

Mining Online Social Data for Detecting Social Network Mental Disorders

Hong-Han Shuai¹, Chih-Ya Shen¹, De-Nian Yang¹, Yi-Feng Lan²,
Wang-Chien Lee³, Philip S. Yu^{4,5}, Ming-Syan Chen⁶

¹Academia Sinica, Taiwan

²Tamkang University, Taiwan

³The Pennsylvania State University, USA

⁴University of Illinois at Chicago, USA

⁵Tsinghua University, China

⁶National Taiwan University, Taiwan

hhshuai@citi.sinica.edu.tw, chihya@citi.sinica.edu.tw, dnyang@iis.sinica.edu.tw
carolyflan@gmail.com, wlee@cse.psu.edu, psyu@uic.edu, mschen@cc.ee.ntu.edu.tw

ABSTRACT

An increasing number of social network mental disorders (SNMDs), such as Cyber-Relationship Addiction, Information Overload, and Net Compulsion, have been recently noted. Symptoms of these mental disorders are usually observed passively today, resulting in delayed clinical intervention. In this paper, we argue that mining online social behavior provides an opportunity to actively identify SNMDs at an early stage. It is challenging to detect SNMDs because the mental factors considered in standard diagnostic criteria (questionnaire) cannot be observed from online social activity logs. Our approach, new and innovative to the practice of SNMD detection, does not rely on self-revealing of those mental factors via questionnaires. Instead, we propose a machine learning framework, namely, *Social Network Mental Disorder Detection (SNMDD)*, that exploits features extracted from social network data to accurately identify potential cases of SNMDs. We also exploit multi-source learning in SNMDD and propose a new SNMD-based Tensor Model (STM) to improve the performance. Our framework is evaluated via a user study with 3126 online social network users. We conduct a feature analysis, and also apply SNMDD on large-scale datasets and analyze the characteristics of the three SNMD types. The results show that SNMDD is promising for identifying online social network users with potential SNMDs.

Keywords

Online social network, mental disorder detection, feature extraction, tensor factorization

1. INTRODUCTION

“As we expect more from technology, do we expect less from each other?” asked Sherry Turkle, the Abby Rockefeller Mauzé

professor of the Social Studies of Science and Technology in MIT.¹ With the explosive growth in popularity of social networking and messaging apps, online social networks (OSNs) have become a part of many people’s daily lives. While OSNs seemingly expand their users’ capability in increasing social contacts, they may actually decrease the face-to-face interpersonal interactions in the real world. Studies show that some people’s behavior is bolder on the OSNs because they can put a mask when communicating with others there, i.e., hide who they really are. However, do those OSN users still know how to connect with others when the masks are off? Lying between receiving positive attention from OSNs and face-to-face interactions may be a great gulf in the real life.

Most research on social network mining focuses on discovering the treasure of knowledge behind the data for improving people’s life. In contrast, much less attention has been drawn to remedy the problems incurred from various social network applications. Indeed, some *social network mental disorders (SNMDs)*, such as Information Overload and Net Compulsion [1], have been recently noted.² For example, studies suggest that 1 in 8 Americans suffers from problematic Internet use.³ Moreover, as reported by the BBC News, 25% of the population in Korea are estimated to suffer from SNMDs.⁴ Due to the epidemic scale of these phenomena, new terms such as Phubbing (Phone Snubbing) and Nomophobia (No Mobile Phone Phobia) have been created to describe those who cannot stop using mobile social networking apps. Moreover, leading journals in mental health, such as American Journal of Psychiatry [2], have reported that the SNMDs may incur excessive use, depression, social withdrawal, and a range of negative repercussions. Indeed, these symptoms are important components of diagnostic criteria for detecting SNMDs [3]: 1) *excessive use* of social networking and messaging apps – usually associated with a loss of the sense of time or a neglect of basic drives; 2) *withdrawal* – including feelings of anger, tension, and/or depression when the computer/apps are inaccessible; 3) *tolerance* – manifesting as the need for more usage; and 4) *negative repercussions* – including arguments, lying, social

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author’s site if the Material is used in electronic media.
WWW 2016, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4143-1/16/04.
<http://dx.doi.org/10.1145/2872427.2882996>.

¹http://www.ted.com/talks/sherry_turkle_alone_together

²<http://phys.org/news/2015-09-social-media-impacts-mental-well-being.html>

³<http://netaddiction.com/faqs/>

⁴<http://www.bbc.com/news/world-asia-33130567>

isolation, and fatigue [2]. SNMDs are social-oriented and tend to happen to users who usually interact with others via online social media. Those with SNMDs usually lack offline interactions, and as a result seek cyber-relationships to compensate.

Today, identification of potential mental disorders often falls on the shoulders of supervisors (such as teachers, employers, or parents) who can observe the aforementioned symptoms better than others but only passively. As the facts that there are very few notable physical risk factors, the patients usually do not actively seek medical or psychological services to reduce these symptoms. Consequently, patients would look for clinical interventions with psychiatrists and medical treatments only when their conditions become very serious. However, a recent study [4] shows a strong correlation between suicidal attempt and SNMDs for students. In this research, 9510 adolescent students aged from 12 to 18 years old are tested using a personality inventory and Internet addiction inventory. The findings indicate that adolescents suffering from social network addictions have a much higher risk of suicidal ideation than non-addictive users. The research also reveals that social network addiction may deteriorate emotional status, causing higher hostility, depressive mood and compulsive behavior. Most importantly, the delay of early intervention may lead to mental illness and thus can seriously damage an individual's social functioning. In short, it is desirable to actively detect potential SNMD users on OSNs at an early stage.

Although previous work in Psychology has identified several crucial mental factors related to SNMDs as standard diagnostic criteria for detecting SNMDs, they are mostly assessed via survey questionnaires by design. To detect potential SNMD cases of OSN users, extracting these factors to assess the mental states of users is very challenging. For example, the extent of loneliness and the effect of disinhibition of OSN users are not easily observable.⁵

There is a need for developing new approaches for detecting SNMD cases of OSN users. We argue that mining social network data of individuals, as a complementary alternative to the conventional psychological approach, provides an excellent opportunity to *actively identify* those cases at an early stage. In this paper, we develop a machine learning framework for detecting SNMDs, namely *Social Network Mental Disorder Detection (SNMDD)*. We formulate the task as a semi-supervised classification problem to detect three types of SNMDs [1], including i) Cyber-Relationship Addiction, which shows addictive behavior for building online relationships; ii) Net Compulsion, which shows compulsive behavior for online social gaming or gambling; and iii) Information Overload, which is related to uncontrollable surfing. By exploiting machine learning techniques with the ground truth obtained via the current diagnostic practice in Psychology [1], we extract and analyze several features of different categories from OSNs, including parasocial relationships, online and offline interaction ratio, social capital, disinhibition, self-disclosure, and bursting temporal behavior. These features capture important factors or serve as proxies for SNMD detection. For example, parasocial relationship represents an asymmetric interpersonal relationship where one party cares more about the other, but the other does not. This asymmetric relationship is related to loneliness, one of the primary mental factors for the users with SNMDs to access online social media excessively [5].

⁵The online disinhibition effect is a loosening (or complete abandonment) of social restrictions and inhibitions that would otherwise be present in normal face-to-face interaction during interactions with others on the Internet.

Therefore, we extract the ratio of the number of actions a user takes to friends and the number of actions friends take to the user as a feature. In this paper, the extracted features are carefully examined through user study.

Users may behave differently on different OSNs, resulting in inaccurate SNMD detection. When the data on different OSNs of a user are available, the accuracy of SNMDD is expected to improve by effectively integrating information from multiple sources for model training. A naïve solution is to concatenate the features from different networks into a feature vector. However, simply increasing the number of features suffers from the curse of dimensionality. Accordingly, we propose an *SNMD-based Tensor Model (STM)* to deal with this multi-source learning problem in SNMDD. There are two advantages of our approach: i) the novel *STM* incorporates the SNMD characteristics into the tensor model; and ii) tensor factorization can capture the structure, latent factors, and correlation of features to derive a full portrait of user behavior.

The contributions of this paper are summarized as follows.

