Yao [43] developed a College Students Stressors Questionnaire. Through 8 questions contained in a semi-structured journal, Lu [42] identified 15 categories of everyday hassles and 4 types of related information behaviors from 133 children in a public elementary school in an urban community in Taiwan. Li *et al.* [45] adopted an organism  $\times$  environment interaction approach to examine the occurrence of PIU (Problematic Internet Use) in adolescent females and males. A mediated moderation model was tested, in which temperament moderated the relationship between stressful life events and PIU, and this moderating effect was mediated by maladaptive cognitions about Internet use. Besides, the sources of gender difference in PIU implied were examined in the model [45]. The findings of the study confirmed the earlier investigation result that there was a positive relationship between stressful life events and PIU in adolescents. You *et al.* [44] conducted an online survey of 5989 Chinese university students, showing that childhood adversity and recent school life stressors were the most important predictors of suicide in this population.

An online survey made by [46] on Facebook examined the relationship between general perceived levels of stress, quality of life, social networking usage, and disclosing important life events

Stressor Event	Duration	Significance <sup>1</sup>	Frequency
School opens	1 day	2	6
Final/mid-term exam	3 days	2	30
Weekly/monthly exam	3 days	2	34
Notification of exam result	1 day	2	64
Contest/competition	1 day	1	13
IT exam	1 day	2	3
Mock exam	3 days	3	9
College entrance application	1 day	1	3
College entrance exam	3 days	3	3
College application	3 days	3	3

Significance = 1: *weak*, 2: *moderate*, 3: *strong*

TABLE 2: Some study-related stressor events at Taicang High School from January 1, 2012 to February 1, 2015

on Facebook in order to better understand the complex relationship between online disclosure and individual well-being status in life.

Through a stress-related questionnaire for 12 weeks of 9 trial participants on Android smart phones, Weppner *et al.* [47] showed that challenges were positively correlated with stress scales, and skills were negatively correlated with stress scales.

### 2.2.2 Public and Private Life Events Detection on Social Media

Taking advantage of internet resources, public and private life events detection on social media has been studied, utilizing burst detection, topic tracking and information diffusion techniques (e.g., bursty news [48], popular events [49], scheduled city events [50], disasters [51], and controversial events [52]).

Different from public events detection, life events detection is an emerging research direction, focusing on private events related to specific users. However, most of the work focus on text analysis of posts and ignore users' emotional behaviors along with the life events. [53] combined a supervised data mining algorithm and an unsupervised distributional semantic model to discover association language patterns from web corpus. Chan *et al.* [54] applied the association rule mining technique to extract frequently appeared word pairs as features to classify user's sentences with negative life events into predefined event categories. Based on users' Twitter streams, Li and Cardie [55] constructed a chronological timeline for personal important events from posts using an unsupervised approach. Tweets talking about personal (as opposed to public) and time-specific (as opposed to time-general) topics were included in the personal timeline. A non-parametric multi-level Dirichlet Process model was introduced to recognize four types of tweets: personal time-specific (PersonTS), personal time-general (PersonTG), public time-specific (PublicTS) and public time-general (PublicTG) topics, which, in turn, were used for further personal event extraction and timeline generation. Furthermore, instead of directly inspecting tweets to determine whether they correspond to major life events, Li *et al.* [56] identified replies corresponding to CONGRATULATIONS or CONDOLENCES, examined the messages they were in response to, and then extracted event properties using Conditional Random Fields (CRF).

## 3 CASE STUDY

### 3.1 Stressor Events and Posting Behaviors

*Stressor Events.* To empirically study the correlation between stressor events and teens' abnormal posting behaviors, we conduct a user study at Taicang High School, a key senior high school

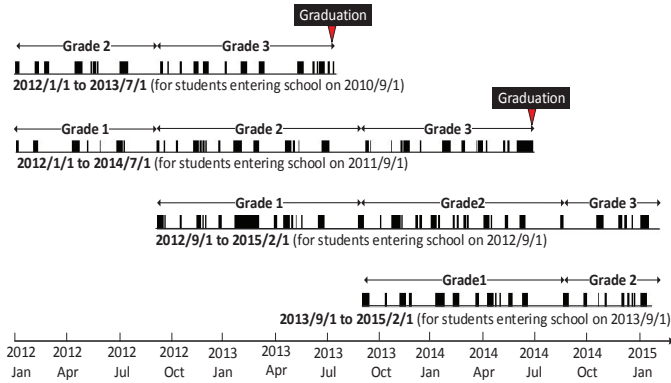


Fig. 1: Distribution of grade-oriented study-related stressful routine events at Taicang High School

	Number of Involved Students	Number of Posts	Number of Stressor Events
Grade 1	88	13311	38
Grade 2	124	10488	48
Grade 3	107	5433	36
Total	-	29232	122

TABLE 3: Number of involved students, posts, and stressor events for three grades in the user study

in Zhejiang Province of China. The school publishes its weekly agenda regularly on the school’s official web site one week in advance to let students, teachers, and parents know school schedules beforehand. We find 273 school events from January 1, 2012 to February 1, 2015, where the majority are study-related. We further filter out 122 stressful study-related events such as examination, contest, result notification, etc. Table 2 lists some typical events, their duration, significance in evaluating students’ school performance, and total occurring frequencies during the observed three years. Fig. 1 shows the distribution of study-related stressor events, which happen to students when they are at different grades. As these events are routine and grade-oriented, each year, students of the same grade experience the same events as those in previous years. On average, 2-3 stressor events take place each month for each grade of students.

*Posts.* We identify 124 students who are active Tencent Weibo users (registered before January 1st, 2012 and posted more than 100 posts on the microblog) in Taicang Senior High School. All students are anonymous, and all posts are public. From January 1, 2012 to February 1, 2015, they posted 29,232 posts in total. The average post number is 236 posts, 1,387 posts maximally, and 104 posts minimally. Table 3 shows for each grade, the number of involved students in the observed three years, number of posts from the students, and number of stressor events.

Applying the stress detection function [11], [14] to each post  $w_t$  posted at time  $t$ , we can sense teen’s stress level in the dimension of school, family, peer-relation, self-cognition, or romantic-relation, denoted as  $stress(t, w) = (t, C, l)$ , where  $C \subset C = \{c^1, c^2, \dots, c^6\}$ , corresponding to stress category *school*, *family*, *peer-relation*, *self-cognition*, *romantic-relation*, and *unknown*, and  $l \in \mathcal{L} = \{0, 1, 2, 3, 4, 5\}$ , corresponding to *none*, *very light*, *light*, *moderate*, *strong*, *very strong* stress level. If no categorical words appear in the post, we assign the detected stress level to an “*unknown*” category. We call  $w$  a **stressful post**, if the detected stress  $l > 0$ , and otherwise, it is called a **non-stressful**

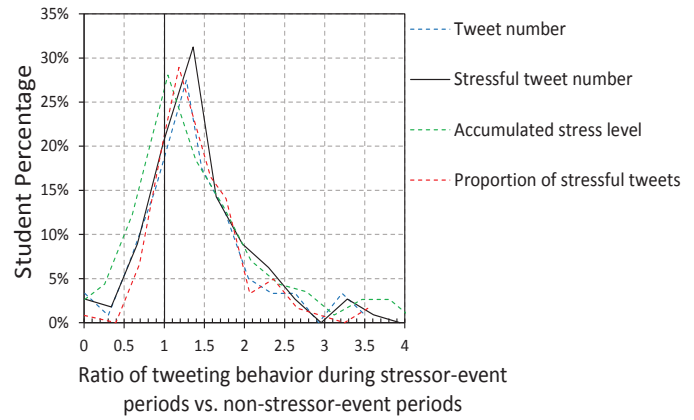


Fig. 2: Histograms of  $R_{\#post}$ ,  $R_{\#stress\_post}$ ,  $R_{\#stress\_level}$  and  $R_{\#stress\_proportion}$  during stressor-event and non-stressor-event periods

### post.

*Relationship between Stressor Events and Stressful Posts.* To examine the influence of stressor events upon students’ posting behavior, we partition the whole observation window into two complementary sets: *stressor-event period set* and *non-stressor-event period set*. Due to the continuity of human’s emotion, we include a  $\Delta_t$ -day window on either side of a stressor-event period. The computation of  $\Delta_t$  is based on the significance of the event and event duration. In this study,  $\Delta_t = \text{Significance} * \text{Duration}$ . The basic time unit is day. If overlap exists between two stressor-event periods, we collapse them into one.

