TeenRead: An Adolescents Reading Recommendation System Towards Online Bibliotherapy

Yunxing Xin, Yongqiang Chen, Li Jin, Yici Cai, Ling Feng
Dept. of Computer Science and Technology
Centre for Computational Mental Healthcare Research, Institute of Data Science, Tsinghua University, Beijing, China
Email: xinyunxing0001@163.com, l-jin12@mails.tsinghua.edu.cn, {chenyongqiang,caiyc,fengling}@tsinghua.edu.cn

***Abstract*—Bibliotherapy is an effective method of psychotherapy. It involves the reading of specific texts with the purpose of healing. It has been proved to be an effective way to deal with adolescents psychological stress. Specific reading materials are provided to patients with physical or mental diseases for the purpose of prevention, healing, and rehabilitation. But traditional bibliotherapy requires professional staff with the background of both psychological and library services, which is quite demanding and labor consuming. Moreover, bibliotherapy based on paper books is getting ill-fitted in the present big data era. To address the limitations, this paper proposes an online reading recommendation system called *TeenRead* to carry out bibliotherapy for adolescents. *TeenRead* involves the management of users and articles, analysis of users' dynamic reading behaviors, as well as the recommendation based on users' stress categories, stress levels, and reading interests. The results of the user study on 10 volunteers show that, the average decrease of users' stress level is significantly dropped by 22% after a period of reading on *TeenRead*, which proves that *TeenRead* performs pretty well as a new method of bibliotherapy.**

***Keywords*-Bibliotherapy; adolescent; stress easing; reading interest; reading recommendation.**

I. Introduction

With the advent of information age and the rapid development of society, mankind faces unprecedented competitive stress. Without exception, adolescents at the growing age are inevitably bearing different psychological stress from academic research, communication, family, emotion, self-awareness, employment, and so on. Chinese Education News reported an investigation of 6000 students from 23 universities in Beijing. The result shows that more than 18% of the respondents have medium or medium above psychological problems. Another statistical result made by the Psychology Consultation Center of Zhejiang University in China shows that the proportion of excellent students with psychological syndrome is more than 40% [1]. Even worse, due to lack of life experiences and anti-pressure capabilities, adolescents are usually unable to deal with stress properly and effectively, leading to more serious consequences.

Many psychological methods have been proposed to help cope with adolescent stress, such as mindfulness meditation [2], decompression Yoga [3], sports [4], and mental health care education [5]. But most of the decompression methods require either people's full attention, which is quite time consuming, or more or less professional cares, which are difficult to be carried out by adolescents. Considering the real condition of adolescents, bibliotherapy (also called as *reading therapy*) could be an effective way to ease stress. Bibliotherapy once served as a psychological treatment for college students, aiming to help them solving problems by reading some specific books or related information under the scientific guidance [6]. Bibliotherapy has many advantages [7]. Firstly, it's convenient, timely, and not subject to space constraints, because books are easy to carry and people can do some reading whenever they want. Secondly, its stress easing effects are far more durable than a single psychological lecture or psychological counseling, for books can be read repeatedly and produce continuous effects. Thirdly, there is a strong confidentiality in bibliotherapy, and adolescents' privacy are protected pretty well. Fourthly, different from those therapeutic measures such as psychological consultation, bibliotherapy integrates the effect of treatment, prevention, and development at the same time. Through reading, bibliotherapy could relieve adolescents' stress at the very beginning, and meanwhile improve readers' knowledge and psychological immunity. Last but not least, bibliotherapy is easy to be carried out in campus, because libraries hold massive reading resources and lots of librarians and teachers of psychological center are paying much attention to the study of students' mental states.

On the other hand, bibliotherapy also faces many difficulties. First of all, it is demanding and labor consuming to make bibliotherapy in practice. As the person, who is responsible for bibliotherapy, is required of the knowledge of not only psychology but also library reader services [7]. For every stressful person, there must be a certain amount of reading materials customized for her/him. Secondly, the traditional methods of bibliotherapy may not be suitable in the current reading environment any more. Nowadays, adolescents are more likely to read through WWW and mobile phones other than paper books.

To deal with these problems of bibliotherapy, this paper builds a reading recommendation system called *TeenRead*, which automatically recommends reading materials to adolescents based on their stress and reading interests. *TeenRead*

serves as a web application so that users can access it through any platform with a web browser.

The major contributions of this work are two-folds.

- We discover a new method of bibliotherapy, which combines reading recommendation and stress easing in the age of information and big data era. To our knowledge, this is the first attempt in the literature aiming at reading recommendation in order to carry our bibliotherapy for adolescents. The developed *TeenRead* system recommends articles based on not only their best stress easing effect to users, but also their attractiveness to the users to make them willing to read.
- We address the cold-start problem of our *TeenRead* system by capturing users implicit reading behaviors and gradually tuning *TeenRead*'s recommendation to tailor to users' interests and meanwhile with stress easing effect. Our user study shows *TeenRead* could be a convenient and effective means of bibliotherapy to assist teens to release stress.

