Evaluating status and value assortativity in Threads

(Discussion Paper)

Gianluca Bonifazi[1,†], Enrico Corradini[1,*,†] and Domenico Ursino[1,†]

[2]DII, Polytechnic University of Marche

Abstract

The concept of assortativity in complex networks indicates the preference of a node to relate to other nodes that are somewhat similar. It is possible to think of different forms of similarity between nodes that can give rise to different forms of assortativity. In this paper, along the lines of homophily (of which assortativity can be seen as a special case), we define two categories of assortativity, namely status assortativity and value assortativity. We then show that all definitions of assortativity introduced in the past belong to one of the two categories. Afterwards, we define and evaluate two forms of status assortativity and one form of value assortativity in Threads. Since this social network is relatively new, we could not use existing datasets related to it, and therefore had to build one from scratch, which we now make available to all interested researchers.

Keywords

Assortativity, Threads, Social Network Analysis, Value and Status Homophily

1. Introduction

Assortativity is a central concept in complex network analysis. It was introduced by Newman [1] and quantifies the propensity of network nodes to connect with other nodes that are similar in some way. Similarities can be of different type; for example they may involve structural aspects or characteristics of the objects/people represented by the nodes. Very often structural aspects are considered, and generally the focus is on degree of nodes, in which case we refer to degree assortativity. Assortativity and its counterpart, disassortativity, play a critical role in determining the structural and dynamic aspects of networks. They influence the coherence of networks, their resilience to perturbations and the efficiency of processes such as information dissemination, epidemic spread, and virus control [2, 3, 4, 5, 1].

In the Social Network Analysis area, assortativity can be seen as a special case of the concept of homophily. This concept was introduced by Lazarsfeld and Merton [6, 7] and indicates the tendency of participants in a community to interact primarily with other participants who have the same characteristics. In sociology, there are two types of homophily, status homophily and value homophily. Status homophily is the tendency of individuals to interact with others who have the same status. Value homophily indicates the tendency of individuals to interact with others with whom they share the same values [8]. When we apply this concept to Social

SEBD 2024: 32nd Symposium on Advanced Database Systems, June 23-26, 2024, Villasimius, Sardinia, Italy

*Corresponding author.

†These authors contributed equally.

g.bonifazi@univpm.it (G. Bonifazi); e.corradini@univpm.it (E. Corradini); d.ursino@univpm.it (D. Ursino)

0000-0002-1947-8667 (G. Bonifazi); 0000-0002-1140-4209 (E. Corradini); 0000-0003-1360-8499 (D. Ursino)

Network Analysis, status homophily mainly refers to the structure of the network where, for example, high-degree nodes tend to interact with high-degree nodes. Value homophily, on the other hand, refers to the content that users associated with nodes exchange and publish; indeed, published content reveals the values of the corresponding authors.

Since assortativity is closely related to homophily in Social Network Analysis, we believe that, at least in this context, and possibly in others that we will analyze in the future, it is possible to define two categories of assortativity by distinguishing between status assortativity and value assortativity. Status assortativity occurs when the similarity between nodes is evaluated based on their structural characteristics (e.g., their degree or betweenness or eigenvector centrality). In contrast, value assortativity occurs when the similarity between nodes is evaluated based on the content published and exchanged by the corresponding users. Accordingly, status assortativity indicates the preference, of the node in a network, to relate to other nodes that are structurally similar. Value assortativity, on the other hand, indicates the preference for the node in a network to relate to other nodes in such a way that the corresponding users publish similar content and thus show similar interests.

Having introduced the distinction between status assortativity and value assortativity, we decided to compute these two measures on Threads[1]. Launched by Meta on July 6, 2023, Threads is a new content-based social platform seen by many experts as a direct competitor to X. Designed to share text updates and foster public conversations, Threads has quickly taken its place in the social media landscape, recently reaching more than 130 million monthly active users[2].Because of its newness, Threads has not yet been extensively studied by social network researchers. The idea of applying our new concepts of status and value assortativity directly to Threads allows us to contribute to better understanding this little explored digital platform. In addition, we extend the scientific community's knowledge of assortativity to this new social platform. In order to study assortativity on Threads, it was crucial to have a dataset derived from it that contained all structural and content information capable of supporting this type of analysis. Unfortunately, we were unable to find a dataset in the literature that would support such an analysis. Therefore, we had to construct one. Once we completed this task, we decided to make such a dataset open by making it available to all researchers who want to perform analyses and studies on Threads.

After building the Threads dataset, we ran our status and value assortativity analyses on it and compared the results with those already known about other social platforms, highlighting similarities and differences.