- Today online SNMDs are usually treated at a late stage. To address this issue, we propose an approach, new to the current practice of SNMD detection, by mining data logs of OSN users to actively identify potential SNMD cases early.
- We develop a machine learning framework for detecting SNMDs, namely *Social Network Mental Disorder Detection (SNMDD)*. Moreover, we design and analyze many features from OSNs, such as disinhibition, parasociality, self-disclosure, etc., which serve as important factors or proxies for identifying SNMDs. The proposed framework can be deployed as a software program to provide an early alert for potential patients and their advisors.
- We study the *multi-source learning* problem for SNMD detection. By leveraging tensor algebra and considering the SNMD characteristics into the tensor model, we propose *STM* to better extract the latent factors from different sources, thus improving the accuracy.
- We conduct a user study with 3126 users to evaluate the effectiveness of the proposed SNMDD framework. To the best of our knowledge, this is the first dataset crawled online for SNMD detection. Also, we apply SNMDD on large-scale real datasets and perform a social network analysis on the detected cases. The result reveals interesting insights on the network structures in SNMD types, which can be of interest to social scientists and psychologists.

The rest of this paper is organized as follows. Section 2 surveys the related work. Section 3 presents *SNMDD*, focusing on feature extraction and the proposed *STM* for multi-source learning. Section 4 reports a user study, various analyses, and the experimental results. Section 5 concludes this paper.

2. RELATED WORK

Recent research in Psychology and Sociology reports a number of mental factors related to social network mental disorders. Specifically, Barbera et al. find that young people with narcissistic tendencies are particularly vulnerable to addiction with OSNs [6]. Moreover, Chak et al. show that the tendency of addiction to the Internet is positively related to the shyness, and negatively related to the faithfulness [7]. Note that

most of these features cannot be directly observed in OSNs. Thus, in our study, we investigate new features in social network data for SNMD detection. On the other hand, statistical analysis on the prevalence and extent of mobile email addictions has been studied [8], and a quantitative analysis of the reasons to leave Facebook with over 400 questionnaires was performed to improve human-computer interactions on social media [9]. Our research in this paper is uniquely different from these prior works since we adopt a data mining approach to detect SNMDs by exploiting various discriminative features to capture mental factors. Moreover, we propose a tensor model to efficiently integrate heterogeneous data from different OSNs and incorporate the domain knowledge, i.e., the SNMD characteristics such as the users with Cyber-Relationship Addiction tend to have friends with Cyber-Relationship Addiction.

A recent line of studies proposes to detect cyberbullying, i.e., harassing other OSN users in a deliberate manner. Dinakar et al. [10] propose a machine learning approach and extract features from content sentiment and context information to detect textual cyberbullying. The detection of SNMDs is more difficult than that of textual cyberbullying because i) it is possible to focus on certain keywords to detect cyberbullying behavior, and ii) users cannot hide the cyberbullying behavior (it is not cyberbullying if users can hide the behavior), whereas users with Net Compulsion may hide their logs.

Our framework is built upon support vector machine, which has been widely used to analyze OSNs in many areas [11, 12], such as business, transportation, and anomaly intrusion detection. In addition, we present a new tensor model that not only incorporates the domain knowledge but also well estimates the missing data and avoids noises to properly handle multi-source data. Crammer et al. propose a PAC-style model for multi-source learning and provide a theory of sampling for learning models [13]. However, the SNMD data from different OSNs may be incomplete due to the heterogeneity. For example, the profiles of users may be empty due to the privacy issue, different functions on different OSNs (e.g., game, check-in, event), etc. We propose a novel tensor-based approach to address the issues of using heterogeneous data and incorporate domain knowledge in SNMD detection.

3. SOCIAL NETWORK MENTAL DISORDER DETECTION

In this paper, we aim to explore data mining techniques to detect three types of SNMDs [1]: 1) *Cyber-Relationship (CR) Addiction*, which includes addiction to social networking, checking and messaging to the point where social relationships to virtual and online friends become more important than real-life ones with friends and families; 2) *Net Compulsion (NC)*, which includes compulsive online social gaming or gambling, often resulting in financial and job-related problems; and 3) *Information Overload (IO)*, which includes addictive surfing of user status and news feeds, leading to lower work productivity and fewer social interactions with families and friends offline.

Accordingly, we formulate the detection of SNMD cases as a classification problem. We detect each type of SNMDs with a binary SVM. In this study, we propose a two-phase framework, called *Social Network Mental Disorder Detection (SNMDD)*, as shown in Figure 1. The first phase extracts various discriminative features of users, while the second phase presents a new SNMD-based tensor model to derive latent factors for training and use of classifiers built upon Transductive SVM

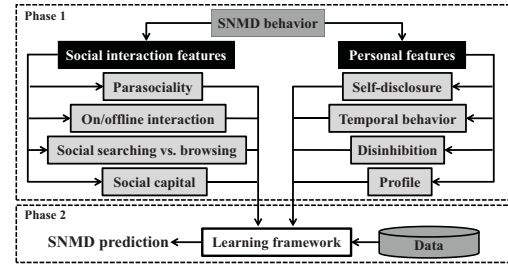


Figure 1: The SNMDD framework.

(TSVM) [14]. Two key challenges exist in design of SNMDD: i) we are not able to directly extract mental factors like what have been done via questionnaires in conventional SNMD detection process for psychiatrists and thus need new features for learning the classification models;⁶ ii) we aim to exploit user data logs from multiple OSNs and thus need new techniques for integrating multi-source data based on SNMD characteristics. We address these two challenges in Sections 3.1 and 3.2, respectively.

3.1 Feature Extraction

We first focus on extracting discriminative and informative features for design of SNMDD. This task is nontrivial for the following three reasons.

1. Lack of mental features. Psychological studies have shown that many mental factors are related to SNMDs, e.g., low self-esteem [3], loneliness [15]. Thus, questionnaires are designed to reveal those factors for SNMD detection. Some parts of Psychology questionnaire for SNMDs are based on subjective comparison of mental states in online and offline status, which cannot be observed from OSN logs. For example:

Q1. How often do you feel depressed, moody, or nervous when you are off-line, which goes away once you are back online?

Q2. How often do you prefer the excitement of the Internet to intimacy with your partner?

Consider Q1. The feel of depression and nervousness offline can not be observed online. To tackle this problem, we have to leverage the knowledge from Psychology, such as withdrawal or relapse patterns, and exploit some proxy features extracted from online social activity logs to approximate them. For Q2, the preference of excitement of the Internet to intimacy with users' partners is important questions for SNMD detection. As it is difficult to directly observe these factors from data collected from OSNs, psychiatrists are not able to directly assess the mental states of OSN users under the context of online SNMD detection.

2. Heavy users vs. addictive users. To detect SNMDs, an intuitive idea is to simply extract the usage (time) of a user as a feature for training SNMDD. However, this feature is not sufficient because i) the status of a user may be shown as "online" if she does not log out or close the social network applications on mobile phones, and ii) heavy users and addictive users all stay online for a long period, but heavy users do not show symptoms of anxiety or depression when they are not using social apps. How to distinguish them by extracting discriminative features is critical.

3. Multi-source learning with the SNMD characteristics. As we intend to exploit user data from different OSNs in SNMDD, how to extract complementary features to draw a

⁶Additional issues in feature extraction will be detailed later.

full portrait of users while considering the SNMD characteristics into the tensor model is a challenging problem.

To address the first two challenges, we identify a number of effective features as proxies to capture the mental states of users, e.g., self-esteem [3] and loneliness [15].⁷ The goal is to distinguish users with SNMDs from normal users. Two types of features are extracted to capture the social interaction behavior and personal profile of a user.

3.1.1 Social Interaction Features

We first extract a number of *social interaction features* to capture the user behavior on social media.

Parasocial relationship (PR). Research shows that the mental factor of loneliness is one of the primary reasons why the users with SNMDs excessively access online social media [5]. As the loneliness of an OSN user is hard to measure, we exploit the parasocial relationship, an asymmetric interpersonal relationship between two people where one party cares more about the other but the other does not, to capture loneliness (as studies show that they are correlated [16]). The feature of parasocial relationship is represented as $|a_{out}|/|a_{in}|$, where $|a_{out}|$ and $|a_{in}|$ denote the number of actions a user takes to friends and the number of actions friends take to the user, respectively.⁸ As the ratio increases, the extent of parasocial relationship also grows.

Online and offline interaction ratio (ONOFF). As observed by mental health professionals, people who indulge themselves in OSNs tend to snub their friends in real life [7]. Therefore, the number of online interactions is inclined to significantly exceed their interactions offline. We extract the number of check-in logs with friends and the number of “going” events as an indicator of the number of offline activities to estimate the online ($|a_{on}|$)/offline ($|a_{off}|$) interaction ratio. Although the number of offline events observed from online is smaller than the actual number, the ratio is relative and is a good indicator (as pointed out in [17]), because the frequent check-in records of a user imply that the user is active in offline activities, which is an indicator of non-SNMD.