The remainder of the paper is organized as follows. Section 2 gives a brief introduction to bibliotherapy and popular recommendation approaches. Section 3 introduces the system framework of *TeenRead*. Section 4 and 5 detail the key reading recommendation module and users reading behavior analysis module, respectively. We conduct a user study to evaluate the effectiveness of *TeenRead* in Section 6. Finally, we give a discussion on *TeenRead* and conclude the paper in Section 7 and Section 8.

II. Related Work

This section consists of two parts. The first part briefly introduces the origin, theory, and history of bibliotherapy. The second part gives a summary of some widely used recommendation approaches.

A. Bibliotherapy

Bibliotherapy is recognized as an effective method of psychotherapy. There are two interpretations of bibliotherapy according to Webster's Third New International Dictionary: 1) the use of selected reading materials as therapeutic adjuvants in medicine and in psychiatry; 2) guidance in the solution of personal problems through directed reading [8]. The first interpretation says bibliotherapy could be used as a treatment of diseases caused by physiology or psychology. The second one explains from a broad perspective that bibliotherapy could deal some personal problems. More specifically, Rongione categorized the applicable scenarios of bibliotherapy into 5 areas, including physical disability, chronic illness, emotional problems, personality disorders, and socioeconomic issues [9].

Research on bibliotherapy has lasted hundreds of years in the west. In 1848, John M. Galt read the paper in the annual convention of the American Psychological Association (APA) about the effects of bibliotherapy, and analyzed the classification of patients and corresponding reading measures. This paper is considered to be the first one in bibliotherapy [10]. The formal and systematic research on bibliotherapy began with Samuel Mc Chord Crothers. He published a paper named "*A Literary Clinic at Atlantic Monthly*" and invented the word "*bibliotherapy*" [11]. After 1930s, bibliotherapy has received more attentions from relevant institutions in Western countries [11]. In 1939, the Hospital Library Branch of American Library Association (ALA) set up a bibliotherapy committee. In 1964, ALA held the first seminar on bibliotherapy. In 1970s, there were more than 4000 hospitals providing bibliotherapy services in the Soviet Union. Later in 1980s, a new wave of research on bibliotherapy occurred with the rapidly rising of people's psychological problems. International Federation of Library Associations and Institutions (IFLA) had affirmed the significance of bibliotherapy in patients' recovery.

To explain the mechanism of bibliotherapy, Hou took reading as a process that not only includes the recognition of characters, languages and image symbols, but also provides a way in which their inner feelings and emotions are deeply communicating with the works [1]. Yang also argued these good mental stimulations from reading are able to regulate people's immune function [12]. Wang *et al.* further explained the mechanism of bibliotherapy based on physiology and psychology. The reason why a book could cure is that the feelings it conveys to the reader have just weaken or offset the hidden unhealthy emotions. This is helpful in easing and alleviate reader's pathogenetic condition [13].

Nowadays, bibliotherapy has formed a fairly complete theory system and become an important research content in both library science and medical rehabilitation.

B. Recommendation Approaches

There are many recommendation approaches proposed in the last 20 years, Beel et al. generally divide them into seven categories, namely, *stereotyping, content-based filtering, collaborative filtering, co-occurrence, graph-based, global relevance,* and *hybrid* [14].

Stereotyping is one of the earliest recommendation approaches and was firstly introduced by Rich[10]. In stereotyping, stereotype are defined as a collection of characteristics. Articles with the maximum matching with the stereotypes are recommended. Content-based filtering (CBF) is one of the most widely used recommendation approaches [15], whose basic process is to match up the attributes of users' preferences and interests with the attributes of the content object. Collaborative filtering (CF) was another popular recommedation method, which is firstly introduced by Resnick *et al.* [16]. Its general idea is that if two users have similar historical ratings, they may like the same items. Co-occurrence recommendation was firstly used in Small's co-citation analysis system [17]. Small proposed that if the two papers are frequently co-cited, they are likely related to each

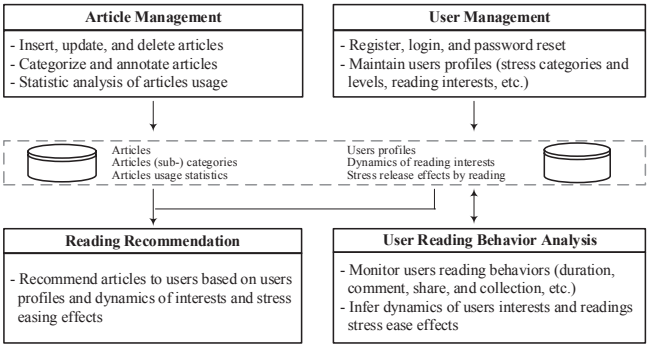


Figure 1. *TeenRead* framework

other. With no need for item features and users' ratings, Co-occurrence works quite well, except for the situation where the number of co-occurrence is extremely small, just like arXiv.org [14]. Graph based approach build grapy network based on the inherent connections, which could be citations, purchases, authorship, relatedness, co-occurrence and so on. Once the graph is built, the recommender system will find the closest items using graph metrics. Global relevance is based on a very simple assumption that people like what most others like. Every single method metioned above has native drawbacks more or less, so hybrid approaches are proposed to combine several recommendation classes to enhance the recommending effect. TechLens is one of the most influential hybrid recommender system, which consists of three CBF variations, two CF variations, and five hybrid approaches [14].