The structure of this paper is as follows: In Section 2, we introduce the concepts of status and value assortativity and propose some specific forms of them. In Section 3, we describe the main features of our Threads dataset, illustrate our experiments to compute different forms of assortativity on Threads and compare our results with those of other social platforms. Finally, in Section 4 we draw our conclusions and highlight some possible developments of our research efforts.

[1] www.threads.net

[2] https://techcrunch.com/2024/02/01/threads-now-reaches-more-130-million-monthly-users-says-meta-up-30m-from-q3

2. Defining status and value assortativity

In the Introduction, we stated that the first goal of this paper is to deepen the concept of assortativity by introducing a distinction between status and value assortativity. In particular: *(i) Status assortativity* indicates the preference, for the node in a network, to relate to other nodes that are structurally similar; *(ii) Value assortativity* denotes the preference, for the node in a network, to relate to other nodes such that the corresponding users publish similar content, thus showing similar interests.

Starting from these two definitions, in this section we propose some forms of status and value assortativity that are applicable to any content-based social platform.

To do this, we must first introduce a model for representing a content-based network. In particular, the latter can be represented as a directed graph:

$$\mathcal{N} = \langle N, A \rangle \tag{2.1}$$

N is the set of nodes in $\mathcal{N}$. A node $n_i \in \mathcal{N}$ is associated with a user u_i in the content-based network. Since there is a biunivocal correspondence between a node and a user, we will use these two terms interchangeably in the following. Each node n_i is associated with a label f_i indicating the number of followers of u_i. A is the set of arcs of $\mathcal{N}$. An arc $a_{ij} = (n_i, n_j) \in A$ represents the set of interactions from a user u_i to a user u_j. An interaction from u_i to u_j indicates that u_i commented on a post by u_j. Each arc a_{ij} is associated with a label TS_{ij} indicating the main topics discussed in the interactions from n_i to n_j. TS_{ij} depends on the content of posts and comments sent from n_i to n_j. Examples of possible topics are “Technology”, “Health”, “Entertainment”, “Politics”, etc.

Starting from this model, we are able to define some possible versions of status and value assortativity.

The first status assortativity we consider is the most classical one, i.e., degree assortativity. Based on our model, it can be defined as:

$$r_D = \frac{\sum_{a_{ij} \in A} (k_i - \overline{k})(k_j - \overline{k})}{\sum_{a_{ij} \in A} (k_i - \overline{k})^2} \tag{2.2}$$

In this case, A is the set of arcs of $\mathcal{N}$, k_i (resp., k_j) is the degree (intended as the sum of indegree and outdegree) of n_i (resp., n_j) and $\overline{k}$ is the average degree of the nodes of $\mathcal{N}$. In this formula, the numerator is the result of the product of the degree deviations from the mean for each pair of connected nodes. It represents the covariance of degrees between all pairs of connected nodes. A positive numerator indicates a tendency for nodes with high degree to connect with other nodes with high degree (assortativity or assortative mixing), while a negative numerator denotes a tendency for nodes with high degree to connect with other nodes with low degree (disassortativity or disassortative mixing). The denominator represents the variance in the degree of all nodes. It acts as a normalization factor and assures us that the values of r_D are within the real range $[-1, 1]$, where 1 denotes perfect assortativity, -1 represents perfect disassortativity and 0 indicates the lack of any assortativity relationships.

Since our model associates each node $n_i \in N$ with the number f_i of its followers, it is possible to think of a second version of status assortativity that we call weighted degree assortativity. It can be defined as follows:

$$r_{WD} = \frac{\sum_{a_{ij} \in A} \alpha(n_i, n_j) \cdot (k_i - \overline{k_w})(k_j - \overline{k_w})}{\sum_{a_{ij} \in A} \alpha(n_i, n_j) \cdot \frac{(k_i - \overline{k_w})^2 + (k_j - \overline{k_w})^2}{2}} \tag{2.3}$$

Here, N is the set of nodes of $\mathcal{N}$. $\alpha(n_i, n_j)$ is an aggregation function of the number of followers of n_i and n_j. It is possible to think of various aggregation functions, for example $\alpha(n_i, n_j) = max(f_i, f_j)$, $\alpha(n_i, n_j) = min(n_i, n_j)$, $\alpha(n_i, n_j) = f_i + f_j$, $\alpha(n_i, n_j) = \frac{f_i + f_j}{2}$. We used the latter function in our experiments. Unlike degree assortativity, weighted degree assortativity takes into account not only the degree of the nodes involved in the interactions but also their influence, as measured by their number of followers.

