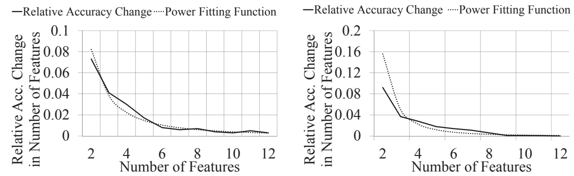
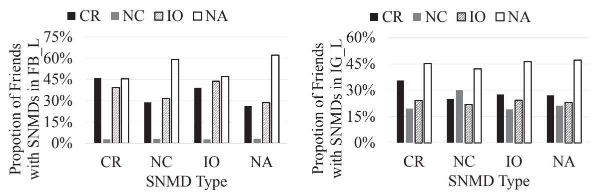
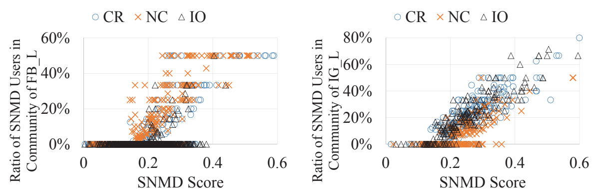


(a) Relative improvement w.r.t. the number of features.
(b) Relative improvement w.r.t. the number of latent features (STM).

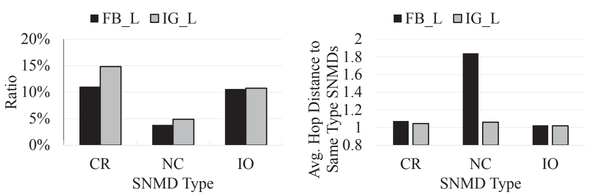
Figure 2: Relative accuracy change with respect to number of features.



(a) SNMD types of friends (FB_L).
(b) SNMD types of friends (IG_L).



(c) Ratios of SNMD users in communities (FB_L).
(d) Ratios of SNMD users in communities (IG_L).



(e) SNMD users.
(f) Hop distance to same type SNMD users.

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 $\langle score, ratio \rangle$, 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 TiiS*, 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. Indiv. 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.