Different from the previous work, this study aims to carry out online bibliotherapy towards adolescents by reading recommendation. In order to make *TeenRead* effective in easing users' stress, the recommended articles should not only be specific to user's stress categories, but also be attractive to users to make them willing to read. The recommendation of *TeenRead* is based on CBF, and we have proposed several ways to deal with the cold-start problem, to balance the stress easing effects and interests of recommended articles, and to update recommendation based on users' dynamic reading behaviors.

III. System Framework

Figure 1 shows the framework of *TeenRead*, which is comprised of four major modules, namely, *article management, user management, user reading behavior analysis*, and *reading recommendation*. They cooperate as follows.

- The **article management module** is responsible for articles uploading, updating, and deleting. It also categorizes each article into one or more blocks, and annotates the content of the article with appropriate labels. The usage status and popularity degree of each article are also computed and dynamically maintained by the module at run-time. As the focus of this study is on articles recommendation, we omit the details of the article management in the paper.
- The **user management module** is in charge of users' accounts, including users' registration, login, and code verification by phone or email. In case users forget their passwords, there is also a way to help them to reset passwords. Besides, *TeenRead* allows users to explicitly input or update their profiles anytime, containing such personal information as gender, address, phone number, etc., as well as the readings interests, endured stress types and stress levels.
- The **users reading behavior analysis module** analyzes users reading behaviors on *TeenRead* to infer the dynamics of users reading interests and readings stress ease effect. Such behaviors include their articles reading durations, reading sequence per login, articles comments, agreement, sharing, and store-up. From these explicit reading activities, users' dynamics of reading interests could be estimated.
- The **reading recommendation module** recommends to users per week four articles around one specific theme, which conveys a positive attitude towards life to ease users' corresponding stress. The recommendation takes the following factors into consideration, namely, user's endured stress types and stress levels (if existing), explicitly indicated reading interests (if existing), dynamics of reading interests and readings stress easing effect, which are to be inferred from the above module. Dynamic adjustment of the recommendation algorithm is performed in order to provide more suitable articles for users' stress easing.

We detail the reading recommendation module and user reading behavior analysis module in the following two sections.

IV. Reading Recommendation Module

A. Guidelines for Reading Recommendation

1) Stress Easing by Reading: The primary goal of *TeenRead* is to alleviate teens psychological stress by reading. Therefore, the recommendation of articles mainly focuses on the effect of easing stress. In the study, six categories of stress that teens usually encounter are considered, i.e., SC={*study, family, affection, inter-personnel, self-recognition, employment*}, and $|SC|$=6. Stress levels take values from 0 to 5, SL={0,1,2,3,4,5}, denoting *no stress* to *very strong stress*, as illustrated in Table I.

2) Reading Interests: On the other hand, users reading interests are also a key factor for consideration. If the recommended articles are not attractive to the users, they won't be read, and thus make no attributions to the stress release. Ideally, the articles with the best effect of stress easing are the same as those with the highest interest degree for the users. Nevertheless, in reality, it's difficult

Table I
STRESS CATEGORIES AND STRESS LEVELS

Stress Category	Stress Level
study	0 - no stress
family	1 - very light stress
affection	2 - light stress
inter-personnel	3 - moderate stress
self-recognition	4 - strong stress
employment	5 - very strong stress

Table II
ARTICLE CATEGORIES AND SUB-CATEGORIES

Category	Sub-Category
study	university, course, resource, competition, rank, dual-degree, final year project, exempt from postgraduate, recommendation, postgraduate entrance exam, study-abroad, etc.
family	family member, kinship, blind date, trifles, etc.
affection	love, gloom, setback, psychology, etc.
inter-personnel	peer, roommate, classmate, society, etc.
self-recognition	self-observation, self-evaluation, etc.
employment	job selection, internship, part-time-job, job-seeking, written-test, interview, hiring, etc.
life	campus life, food, entertainment, sports, outdoors, life experience, encouragement, realization, travel,
art	literature, painting, music, dance, movie, drama, building, sculpture, poem, etc.
humor	joke, anecdote, satire, gossip, etc.
book	history, figure, reading, philosophy, etc.
.....	

to reach the maximum on both aspects, so it is necessary and challenging for *TeenRead* to make a tradeoff between these two.

3) Unification of Stress Easing and Reading Interest: To unify the two factors (stress easing and reading interest) into recommendation, we classify each article into different categories corresponding to stress categories. Due to the diversity and huge volume of articles, a finer grained categorization is performed, as reflected from the following two perspectives.

- Article's category set (denoted as AC) encloses stress category set SC, where $SC \subset AC$.
- An article categorization is further divided into a number of sub-categories, as shown in Table II. Let ACs denote the set of all sub-categories, and $|ACs|$ denotes the total number of articles sub-categories.

An article may belong to multiple categories and multiple sub-categories. For example, an article entitled "*Advice into Society*" gives suggestions on job, personnel relation, self-cognition, etc., falling into multiple categories and sub-categories.