Social capital (SC). Two types of friendship ties are usually involved in the theory of *social capital* [18]: i) Bond strengthening (strong-tie), which represents the use of OSNs to strengthen the relationships; and ii) Information seeking (weak-tie), which corresponds to the use of social media to find valuable information. The first type usually creates more interactions in order to increase the social tightness and is related to Cyber-Relationship (CR) Addiction, while the second type concentrates more on finding and reading the information and is thus related to Information Overload (IO) [19]. Therefore, the ratio between the number of strong ties (n_{strong}) and weak ties (n_{weak}) could be used for differentiating the CR and IO types. Moreover, since the number of strong ties is much smaller than that of weak ties [20], and the number of friends that CR users frequently interact with is less than that of IO users, we exploit the ratio between the number of friends the user interacts online (likes, comments, and posts) and the total number of the user’s friends as proxy features to differentiate the CR and IO types.

Social searching vs. browsing (SSB). The human appetitive system is in charge of the addictive behavior. A recent study has shown that *social searching* (actively reading news feeds from friends’ walls) creates more pleasure than

social browsing (passively reading personal news feeds) [21]. This finding indicates that goal-directed activities of social searching are more likely to activate the appetitive system of a person as drug rewards do, and it is more related to SNMDs because the appetitive system is responsible for finding things in the environment that promote species survival (i.e., food, sexual mates) and thus is inclined to form addictive behavior after several rewards. While users with SNMDs perform social searching more frequently than non-SNMDs, it is not easy to distinguish these two behavior on social media. Let n_i denote the total number of the i -th action for posts among friends. For example, if a user is the second one among her friends who click “likes” on a post, the n_2 increases 1 for the user. As most social media provide friends’ comments and “likes” in the form of news feeds to users, we consider the number of likes/comments on news feeds from friends as social browsing ($\sum_{i=2}^{\infty} n_i$). On the other hand, if users take an initiative to search for someone’s profile and like/comment on it, we consider this as a social searching (i.e., the number of likes/comments on others’ news feeds that are not liked/commented by his friend before (n_1)). Therefore, we use $\frac{n_1}{\sum_{i=2}^{\infty} n_i}$ as a feature. The social searching features are related to CR because CR users tend to find social supports, whereas social browsing is more related to IO. Compared with social capital, SSB focuses on different behavior in reading news feeds, rather than the different types of friend ties.

3.1.2 Personal features

In the following, we present the features extracted from the personal profiles of OSN users.

Self-disclosure based features (SD). Researchers from Harvard University point out that a person’s *self-disclosure communications* (i.e., describing the personal feeling) stimulate the brain’s pleasure center [22], similar to sex and food. However, to conduct a sentiment analysis on the contents associated with a user is very complicated and computationally expensive. Inspired by emotional signal detection, which finds that when the users use emoticons, they are effectively expressing an emotional state [23], we retrieve and exploit the numbers of emoticons, stickers, and selfies in each post as the features for self-disclosure [24].

Temporal behavior features (TEMP). *Relapse* is the state that a person is inclined to quickly revert back to the excessive usage of social media after an abstinence period, while *tolerance* is the state that the time spent by a person with SNMDs tends to increase due to the mood modification effect.⁹ It is worth noting that the above two mental states have been exploited to evaluate clinical addictions [15]. We aim to use them to distinguish *heavy users* and *addictive users* because heavy users do not suffer from relapse and tolerance in use of OSNs. An issue arising here is how to assess relapse and tolerance quantitatively.

It is observed that the use of social media by an SNMD patient is usually in the form of *intermittent bursts* [3]. Therefore, given a stream of a user’s activities on an OSN, e.g., “likes”, “comments”, “posts”, we exploit Kleinberg’s burst detection algorithm [25], which is based on an infinite Markov model, to detect periods of the user’s activities as bursty and non-bursty periods. The bursty period refers to a period during which the activities significantly increase. A bursty period is modeled as a bursty state q_1 in the Markov model, while

⁷The third challenge is addressed in Section 3.2.

⁸The actions include like, comment, and post in our work.

⁹A patient may need to spend more time on social media to reach the happiness/excitement than before.

a non-bursty period is correspondingly modeled as a normal state q_0 . The burst detection algorithm finds a state transition sequence \mathbf{q} for each user to divide the corresponding log (stream of activities) into bursty and non-bursty periods. Specifically, let $\mathbf{x} = (x_1, x_2, \dots, x_n)$ denote a sequence of n time intervals between $n + 1$ consecutive activities, with the intervals distributed according to a density function, such as $f_{i_t}(x_t) = \alpha_{i_t} e^{-\alpha_{i_t} x_t}$, where α_{i_t} is either α_0 or α_1 , and α_0 and α_1 are parameters that correspond to normal and burst states, respectively, $\alpha_1 > \alpha_0$. A time interval x_t is in a burst state q_1 if $f_0(x_t) < f_1(x_t)$. Otherwise, it is in a normal state q_0 . However, simply deciding the state sequence \mathbf{q} based on this criteria results in numerous small periods. Therefore, a cost $\tau(q_i, q_j)$ is associated with a state transition from q_i to q_j to filter out noises and to ensure that each bursty or non-bursty period is sufficiently long. Therefore, the remaining issue is to find an optimal state-transition sequence \mathbf{q} to minimize the following cost function [25],

$$c(\mathbf{q}|\mathbf{x}) = \sum_{t=1}^{n-1} \tau(q_i, q_{i+1}) + \sum_{t=1}^n (-\ln f_{i_t}(x_t)),$$

where $\tau(q_i, q_{i+1}) = 0$ if the state q_i and q_{i+1} is the same. $\tau(q_i, q_{i+1})$ is $\gamma \ln n$ otherwise, where γ is an algorithm parameter larger than 0. Notice that the state sequence that minimizes the cost depends on 1) how easy it is to jump from one state to another and 2) how well it is to comply to the rates of arrivals. After identifying the bursts, we measure their intensity (the number of activities within a burst) and length (the time period of a burst) as the proxy features for *relapse* and *tolerance*, respectively. The (average, median, standard deviation, maximum, minimum) of both burst intensity and burst length are included in our feature set, because they capture the characteristic of bursts. For instance, the standard deviation of the burst length for SNMD patients is usually larger than that for heavy users since heavy users constantly use OSNs while users with SNMDs increase the usage time due to tolerance.

Usage time (UT). In addition to the above, two features regarding usage of an OSN are also adopted. i) *Duration*. The duration that a user spends on social media a day is estimated by consecutive activity logs.¹⁰ ii) *Number of online states*. The number of online states during a day is also important. It was widely believed that a person spending a lot of time on OSNs usually belongs to CR or IO. However, a recent study points out that the usage time is only moderately correlated to CR and IO [26]. Indeed, both heavy users and the users with SNMDs tend to stay online for lengthy time periods, but heavy users do not feel anxious and depressed when they are not using social apps. With the bursts detected earlier, we aim to use the relapse and tolerance to distinguish users with SNMDs from heavy users. For example, given a heavy user spends the same amount of time online as a user with SNMDs does, the standard deviation of burst length for the heavy user is expected to be smaller than that of the user with SNMDs since the heavy user explores OSNs more regularly and does not suffer from the tolerance of increasing usage.

Disinhibition based features (DIS). The mental factor of disinhibition is also one of the primary reasons the users excessively access online social media [5]. When surfing online,

some people tend to act out more frequently or intensely than they act offline due to the dissociative anonymity, asynchronicity, and solipsistic introjection, which is called the *online disinhibition effect* [28]. When user identities can be anonymous or the conversation is not face-to-face (e.g., Whisper, Snapchat), offline-shy users are more inclined to addict to cyberspace relationships due to disinhibition. Although we know which OSNs are anonymous and thus expect a stronger disinhibition effect on them, how to detect the users experiencing the disinhibition effect from non-anonymous OSNs is challenging. As reported in [29], the average clustering coefficient (CC) on anonymous OSNs is smaller than that on non-anonymous ones, e.g., 0.033 on Whisper but 0.059 on Facebook. Inspired by this observation, we use CC as a proxy for disinhibition on non-anonymous websites since a user with disinhibition tends to have a small CC. The disinhibition effect is more related to CR and NC because they both show stronger intensity of usage under the anonymity. Also, the clustering coefficient may potentially be effective for detecting CR and NC because it depends on the links among friends which cannot be hidden by users.

