Additional file 1 for:

Characterizing dynamic communication in online eating disorder communities: a multiplex network approach

Tao Wang, Markus Brede, Antonella Ianni, and Emmanouil Mentzakis

University of Southampton, UK.

Contents

List of Figures

List of Tables

1   ED-related hashtags

We identify ED-related users by searching for users who posted an ED-related hashtag in tweets. The ED-related hashtags are obtained by (i) detecting clusters of hashtags that frequently co-occur in a tweet posted by 3,380 ED users, using a similar method that we used to detect topics of conversations in the main text; (ii) selecting ED-related clusters of tags based on prior evidence from language use in online ED-related content [1, 2, 3]; and (iii) removing generic tags (e.g., "#skinny" and "#food") from the selected clusters. Ref. [4] for details. We obtain 375 ED-related hashtags in total and Table S1 lists examples of these hashtags.

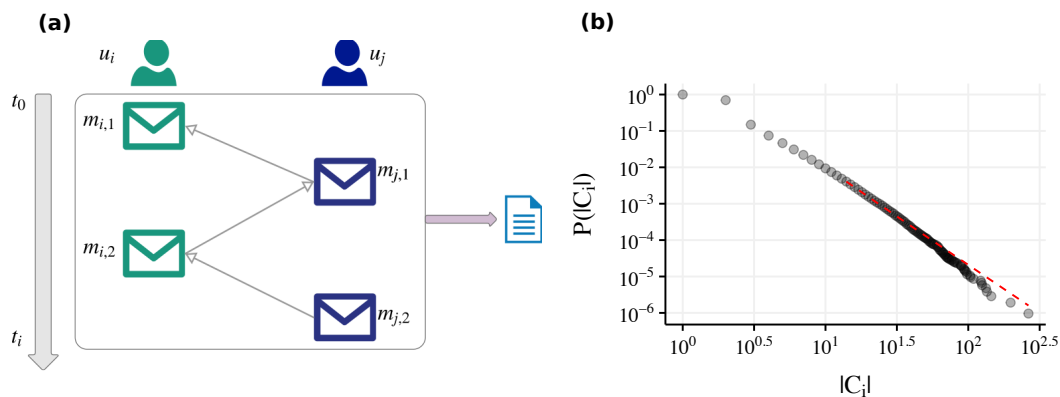| thinspo, edproblems, thinspiration, proana, ana, thighgap, edprobs, ed, eatingdisorder, anorexia, mia, skinny4xmas, bonespo, hipbones, proed, bulimia, ednos, edfamily, edlogic, thinkthin, legspo, promia, edthoughts, mythinspo, anorexic, edgirlprobs, edprobz, anamia, eatingdisorders, internationaledmeetup, edlife |
| --- |

2 Statistics of conversations

Figure S1: (a) Aggregating tweets by conversations. If user u_i posts tweet message $m_{i,1}$ at time t_0, user u_j posts tweet $m_{j,1}$ at time t_1 to reply to $m_{i,1}$, and the two users further have two subsequent interactions through $m_{i,2}$ and $m_{j,2}$, then the conversation D_i between u_i and u_j is represented as $D_i = \langle m_{i,1}, m_{j,1}, m_{i,2}, m_{j,2} \rangle$. (b) Distribution of conversation sizes. The red line fits a power-law distribution with exponent $\lambda = 3.65 \pm 0.04$.

Figure S1(a) shows an example how we aggregate tweets into conversations. We obtain 1,044,573 conversations consisting of 2,206,919 tweets. The average number of tweets in a conversation is 2.11, with a standard deviation of 1.93. Figure S1(b) shows the distribution of numbers of tweets $|C_i|$ in each conversation C_i. We see that many conversations contain a small number of tweets while a few have very large numbers of tweets, suggesting the heterogeneity of conversation sizes. Also, the straight line on the logarithmic histogram indicates a power law in the distribution [5]. To quantify this pattern, we fit a power-law function $P(|C_i|) = |C_i|^{-\lambda}$ using the maximum likelihood estimator and calculate p-value for the goodness of fit via a bootstrapping procedure [6]. From $N = 1,000$ bootstrap replications, we obtain the fitted values of exponent λ with the mean value $\mu = 3.65$ ($\sigma = 0.04$) and $p = 0.18$, confirming a power-law distribution in the sizes of conversations. This may imply a preferential attachment process that people tend to follow hot conversations in which many users have already involved.