B. Method of Reading Recommendation

1) Notions of Stress Easing Effect Vector and Interest Vector: Users may experience different types of stress. For each user, *TeenRead* computes and maintains *a stress easing effect vector* $\overline{E}$, together with *an interest vector* $\overline{I}$. The former records the stress easing effects by reading different categories of articles, while the later records the interests of reading different categories and sub-categories of articles.

Definition 1: Let SC be the set of stress categories. Corresponding to the six typical stress categories, a **user's stress easing effect vector** is defined as $\overline{E}$=$(c_1.ease, c_2.ease, \cdots, c_{|SC|}.ease)$, where for $\forall i\ (1 \leq i \leq |SC|) \wedge (c_i \in SC) \wedge (c_i.ease \in [0,5])$. Initially, $c_1.ease$=$\cdots$=$c_{|SC|}.ease$=0. □

Definition 2: Let ACs be the set of articles' sub-categories. A **user's interest vector** is defined as $\overline{I} = (cs_1.int, cs_2.int, \cdots, cs_{|ACs|}.int)$, where for $\forall i\ (1 \leq i \leq |ACs|) \wedge (cs_i \in ACs) \wedge (cs_i.int \in [0,1])$.

Initially, $cs_1.int$=$\cdots$=$cs_{|ACs|}.int$=0. □

2) Recommendation based on Stress Easing Effect Vector and Interest Vector: For simplicity, assume $SC(article)$ and $ACs(article)$ return the set of categories in SC and set of sub-categories in ACs which $article$ falls into.

We measure a user's interest in $article$ by summing up the interest values of article's sub-categories.

$$Interest(\overline{I}, article) = \sum_{cs \in ACs(article)} cs.int$$

Also, $article$'s stress easing effect is measured by summing up its categories's stress easing effects.

$$Ease(\overline{E}, article) = \sum_{c \in SC(article)} c.ease$$

Taking both interest and stress easing into account, the computation of an article 's recommendation score can be defined as follows.

Definition 3: Given a user's stress easing effect vector $\overline{E}$ and an interest vector $\overline{I}$. the **recommendation score of** $article$ **to the user** is defined as:

$$Recommend(\overline{E}, \overline{I}, article) = \rho * Ease(\overline{E}, article) + (1-\rho) * Interest(\overline{I}, article)$$

where parameter $\rho \in [0,1]$ is a coefficient for adjusting the weights of stress easing effect and reading interest. In the study, ρ=0.5. □

Note that at the initial start-up phase when $\overline{E}$=(0, $\cdots$, 0) and $\overline{I}$=(0, 0, $\cdots$, 0), all articles's recommendation scores are 0. In this case, *TeenRead* will randomly select articles to recommend.

According to Definition 3, good recommendation is subject to the preciseness of the user's stress easing vector $\overline{E}$

and reading interest vector $\overline{I}$. Despite users can manually explicitly indicate their endured stress and interested (sub-)categories via the *TeenRead*'s personal profile interface, the dynamics of their reading interests and stress status, plus users' non-real-time indication, demands for consistent maintenance of the two vectors. This will be taken care of by *TeenRead*'s **users reading behavior analysis module**.

V. Users Reading Behavior Analysis Module

This module is responsible for dynamically tuning user's interest vector and stress easing effect vector to tailor to users' interests and stress status.

A. Maintenance of User's Interest Vector $\overline{I}$

$\overline{I} = (cs_1.int, cs_2.int, \cdots, cs_{|ACs|}.int)$ reflects user's interested (sub-)categories. Each element cs_i (where $1 \leq i \leq |ACs|$) of $\overline{I}$ is a sub-category in ASc, and $cs_i.int \in [0, 1]$ is the interest degree of the user in sub-category cs_i, where cs_i=0 indicates no interest at all, and cs_i=1 indicates a complete interest. Initially, $\overline{I}$=$(0, 0, \cdots, 0)$.

Maintenance of user's interest vector $\overline{I}$ is conducted under the following two situations.

1) Re-Setting Up $\overline{I}$ Upon User's Explicit Selection: When the user explicitly selects/updates the interested (sub-)categories anytime in his/her personal profile, *TeenRead* will firstly initialize $\overline{I}$=$(0, 0, \cdots, 0)$. For the non-selected sub-categories, their interest degrees are set to 0.

Beyond that, considering the content similarity among sibling sub-categories under the same category (e.g., (*post-graduate study* and *study-abroad*), if a user is interested in cs_i, s/he may probably like other similar sibling sub-categories. Therefore, followed by each user's manual selection, we increase those non-selected sub-categories' interest degrees based on their similarity with the selected sibling sub-categories.

Let $cs_i, cs_k \in ACs$ be two sibling sub-categories under the same category. Their content similarity is measured by function $Sim(cs_i, cs_k) \in [0, 1]$. Assume cs_i is not manually selected by the user as an interested sub-category, but its sibling cs_k is selected, we refine $cs_i.int$=$arg\ max_{cs_k}\ Sim(cs_i, cs_k)$, and re-compute $\overline{I}$ as follows.

$$cs_i.int = \begin{cases} 1 & \text{if } cs_i \text{ is selected;} \\ max_{cs_k}\ Sim(cs_i, cs_k) & \text{if } cs_i \text{ is not selected, and its sibling sub-category } cs_k \text{ is selected;} \\ 0 & \text{otherwise} \end{cases}$$

In the current study, $Sim(cs_j, cs_k)$=0.2 for any two sibling sub-categories. After re-setting up all the elements of $\overline{I}$, we normalize each element value into [0, 1].

$$\overline{I} = (\frac{cs_1.int}{v}, \frac{cs_2.int}{v}, \cdots, \frac{cs_{|ASc|}.int}{v}),$$

where v=$\sqrt{(cs_1.int)^2 + (cs_2.int)^2 + \cdots + (cs_{|ASc|}.int)^2}$.