Profile features (PROF). We also extract some demographic features commonly adopted in questionnaires from user profiles, such as *age* and *gender*. A study [15] has shown that the age when a user logs in Facebook the first time is correlated to the intensity of SNMDs. For children growing up with the Internet, this is a key position in their lives. Empirical studies [15] also suggest that gender difference results in varying degrees of SNMDs because the goals of using OSNs are different for different genders, e.g., females are more inclined to use online communication whereas males are inclined to follow news and play online games. Also, the number of game posts is extracted.

Feature Effectiveness. It is worth noting that each individual feature cannot precisely classify all cases, as research shows that exceptions may occur. Therefore, it is necessary to exploit multiple features to effectively remove exceptions. For example, the social searching vs. browsing (SSB) feature may be affected by the gender difference of using OSNs, i.e., males are more inclined to social searching than females [30]. Moreover, when using the number of game logs as an indicator for NC, a special case happens when the users with NC attempt to hide the game logs by not allowing games to publish news feeds. However, we are able to address the above issue by exploiting other features, e.g., CC for disinhibition effect, which is dependent on the user's friends so that the user cannot change it. Therefore, the special cases on existing individual features can be overcome by fusing multiple features together.

3.2 Multi-Source Semi-Supervised Learning

Many users are inclined to use different OSNs, and it is expected that data logs of these OSNs could provide enriched and complementary information about the user behavior. Thus, we aim to explore multiple data sources (i.e., OSNs) in SNMDD, in order to derive a more complete portrait of users' behavior and effectively deal with the data sparsity problem. To exploit multi-source learning in SNMDD, one simple way is to directly concatenate the features of each person derived from different OSNs as a huge vector. However, the above approach tends to miss the correlation of a feature in different OSNs and introduce interference. Thus, we explore tensor techniques which have been used increasingly to model multiple data sources because a tensor can naturally represent multi-source data. We aim to employ tensor decomposition

¹⁰If the difference of timestamps of two logs is smaller than a few minutes, the user is regarded as online during this period [27].

to extract common latent factors from different sources and objects.

In this paper, given D SNMD features of N users extracted from M OSN sources, we construct a three-mode tensor $\mathcal{T} \in \mathbb{R}^{N \times D \times M}$ and then conduct Tucker decomposition, a renowned tensor decomposition technique, on \mathcal{T} to extract a latent feature matrix \mathbf{U} , which presents the latent features of each person summarized from all OSNs. We aim to feed these latent features for SNMD detection. Matrix \mathbf{U} effectively estimates a deficit feature (e.g., a missing feature value unavailable due to privacy setting) of an OSN from the corresponding feature of other OSNs, together with the features of other users with similar behavior. Based on Tucker decomposition on \mathcal{T} , we present a new *SNMD-based Tensor Model (STM)*, which enables \mathbf{U} to incorporate important characteristics of SNMDs, such as the correlation of the same SNMD sharing among close friends.¹¹ Finally, equipped with the new tensor model, we conduct semi-supervised learning to classify each user by exploiting Transductive Support Vector Machines (TSVM).

We first summarize the notations introduced in this section. Here scalars are denoted by lowercase letters, e.g., u , while vectors are denoted by boldface lowercase letters, e.g., \mathbf{u} . Matrices are represented by boldface capital letters, e.g., \mathbf{U} , and tensors are denoted by calligraphic letters, e.g., \mathcal{T} . Each element (i, j, k) of a three-mode tensor \mathcal{T} is denoted by t_{ijk} , whereas the i -th row and the j -th column of a two-dimensional matrix \mathbf{U} are respectively denoted by \mathbf{u}_i and \mathbf{u}_j . Specifically, Tucker decomposition [31] of a tensor $\mathcal{T} \in \mathbb{R}^{N \times D \times M}$ is defined as:

$$\mathcal{T} = \mathcal{C} \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}, \quad (1)$$

where $\mathbf{U} \in \mathbb{R}^{N \times R}$, $\mathbf{V} \in \mathbb{R}^{D \times S}$ and $\mathbf{W} \in \mathbb{R}^{M \times T}$ are latent matrices. In this paper, the matrix of users' latent features \mathbf{U} plays a crucial role.

In Tucker decomposition, R , S , and T are parameters to be set according to different criteria [31]. The 1-mode product of $\mathcal{C} \in \mathbb{R}^{R \times S \times T}$ and $\mathbf{U} \in \mathbb{R}^{N \times R}$, denoted by $\mathcal{C} \times_1 \mathbf{U}$, is a matrix with size $N \times S \times T$, where each element $(\mathcal{C} \times_1 \mathbf{U})_{nst} = \sum_{r=1}^R c_{rst} u_{rn}$. Given the input tensor matrix \mathcal{T} that consists of the features of all users from every OSN, Tucker decomposition derives \mathcal{C} , \mathbf{U} , \mathbf{V} , and \mathbf{W} to meet the above equality on \mathcal{T}_{ndm} for every n , d , and m , where \mathcal{C} needs to be diagonal, and \mathbf{U} , \mathbf{V} , and \mathbf{W} are required to be orthogonal [31]. By regarding \mathbf{u}_i in \mathbf{U} as the latent features of user i , we can efficiently integrate the information from different networks for i .

Equipped with tensor decomposition on \mathcal{T} , we propose a new *SNMD-based Tensor Model (STM)* to minimize the following objective function \mathcal{L} ,

$$\begin{aligned} \mathcal{L}(\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathcal{C}) = & \frac{1}{2} \|\mathcal{T} - \mathcal{C} \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}\|^2 \\ & + \frac{\lambda_1}{2} \text{tr}(\mathbf{U}^T \mathbf{L}_a \mathbf{U}) + \frac{\lambda_2}{2} \|\mathbf{U}\|^2, \quad (2) \end{aligned}$$

where $\text{tr}(\cdot)$ denotes the matrix traces, the Frobenius norm of a tensor \mathcal{T} is defined as $\|\mathcal{T}\| = \sqrt{\langle \mathcal{T}, \mathcal{T} \rangle}$, and λ_1 and λ_2 are parameters controlling the contribution of each part during the above collaborative factorization. \mathcal{L} first minimizes the decomposition error, i.e., $\|\mathcal{T} - \mathcal{C} \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}\|^2$, for \mathcal{T} . Note that Eq. (1) does not always need to hold since other crucial goals are also incorporated in the model. For example, the term that minimizes $\|\mathbf{U}\|^2$ is to derive a more concise latent

feature matrix and avoid overfitting, where $\|\mathbf{V}\|^2$, $\|\mathbf{W}\|^2$, and $\|\mathcal{C}\|^2$ are not necessary to be reduced since only the latent feature matrix $\|\mathbf{U}\|^2$ will be employed in the semi-supervised learning later in this section.

The *STM* is different from the conventional tensor models in the second term of Eq. (2), where important characteristics of SNMDs are incorporated. For example, the probability of finding CR cases around a CR patient is higher than that around a non-CR user due to the loneliness propagation [16]. That is, CR users usually feel lonely and are more likely to establish friendships in cyberspace with other users with similar behavior. Since the nearby nodes with a great quantity of interactions tend to be the same (either CR or non-CR), it is envisaged that the distance of \mathbf{u}_i and \mathbf{u}_j will be small if the edge weight $a_{i,j}$ of $e_{i,j}$ is sufficiently large. Therefore, a regularization (smoothing) term $\frac{1}{2} \text{tr}(\mathbf{U}^T \mathbf{L}_a \mathbf{U})$ is included in the model to achieve the above goal,

$$\begin{aligned} \frac{1}{2} \text{tr}(\mathbf{U}^T \mathbf{L}_a \mathbf{U}) &= \text{tr}(\mathbf{U}^T (\mathbf{D} - \mathbf{A}) \mathbf{U}) \\ &= \frac{1}{2} \sum_{i,j} \|\mathbf{u}_i - \mathbf{u}_j\|_2^2 a_{ij} \\ &= \sum_{i,j} \mathbf{u}_i a_{ij} \mathbf{u}_i^T - \sum_{i,j} \mathbf{u}_i a_{ij} \mathbf{u}_j^T \\ &= \sum_i \mathbf{u}_i d_{ii} \mathbf{u}_i^T - \sum_{i,j} \mathbf{u}_i a_{ij} \mathbf{u}_j^T. \end{aligned}$$

Notice that $\text{tr}(\mathbf{U}^T \mathbf{L}_a \mathbf{U})$ decreases when the distance of latent factors for any two close friends is small. However, this term does not enforce the distance of features between every two friends to be small since the objective function \mathcal{L} significantly increases in this case due to a larger decomposition error appearing in the first term. The distance of \mathbf{u}_i and \mathbf{u}_j tend to be smaller if users i and j interact frequently, i.e., a_{ij} is large. Therefore, the above model is able to incorporate the characteristics of SNMD and alleviate the data sparsity problem, i.e., the users with fewer social activities may benefit from the auxiliary information of their friends with abundant features.