Then, for each student, we look at how different his/her posting behavior is during stressor-event periods and non-stressor-event periods through the following four indicators:

- $R_{\#post}$ : ratio of average number of posts per day during stressor-event periods and non-stressor-event periods;
- $R_{\#stress\_post}$ : ratio of average number of stressful posts per day during stressor-event periods and non-stressor-event periods;
- $R_{\#stress\_level}$ : ratio of average accumulated stress level per day during stressor-event periods and non-stressor-event periods;
- $R_{\#stress\_proportion}$ : ratio of average stressful posts proportion per day during stressor-event periods and non-stressor-event periods, i.e.,  $R_{\#stress\_proportion} = \frac{R_{\#post}}{R_{\#stress\_post}}$ .

If an indicator value is close to 1.0, the difference under this indicator during stressor and non-stressor-event periods is small.

## 3.2 Empirical Findings

We plot the average histograms of the four ratios  $R_{\#post}$ ,  $R_{\#stress\_post}$ ,  $R_{\#stress\_level}$ , and  $R_{\#stress\_proportion}$  for all the 124 students in Fig. 2, and find that all the four indicator values are more to the right of the ratio value 1.0, especially for indicator  $R_{\#stress\_post}$  (whose black curve skews to the right of ratio value 1.0 the most). It shows that the students tend to post more stressful posts per day during stressful-event periods than non-stressful-event periods. This phenomenon particularly holds for the final-year Grade 3’s students, whose black curve skews to the right of ratio value 1.0 the most, as shown in Fig. 3. Close to graduation and national college entrance exam, apparently, the study-related events pressurize them the most than Grade 1 and 2’s students.

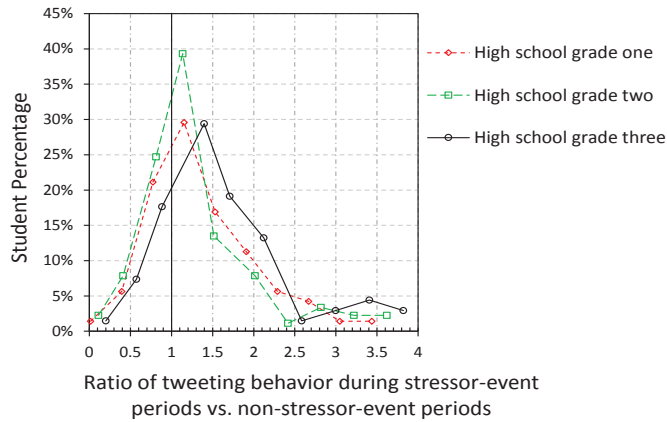


Fig. 3: Histograms of  $R_{\#stress\_post}$  for different grades' students during stressor-event and non-stressor-event periods

This empirical finding helps us shape our problem solution: a teen's posting behavior is different during stressor-event periods and non-stressor-event periods, and higher ratio values of  $R_{\#post}$ ,  $R_{\#stress\_post}$ ,  $R_{\#stress\_level}$ , and  $R_{\#stress\_proportion}$  exist during stressor-event periods than non-stressor-event periods. By mining the information statistically, we may identify teen's stressful periods, from which stressor events can be extracted.

## 4 THE FRAMEWORK

Based on the empirical finding that a teen posts more stressful posts per day during stressor-event periods than non-stressor-event periods (referred to as *stressful posting rate*), we present a statistic model of teen's stressful posting rate. With the model, we can extract stressful periods and stressor events to be extracted from each stressful period.

### 4.1 A Statistic Model of Teens Stressful Period

#### 4.1.1 Poisson Process of Stressful Posting Rates Over a Time Period

We model teen's posting behaviors during stressor-event and non-stressor-event periods as two independent homogeneous Poisson processes. The number of stressful posts created during a stressor-event period is modeled as a homogeneous Poisson process of stressful posting rate  $\lambda_1$  per unit time (say, per day), and the number of stressful posts created during a non-stressor-event period as another independent homogeneous Poisson process of stressful posting rate  $\lambda_0$  per unit time. From Fig. 3, we observe that  $\lambda_1 > \lambda_0$ .

In general, for a Poisson process of a constant stressful posting rate  $\lambda$ , the number of stressful posts  $N$  posted during a period of length  $T$  (which needs not to be contiguous) is a Poisson random variable with mean  $\lambda \times T$ . That is, the probability that a teen posts  $n$  stressful posts within  $T$  time units can be computed as:  $Pr[N = n|\lambda] = \frac{e^{-\lambda T} (\lambda T)^n}{n!}$ , where  $n = 0, 1, \dots, \infty$ . Under the Poisson process model, the number of stressful posts created by a teen in non-overlapping time periods are independent random variables.

#### 4.1.2 Inferring Stressful Periods From Stressful Posting Rates

To measure the confidence that a teen has a higher stressful posting rate during stressor-event periods than non-stressor-event periods, we use teens historic posting information to estimate the probability distributions of the stressful posting rates  $\lambda_1$  and  $\lambda_0$ , and then compute a statistic based on these posterior distributions. This statistic is our metric for stressor-event periods, i.e., stressful periods.

Let  $N_1$  be the number of stressful posts during stressor-event periods of length  $T_1$ , and  $N_0$  be the number of stressful posts during non-stressor-event periods of length  $T_0$ . In estimation theory,  $N_1$  and

$N_0$  are minimal sufficient statistics for estimation of  $\lambda_1$  and  $\lambda_0$ , respectively. Assume a Jeffreys non-informative prior on the stressful posting rate  $\lambda_1$  and  $\lambda_0$  [57]. That is, for  $\lambda_1, \lambda_0 > 0$ , the prior density  $p(\lambda_1, \lambda_0) \propto \frac{1}{\sqrt{\lambda_1 \lambda_0}}$ , and the corresponding priors are  $p(\lambda_1) \propto \frac{1}{\sqrt{\lambda_1}}$  and  $p(\lambda_0) \propto \frac{1}{\sqrt{\lambda_0}}$ . According to the Bayes Rule, the posterior distribution of  $\lambda_i$  (for  $i=0,1$ ) is

$$p(\lambda_i | N_i) \propto Pr[N_i | \lambda_i] p(\lambda_i) = \left( \frac{1}{\sqrt{\lambda_1 \lambda_0}} \times \frac{e^{-\lambda_i T_i} (\lambda_i T_i)^{N_i}}{N_i!} \times \frac{1}{\sqrt{\lambda_i}} \right) \propto \lambda_i^{N_i - 0.5} e^{-T_i \lambda_i} \quad (1)$$

Normalizing the posteriors into 1, we have

$$p(\lambda_i = x | N_i) = \begin{cases} \frac{T_i (T_i x)^{N_i - 0.5} e^{-T_i x}}{\gamma(N_i + 0.5)} & \text{if } x \geq 0, \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where  $\gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$  is the gamma function.

We quantify the probability that a teen is undergoing a stressor-event period as:

$$p(\lambda_1 > \lambda_0 | N_1, N_0) = \iint_{x>y} p(\lambda_1 = x | N_1) p(\lambda_0 = y | N_0) dx dy = \iint_{x>y} \frac{T_1 (T_1 x)^{N_1 - 0.5} e^{-T_1 x}}{\gamma(N_1 + 0.5)} \times \frac{T_0 (T_0 y)^{N_0 - 0.5} e^{-T_0 y}}{\gamma(N_0 + 0.5)} dx dy \quad (3)$$

When  $p(\lambda_1 > \lambda_0 | N_1, N_0)$  is over a threshold  $\tau$ , we are confident enough that the teen's stressful posting rate during stressor-event periods is larger than that during non-stressor-event periods. We thus declare that the corresponding period is a *stressful period*.

### 4.2 Discovery of Teens Stressful Periods

With the model, we can now formally define and extract stressful periods.

#### 4.2.1 Definitions of Stressful Periods

Assume a teen posts a list of posts  $w_1, w_2, \dots, w_m$  chronologically at time  $t_1, t_2, \dots, t_m$  on the microblog, respectively, denoted as  $W[t_1, t_m] = \{(t_1, w_1), (t_2, w_2), \dots, (t_m, w_m)\}$ . From each individual post  $w_i$  ( $1 \leq i \leq m$ ), teen's stress categories  $C_i \in \mathcal{C}$  and stress level  $l_i \in \mathcal{L}$  are detected, denoted as  $tSW[t_1, t_m] = \{(t_1, C_1, l_1), (t_2, C_2, l_2), \dots, (t_m, C_m, l_m)\}$ . Assume during  $[t_1, t_m]$  there are  $z \in \mathbb{Z}$  unit time intervals (say, day)  $T_1, T_2, \dots, T_z$ . We collapse the *point-wise stress representation*  $tSW[t_1, t_m]$  of  $W[t_1, t_m]$  into *interval-wise representation*  $TSW[T_1, T_z] = \{(T_1, n_{T_1}), (T_2, n_{T_2}), \dots, (T_z, n_{T_z})\}$ , where  $n_{T_i}$  is the aggregate number of stressful posts in the  $i$ -th unit time interval  $T_i$  (for  $1 \leq i \leq z$ ). When no stressful posts are detected in  $T_i$ ,  $n_{T_i} = 0$ .