2) Adjusting $\overline{I}$ Based On User's Implicit Reading Behaviors: If the user does not explicitly select interested sub-categories, *TeenRead* dynamically updates $\overline{I}$ every w days according to user's reading behaviors during the past w days. User's interest in a sub-category is updated based on the following three factors:

(1) reading priority f_{seq} of articles containing this sub-category among the m recommended articles. Reading priority is an important measurement of user's interest. For simplicity of description, let $seq(article_{cs_i,k})$ indicates the reading order of the kth article which belongs to the ith sub-category cs_i, where $k \in [1, m]$, and $cs_i \in ACs$. If $article_{cs_i,k}$ is not read, we set $seq(article_{cs_i,k})$ to $m + 1$. We assume that, the earlier $article_{cs_i,k}$ is read, the larger its reading priority should be. So $f_{seq}(article_{cs_i,k})$ can be defined as follows:

$$f_{seq}(article_{cs_i,k}) = \frac{(m + 1 - seq(article_{cs_i,k}))}{m}$$

Bacause one article may belongs to many sub-categories, we make the maximum value of $f_{seq}(article_{cs_i,k})$ to be the real f_{seq} value of sub-category cs_i, which is:

$$f_{seq}(cs_i) = Max_{k \in [1,m]} f_{seq}(article_{cs_i,k}).$$

(2) whether the user has supported, commented, collected and shared articles containing this sub-category or not f_{act}. We take the above four actions into account for the reason that they are positively related to user's interests. We use four vectors SupportVec, CommentVec, CollectVec and ShareVec to store the frequency of the mentioned four reading actions to articles belong to the sub-category during the past w days. Taking SupportVec for example, the element of it is $support_{cs_i}$, indicating his/her supporting number to articles belonging to cs_i. So $support_{cs_i}$ is calculated like:

$$support_{cs_i} = \sum_{k \in [1,m]} doSupport(article_{cs_i}, k).$$

Where $doSupport(article_{cs_i,k}) = 1$ if the user have supported $article_{cs_i,k}$; otherwise it's 0;

The calculation of elements in CommentVec, CollectVec and ShareVec is similar to that of SupportVec. Let $comment_{cs_i,k}$, $collect_{cs_i,k}$ and $share_{cs_i,k}$ respectively denote the element of each vector above. To combine the effects of these four vectors, we standardize each elements to be $0 \sim 1$, so f_{act} can be calculated as:

$$f_{act}(cs_i) = (\frac{support_{cs_i}}{Max(SupportVec)} + \frac{comment_{cs_i}}{Max(CommentVec)} + \frac{collect_{cs_i}}{Max(CollectVec)} + \frac{share_{cs_i}}{Max(ShareVec)})/4$$

(3) time spent on reading articles containing this sub-category f_{time}. The longer the time spent on $article_{cs_i,k}$ is, the more the user is interested in sub-category cs_i. Even though this assumption may not be valid in all situations, it can be applied to most situations. In addition, the time spent on reading is also influenced by the length of the article. So we take the ratio of time spent on reading to the length of

this article to illustrate f_{time}, whose value for the past w days can be calculated as

$$f_{time}(int_i) = \sum_{k\in[1,m]} \frac{ReadTime(article_{cs_i,k})}{Len(article_{cs_i,k})}$$

where $ReadTime(article_{cs_i,k})$ is the time spent on reading $article_{cs_i,k}$, and $Len(article_{cs_i,k})$ is the length of $article_{cs_i,k}$ represented by the number of characteristics it.

After the values of f_{seq}, f_{act} and f_{time} are obtained, we can update $\overline{I}$ with their comprehensive effects. Let $\triangle(cs_i) = (\frac{f_{seq}(cs_i)}{Max(f_{seq})} + \frac{f_{act}(cs_i)}{Max(f_{act})} + \frac{f_{time}(cs_i)}{Max(f_{time})})/3$, $\overline{I}$ can be updated as following:

$$\overline{I}(cs_i.int) = \overline{I}(cs_i.int) + \alpha \cdot \triangle(cs_i)$$

where $i \in [1, |ACs|]$ and α is coefficient controlling how fast the reading behaviors could change $\overline{I}$. In *TeenRead*, we set α to 0.1. Afterwards, we should normalize $\overline{I}$ to 1 again.