To properly find $\mathcal{L}(\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathcal{C})$ in *STM*, let \mathbf{A} denote the weighted adjacency matrix of graph G .¹² $\mathbf{L}_a = \mathbf{D} - \mathbf{A}$ is the Laplacian matrix of the weighted adjacency matrix \mathbf{A} , where \mathbf{D} is a diagonal matrix with the entries $d_{ii} = \sum_j a_{ij}$. We present a gradient-descent algorithm to iteratively improve each element in the matrices according to the corresponding gradient, where the gradient for each variable is derived as follows:

$$\begin{aligned} \nabla_{\mathbf{u}_i} \mathcal{L} = & (\mathcal{C} \times_1 \mathbf{u}_i \times_2 \mathbf{v}_j \times_3 \mathbf{w}_k - t_{ijk}) \mathcal{C} \times_2 \mathbf{v}_j \times_3 \mathbf{w}_k \\ & + \lambda_1 (\mathbf{L}_a \mathbf{U})_{i:} + \lambda_2 \mathbf{u}_i; \end{aligned}$$

$$\nabla_{\mathbf{v}_j} \mathcal{L} = (\mathcal{C} \times_1 \mathbf{u}_i \times_2 \mathbf{v}_j \times_3 \mathbf{w}_k - t_{ijk}) \mathcal{C} \times_1 \mathbf{u}_i \times_3 \mathbf{w}_k;$$

$$\nabla_{\mathbf{w}_k} \mathcal{L} = (\mathcal{C} \times_1 \mathbf{u}_i \times_2 \mathbf{v}_j \times_3 \mathbf{w}_k - t_{ijk}) \mathcal{C} \times_1 \mathbf{u}_i \times_2 \mathbf{v}_j;$$

$$\nabla_{\mathcal{C}} \mathcal{L} = (\mathcal{C} \times_1 \mathbf{u}_i \times_2 \mathbf{v}_j \times_3 \mathbf{w}_k - t_{ijk}) \mathcal{C} \mathbf{u}_i \circ \mathbf{v}_j \circ \mathbf{w}_k;$$

Algorithm 1 presents the pseudo code. By integrating the features from different OSNs and exploiting the information of the users with similar behavior, \mathbf{U} serves as the latent feature factor of all users from every data source for semi-supervised learning.

¹¹Note that D does not capture the social correlations among friends.

¹²The edge weight can be derived according to the number of interactions that represents the proximity [32].

Algorithm 1 Tensor Factorization for SNMDD

Input: Tensors \mathcal{T} , an error threshold ϵ , and the max iteration times I_{Max}

Output: Low rank matrix \mathbf{U} , and matrices \mathbf{V} , \mathbf{W} , core tensor \mathcal{C}

- 1: Initialize $\mathbf{U} \in \mathbb{R}^{N \times R}$, $\mathbf{V} \in \mathbb{R}^{D \times S}$, $\mathbf{W} \in \mathbb{R}^{D \times T}$, and $\mathcal{C} \in \mathbb{R}^{R \times S \times T}$ with small random values.
- 2: Set η as step size
- 3: $d_{ii} = \sum_i z_{ij}$
- 4: $\mathbf{L}_z = \mathbf{D} - \mathbf{Z}$
- 5: **while** $t < I_{Max}$ and $Loss_t - Loss_{t+1} > \epsilon$ **do**
- 6: **for** each t_{ijk} **do**
- 7: Get $\nabla_{u_i} \mathcal{L}$, $\nabla_{v_j} \mathcal{L}$, $\nabla_{w_k} \mathcal{L}$, $\nabla_{\mathcal{C}} \mathcal{L}$
- 8: $\mathbf{u}_i^{t+1} = \mathbf{u}_i^t - \eta \nabla_{u_i} \mathcal{L}$
- 9: $\mathbf{v}_j^{t+1} = \mathbf{v}_j^t - \eta \nabla_{v_j} \mathcal{L}$
- 10: $\mathbf{w}_k^{t+1} = \mathbf{w}_k^t - \eta \nabla_{w_k} \mathcal{L}$
- 11: $\mathcal{C}^{t+1} = \mathcal{C}^t - \eta \nabla_{\mathcal{C}} \mathcal{L}$

In the following, we briefly summarize semi-supervised learning for SNMDD. Let \mathbf{x}_i denote the feature vector of user i , and $\mathbf{x}_i = \mathbf{u}_i$. Let \mathbf{y}_i denote the class label vector of user i with length $K = 3$ (i.e., three types of SNMDs), and $\hat{y}_{ik} = 1$ indicates that user i suffers from type k of SNMD. For example, $\{+1, -1, -1\}$ represents a user with Cyber-Relationship Addiction but without Net Compulsion and Information Overload. Given a vector set \mathcal{D} of L labeled samples $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^L$ and L' unlabeled samples $\{\mathbf{x}_j, \hat{\mathbf{y}}_j\}_{j=L+1}^{L+L'}$, the optimization problem of TSVM can be formulated as follows:

$$\begin{aligned}
\min_{\theta, \xi_{ik}} \quad & \frac{1}{2} \|\mathbf{w}\| + C \sum_{i=1}^N \xi_{ik} + C^* \sum_{i=L+1}^{L+L'} \xi_{ik} \\
\text{s.t.} \quad & \mathbf{y}_{ik} (\mathbf{w}_k^T \mathbf{x}_i + b_k) \geq 1 - \xi_{ik}, \quad 1 \leq i \leq L, \\
& |\hat{\mathbf{y}}_{ik} (\mathbf{w}_k^T \mathbf{x}_i + b_k)| \geq 1 - \xi_{ik}, \quad L+1 \leq i \leq L+L', \\
& \xi_{ik} \geq 0, \quad 1 \leq i \leq L+L' \text{ and } \hat{\mathbf{y}}_{ik} \in \{+1, -1\},
\end{aligned}$$

where ξ_{ik} is the slack variable set, and C and C^* are the trade-off parameters between the classification margin and misclassification errors for labeled samples and unlabeled samples, respectively. The model parameters \mathbf{w}_k and b_k returned by this binary learning problem represent a binary classifier associated with the k -th class: $f_k(\mathbf{x}_i) = \mathbf{w}_k^T \mathbf{x}_i + b_k$. The binary classifiers for the three SNMD types are trained and used independently to predict the label vector $\hat{\mathbf{y}}$ for an unlabeled instance \mathbf{x} . The experimental results in the next section discover that the accuracy of the semi-supervised learning without tensor is 78.3% and 83.1% for Instagram and Facebook, while *STM* increases the accuracy to 89.7% by integrating data from Facebook and Instagram.

4. EXPERIMENTAL RESULTS

In this section, we evaluate SNMDD with real datasets. A user study with 3126 people is conducted to evaluate the accuracy of SNMDD. Moreover, a feature study is performed. Finally, we apply SNMDD on large-scale datasets and analyze the detected SNMD types.

4.1 Data Preparation and Evaluation Plan

In the following, we detail the preparation of the datasets used in our evaluation.

Table 1: Details of the datasets

Dataset	Description
FB_US	User profile, the friends of each user, the news feeds created by users with metadata (who likes, who comments, stickers, and geotag), the news feeds users like or comment (stickers also), events (join/decline), joined groups with events, and game posts created by game apps
IG_US	User profile, the followers/followees of each user, the media created by users with metadata (who likes, who comments, and geotag), and the contents users like or comment
FB_L	Anonymized user ID that performs the action, anonymized user ID that receives the action, and timestamp of action creation
IG_L	Anonymized media ID, anonymized ID of the user who created the media, timestamp of media creation, set of tags assigned to the media, number of likes and number of comments received

4.1.1 User Study

We recruit 3126 OSN users around the world via Amazon Mechanical Turk (MTurk) to obtain data for training and testing the classifiers in SNMDD. The participants include 1790 males and 1336 females. Their professions are very diverse, affiliating with universities, government offices, technology companies, art centers, banks, and businesses. Each user is first invited to fill out the standard SNMD questionnaires [1, 33].¹³ Then, a group of professional psychiatrists participating in this project assess and manually label the users as *potential SNMD cases* (and identify their types of SNMDs) or *normal users*.¹⁴ There are 389 users labeled as *SNMD*, including 246 Cyber-Relationship (CR) Addictions, 267 Information Overload (IO), and 73 Net Compulsion (NC).¹⁵ The result obtained by the psychiatrists serves as the ground truth for our evaluation. We also crawl the Facebook (denoted as FB_US) and Instagram (denoted as IG_US) data of the participants in the user study for training and testing of our machine learning models (based on features detailed in Section 3.1). All the data are collected with the Facebook and Instagram APIs as listed in Table 1.