Note that about 7% of tweets in the 1,044,573 conversations had already been deleted at the time of our data collection. Only content of 2,062,690 tweets were retrieved and used in the content analysis of main text. Moreover, these tweets are posted by 66,316 distinct users, where 24,860 users are absent in our user sample of 41,456 ED-related users. Although not all the new users are strongly related to ED, to preserve the integrity of communication flow, we included these new users in the network analysis of the main text.

3 Statistics of topics

Figure S2 gives descriptive statistics for the 26 topics identified above. As shown in Figure S2(a), most hashatags (83%) are classified into five topics with IDs 2, 4, 8, 16 and 22 respectively. These topics have been talked about by a large number of users in tweets (Figures S2(b) and (c)), indicating their popularity among users. By inspecting the numbers of tweets that contain a hashtag labeled with a topic per month, we find that users have consistently high levels of engagement in sharing these five topics over time (Figure S2(d)). In contrast, other topics are much less popular. To avoid analyzing topics of interest to a specific subgroup of online ED communities, we focus on analyzing the five popular topics.
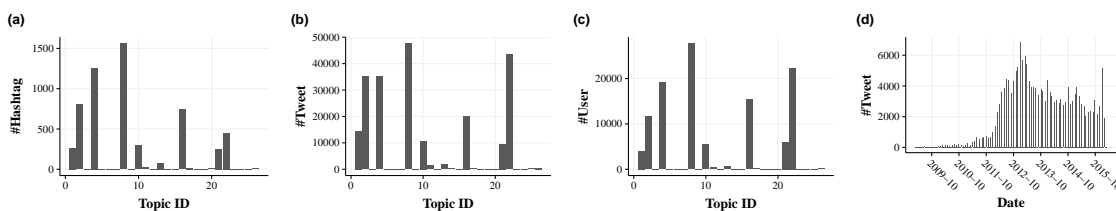

Figure S2: Characterization of the 26 topics found in hashtag co-occurrence networks. (a) The numbers of hashtags in each topic labeled by an ID in x-axis. (b) The number of tweets containing a hashtag in each topic. (c) The number of users who posted a hashtag of each topic. (d) The numbers of tweets on the five most popular topics per month.

4 Topic validation

To ensure the validity of our results, we check the reliability of the topic structure found in users' conversations. First, we check if the relationships of topics aligns with findings in prior qualitative studies on online ED content [7, 8]. To this end, we project hashtags to their associated topics and measure the relatedness of topics based on the co-occurrences of topics in tweets, where topics often co-occurring in the same tweets tend to correlate [9]. To avoid bias that popular topics tend to have frequent co-occurrences, we quantify the relatedness of pairwise topics using the Jaccard coefficient, i.e., the ratio of the number of tweets mentioning both topics over the number of tweets mentioning at least one topic, rather than the absolute numbers of co-occurrences. Figure 1(f) of the main text shows relatedness of topics, where we notice that the *thinspo* topic is highly related to the *body* and *fitness* topics. This confirms prior qualitative studies showing that pro-ED content often contains graphic material inspiring the adoration of a thin body, and exercising or dieting tips on losing weight [7, 8]. In contrast, the *mental* topic is less related to *body* and more related to *rdchat* which is about online charting on nutrition with registered dietitians[1]. This confirms the recovery-oriented feature of *mental*, as a shift of focus from physical appearance to healthy diet is an important movement into ED recovery [10] and talking to professionals is a useful way to cope with the disease [11].

Second, we check if these topics cover real world events in ED communities. Two well-known events are "Eating Disorders Awareness Week (EDAW)", a campaign run by pro-recovery communities to raise awareness of risks of ED from February to March[2], and the "Skinny4Xmas" challenge which is run by pro-ED communities to achieve a net calorie goal (i.e., the total number of calories consumed minus that burned by exercises) from October to December[3]. Inspecting the