B. Maintenance of Users Stress Easing Effect Vectors

$\overline{E} = (c_1.ease, c_2.ease, \cdots, c_{|SC|}.ease)$ reflects the stress easing effects of categories to the user. Each element c_i (where $1 \le i \le |SC|$) of $\overline{E}$ is a category in SC, and $c_i.ease \in [0, 5]$ is the easing effect of category c_i for the user, where $c_i = 0$ indicates no stress easing effect at all, and $c_i = 5$ indicates a very strong effect. Initially, $\overline{E} = (0, 0, \cdots, 0)$.

The update of $\overline{E}$ is manually conducted by users. While using *TeenRead*, users can update their stress categories and the corresponding stress level according to their current psychological states. Each time the update occurs, *TeenRead* updates user's $\overline{E}$ simultaneously. We expect the new $\overline{E}$ could not only indicates user's present stress, but also reflects the changes of stress and provides a more comprehensive representation for user's stress status.

Assume the user's last stress level PL_1 = (d_1, d_2, …, $d_{|SC|}$). After an update, stress level becomes PL_2 = (ℓ_1, ℓ_2, …, $\ell_{|SC|}$). The change of stress level is $\triangle P = PL_2 - PL_1$ = $\ell_i - d_i$, $i \in [1, |SC|]$. For each element of the new $\overline{E}$, let $\overline{E}(i)$ denote $c_i.ease$, $c_i \in CS$, we add a delta item to reflect the changes of stress, so: $\overline{E}(i) = \ell_i + \lambda \cdot \ell_i \cdot \frac{\triangle \ell_i}{5}$, where λ is an update coefficient controlling the weight of historical stress in new $\overline{E}$. In order to fuse with $\overline{I}$, $overline E$ also needed to be normalized to 1 in the same way as the normalization of $\overline{I}$. In the next two sections, we discuss how to choose a suitable λ for *TeenRead*, and evaluate the effect of λ.

(1) How to choose the value of λ?

TeenRead recommends articles with positive effects for easing users' stress based on psychology reading therapy, so we make an assumption that any increment of user's stress is not caused by reading these recommended articles. So we only consider the non-positive elements of $\triangle P$ and turn the positive to zero. Afterwards, in order to represent the magnitude of changes, we make $\triangle P = -\triangle P$ so each element of $\triangle P$ is either positive or zero.

Firstly, $\overline{E}(i)$ must be smaller than 5 according to Definition 1.

In $\triangle P$, the element $delta_i = -(\ell_i - d_i) = d_i - \ell_i \le 5 - \ell_i$. Bring $delta_i$ into $\overline{E}(i)$, we get its maximum:

$$Max(\overline{E}(i)) = \ell_i - \frac{\lambda \cdot \ell_i \cdot (5 - \ell_i)}{5} = -\frac{\lambda \cdot \ell_i^2}{5} + \ell_i \cdot (1 + \lambda)$$

This is a concave quadratic function with two fixed point (0, 0) and (5, 5). And it reaches its maximum $\rho_{max}(\lambda) = \frac{5}{4}(\lambda + \frac{1}{\lambda} + 2)$ when $\ell_i = f(\lambda) = \frac{5}{2} + \frac{5}{2 \cdot \lambda}$. Its shape is shown as Figure 2.

If $\lambda = 1$, $f(\lambda) = 5$ and $\rho_{max}(\lambda)$ get the minimum value 5. So (5, 5) is exactly the maximum point of $\overline{E}(i)$;

If $\lambda < 1$, $f(\lambda) > 5$ and $\rho_{max}(\lambda) > 5$. In the interval [0, 5], $Max(\overline{E}(i))$ increase from 0 to 5;

If $\lambda > 1$, $f(\lambda) < 5$ and $\rho_{max}(\lambda) > 5$. $Max(\overline{E}(i))$ gets the maximum value bigger than 5 in interval [0, 5], which contradicts with our assumption that $\overline{E}(i) \in [0, 5]$. So λ should be smaller than 1.

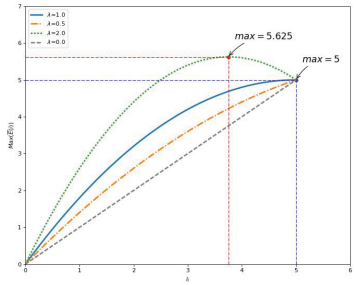


Figure 2. The shape of $Max(\overline{E}(i))$ with different λ

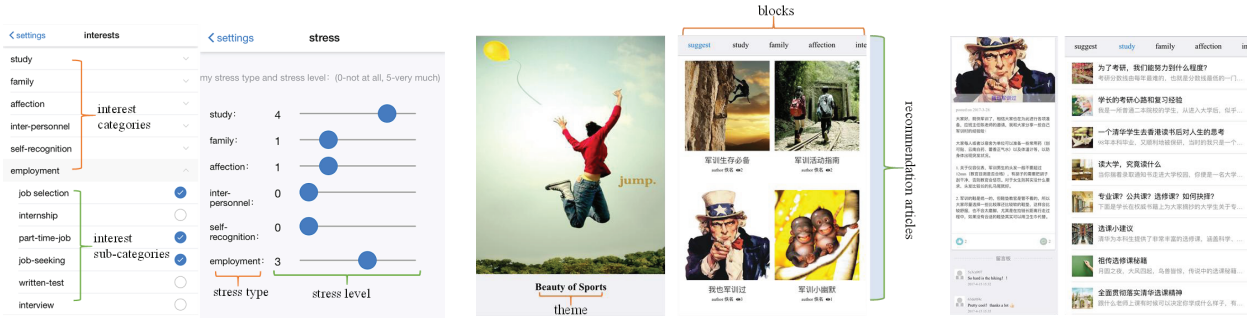
Secondly, in order to make the recommendation more sensitive and precise, the historical changes of stress should play a more significant role in the update process of $\overline{E}$. The significance of the historical changes can be reflected by the gradient of $Max(\overline{E}(i))$:

$$\nabla Max(\overline{E}(i)) = -\frac{2 \cdot \lambda}{5} \cdot \ell_i + \lambda + 1$$