**Definition 1.**  $TSW[T_1, T_z]$  is a **stress wave**, if there is at least one stressful post in every unit time interval throughout  $[T_1, T_z]$ , i.e.,  $\forall i (1 \leq i \leq z) (n_{T_i} > 0)$ .  $TSW[T_1, T_z]$  is a **big stress wave**, if it is a stress wave, and meanwhile the probability that the *stressful posting rate*  $\lambda_1$  during  $[T_1, T_z]$  is bigger than  $\lambda_0$  during historic non-stressor-event periods is over a confidence threshold  $\tau$ . If a stress wave is not a big stress wave, then it is a **small stress wave**.  $\square$

Fig. 4 shows six big stress waves  $A, D, E, F$  and two small waves  $B, C$ . Although  $B, C$  are not prominent and counted as big stress waves, they could be some late-time oscillations of big stress wave  $A$  or prelude of big stress wave  $D$ , caused by the same stressor event(s). Also, big stress wave  $D$  and  $E$  may arise due to the same stressor event(s). To deal with such intermittence and continuous characteristics of teens stress and restore an integral stressful period



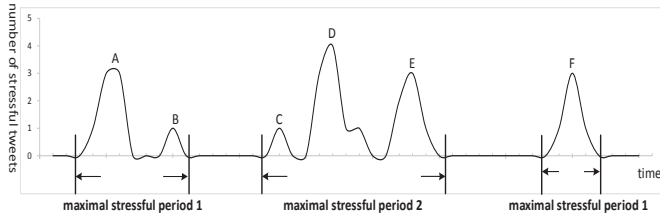


Fig. 4: Illustration of 4 big stress waves (A,D,E,F), 2 small stress waves (B,C), and 3 maximal stressful periods

triggered by the same stressor event(s), we merge these semantically related waves together based on common stress categories.

Given  $tSW[t_1, t_m] = \{(t_1, C_1, l_1), \dots, (t_m, C_m, l_m)\}$ , we define the proportion of stress category  $c \in \mathcal{C}$  in  $tSW[t_1, t_m]$  over the period  $[t_1, t_m]$  as

$$cRatio(tSW[t_1, t_m], c) = \frac{|\{C_i | i \in [1, m] \wedge (c \in C_i)\}|}{\sum_{i \in [1, m]} |C_i|} \quad (4)$$

Extending the single time period parameter of the function to allow multiple time periods, we have

$$cRatio(tSW[t_1, t_{m_1}], \dots, tSW[t_k, t_{m_k}], c) = \frac{|\{C_i | i \in ([t_1, t_{m_1}] \cup \dots \cup [t_k, t_{m_k}]) \wedge (c \in C_i)\}|}{\sum_{i \in [t_1, t_{m_1}] \cup \dots \cup [t_k, t_{m_k}]} |C_i|}, \quad (5)$$

$$\sum_{c \in \mathcal{C}} cRatio(tSW[t_1, t_m], \dots, tSW[t_k, t_{m_k}], c) = 1$$

**Definition 2.** Let  $TSW[T_1, T_z]$  and  $TSW'[T'_1, T'_z]$  be two stress waves, whose point-wise stress representation are  $tSW[t_1, t_m]$  and  $tSW'[t'_1, t'_m]$ . We call  $TSW[T_1, T_z]$  and  $TSW'[T'_1, T'_z]$  **category distribution similar**, if and only if the Euclidean Distance of their category distributions  $cDist(\cdot) \leq \beta$ , where  $\beta$  is set to 0.2 in the study, and

$$cDist(tSW[t_1, t_m], tSW'[t'_1, t'_m]) = \sqrt{\sum_{c \in \mathcal{C}} (cRatio(tSW[t_1, t_m], c) - cRatio(tSW'[t'_1, t'_m], c))^2} \quad (6)$$

Assume from a sequence of teen's intermittent posts created throughout  $[T_1, T_z]$ , a sequence of big stress waves  $B$  and small stress waves  $S$  are detected. We call  $TSW[T_1, T_z]$  a **stressful period**, if and only if the following four conditions hold.

- **i)** There is at least one stressful post in the start interval  $T_1$  and end  $T_z$ .
- **ii)** Every two big stress waves in  $B$  are category distribution similar.
- **iii)** The category distribution of each small wave in  $S$  is similar to the *centroid* category distribution of the big waves covering all the big wave time periods  $[t_1, t_{m_1}], \dots, [t_k, t_{m_k}]$  based on the extended ratio function  $cRatio(tSW[t_1, t_{m_1}], \dots, tSW[t_k, t_{m_k}], c)$ , where  $c \in \mathcal{C}$ .
- **iv)** The probability that the stressful posting rate  $\lambda_1$  throughout  $[T_1, T_z]$  is bigger than  $\lambda_0$  during historic non-stressor-event periods is over a confidence threshold  $\tau$ .

We call  $TSW[T_1, T_z]$  a **maximal stressful period**, if and only if it is a stressful period, and there exists none stressful period  $TSW[T_x, T_y]$ , whose time period encloses  $[T_1, T_z]$ , i.e.,  $(T_x \leq T_1) \wedge (T_y \geq T_z)$ . According to the definition, a big stress wave is also a stressful period. Among the six stress waves in Fig. 4, (A, B), (C, D, E), and (F) form three maximal stressful periods.

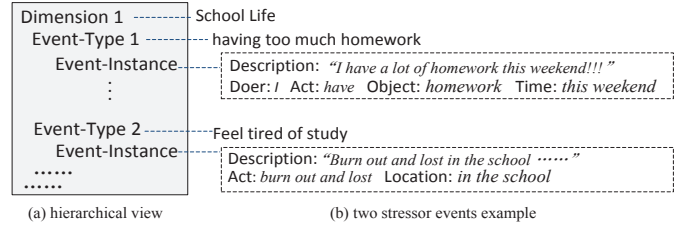


Fig. 5: Example of stressor events.

#### 4.2.2 Method of Maximal Stressful Periods Extraction

Discovery of maximal stressful periods proceeds in three steps:

**Step 1:** Identify all the big stress waves  $B$  and small stress waves that exist in the detected stress list. According to the definition of stressful period, each big stress wave in  $B$  is a stressful period. Let  $SP$  be a set of stressful periods. Initially,  $SP=B$ .

**Step 2:** Merge precedent or successive neighbouring small stress waves with each big stress wave  $b$  in  $SP$ , if the result is also a stressful period. Each time we probe the merging, starting from the neighbour with the more similar topic distribution as  $b$ . We repeat step 2 until the obtained stressful period cannot be longer. We drop  $b$  out of  $SP$ , and put the obtained longer stressful period into  $SP$ .

**Step 3:** Combine every two neighbouring stressful periods (without small stress wave gaps in between) in  $SP$  if the result also constitutes a stressful period as well. The procedure is the same as Step 2. We repeat Step 3 until no more combination is possible. We eliminate those successfully combined stressful periods from  $SP$  and put the obtained longer stressful period into  $SP$ . The result  $SP$  keeps all the maximal stressful periods.

### 4.3 Extraction of Stressor Events from Maximal Stressful Periods

#### 4.3.1 Definition of Stressor Events

A stressor event  $se$  is of the form [Dimension, Event-type, Event-instance (description, doer, act, object, time, location), influence], Element doer, act, object, time, or location could be empty. Table 1 gives different event types in each of the five dimensions - *school life, family life, peer relation, self-cognition, and romantic relation*. Event-instance refers to the happening of a specific event with the involved doer, act, object, time, location, and a detailed linguistic description. An example of two stressor events of type "having too much homework" and type "feel tired of study" in the "school life" dimension is illustrated in Fig. 5.