---

[1]https://www.symplur.com/healthcare-hashtags/rdchat/

[2]https://www.beateatingdisorders.org.uk/edaw

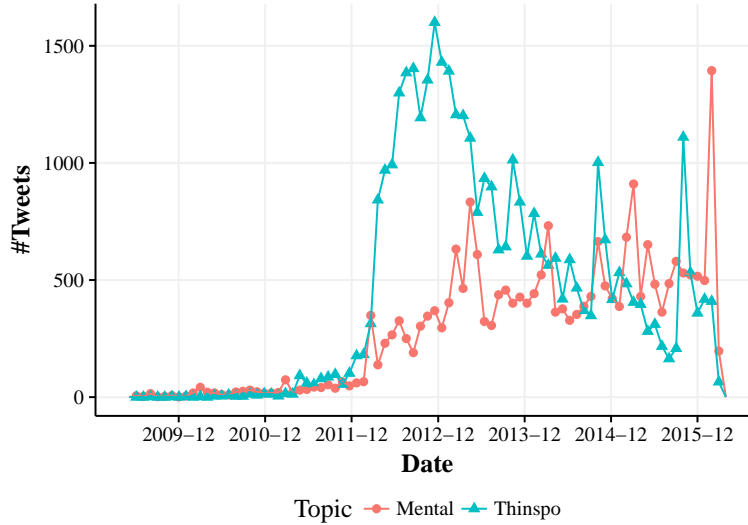[3]http://letters-from-ana.blogspot.co.uk/2013/10/skinny4xmas.html

Figure S3: Numbers of tweets on *mental* and *thinspo* per month.

numbers of tweets on *mental* and *thinspo* topics over time (Figure S3), we find that our identified topics indeed relate to these two events, as a large number of tweets on *mental* appear around March (i.e., the time period of "EDAW") and many tweets on *thinspo* appear around October (i.e., the period of "Skinny4Xmas"). This further confirms that the topics found by the clustering algorithms give a reliable picture on the types of content discussed in online ED communities.

5 Null models for testing inter-layer correlations

We use the following null models to evaluate the significance of inter-layer correlations.

Hypergeometric model: a null model testing the correlations of nodes' activities across layers in a multilayer network [12]. In this model, the number $N^{[\alpha]}$ of active nodes at each layer α is fixed to be that in the original multilayer network and $N^{[\alpha]}$ nodes are randomly sampled to be active at a layer α by a uniform probability from all N nodes of the network. The null hypothesis of this model is that the activity of a node at a layer is uncorrelated from its activities at other layers. We use this model to assess the correlations of users' activities in different communication, i.e., empirical results of multiplexity.

Independent multilayer node-permutation model: an extension of the node-label permutation model [13], in which we randomly reshuffle the identities of nodes (i.e., IDs of the corresponding users) while keeping the topology (i.e., the degrees of nodes, the edges and the weights attached to edges) at a layer α intact. However, reshuffling the identities of all nodes (i.e., both active and inactive nodes) at a layer can lead to variations in the activities of nodes at the layer, which makes randomized networks not comparable to the original multilayer network. To maintain the activities of nodes across layers, we only reshuffle nodes that are active at each layer of the original multilayer network. The null hypothesis is that individuals can occupy any network position at a layer and their positions at the layer are unrelated to those at other layers. We use this model to evaluate the correlations of users' roles in different communication, i.e., empirical results of in-/out-strength correlations of nodes across layers.

4

Independent multilayer configuration model: a model testing the relations of link structures across layers in a multilayer network [14]. In this model, we fix the degree sequences of nodes at each layer α and randomly rewire the edges at α. The null hypothesis is that nodes' interconnections at a layer are independent from those at other layers. We use this model to assess the correlations of users' connectivities in different communication, i.e., empirical results of link overlaps.

Independent directed-weight reshuffling model: an extension of the directed-weight reshuffling model [15]. For each layer α, we fix the network structure at α (i.e., the degrees of nodes and the links), and reshuffle weights locally for each node across its out-links. That is, weights are reshuffled within links sourced from the same node at each layer. In this way, randomized networks preserve not only nodes' activities and contacts, but also their engagement levels (i.e., out-strengths) across layers in the original multilayer network. The null hypothesis is that the strength of a link between two nodes at a layer is unrelated to those at other layers. We use this model to test the correlations of users' interaction strengths in different communication, i.e., empirical results of link strengths across layers.
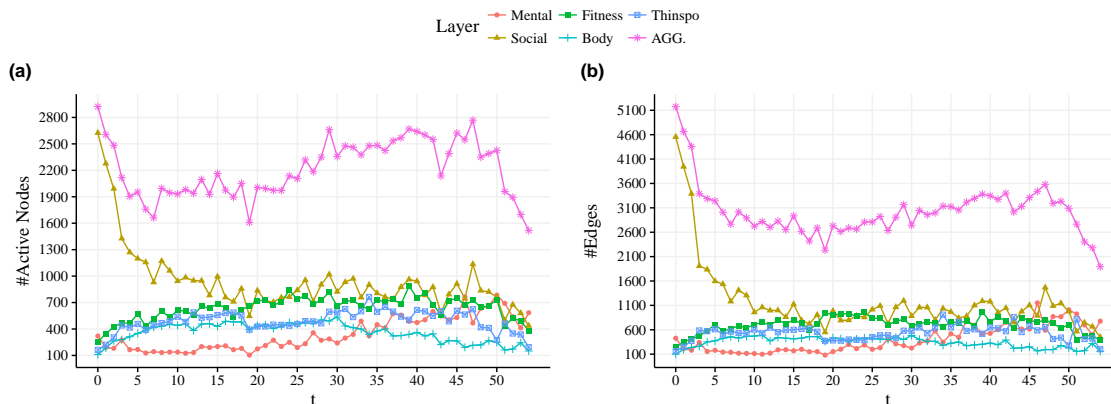
6 Statistics of temporal multilayer networks