In the experiment, we first evaluate the effectiveness of proposed features, including all features (**All**), social interaction features (**Social**) and personal profile features (**Personal**), with a baseline feature **Duration**, i.e., the total time spent online, using TSVM [14] for semi-supervised learning in the user study. We also collect two large-scale datasets, including Facebook (denoted as FB_L) with 63K nodes, 1.5M edges, and 0.84M wall posts [34], and Instagram (denoted as IG_L) with 2K users, 9M tags, 1200M likes, and 41M comments [35]. Note that some proposed features cannot be extracted from certain large-scale datasets, e.g., game posts and stickers are not available in IG_L, which is handled by using the imputation technique [36]. The details of the data crawled from each social media are listed in Table 1.

¹³The IRB number of this project is AS-IRB-HS 15003 v.1.

¹⁴They are from California School of Professional Psychology, Taipei City Hospital, Nat'l Taipei Univ., psychiatric clinics, etc.

¹⁵Note that a person may have multiple types of SNMDs simultaneously.

With labeled (IG_US and FB_US) and unlabeled data (IG_L and FB_L) described above, we divide data into 5 folds with the labeled/unlabeled ratio is preserved. We use 5-fold cross validation, i.e., take 4 folds for training and 1 fold for testing, to evaluate the performance of proposed features using semi-supervised TSVM [37]. A number of supervised approaches, including J48 Decision Tree Learning [38], SVM [39], and Logistic Regression, and DT SVM [40] which do not use unlabeled data, are also compared to justify our choice of using TSVM in SNMDD. Next, we compare the proposed *SNMD-based Tensor Model (STM)* with two baseline algorithms. The first baseline algorithm is to simply concatenate the features from different networks together (denoted as *CF*), while the second baseline algorithm is to use the existing Tucker model (denoted as *Tucker*) which does not incorporate prior knowledge regarding the characteristics of SNMD cases (observed from our analysis). Finally, the effectiveness of each feature is carefully analyzed in Section 4.4.

4.1.2 Large-Scale Experiments

To discover new insights, we apply our semi-supervised SNMDD on IG_L and FB_L to classify their users and then analyze the detected cases of different SNMD types. Notice that the goal of this analysis is exploratory-oriented as we do not have the ground truth for the large datasets. We examine whether friends of SNMD cases tend to be potential SNMD cases as well. Also, we apply community detection on FB_L and IG_L to derive the relationships between different types of SNMD users in their communities. Finally, the average hop distance between the SNMD users of the same type is reported.

4.2 Evaluation of the Proposed Features

In the following, we first evaluate the performance of the proposed features using TSVM. We adopt Accuracy (Acc.) and Area Under Curve (AUC) for evaluation of SNMDD. Moreover, Microaveraged-F1 (Micro-F1) and Macroaveraged-F1 (Macro-F1) are also compared for multiple-label classification. Table 2 summarizes the average results and standard deviations, where the examined feature sets are denoted by self-explained labels.

The results on the IG_US and FB_US datasets in the user study show that **Duration** leads to the worst performance, i.e., the results of accuracy are 34% and 36%, and the AUC are 0.362 and 0.379, respectively. Using all (**All**) or parts (**Social/ Personal**) of the features proposed in the paper outperform **Duration** significantly (see Table 2). **All** achieves the best performance (78% and 83% accuracy on the IG_US and FB_US datasets, respectively) because SNMDD is able to capture the various features extracted from data logs to effectively detect SNMD cases. Between **Social** and **Personal**, **Personal** outperforms **Social** because the features of temporal behavior (**TEMP**) in **Personal** are very effective. Since the F1 measure ignores true negatives, its magnitude is mostly determined by the number of true positives, i.e., large classes dominate small classes in microaveraging. As shown in Table 2, Micro-F1 of **Duration**, **Social**, and **Personal** are larger than Macro-F1 using both IG_US and FB_US datasets, indicating that using parts of features performs better on IO and CR (large classes) than NC. In contrast, the performance of SNMDD is almost the same in Micro-F1 and Macro-F1, which indicates its robustness. The results from FB_US are better than those from IG_US because IG_US is sparser, e.g., there are no event and game posts on Instagram.

After comparing the results from SNMDD with the ground truth obtained via user study, we observe that some false-positive users are detected as NC, probably because people with NC are more likely to hide their real usage time, e.g., the game logs of some people with NC are hidden. As a result, a few normal users may be incorrectly detected as NC. However, SNMDD generally performs very well for NC due to some effective features. For example, users of NC are usually less parasocial since they are less frequent to interact with friends. Moreover, since the NC users' friends with game benefits usually do not know the NC users' other friends (e.g., colleagues), their clustering coefficients are lower than the normal users.

4.3 Evaluation of Classification Techniques and STM

In the following, given all the proposed features, we first evaluate TSVM in comparison with some representative supervised learning approaches in SNMDD. As shown in Table 3, the accuracy of semi-supervised TSVM (83.1%) outperforms all the supervised algorithms, including 76.4% for ℓ_2 -regularized logistic regression, 77.9% for ℓ_2 -regularized ℓ_2 -loss SVM, since TSVM effectively uses unlabeled data to address the issues of overfitting and data sparsity. It is worth noting that the accuracy and AUCs of the supervised learning methods, ranging from 74.4% to 77.9% and 0.75 to 0.78, respectively, are not significantly different. This is because the proposed features provide robust information so that the accuracy is not sensitive to the choice of learning algorithms.

Next, we compare the proposed *STM* with two baseline algorithms, i.e., *CF* and *Tucker*, to integrate features extracted from both of the IG_US and FB_US datasets for learning of classification models using TSVM, as described in Section 4.1. Also as shown in Table 3, the accuracy and AUC of *STM* are 89.7% and 0.926. The results indicate that *STM*, via the decomposed latent factor matrix **U**, is able to recover some missing features and provide extra latent information to better characterize the users. In contrast, *CF*, which simply concatenates features from FB_US and IG_US, performs the worst. Its accuracy and AUC are 75.5% and 0.759, respectively, even worse than those of some single-source learning algorithms using only the FB_US dataset, e.g., 77.9% and 0.783 for ℓ_2 -regularized ℓ_2 -loss SVM. This is because *CF* loses the correlations in some features and thus introduces noises. On the other hand, while *Tucker* is able to achieve 85.6% accuracy and 0.872 AUC, its performance is still not as great as *STM*. This result indicates that *STM*, by incorporating important prior knowledge about characteristics of SNMD cases (see discussion regarding Eq. (2)), is able to derive more precise and accurate latent features than *Tucker* to achieve the best performance in SNMDD.

4.4 Feature Study

To observe the differences among the three types of SNMDs, Table 4 lists the top-5 discriminative features and corresponding accuracy on the FB_US dataset by TSVM, where CC, BI, BL, and SD respectively denote the clustering coefficient, burst intensity, burst length, and standard deviation. It is worth noting that the number of selfies, an indicator of self-disclosure, is not useful for detecting CR and IO, but it is effective for NC. This is because NC users are usually less socially active as compared to CR and IO users. Moreover, the online/offline interaction ratio of NC is much higher than the ratios of the other two types, probably because NC users usually show less willingness to join offline activities. In con-

Table 2: Performance comparisons on the IG_US and FB_US datasets.

Measure	Instagram				Facebook			
	Duration	Social	Personal	All	Duration	Social	Personal	All
Acc.	0.34±0.02	0.59±0.01	0.69±0.03	0.78±0.02	0.36±0.01	0.65±0.05	0.73±0.02	0.83±0.02
AUC	0.36±0.02	0.61±0.01	0.74±0.01	0.79±0.01	0.37±0.01	0.68±0.01	0.77±0.02	0.84±0.01
Micro-F1	0.42±0.02	0.71±0.01	0.78±0.04	0.85±0.01	0.44±0.04	0.74±0.02	0.81±0.01	0.89±0.01
Macro-F1	0.33±0.01	0.64±0.01	0.73±0.02	0.85±0.01	0.35±0.02	0.68±0.02	0.77±0.03	0.90±0.01

Table 3: Comparisons of SNMDD with different classification techniques.