As multiple event instances could be extracted through linguistic analysis during the teen's stressful period, we rank them based on their *stress influence*. Considering that a stressor event usually stimulates the increase of stress levels, the higher the stress level increases, and the higher the impact is. We measure the stress influence of the event as

$$influence = w_1 * \frac{level(se)}{\arg \max_{se^* \in SE} level(se^*)} + w_2 * \frac{increase(se)}{\arg \max_{se^* \in SE} increase(se^*)} \quad (7)$$

where 1)  $level(se)$ : the accumulated stress level in the time unit when the event's influence starts; To identify the event's starting time, we look for linguistic words/phases that indicate time (e.g., exact date, today, tomorrow, yesterday, the day after tomorrow, on the weekend, etc.). When no such words are found or the starting time is later than the posting time, we take the posting time as the event's influence starting time. 2)  $increase(se)$ : the increase rate of the accumulated stress level in the time unit when event's influence starts. 3)  $SE$  is the set of event instances extracted from within the examined maximal stressful period, and 4)  $w_1=w_2=0.5$  in the study.

When multiple event instances referring to the same occurrence of an event are extracted, we take their maximal influence value. Due to the casuality of the posts, we may not be able to obtain all the details (i.e., role and act) of an event instance. In this case, we use symbol “-” to represent.

### 4.3.2 Method of Stressor Events Extraction

From each post in an identified *maximal stressful period*, we extract stressor events, combine and rank the extracted stressor events from all the posts in the period, and finally present them in a 3-leveled stressor event hierarchy (Dimension, Event-type, Event-instance).

**Step 1:** Extract stressor events from a post within a maximal stressful period.

- We first build five stress-related lexicons corresponding to the five stress dimensions in Table 3. The *school life* dimension contains 300 phrases, the *family life* dimension contains 296 phrases, the *peer relation* dimension contains 222 phrases, the *self-cognition* dimension contains 784 phrases, and the *romantic relation* dimension contains 350 phrases. Each phrase in a dimension may describe one or more event types. For example, phrase “*mid-term exam*” in the school-life dimension lexicon depicts event type “*Having exams or tests*”. For a phrase referring to a personal noun like “*teacher*” and “*classmate*”, we mark it as possible *role* and *object*, and if the phrase represents an action like “*change school*”, we mark it *act*.
- We then extract stressor events from a post within a maximal stressful period. We apply the Chinese natural language processing tool LTP [58], [59] to each sentence in the post. The aim of LTP is to linguistically localize the central verb/verb-phrase and its associated semantic roles like the *doer*, *object*, *time*, and *location* of the action in a sentence. Two cases exist here. 1) The verb/verb-phrase exists in one or more stress dimension lexicons, and may or may not be annotated with certain event types. We take this verb/verb-phrase as the *act*, its associated *doer*, *object*, *time*, and *location* (if found) as its *doer*, *object*, *time*, and *location*. The whole sentence forms the *description* of the event instance. 2) The verb/verb-phrase does not exist in any stress dimension lexicon. However, if the found *doer* or *object* exists in one or more stress dimension lexicons as the corresponding role, we also output the event instance whose *act* is empty.

**Step 2:** Combine and rank stressor events from all the posts within the maximal stressful period.

From each post, we may extract multiple event instances of multiple event types. We count the occurrence frequency of event instances for each event type, and rank the event type by the frequency from high to low. For each event type, event instances with complete {*doer*, *act*, *object*, *time*, *location*} are ranked ahead. In each maximal stressful period, event-instances under the same event-type and dimension are sorted based on their *influence* values. Event-types under the same dimension are sorted according to the accumulated influence of its event-instances, and dimensions are sorted according to the accumulated influence of its event-types.

## 5 PERFORMANCE STUDY

### 5.1 Experimental Setup

**Dataset.** We divide the 124 high-school students’ posts in the case study in Section 3 into two parts. For each grade, the first training part covers the former 60% period for each student, and the second testing part covers the latter 40% period. There are 12,139 posts in the first part, and 17,093 posts in the second part.

**Ground truth.** The 122 stressful study-related events with their duration published in the agenda of Taicang High School on its official website are taken as the ground truth for the analysis of detected stressful periods and stressors. Due to the continuity of human’s

TABLE 4: Performance of detecting maximal stressful periods ( $\tau=0.59$ )

	Recall	Precision	$F_1$ -Measure
Grade 1	0.707	0.705	0.706
Grade 2	0.738	0.726	0.711
Grade 3	0.838	0.778	0.784
Average	0.761	0.737	0.734

emotion, we add a  $\Delta_t$ -day on either side of an event period, based on the significance of the event and event duration (shown in Table 2). In the experiment, we set  $\Delta_t = \text{Significance} * \text{Duration}$ . If two events overlap temporally, we collapse them into one. From the training data, we derive  $\lambda_0$  (the stressful posting rate during historic non-stressor-event periods) for each teen, and apply it to detect maximal stressful periods and stressor events on the teen’s testing data.

**Metrics.** We evaluate the detection performance of maximal stressful periods and stressor events through *precision*, *recall*, and the weighted average performance of precision and recall  $F_1$ -measure, where  $F_1\text{-measure} = 2 * \text{precision} * \text{recall} / (\text{precision} + \text{recall})$ .

If a detected maximal stressful period overlaps with a study-related event period in the ground truth, we count this detection result correctly. We evaluate our detected stressor events from all the correctly identified maximal stressful periods. If a stressor event in the *school life* dimension is returned and ranked within the top-1/top-3 result list, we count this event detection correctly.

## 5.2 Experiment Results

Applying the stress detection function [11] to each post, we firstly sense the stress level in the category of school life, family life, peer relation, self-cognition and romantic-relation. We then discover maximal stressful periods and rank stressor events within each maximal stressful period.

### 5.2.1 Detection of Maximal Stressful Periods

As the ground truth includes only school life related events, we drop out detected maximal stressful periods whose stressors are not in the *school life* dimension.

Fig. 6 shows the average stressful period detection performance under different confidence thresholds  $\tau$ s. When  $\tau$  is less than 0.5, the recall is stable, and around 0.8 on average for the three grades students. It starts to drop after  $\tau > 0.5$ . This is because a higher  $\tau$  poses a more strict requirement on the stressful posting rate during stressor event periods. On the contrary, the precision increases greatly along with  $\tau$  till 0.5, remains relatively stable around 0.745 until  $\tau = 0.6$ , and then decreases. This verifies the statistical model that through the difference of stressful posting rates during stressor and non-stressor event periods, we can distinguish the stressful and non-stressful periods. When  $\tau$  is around 0.59, we can achieve the highest  $F_1$ -measure 0.734.

Comparing different grades of students, we find that the average detection performance on Grade 3’s students is the best, as illustrated in Table 4. This might be because Grade 3 faces the strongest stress due to the final college entrance examination than Grade 1 and 2. In its training and testing data sets, the number of scheduled stressor events and stressful periods is the most, leading to the best result.

Another interesting observation we made is that the precision rate is less than the recall rate. This is due to the fact that the ground truth contains only various examination events. However, some other stressful periods like “*Having too much homework*” in the *school life* dimension are also returned, as the school normally assigns a lot more homework to its students before an exam. As these periods do not exist in the ground truth, they are regarded as wrong results, negatively influencing the detection precision.



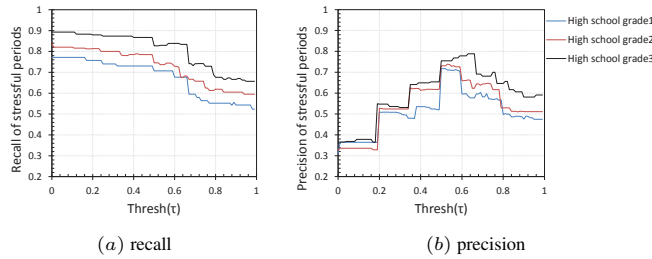


Fig. 6: Performance of detected maximal stressful periods for the 124 students in three different grades.

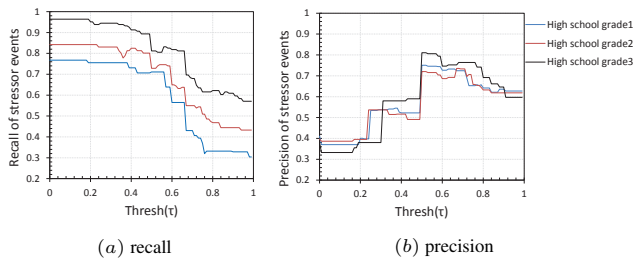


Fig. 7: Performance of detected stressor events (in top-3 rank) for the 124 students in three different grades.

### 5.2.2 Identification of Stressor Events

The performance of stressor events detection exhibits a similar trend as that of stressful periods detection, as shown in Fig. 7. When  $\tau$  is around 0.56, the highest  $F_1$ -measure 0.759 can be obtained. Table 5 compares the performance when we require the extracted stressor event is top-1 or top-3 ranked in the identified maximal stressful period. As the former is more strict, its  $F_1$ -Measure performance is 14.4% less than the later one's, mainly caused by the 8.59% decrease of the recall.