Figure S4: Statistics of temporal multilayer networks. (a) Numbers of active nodes and (b) numbers of directed links at each individual layer of a temporal multilayer network and the aggregated network of all layers (AGG.) at time t.

Figure S4 shows the numbers of active nodes and edges at each layer of the generated temporal multilayer networks, as well as the numbers of nodes and edges in the aggregated networks over time. While different single-layer networks show different trends in the numbers of active nodes and links, the total numbers of actors and connections in each temporal multilayer network (i.e., statistics for the aggregated networks) are sufficiently large to provide reasonably statistical power.

References

[1] S. Chancellor, T. Mitra, and M. De Choudhury, "Recovery amid pro-anorexia: Analysis of recovery in social media," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, (San Jose, California, USA), pp. 2111–2123, ACM, 2016.

[2] M. De Choudhury, "Anorexia on tumblr: A characterization study," in *Proceedings of the 5th international conference on digital health 2015*, (Florence, Italy), pp. 43–50, ACM, 2015.

[3] E. Yom-Tov, L. Fernandez-Luque, I. Weber, and S. P. Crain, "Pro-anorexia and pro-recovery photo sharing: a tale of two warring tribes," *Journal of medical Internet research*, vol. 14, no. 6, p. e151, 2012.

[4] T. Wang, M. Brede, A. Ianni, and E. Mentzakis, "Social interactions in online eating disorder communities: A network perspective," *PloS one*, vol. 13, no. 7, p. e0200800, 2018.

[5] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *science*, vol. 286, no. 5439, pp. 509–512, 1999.

[6] A. Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," *SIAM review*, vol. 51, no. 4, pp. 661–703, 2009.

[7] A. S. Juarascio, A. Shoaib, and C. A. Timko, "Pro-eating disorder communities on social networking sites: a content analysis," *Eating disorders*, vol. 18, no. 5, pp. 393–407, 2010.

[8] D. L. Borzekowski, S. Schenk, J. L. Wilson, and R. Peebles, "e-ana and e-mia: A content analysis of pro–eating disorder web sites," *American journal of public health*, vol. 100, no. 8, pp. 1526–1534, 2010.

[9] L. Weng and F. Menczer, "Topicality and impact in social media: diverse messages, focused messengers," *PloS one*, vol. 10, no. 2, p. e0118410, 2015.

[10] J. R. Shapiro, N. D. Berkman, K. A. Brownley, J. A. Sedway, K. N. Lohr, and C. M. Bulik, "Bulimia nervosa treatment: a systematic review of randomized controlled trials," *International Journal of Eating Disorders*, vol. 40, no. 4, pp. 321–336, 2007.

[11] D. Linville, T. Brown, K. Sturm, and T. McDougal, "Eating disorders and social support: perspectives of recovered individuals," *Eating Disorders*, vol. 20, no. 3, pp. 216–231, 2012.

[12] V. Nicosia and V. Latora, "Measuring and modeling correlations in multiplex networks," *Physical Review E*, vol. 92, no. 3, p. 032805, 2015.

[13] D. P. Croft, J. R. Madden, D. W. Franks, and R. James, "Hypothesis testing in animal social networks," *Trends in Ecology & Evolution*, vol. 26, no. 10, pp. 502–507, 2011.

[14] S. Paul and Y. Chen, "Null models and modularity based community detection in multi-layer networks," *arXiv preprint arXiv:1608.00623*, 2016.

[15] T. Opsahl, V. Colizza, P. Panzarasa, and J. J. Ramasco, "Prominence and control: the weighted rich-club effect," *Physical review letters*, vol. 101, no. 16, p. 168702, 2008.