Technique	Acc.	AUC
Single-source (FB)		
J48 Decision Tree Learning	74.4%	0.750
ℓ_1 -regularized ℓ_2 -loss SVM	77.6%	0.781
ℓ_2 -regularized ℓ_2 -loss SVM	77.9%	0.783
ℓ_1 -regularized logistic regression	76.3%	0.776
ℓ_2 -regularized logistic regression	76.4%	0.777
DTSVM	76.4%	0.774
TSVM	83.1%	0.842
Multi-source (FB+IG)		
CF	75.5%	0.759
Tucker	85.6%	0.872
STM	89.7%	0.926

trast, users of CR and IO prefer to use social media instead of playing games alone. Moreover, people with compulsive personality are more introverted. In contrast, people with CR usually create virtual bonds to develop pathological relationships for compensation of their (missing) offline relationship.

The parasociality, effective for detecting all SNMD types, is especially useful for detecting CR cases. For example, in our user study, we find user A, 21-year-old male, frequently posting news feeds, such as “I’m so bored :((((...Ahhhhhh!!”, and his cross-dressing photos on his Facebook timeline, more than 3 times a week, which usually get less than 5 likes. At the same time, he “likes” a large number of posts from others. SNMDD classifies him as a potential CR case and his questionnaire reveals that he constantly blocks out disturbing thoughts about life and finds himself anticipating when he goes online again.

Burst intensity and length seem to be quite useful for detecting IO cases. For example, user B, 36-year-old male, is detected as IO since the behavior of clicking “likes” fits the pattern of bursts, i.e., the median of his burst intensity is high, equal to 31. His answers to the standard questionnaire reveal that he loses sleep due to late-night access on Facebook to check others’ news feeds. Through interview, user B explains that he cannot stop checking for new posts and e-mails even when all his news feeds and emails are read. Some of his friends reply him: “are you a robot? no sleep needed?!!?”, indicating that user B is indulged in finding social news.

Next, we analyze the importance of different features to our classifiers. The information gain is exploited to measure the importance of each feature. In summary, the top 5 important features overall are : 1) median of the intensity of bursts, 2) online/offline interaction ratio, 3) parasociality, 4) number of used stickers, and 5) standard deviation of the length of bursts. It is worth noting that TSVM using only these 5 features in SNMDD achieves an accuracy of 76.4% and 80.7% for IG_US and FB_US, respectively, close to that of using all features (A11). In other words, integrating important social and personal features provides good results because effective

Table 4: Top features and Acc. on the FB_US dataset.

CR	NC	IO
Parasociality	Game posts	Median of BI
Median of BI	Online/offline ratio	Online/offline ratio
Sticker number	Parasociality	SD of BL
Online/offline ratio	Number of selfies	Sticker number
CC.	CC.	Parasociality
Acc.: 80.2%	Acc.: 76.8%	Acc.: 82.7%

Table 5: Feature effectiveness analysis: SNMDD accuracy on the FB_US dataset.

Used Features	Accuracy	Used Features	Accuracy
PR	56.9%	A11-PR	78.2%
ONOFF	60.3%	A11-ONOFF	75.1%
SC	40.1%	A11-SC	78.8%
SSB	44.4%	A11-SSB	79.3%
SD	58.9%	A11-SD	73.2%
TEMP	67.5%	A11-TEMP	68.1%
UT	36.4%	A11-UT	82.6%
DIS	54.0%	A11-DIS	75.9%
PROF	18.2%	A11-PROF	81.5%
A11	83.1%		

personal features, e.g., the temporal behavior features, can be used to differentiate the users suffering from withdraw or relapse symptoms and heavy users, while social features capture the interactions among users to differentiate different SNMDs.

Finally, we carefully examine the effectiveness of each feature on the FB_US dataset. Table 5 compares the performance of different feature combinations using TSVM, where A11- X means all features *excluding* category X . Unsurprisingly, combining all features leads to the best performance with the accuracy of 83.1%. In terms of individual category of features, the temporal behavior features are the most effective, whereas the profile features are the least effective. The accuracy obtained by excluding one category of features is at least 68.1%, which shows that the features are generally robust even when one feature set is missing. The accuracy of A11-PROF is close to A11, indicating that PROF is the least important.

Figs. 2(a) and 2(b) show the improvement of adding different features in TSVM on the FB_US dataset and the proposed STM on multi-source data (i.e., FB_US and IG_US). The feature selection of TSVM is based on the information gain (the top-5 features mentioned earlier), while the tensor approach automatically extracts important latent features. We observe a diminishing return property on both figures, where the improvement becomes marginal as more features are included. Fig. 2(a) shows a power fit function ($p(x) = 0.3087x^{-1.89}$) of the curve with $R^2 = 0.9512$. The exponent -1.89 denotes that the improvement by adding n -th feature is $n^{-1.89}$ times smaller than that by adding the first feature. On the other hand, the results of the tensor-based approach in Fig. 2(b) show that the accuracy increment for adding a single feature drops faster ($p(x) = 1.06x^{-2.82}$) since the proposed STM can extract much more important and concise features.

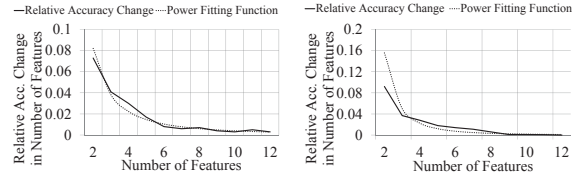


Figure 2: Relative accuracy change with respect to number of features.

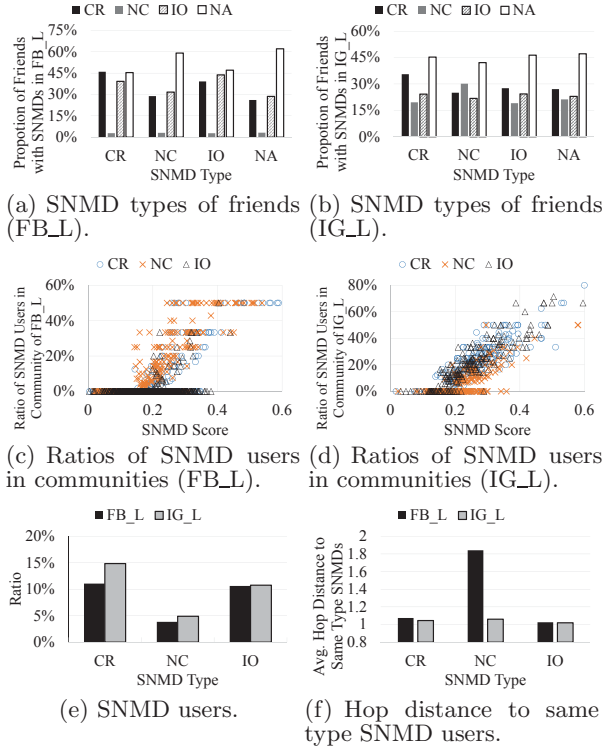


Figure 3: Comparisons of different datasets.

4.5 Analysis of SNMD Types in Large Datasets

In this analysis, we first apply the proposed SNMDD framework (with TSVM) on some large-scale OSN datasets, i.e., FB_L and IG_L, to classify their users. In Figs. 3(a) and 3(b), we analyze the detected SNMD cases among the friends of an SNMD user. In Fig. 3(a), the leftmost bar indicates that in FB_L, among all CR users, about 45% of their friends are also CR users, which is greater than the percentage of other SNMD types. On the other hand, the 8th bar from the left in Fig. 3(a) indicates that in FB_L, about 59% of NC users' friends are NA (non-SNMD users). Figs. 3(a) and 3(b) show that, in FB_L and IG_L, CR and IO users have similar friend types. This is because CR and IO cases, by their nature, are similar, i.e., they are both seeking social satisfaction (e.g., relationships and information) from the OSNs. Moreover, among different SNMD cases, CR and IO users are likely to be friends with other CR and IO users. For CR users, this phenomenon has been described as “loneliness propagates” [16].

Furthermore, Infomap community detection [41] is performed on FB_L and IG_L to derive the relationships between different types of SNMD users in their communities. Figs. 3(c) and 3(d) analyze the community structures of SNMD users

with different SNMD scores, where each point represents the characteristic of a community. Specifically, each community in the dataset is represented by three different types of points, i.e., CR, NC, and IO. For example, each CR point is represented as $\langle \text{score}, \text{ratio} \rangle$, where *score* is the average CR score in that community, and *ratio* indicates the proportion of CR users in the community. It is similar for each IO/NC point. As Figs. 3(c) and 3(d) show, for each SNMD type, when the average SNMD score is higher, it is likely to have more SNMD users in the community. Moreover, there are many communities with large IO scores in IG_L that have IO ratios close to 1. This implies that the users with large IO scores in IG_L are more inclined to form homogeneous groups. At the first glance, one may feel that NC users frequently appear in many communities, and there seems to be a large number of NC users, especially in FB_L (i.e., Fig. 3(c)). However, after carefully examining these communities, we find that those communities (with large ratios of NC users) are usually very small (usually with the size around 5) because NC users are less-active. On the other hand, in IG_L, when SNMD scores are larger, the ratios of IO users in communities are also larger. This is because IO users can view, like, or follow others in Instagram more easily (not necessary to be friends first).