### 5.2.3 Comparison with the State-of-Art Personal Life Event Detection Approach

While there was none approach of identifying teens stressful periods and stressor events within the stressful periods as a whole in the literature, we compare the performance of our linguistics-based stressor events detection approach with that of the state-of-art statistics-based personal life event detection approach proposed by Li *et al.* [56]. The steps for major personal life events detection are the following. 1) *Personal life event clustering*. The authors defined a small set

TABLE 5: Performance of detecting stressor events ( $\tau=0.56$ )

	Top-1			Top-3		
	Recall	Precision	$F_1$ -Measure	Recall	Precision	$F_1$ -Measure
Grade1	0.608	0.723	0.661	0.711	0.746	0.728
Grade2	0.671	0.69	0.680	0.745	0.715	0.731
Grade3	0.722	0.76	0.741	0.831	0.806	0.818
Average	0.667	0.724	0.694	0.763	0.756	0.759

TABLE 6: Comparison of the linguistics-based stressor event extraction method with the statistics-based life event detection approach

	Stressor Event Extraction (Top-1)			Life Event Detection		
	Recall	Precision	$F_1$ -Measure	Recall	Precision	$F_1$ -Measure
Grade1	0.608	0.723	0.661	0.521	0.614	0.564
Grade2	0.671	0.69	0.680	0.562	0.641	0.599
Grade3	0.722	0.76	0.741	0.596	0.655	0.624
Average	0.667	0.724	0.694	0.560	0.637	0.596

of seed responses which capture common CONGRATULATIONS and CONDOLENCE, including such phrases as “Congratulations”, “Congrats”, “Sorry to hear that”, “Awesome”, and gather tweets that were observed with seed responses. Next, an LDA based topic model [60] is used to cluster the gathered tweets to automatically identify important categories of major life events. One of the authors manually inspected the resulting major life event types inferred by the model, and manually assigned them labels such as “getting a job”, “graduation”, or “marriage”, and discarded incoherent topics. A semi-supervised bootstrapping approach [61] was further used to expand the set of response seeds and event-related posts, with totally 42 major life events being manually identified. 2) *Event life identification*. A 43-class maximum entropy classifier based on tweet's four features (*Word*, *NER*, *Dictionary*, and *Window around the dictionary term*) was trained and built to judge whether a given tweet corresponds to one of the 42 predefined life events. 3) *Self-reported information identification*. Then they trained a SVM classifier to further identify whether each post referred to an event directly involving the user who published it. Features used in this step were *Bigram*, *Window around the topic word*, *Tense*, *Factuality*, *First person singular*, and *Dependency between the subject and topic word*. 4) *Event property Extraction*. They manually assigned one property to 13 life event categories (e.g., “Name of Enterprise” was assigned as the property for the life event “Getting a job”), and extracted the property of the detected life event using the CRF model [62]. Main features used in this step were *Word*, *POS*, *NER* and *The dictionary of universities and employers*.

We apply the above approach to extract stressor events from each post. Instead of using a LAD-CLUSTERING + HUMAN-IDENTIFICATION strategy to obtain major life event types at step 1), we base the stressor event detection on the five stress lexicons built according to psychological questionnaires for 69 stressor events. As the targets are stressor events rather than personal life events, and a stressor event might be an external event, we omit the self-reported information identification step 3). Also, the above event property extraction step 4) is limited to the manually assigned properties for specific life events, and is thus omitted as well.

We train a 70-class maximum entropy classifier for each of the three grades students respectively using four types of features (*Word*, *NER*, *Dictionary* and *Window around the dictionary term*). For the data set of each grade, we took 70% posts for training and the rest 30% posts for testing. The results are listed in Table 6. We can see that our linguistic-based method is 13.72% higher in *precision*, 19.18% higher in *recall*, and 16.50% higher in  $F_1$ -measure than the statistic-based major life event detection approach. The reason may be that our method firstly identifies maximal stressful periods, then in each maximal stressful period, we extract stressor events. This enables to filter noisy posts and localize stressor events more accurately, resulting in higher precision and recall.

Furthermore, we look into the detected stressor events from all the detected maximal stressful periods (not constrained to *school life* dimension). For each student, we count the occurrence number *occur\_num* of each stressor dimension across all the detected maximal stressful periods, and compute its occurrence ratio by  $occur\_num / max\_stressful\_period\_num$ . Table 10 shows the average occurrence ratios for students at Grade 1, 2, and 3, respectively. Stress in *School life* is the most common among the grade 3's students, who have to face the most frequent exams and college entrance examination. For all grades, the top stressor dimensions come from *self-cognition* and *school life*. This conforms to the adolescent psychological investigation result that problems in school life usually accompanied with teens inner cognition problems [63].

Table 7,8,9 list the occurrence numbers of stressor event types among students at grade 1, 2, and 3, respectively. More stressors in *school-life* dimension exist for students at Grade 3, and students at Grade 1 are faced with more stressors in *self-cognition* dimension, compared with students at other two grades.

TABLE 7: Occurrence number of stressor event types among grade 1’s students

Dimension	Event-Type	Student number	Event-Instance
<i>Self-</i>	Concern about your future	43	<i>I'm feeling hopeless, experiencing so much disappointment.</i> (Doer: I, Act:feel hopeless, experiencing, Object:disappointment)
<i>Cognition</i>	Depression	34	<i>No one can understand my feelings, so tired and bad.</i> (Doer:no one, Act:understand, Object: my feelings)
	Getting sick or hurt	33	<i>I had to do math with the 38 degrees fever in the stuffy classroom.</i> (Doer: I, Act: do, Object: math, Location: classroom)
	Unsatisfied with appearance/ weight	29	<i>My hair was cut off today, such bad hair!</i> (Doer: my hair, Act: cut off, Time: today)
	Changes of daily routine	27	<i>I have tried hard to adjust myself into the dorm life.</i> (Doer: I, Act: try hard to adjust, Object: myself)
	Bad sleep conditions	26	<i>I hardly fell into sleep recently until midnight.</i> (Doer: I, Act: hardly fall into sleep, Time: recently)
	Doubt the meaning of life	25	<i>My life seems ruined, the moment I heard the score.</i> (Doer: my life, Act: ruin)
	Having to take on new family responsibilities	20	<i>You should fight for your dream and your parents.</i> (Doer: you, Act: fight for, Object: your dream and your parents)
	Lost things or being stolen	16	<i>My phone was lost and the number has been written off.</i> (Doer: my phone, Act: lose)
	Internet addiction	13	<i>I play all kinds of games now to feel not so lonely.</i> (Doer: I, Act: feel not, Object: so lonely)
	Having to make decisions about future work	13	<i>I felt lost with so many options.</i> (Doer: I, Act: feel lost)
	Being frightened/threatened unexpectedly	11	<i>The old man terrified me out of my wits on my way home.</i> (Doer: old man, Act: terrify, Object: me, Location: on my way home)
<i>School</i>	Having too much homework	29	<i>I am driven mad by too much homework!</i> (Doer: I, Act: am driven mad, Object: too much homework)
<i>Life</i>	Unsatisfied with exam results	25	<i>Seeing the score, my efforts were in vain.</i> (Doer: my efforts, Act: are, Object: in vain)
	Keeping up with schoolwork	23	<i>When can I listen to the maths classes without sleepiness.</i> (Doer: I, Act: listen to, Object: maths classes)
	Having to study things you do not understand	18	<i>This course is too difficult.</i> (Doer: course, Act: is, Object: too difficult)
	Getting along with your teachers	14	<i>I prefer a gentle, kind teacher, instead of this shrew talking dirty.</i> (Doer: shrew, Act: talk dirty)
	Going to school	12	<i>I am really scared of going to school.</i> (Doer: I, Act: am scared of, Object: going to school)
	Feeling tired of study	11	<i>I will never know my course.</i> (Doer: I, Act: never know, Object: my course)
	Wasting time	11	<i>Time is fair, but time is really not enough for me!</i> (Doer: time, Act: is not, Object: enough)
	Fail to be admitted to a university	9	<i>The score is hard to enter my ideal university.</i> (Doer: score, Act: is hard to enter, Object: university)
<i>Romantic</i>	Getting along with your boy/girl-friend	40	<i>You are so busy that our love can not receive any concern.</i> (Doer: our love, Act: can not receive, Object: concern)
<i>Relation</i>	Breaking up with your boy/girl-friend	23	<i>It's funny that I was wondering to save something between us.</i> (Doer: I, Act: save, Object: us)
	Secret adoration	20	<i>The girl you cannot reach always attracts you.</i> (Doer: girl, Act: attract, Object: you)
	Being ignored or rejected by the person you like	13	<i>I am always waiting for you, no matter how you rocking my love.</i> (Doer: you, Act: rock, Object: my love)
<i>Family</i>	Family financial difficulties	17	<i>Huh. Money is always more important than me.</i> (Doer: money, Act: is, Object: important)
<i>Life</i>	Parents expecting too much from you	13	<i>Have you ever considered my feelings to force your dream on me.</i> (Doer: you, Act: force, Object: your dream)
	Suffering beating/scolding from parents	12	<i>My parents criticized me to speechlessness.</i> (Doer: my parents, Act: criticize, Object: me)
	Lack of understanding by your parents	11	<i>If you cannot support me, please don't deny me without your brain.</i> (Act: do not deny, Object: me)
<i>Peer</i>	Being hassled for not fitting in with peers	15	<i>Now I know what it feels like when past friends going away one by one.</i> (Doer: past friends, Act: go away)
<i>Relation</i>	Fierce competition among peers	8	<i>I am a love loser.</i> (Doer: I, Act: am, Object: love loser)