Its shape is shown as Figure 4(a). We can see that in interval [0, 2.5], the bigger the λ is, the bigger the $\nabla Max(\overline{E}(i))$ is, so $Max(\overline{E}(i))$ gets its maximum much faster. So considering the purpose mentioned above, a bigger λ is suggested for *TeenRead*. Synthesizing both the maximum restriction and the bigger gradient, the best λ should be 1.0.

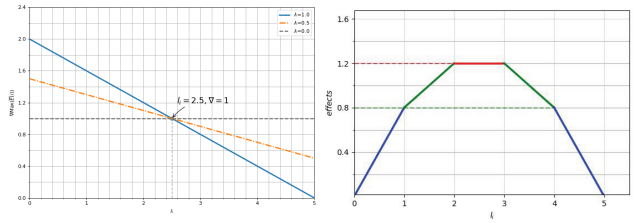
(2) How λ works in *TeenRead*?

Let's take a look at what the exact effect the efficient λ makes in *TeenRead*. We separately take λ=1 and λ=0 (means the historical changes of user's stress is neglected) into $Max(\overline{E}(i))$ and obtain two different values, namely $Max(\overline{E}(i))_1$ and $Max(\overline{E}(i))_0$. Effects of λ can naturally be defined as the difference value, which is $Max(\overline{E}(i))_1$ - $Max(\overline{E}(i))_0$. So the effects of λ can be shown as



(a) users' manual settings of stress and interests. (b) the recommended articles around a theme (c) an article and article categories

Figure 3. System interface of *TeenRead*. In (b), articles are divided into several blocks based on the categories shown in Table II



(a) The gradient of $Max(\overline{E}(i))$ with different λ (b) effects of λ in increasing user's stress level when $\lambda = 1$

Figure 4. Influence of λ to gradient and increment of $Max(\overline{E}(i))$

Figure 4(b), through which we see that λ could provide a suitable increment to user's stress level based on user's present stress level and its historical changes.

VI. User Study

We did two user studies to guide the building of *TeenRead* and the test of its performance.

The first user study was conducted in the incipient stage. We did a survey among college students on questions about what articles they prefer to read, what articles can ease their stress, how often the recommendation should be updated, etc. We collected 17 valid questionnaires, the specific questions and results are shown in Table III. We find that *TeenRead* is pretty expected and praised by the respondents, whose mean scores are all greater than 3. Besides, what articles they read most are about literature, culture and popularity, whose total proportion exceeds 77%. Based on the study result, *TeenRead* adjusts the article categories and system interface (shown in Figure 3.), and recommends 4 articles every 7 days.

The other study is conducted in order to test the stress easing effects and interestingness of articles recommended by *TeenRead*. We conduct a user study on 10 volunteers varying from the senior undergraduates to the senior doctoral students. To meet the different needs of volunteers, we have prepared 11 reading themes, including career planning, the art of interviewing, beauty of sports and so on, from which volunteers are required to choose six most interested themes to read. The recommendation is updated every 7 days, which requires volunteers to switch themes to read every 7 days. Each time a theme is read, the user fill out the questionnaire in Table IV. After six weeks' tracking survey, we collect 45 valid questionnaires. The survey results are also shown in Table IV (the first data in table cell is the mean, the last is the standard deviation). From the study statistics, we find a significant decrease in stress level after a period of reading, which is up to 22% (1.11 compared to 5). Moreover, the interest degrees of the four recommended articles all exceed 3.22 and the stress easing effects all exceed 2.89, showing the great performance of *TeenRead* in dealing with people's psychological problems as a method a bibliotherapy.

VII. Discussion

A further discussion on the choice of the 6 coefficients, including w, m, ρ, $Sim(\cdot)$, α and λ, is needed to adjust the recommending process.

As for w, the recommending cycle and m, the number of recommended articles, they varies from target readers, so it's necessary to adjust w and m according to the specific scenarios. ρ adjust the weight of stress easing effect in recommendation. The exact value of ρ contributing most to the performance of *TeenRead* should be tested through much observation and investigation. $Sim(\cdot)$ is the similarity measurement between two sibling sub-categories. It's hard and labor consuming to determine every two categories' similarity, so a similarity measurement algorithm is needed in the future work. α determines how fast the reading behaviors influence users' interest vector. If α is too big, the recommendation gets unstable; Otherwise if λ is too small, the recommendation becomes hysteretic and insensitive. Correspondingly, λ controls the weight of historical changes of stress in present easing vector. But we only take into account the last one changes, a more comprehensive consideration may be helpful.