Fig. 3(e) compares the ratios of different types of SNMD users identified in FB_L and IG_L. There are more CR users in IG_L probably because CR users seek social supports online to compensate the loneliness in real life. We argue that the Instagram platform makes it easy to freely create social relationships with strangers. In contrast, it is not that easy to create new social relationships on Facebook since the friend requests need to be approved. Finally, Fig. 3(f) compares the average number of hops from each SNMD user to the nearest user with the same type of SNMDs. The leftmost bar shows that the average hop distance from each CR user to the closest CR user is 1.07 hop, indicating that CR and IO users are close to other same-type users, i.e., average hop distances are within 1.15, where Figs. 3(a) and 3(b) also report similar results.

5. CONCLUSION

In this paper, we make an attempt to automatically identify potential online users with SNMDs. We propose an SNMDD framework that explores various features from data logs of an OSN and a new tensor technique for deriving latent features from multiple OSNs for SNMD detection. This work represents a collaborative effort between computer scientists and mental healthcare researchers to address emerging issues in SNMDs. As for the next step, we plan to study the features extracted from multimedia contents by techniques on NLP and computer vision. We also plan to further explore new issues from the perspective of a social network service provider, e.g., Facebook or Instagram, to improve the well-beings of OSN users without compromising user engagement.

6. ACKNOWLEDGMENTS

This work is supported in part by NSF through grants III-1526499, CNS-1115234, and OISE-1129076, USA, and by Ministry of Science and Technology through grants MOST 104-2221-E-002-214-MY3, 103-2221-E-001-005-MY2, and 104-2221-E-001-005-MY2, Taiwan.

7. REFERENCES

- [1] K. Young, M. Pistner, J. O'Mara, and J. Buchanan. Cyber-disorders: The mental health concern for the new millennium. *Cyberpsychol. Behav.*, 1999.

- [2] J. Block. Issues of DSM-V: internet addiction. *American Journal of Psychiatry*, 2008.
- [3] K. Young. Internet addiction: the emergence of a new clinical disorder. *Cyberpsychol. Behav.*, 1998.
- [4] I.-H. Lin, C.-H. Ko, Y.-P. Chang, T.-L. Liu, P.-W. Wang, H.-C. Lin, M.-F. Huang, Y.-C. Yeh, W.-J. Chou, and C.-F. Yen. The association between suicidality and Internet addiction and activities in Taiwanese adolescents. *Compr. Psychiat.*, 2014.
- [5] Y. Baek, Y. Bae, and H. Jang. Social and parasocial relationships on social network sites and their differential relationships with users' psychological well-being. *Cyberpsychol. Behav. Soc. Netw.*, 2013.
- [6] D. La Barbera, F. La Paglia, and R. Valsavoia. Social network and addiction. *Cyberpsychol. Behav.*, 2009.
- [7] K. Chak and L. Leung. Shyness and locus of control as predictors of internet addiction and internet use. *Cyberpsychol. Behav.*, 2004.
- [8] O. Turel and A. Serenko. Is mobile email addiction overlooked? *CACM*, 2010.
- [9] E. Baumer, P. Adams, V. Khovanskaya, T. Liao, M. Smith, V. Sosik, and K. Williams. Limiting, leaving, and (re)lapsing: an exploration of Facebook non-use practices and experiences. *CHI*, 2013.
- [10] K. Dinakar, B. Jones, C. Havasi, H. Lieberman, and R. Picard. Common sense reasoning for detection, prevention, and mitigation of cyberbullying. *ACM TiS*, 2012.
- [11] R. Jain and N. Abouzakhar. A comparative study of hidden markov model and support vector machine in anomaly intrusion detection. *JITST*, 2013.
- [12] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li. User-level sentiment analysis incorporating social networks. *KDD*, 2011.
- [13] K. Crammer, M. Kearns, and J. Wortman. Learning from multiple sources. *JMLR*, 2008.
- [14] R. Collobert, F. Sinz, J. Weston, and L. Bottou. Large scale transductive svms. *JMLR*, 2006.
- [15] L. Leung. Net-generation attributes and seductive properties of the internet as predictors of online activities and internet addiction. *Cyberpsychol. Behav. Soc. Netw.*, 2004.
- [16] J. Cacioppo, J. Fowler, and N. Christakis. Alone in the crowd: the structure and spread of loneliness in a large social network. *J. Pers. Soc. Psychol.*, 2009.
- [17] R. Schwartz. Mobility and locative media: mobile communication in hybrid spaces. *Routledge*, 2014.
- [18] C. Steinfield, N. Ellison, C. Lampe, and J. Vitak. Online social network sites and the concept of social capital. *Frontiers in new media research, New York: Routledge*, 2012.
- [19] R. Davis. A cognitive-behavioral model of pathological internet use. *Comput. Hum. Behav.*, 2001.
- [20] V. Arnaboldi, M. Conti, A. Passarella, and R. Dunbar. Dynamics of personal social relationships in online social networks: a study on twitter. *ACM COSN*, 2013.
- [21] K. Wise, S. Alhabash, and H. Park. Emotional responses during social information seeking on Facebook. *Cyberpsychol. Behav. Soc. Netw.*, 2010.
- [22] D. Tamir and J. Mitchell. Disclosing information about the self is intrinsically rewarding. *Natl. Acad. Sci.*, 2012.
- [23] K.-L. Liu, W.-J. Li, and M. Guo. Emoticon smoothed language models for twitter sentiment analysis. *AAAI*, 2012.
- [24] G. Seidman. Self-presentation and belonging on Facebook: How personality influences social media use and motivations. *Pers. Individ. Differ.*, 2013.
- [25] J. Kleinberg. Bursty and Hierarchical Structure in Streams. *KDD*, 2002.
- [26] C.-M. Lai, K.-K. Mak, H. Watanabe, R. Ang, J. Pang, and R. Ho. Psychometric properties of the internet addiction test in Chinese adolescents. *J. Pediatr. Psychol.*, 2013.
- [27] R. Jones and K. Klinkner. Beyond the session timeout: automatic hierarchical segmentation of search topics in query logs. *CIKM*, 2008.
- [28] J. Suler. The online disinhibition effect. *CyberPsychol. Behav.*, 2004.
- [29] G. Wang, B. Wang, T. Wang, A. Nika, H. Zheng, and B. Y. Zhao. Whispers in the dark: analysis of an anonymous social network. *IMC*, 2014.
- [30] M. Makashvili, B. Ujmajuridze, T. Amirejibi, B. Kotetishvili, and S. Barbakadze. Gender difference in the motives for the use of Facebook. *TOJET*, 2013.
- [31] T. Kolda and B. Bader. Tensor decompositions and applications. *SIAM review*, 2009.
- [32] M. Roth, A. Ben-David, D. Deutscher, G. Flysher, I. Horn, A. Leichtberg, N. Leiser, Y. Matias, and R. Merom. Suggesting friends using the implicit social graph. *KDD*, 2010.
- [33] C. Andreassen, T. Torsheim, G. Brunborg, and S. Pallesen. Development of a Facebook addiction scale. *Psychol. Rep.*, 2012.
- [34] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi. On the evolution of user interaction in Facebook. *WOSN*, 2009.
- [35] E. Ferrara, R. Interdonato, and A. Tagarelli. Online popularity and topical interests through the lens of Instagram. *HT*, 2014.
- [36] M. Saar-Tsechansky and F. Provost. Handling Missing Values when Applying Classification Models. *JMLR*, 2007.
- [37] F. Shang, L. C. Jiao, and F. Wang. Semi-Supervised Learning with Mixed Knowledge Information. *KDD*, 2012.
- [38] I. Witten and E. Frank. Data mining: practical machine learning tools and techniques with Java implementations. Morgan-Kaufmann, San Francisco, 2000.
- [39] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines, 2001.
- [40] F. Chang, C.-Y. Guo, X.-R. Lin, and C.-J. Lu. Tree decomposition for large-scale SVM problems. *JLMR*, 2010.
- [41] M. Rosvall and C. Bergstrom. Maps of random walks on complex networks reveal community structure. *Natl. Acad. Sci.*, 2008.