TABLE 8: Occurrence number of stressor event types among grade 2's students

Dimension	Event-Type	Student number	Event-Instance
Self-cognition	Concern about your future	55	<i>Life is anything but smooth.</i> (Doer: Life, Act: is, Object: anything but smooth)
	Changes of daily routine	39	<i>I must keep up with the pace of new term.</i> (Doer: I, Act: keep up, Object: new term)
	Depression	39	<i>The feel of suicide tells me that I will go death again.</i> (Doer: I, Act: go death)
	Getting sick or hurt	38	<i>Oh,God! My cramp were killing me!</i> (Doer: my cramp, Act: kill, Object: me)
	Having to take on new family responsibilities	38	<i>#Monthly salary# How much salary can let me feel safety.</i> (Doer: salary, Act: feel safety, Object: me)
	Unsatisfied with appearance/ weight	37	<i>From today I'm going to become short, poor and ugly.</i> (Doer: I, Act: become short, poor and ugly, Time: today)
	Doubt the meaning of life	35	<i>Finally the disappointment will swallow me.</i> (Doer: disappointment, Act: swallow, Object: me)
	Bad sleep conditions	34	<i>My own insomnia walked through my own world.</i> (Doer:my own insomnia, Act:walk through, Object:my own world)
	Regret for past behaviors	24	<i>If I have the chance to restart my high school three years ago.</i> (Doer:I, Act:restart, Object:high school, Time: three years ago)
	Lost things or being stolen	24	<i>My one-hundred hat was lost home this time.</i> (Doer:hat, Act:lose, Location:home)
School	Having too much homework	40	<i>Recently I am used to making homework up at midnight.</i> (Doer:I, Act:make homework up, Time: Recently, midnight)
Life	Unsatisfied with exam results	32	<i>No one can save my transcripts!</i> (Doer:no one, Act:save, Object:my transcripts)
	Having to study things you do not understand	29	<i>I can only play the game rather than the elusive course.</i> (Doer:I, Act:play the game)
	Keeping up with schoolwork	26	<i>Bless me passing the Chinese and chemistry tests the day after tomorrow.</i> (Act:pass, Object:Chinese and chemistry tests, Time: the day after tomorrow)
	Wasting time	22	<i>Half of the summer holiday has passed.</i> (Doer:half of the summer holiday, Act:pass)
	Getting along with your teachers	20	<i>The stupid teacher can destroy my whole course!</i> (Doer:teacher, Act:destroy, Object:my whole course)
	Feeling tired of study	17	<i>It's really hard to preview the course which has already explained by the teacher.</i> (Act:is hard to preview, Object:course)
	Fail to be admitted to a university	15	<i>The score of each university in Beijing is higher than the other.</i> (Doer:score, Act:is, Object:higher, Location: Beijing)
	Fail to reach expectations	12	<i>It's boring that the head teacher always expects the score.</i> (Doer:head teacher, Act:expect, Object:score)
Romantic	Getting along with your boy/girl-friend	52	<i>I should say sorry to you that I killed the faith.</i> (Doer:I, Act:kill, Object:faith)
	Breaking up with your boy/girl-friend	30	<i>I was left behind and cannot stopped missing your back.</i> (Doer:I, Act:leave behind)
Relation	Being ignored or rejected by the person you like	19	<i>My love seems to be transparent for you.</i> (Doer:my love, Act:seem to, Object:transparent)
Family	Lack of understanding by your parents	25	<i>You have never tried to understand me like a mum.</i> (Doer:you, Act:never try to understand, Object:me)
Life	Family financial difficulties	22	<i>Finally a day I can buy all the books in my shopping cart.</i> (Doer:I, Act:buy, Object:books)
	Suffering beating/scolding from parents	19	<i>Every time he defeated me by his identity as the parent.</i> (Doer:he, Act:defeated, Object:me)
	Parents expecting too much from you	17	<i>I'm too tired to tell parents that I cannot live up to their expectations.</i> (Doer:I, Act:live up, Object:their expectations)
	Not being taken seriously by your parents	10	<i>Adults always cannot keep faith.</i> (Doer:adults, Act:cannot keep faith)
Peer	Being hassled for not fitting in with peers	24	<i>But all of you make friends with me and then deceive me.</i> (Doer:all of you, Act:deceive, Object:me)
Relation	Fierce competition among peers	10	<i>I cannot lose my future.</i> (Doer:I, Act:cannot lose, Object:my future)
	Rumor or satire from peers	8	<i>Don't point fingers at me.</i> (Act:point fingers)
	Trying to keep up with the Joneses	7	<i>I took part in this foolish farce full of vanity and envy.</i> (Doer:I, Act:take part in, Object:farce)

TABLE 9: Occurrence number of stressor event types among grade 1’s students

Dimension	Event-Type	Student number	Event-Instance
Self-cognition	Concern about your future	16	<i>Suddenly I feel that I struggle to grow up to suffer more in the future.</i> (Doer:I, Act:struggle to grow up)
	Unsatisfied with appearance/weight	11	<i>I am not the dwarf.</i> (Doer:I, Act:am not, Object:dwarf)
	Doubt the meaning of life	11	<i>My view of life was pushed to the verge of collapse once again.</i> (Doer:view, Act:push, Object:collapse)
	Getting sick or hurt	10	<i>Knocked over by fever yesterday.</i> (Doer:fever, Act:knock over, Time: yesterday)
	Changes of daily routine	9	<i>I gradually sank into the endless loop of staying up.</i> (Doer:I, Act:sink, Object:staying up)
	Depression	8	<i>If I cry at this moment, will anyone come to me?</i> (Doer:I, Act:cry)
	Lost things or being stolen	8	<i>My money was lost again!!!</i> (Act:lose, Object:my money)
	Bad sleep conditions	7	<i>I slept early to adjust my disorder biological clock.</i> (Doer:I, Act:adjust, Object:biological clock)
School	Unsatisfied with exam results	14	<i>The exam and score almost drive me to death.</i> (Doer:exam and score, Act:drive, Object:me)
Life	Fail to be admitted to a university	9	<i>The score was estimated, and the bomb is going to explode.</i> (Doer:-, Act:estimate, Object:score)
	Having to study things you do not understand	8	<i>I Really feel difficult learning physics.</i> (Doer:I, Act:feel difficult, Object:learning physics)
	Having too much homework	8	<i>We suffer the most homework in the whole school.</i> (Doer:we, Act:suffer, Object:homework, Location: in the whole school)
	Wasting time	7	<i>It vaguely seems to me that I don't have enough time.</i> (Doer:I, Act:do not have, Object:time)
	Worry about cannot keeping ahead	6	<i>Taking the second place means I am the loser!</i> (Doer:I, Act:am, Object:loser)
	Keeping up with schoolwork	5	<i>I cannot bear the torture of math class.</i> (Doer:I, Act:cannot bear, Object:torture of math class)
	Getting along with your teachers	4	<i>I am going to kill the English teacher!</i> (Doer:I, Act:kill, Object:English teacher)
	Scolded by teachers	3	<i>I just felt hurt, depressed, self-abased and sad.</i> (Doer:I, Act:feel hurt, depressed, self-abased and sad)
	Feeling tired of study	3	<i>My holiday is filled with all kinds of homework.</i> (Doer:My holiday, Act:fill with, Object:homework)
	Going to school	3	<i>Unescapably, it's time to go back to school.</i> (Act:go back, Object:school)
Romantic	Getting along with your boy/girl-friend	14	<i>When can you be aware of my heart-broken feeling again and again?</i> (Doer:you, Act:be aware of, Object:heart-broken feeling)
Relation	Secret adoration	6	<i>I choose to stay silent and hide my throb.</i> (Doer:I, Act:hide, Object:throb)
	Breaking up with your boy/girl-friend	6	<i>I tried to forget you and tried to live by myself.</i> (Doer:I, Act:forget, Object:you)
Family	Family financial difficulties	7	<i>It seems that I cannot afford to buy all of them even years later.</i> (Doer:I, Act:cannot afford, Time: years later)
Life	Suffering beating/scolding from parents	4	<i>Why always talk about it to disgust your son?</i> (Act:disgust, Object:your son)
	Parents expecting too much from you	4	<i>I cannot reach the ambitious goal my parents set for me.</i> (Doer:I, Act:cannot reach, Object:goal)
	Parents hassling you about the way you look	3	<i>I don't know how long can I bear the nag.</i> (Doer:I, Act:bear, Object:nag)
	Not being taken seriously by your parents	3	<i>Parents like to judge everything around me with their emotion.</i> (Doer:parents, Act:judge, Object:everything)
	Serious disease of families or friends	2	<i>Hope that my uncle could revive earlier.</i> (Doer:my uncle, Act:revive)
Peer	Being hassled for not fitting in with peers	5	<i>Every one betrayed me.</i> (Doer:every one, Act:betray, Object:me)
Relation	Fierce competition among peers	4	<i>I'm too weak to handle such a fierce competition.</i> (Doer:I, Act:too weak to handle, Object:competition)