VIII. Conclusion

In this paper, we address the problem of traditional bibliotherapy in coping with stress of adolescents and propose an

Table III
QUESTIONNAIRE RESULT OF USER STUDY 1

Question	Mean	Standard Deviation
Do you think *TeenRead* will help your school life? (Answer: 1-5, 1 for worst and 5 for best)	3.06	1.30
Do you think *TeenRead* categorizes articles properly? (Answer: 1-5, 1 for worst and 5 for best)	3.41	0.87
Does *TeenRead* recommend appropriate themes? (Answer: 1-5, 1 for worst and 5 for best)	3.19	1.00
How frequently shall *TeenRead* recommend a theme? (Answer: per day, per week, others)	per week:70.6% per day:17.6% others:11.8%	
How many articles shall a theme contain? (Answer: 4, 6, others)	4:41.2% 6:17.6% 10:5.9% 5:5.9% others:29.4%	
Which articles do you spend a lot of time in reading? (Answer: literature, culture, life, popular books, magazine, news, others)	popular books:32.3% literature:22.6% culture:22.6%	
Which articles can help you release stress a bit? (Answer: humor, figure, encouraging story, history, philosophy, poem, novel, others)	humor: 41.7% history: 25% philosophy: 16.7%	

Table IV
QUESTIONNAIRE RESULT OF USER STUDY 2 (LIKERT SCALE, 1-5 POINT, 5 FOR BEST AND 1 FOR WORST)

Question	article1	article2	article3	article4
What's your stress level before reading the recommended articles?	3.53/0.97 (the mean is 3.53, the standard deviation is 0.97)			
What's your stress level after reading the recommended articles?	2.42/0.89			
What's your interest degree for the recommended articles? (Answer: 1-5, 0 for worst and 5 for best)	3.31/1.44	3.38/1.25	3.51/1.14	3.22/1.33
What's the easing effect of the recommended articles for you? (Answer: 1-5, 0 for worst and 5 for best)	3.13/1.20	2.96/1.30	3.22/1.31	2.89/1.27

online reading recommendation system as a new exploration of bibliotherapy. The major advantage of *TeenRead* is that articles with good stress easing effects as well as much attraction are automatically recommended to users without human intervention. Two user studies involving about 30 volunteers have also demonstrated the good effectiveness of *TeenRead* as a new method of bibliotherapy, in which users' stress level decreases by 22% as well as users' interests in reading is pretty high.

REFERENCES

[1] Yuping hou. Promoting bibliotherapy and building sunshine of soul for college students. *Journal of Business, Education and Economy Research*, 1:9–10, 2008.

[2] Yu Chen, Xin Zhao, and Junhong Huang et al. Effects of mindfulness meditation on emotion regulationtheories and neural mechanismse. *Advances in Psychological Science*, 19(10):1502–1510, 2011.

[3] Jiao Guan, Jie Li, and Qiangjun Zhao. A brief analysis of the influence of yoga on the decompression of female college students. *Sport*, 16:54–55, 2013.

[4] Jun Yan, Aiguo Chen, and Fengshu Zhu. The development on college students' psychological stress of exercise intervention. *Journal of Sports and Science*, 31(5):90–93, 2010.

[5] Yu Zhang and Lijiang Tao. A survey of college students' psychological crisis. *Ideological and Political Education*, 13:295–296, 2011.

[6] Peilin Li and Jiangli Hao. The application of bibliotherapy for the mental health education of college students. *Beijing Education*, Z1:72–74, 2010.

[7] Boxing Zhao. A history, present condition and development of bibliotherapy. *Libray*, 3:13–15, 2003.

[8] Noah. Webster. *Webster's third new international dictionary of the english language*. Merriam Webster, 1961.

[9] Louis A Rongione. *Bibliotherapy: its nature and uses*. Phoenix: Oryx Press, 1978.

[10] E. Rich. User modeling via stereotypes. *Cognitive Science*, 3(4):329–354, 1979.

[11] Fen Li, Yinman Chen, and Jianxin Zhang. Review of the bibliotherapy study in china. *Journal of Modem Information*, 30(2):173–177, 2010.

[12] Bo Wang and Xin Fu. The principle of bibliotherapy. *Library*, 3:1–12, 2003.

[13] Zhi Yang. Bibliotherapy in china: a review. *Chinese journal of medical library and information science*, 16(3):1–4, 2007.

[14] Joeran Beel, Bela Gipp, and Stefan Langer et al. Research-paper recommender systems: a literature survey. *International Journal on Digital Libraries*, 17(4):305–338, 2016.

[15] P. Lops, M. Gemmis, and G. Semeraro. Content-based recommender systems: state of the art and trends. *Recommender Systems Handbook*, pages 73–105, 2011.

[16] Paul Resnick, Neophytos Iacovou, and Mitesh Suchak et al. Grouplens: an open architecture for collaborative filtering of netnews. In *Proc. of CSCW*, pages 175–186, 1994.

[17] H. Small. Co-citation in the scientific literature: a new measure of the relationship between two documents. *Journal of the American Society for Information Science and Technology*, 24(4):265C–269, 1973.