TABLE 10: Occurrence ratios of the five stress dimensions

Dimension	Grade 1	Grade 2	Grade 3	Average
Self-Cognition	0.428	0.460	0.464	0.451
School Life	0.206	0.214	0.255	0.225
Romantic Relation	0.178	0.153	0.249	0.193
Family Life	0.084	0.119	0.152	0.118
Peer Relation	0.032	0.053	0.081	0.055

## 6 DISCUSSION

A microblog-based framework towards the detection of adolescent psychological stressful periods and stressor events is presented. The underlying poisson-based probability model for the correlation between stressor events and stressful posting behaviors, together with the event-oriented stress detection method proposed in this paper, targets at sensing teens' psychological stressful periods and stressor events from appearance to essence.

### Advantage and Disadvantage of the Probability-based Model.

As this probability model only depends on stressful tweeting rates rather than on specific stressful tweeting time during event and non-event periods, the model is attractive for statistic inference in both measurement and computation. The shortcoming of the model is that it needs teen's stressful tweeting rates  $\lambda_0$  during non-stressor-event periods from historic training data. Also, the method relies on a confidence threshold  $\tau$  to distinguish between a stressful period and a non-stressful period. When  $\tau$  is around 0.5, we can obtain the best performance in this study. However, when a teen has no historic data to derive appropriate values of  $\lambda_0$  and  $\tau$ , we may have to pre-set the values by consulting with other teens with similar backgrounds and characteristics, and dynamically tune the values based on the feedback. Here, accurately identifying similar teens in stress detection is critical to the overall performance.

**Extension to Other Microblogs.** Although the current study is carried on Chinese posts on Tencent Weibo, the general framework (consisting of identifying stressful periods and then extracting stressor events) is applicable to other microblogs like Twitter, Tumblr, and Sino Weibo. However, for microblogs in different languages, different linguistic processing techniques are needed in building stress-related lexicons, sensing teens stressful emotion from each post, and finally extracting details of stressor events.

## 7 CONCLUSION

In this paper, we present techniques to analyze and identify teens stressful periods and stressor events from the microblog. A poisson-based probability approach is developed to model the correlation between stressor events and stressful posting behaviors. With the model, we discover maximal stressful periods and extract stressor events from each stressful period. We examine 124 high-school students' posts from January 1, 2012 to February 1, 2015, and detect stressful periods with (*recall* 0.761, *precision* 0.737, and  $F_1$ -measure 0.734), and top-3 stressor events with (*recall* 0.763, *precision* 0.756, and  $F_1$ -measure 0.759). The performance of our top-1 stressor event detection is better than the state-of-art personal life event detection approach, which is about 13.72% higher in *precision*, 19.18% higher in *recall*, and 16.50% higher in  $F_1$ -measure.

Due to the complexity of stressor theory, it can be said that identifying stressor events and incurred stressful periods is interesting and challenging, and much more related issues need further investigation. First, people's response to the stressor, defined as coping mechanism in psychology, is also an important element in stress theory. We will take coping mechanisms into consideration as a key factor in our future work. Second, while this study shows a promising direction to detect stressful periods and stressor events for adolescents, the microblog is still an insufficient source with a lot of data missing. Integrating multiple sources (e.g., GPS trajectory outlier data and sleep outlier data) could be helpful to provide more comprehensive and continuous information, enabling further improvement of the detection performance.

## ACKNOWLEDGMENTS

The authors would like to thank anonymous reviewers for the constructive comments and suggestions, enabling the great improvement of the manuscript. The work is funded by National Natural Science Foundation of China (61373022, 61532015, 71473146).

## REFERENCES

- [1] R. Hoffmann, "There is a difference between stress and stressors," NBC News, <http://www.tm.org/blog/meditation/there-is-a-difference-between-stress-and-stressors-and-you-need-to-know-what-it-is/>, 2015.
- [2] J. Aleccia, "Teens more stressed-out than adults," NBC News, <http://www.nbcnews.com/health/kids-health/teens-more-stressed-out-adults-survey-shows-n26921>, 2014.
- [3] B. F. Jeronimus, J. Ormel, A. Aleman, B. W. J. H. Penninx, and H. Riese, "Negative and positive life events are associated with small but lasting change in neuroticism," *Psychological Medicine*, vol. 43, no. 11, pp. 2403–2415, 2013.
- [4] F. Bousefsaf, C. Maaoui, and A. Pruski, "Remote assessment of the heart rate variability to detect mental stress," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on*, 2013, pp. 348–351.
- [5] F. Sun, C. Kuo, H. Cheng, S. Buthpitiya, P. Collins, and M. Griss, "Activity-aware mental stress detection using physiological sensors," in *Proc. of Social Informatics and Telecommunications Engineering*, 2012, pp. 211–230.
- [6] A. de Santos Sierra, C. Avila, J. Guerra Casanova, and G. del Pozo, "A Stress-Detection System Based on Physiological Signals and Fuzzy Logic," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4857–4865, 2011.
- [7] G. Rigas, Y. Goletsis, and D. Fotiadis, "Real-time driver's stress event detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 221–234, 2012.
- [8] G. Bauer and P. Lukowicz, "Can smartphones detect stress-related changes in the behaviour of individuals?" in *Proc. of PERCOM Workshop*, 2012, pp. 423–426.
- [9] R. Wang, F. Chen, Z. Chen, T. Li, and et al., "Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proc. of Ubicomp*, 2014.
- [10] Y. Xue, Q. Li, L. Feng, G. Clifford, and D. Clifton, "Towards a microblog platform for sensing and easing adolescent psychological pressures," in *Proc. of Ubicomp*. poster, 2013.
- [11] Y. Xue, Q. Li, L. Jin, L. Feng, D. Clifton, and G. Clifford, "Detecting adolescent psychological pressures from micro-blog," in *Proc. of HIS*, 2014.
- [12] Y. Xue, Q. Li, J. Jia, and L. Feng, "Helping teenagers relieve psychological pressures: A micro-blog based system," in *Proc. of EDBT*, 2014.
- [13] 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," in *Proceedings of the ACM International Conference on Multimedia, MM'14, Orlando, FL, USA, November 03 - 07, 2014*, 2014, pp. 507–516.
- [14] H. Lin, J. Jia, Q. Guo, Y. Xue, and et al., "Psychological stress detection from cross-media microblog data using deep sparse neural network," in *Proc. of ICME*, 2014, pp. 1–6.
- [15] L. Zhao, J. Jia, and L. Feng, "A category-dependent time-aware feature space for teenagers' stress detection from micro-blog," in *Proc. of the 4th IFIP Intl. Conf. on Artificial Intelligence in Theory and Practice*, 2015.
- [16] "The number of users in Tencent micro-blog has exceeded 500 million," <http://t.qq.com/p/t/173717058821583>.
- [17] 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.
- [18] 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.
- [19] G. Jiang, "The development of the chinese adolescent life events checklist," *Chinese Journal of Clinical Psychology*, vol. 8, no. 1, pp. 10–14, 2000.
- [20] 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.
- [21] S. Cohen, T. Kamarck, and R. Mermelstein, "A global measure of perceived stress," *Journal of Health and Social Behavior*, pp. 385–396, 1983.

- [22] Z. Ting and H. Huang, "An epidemiological study on stress among urban residents in social transition period," *Chinese Journal of Epidemiology*, pp. 760–764, 2003.
- [23] B. E. Pozos-Radillo, M. de Lourdes Preciado-Serrano, M. Acosta-Fernández, M. de los Angeles Aguilera-Velasco, and et al., "Academic stress as a predictor of chronic stress in university students," *Psicología Educativa*, vol. 20, no. 1, pp. 47–52, 2014.
- [24] S. Michie, "Causes and management of stress at work," *Occup Environ Med*, vol. 59, pp. 67–72, 2002.
- [25] D. Wohn and R. LaRose, "Effects of loneliness and differential usage of facebook on college adjustment of first-year students," *Computers and Education*, vol. 76, p. 158C167, 2014.
- [26] R. Kocielnik, N. Sidorova, F. Maggi, M. Ouwerkerk, and J. West-erink, "Smart technologies for long-term stress monitoring at work," in *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*, 2013, pp. 53–58.
- [27] K. Plarre, A. Raij, S. M. Hossain, A. A. Ali, and et al., "Continuous inference of psychological stress from sensory measurements collected in the natural environment," in *Proc. of Int. Conf. Information Process. Sens. Netw.*, 2011, pp. 97–108.
- [28] S. Fairclough and L. Venables, "Prediction of subjective states from psychophysiology: A multivariate approach," *Biol. Psychol.*, vol. 71, no. 1, pp. 100C–110, 2006.
- [29] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156C–166, 2005.
- [30] N. Hamid, F. of Electr. Eng., U. T. MARA, M. Shah Alam, N. Sulaiman, and et al., "Evaluation of human stress using EEG power spectrum," in *2010 6th International Colloquium on Signal Processing and Its Applications (CSPA)*, 2010, pp. 1–4.
- [31] Q. Xu, T. Nwe, and C. Guan, "Cluster-based analysis for personalized stress evaluation using physiological signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 275–281, 2015.
- [32] M. Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Prediction depression via social media," in *Proc. of ICWSM*, 2013, pp. 128–137.
- [33] M. Choudhury, S. Counts, and E. Horvitz, "Social media as a measurement tool of depression in populations," in *Proc. of ACM Web Science*, 2013, pp. 47–56.
- [34] Y. Shen, T. Kuo, I. Yeh, T. Chen, and S. Lin, "Exploiting temporal information in a two-stage classification framework for content-based depression detection," in *Proc. of PAKDD*, 2013, pp. 278–288.
- [35] M. A. Moreno, L. A. Jelenchick, K. G. Egan, E. Cox, and et al., "Feeling bad on facebook: depression disclosures by college students on a social networking site," *Depress and Anxiety*, vol. 28, no. 6, pp. 447–455, 2011.
- [36] J. Campisi, P. Bynog, H. McGehee, J. C. Oakland, S. Quirk, and et al., "Facebook, stress, and incidence of upper respiratory infection in undergraduate college students," *Cyberpsychology, behavior, and social networking*, vol. 15, no. 12, pp. 675–681, 2012.
- [37] J. M. Whitehill, L. N. Brockman, and M. A. Moreno, "'just talk to me': Communicating with college students about depression disclosures on facebook," *Journal of Adolescent Health*, vol. 52, no. 1, pp. 122–127, 2013.
- [38] D. R. George, C. Dellasega, M. M. Whitehead, and A. Bordon, "Facebook-based stress management resources for first-year medical students: A multi-method evaluation," *Computers in Human Behavior*, vol. 29, no. 3, pp. 559–562, 2013.
- [39] L. Qiu, H. Lin, A. K. Leung, and W. Tov., "Putting their best foot forward: Emotional disclosure on facebook," *Cyberpsychology, Behavior, and Social Networking*, vol. 15, no. 10, pp. 569–573, 2012.
- [40] R. B. Clayton, R. E. Osborne, B. K. Miller, and C. D. Oberle, "Loneliness, anxiousness, and substance use as predictors of facebook use," *Computers in Human Behavior*, vol. 29, no. 3, p. 687C693, 2013.
- [41] Bobo, Gilchrist, Elmer, and Schinke, "Hassles, role strain, and peer relations in young adolescents," *Journal of Early Adolescence*, vol. 6, pp. 339–352, 1986.
- [42] Y. L. Lu, "Everyday hassles and related information behaviour among youth: a case study in taiwan," *Information Research*, vol. 16, no. 1, pp. 95–116, 2011.
- [43] D. Y. Wang and B. Yao, *The Development of College Students Stressors Questionnaire*. Springer Berlin Heidelberg, 2011.
- [44] Z. You, M. Chen, S. Yang, Z. Zhou, and P. Qin, "Childhood adversity, recent life stressors and suicidal behavior in chinese college students," *Adversity, Life Stressors and Suicide*, vol. 9, no. 3, pp. 1–7, 2014.
- [45] D. Li, W. Zhang, X. Li, S. Zhen, and Y. Wang, "Stressful life events and problematic internet use by adolescent females and males: A mediated moderation model," *Computers in Human Behavior*, vol. 26, no. 5, pp. 1199–1207, 2010.
- [46] J. L. Bevan, R. Gomez, and L. Sparks, "Disclosures about important life events on facebook: Relationships with stress and quality of life," *Computers in Human Behavior*, vol. 39, pp. 246–253, 2014.
- [47] J. Weppner, P. Lukowicz, S. Serino, P. Cipresso, and et al., "Smartphone based experience sampling of stress-related events," in *Proc. of Pervasive Health*, 2013, pp. 464–467.
- [48] A. D. Sarma, A. Jain, and C. Yu, "Dynamic relationship and event discovery," in *Proc. of WSDM*, 2011, pp. 207–216.
- [49] A. Ritter, Mausam, O. Etzioni, and S. Clark, "Open domain event extraction from twitter," in *Proc. of KDD*, 2012, pp. 1104–1112.
- [50] E. Benson, A. Haghighi, and R. Barzilay, "Event discovery in social media feeds," in *Proc. of ACL*, 2011, pp. 389–398.
- [51] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in *Proc. of WWW*, 2010, pp. 851–860.
- [52] A. Popescu and M. Pennacchiotti, "Detecting controversial events from twitter," in *Proc. of CIKM*, 2010, pp. 1873–1876.
- [53] L. C. Yu, C. L. Chan, C. C. Lin, and I. C. Lin, "Mining association language patterns using a distributional semantic model for negative life event classification," *Journal of Biomedical Informatics*, vol. 44, no. 4, pp. 509–518, 2011.
- [54] L. C. Chan, C. L. Lin, C. C. Lin, and et al., "Mining association language patterns using a distributional semantic model for negative life event classification," *Journal of Biomedical Informatics*, vol. 44, no. 4, pp. 509–518, 2011.
- [55] J. Li and C. Cardie, "Timeline generation: tracking individuals on twitter," in *Proc. of WWW*, 2014, pp. 643–652.
- [56] J. Li, A. Ritter, C. Cardie, and E. H. Hovy, "Major life event extraction from twitter based on congratulations/condolences speech acts," in *Proc. of EMNLP*, 2014, pp. 1997–2007.
- [57] H. Jeffreys, *Theory of probability*. Oxford University Press, 1939.
- [58] 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.
- [59] W. Che, Z. Li, and T. Liu, "Ltp: A chinese language technology platform," in *Proc. of ACL*, 2010, pp. 13–16.
- [60] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.
- [61] Z. Kozareva and E. H. Hovy, "Not all seeds are equal: Measuring the quality of text mining seeds," in *Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 2-4, 2010, Los Angeles, California, USA*, 2010, pp. 618–626.
- [62] J. D. Lafferty, A. McCallum, and F. C. N. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proceedings of the Eighteenth International Conference on Machine Learning (ICML 2001), Williams College, Williamstown, MA, USA, June 28 - July 1, 2001*, 2001, pp. 282–289.
- [63] D. G. Byrne, S. C. Davenport, and J. Mazanov, "Profiles of adolescent stress: the development of the adolescent stress questionnaire (asq)," *Journal of Adolescence*, vol. 30, no. 3, pp. 393–416, 2007.