In this formula, the numerator computes a weighted degree covariance between pairs of connected nodes, where $\alpha(n_i, n_j)$ is an aggregation function of the number of followers of n_i and n_j, and acts as a weight. Therefore, the numerator captures the tendency of influential nodes to connect with other influential nodes and takes into account both the number of connections and the level of influence of connected nodes. The denominator computes the variance of the nodes' degrees, weighted against the nodes' level of influence, and provides a basis for comparison with the numerator that takes into account both the degree and the level of influence of connected nodes. It acts as a normalization factor by ensuring that the values of r_{WD} are within the real range $[-1, 1]$, where 1 denotes complete assortativity, -1 indicates complete disassortativity, and 0 represents the lack of any assortativity relationship.

Weighted degree assortativity r_{WD} is useful in networks where the number of connections (degree) alone does not fully capture the importance or influence of a node. For example, in the case of content dissemination, a user with few connections but a large number of followers can have a significant impact. In a case like this, r_{WD} provides a more nuanced version of assortativity, reflecting not only how nodes are connected but also how their influence or popularity contribute to the network's potential. In particular, if we consider high-influence networks, where influence plays a critical role in shaping interactions, this assortativity can better capture the dynamics of interactions between nodes than degree assortativity. Moreover, in heterogeneous networks, with a high variance of influence of nodes, r_{WD} can show connectivity patterns that r_D could overlook. This is especially relevant when we want to analyze how content dissemination or engagement patterns are correlated with user influence.

Having seen two examples of status assortativity, we now introduce a definition of value assortativity. As in the case of status assortativity, the following is not the only possible definition of value assortativity, but in the future other definitions could be introduced based on research needs. Since value assortativity concerns content, the topic set TS_{ij} will play a key role in its definition. The formula we propose for value assortativity is the following:

$$r_V = \frac{\sum_{a_{ij} \in A} |TS_{ij}| \cdot (T_i - \overline{T})(T_j - \overline{T})}{\sum_{a_{ij} \in A} |TS_{ij}| \cdot \frac{(T_i - \overline{T})^2 + (T_j - \overline{T})^2}{2}} \tag{2.4}$$

As can be seen, the structure of this formula is similar to that of the formula of r_{WD} except that, in this case, instead of the function $\alpha(n_i, n_j)$ and node degrees (which are all structural measures), we have the cardinality of the set TS_{ij} and the variables T_i, T_j and $\overline{T}$. In particular, recall that TS_{ij} denotes the set of topics related to the comments from n_i to n_j. Instead, T_i (resp., T_j) is a variable indicating the number of topics characterizing all messages posted by n_i (resp., n_j), while $\overline{T}$ is the average number of topics of the messages posted by all users of $\mathcal{N}$.

In this formula, the numerator is given by the product of the differences between the average number $\overline{T}$ of topics handled by the connected nodes of $\mathcal{N}$ and the number of topics handled by each pair of connected nodes in the network. Each of these differences is weighted by the cardinality of the set TS_{ij} of topics in common between each pair of connected nodes in the network. $|TS_{ij}|$ acts as a weight indicating how extensive the shared interests are between the users u_i and u_j associated with the nodes n_i and n_j. Each component of the sum at the numerator measures how similarly or dissimilarly the nodes engage in topics compared with the average engagement level of the nodes in the network. A positive, high numerator indicates that nodes with similar levels of engagements in topics (both in terms of number of topics and number of common topics) tend to interact, implying value assortativity. A negative numerator with a high absolute value denotes that nodes with different levels of engagement on topics tend to interact, implying value disassortativity. The denominator is used to normalize the numerator by taking into account the variance of the number of topics related to each node, again weighted by the number of topics in common between each pair of nodes. This ensures that the value of r_V varies in the real range $[-1, 1]$, where 1 indicates total value assortativity, -1 denotes complete value disassortativity, and 0 represents the lack of any assortativity relationship.

3. Experiments

3.1. Dataset description

As we mentioned in the Introduction, in order to test our definitions of status and value assortativity on Threads, we had to build an appropriate dataset, since we did not find any available open dataset that could support our work. Once we built such a dataset, we decided to make it available for all researchers who want to perform analyses on Threads. It can be found at the following address: https://github.com/ecorradini/ThreadsDataset. The dataset is anonymized to preserve user privacy. In this section, we describe in detail how we obtained it and its main features.

To collect data from Threads we used a server with a 16 core CPU, 96 GB of RAM and Ubuntu 22.04 operating system. Threads allows access to its feed in the European Union without the need for an account. We organized all collected data into two main files, namely: *(i)* `posts.csv`, which records all data related to posts, and *(ii)* `users.csv`, which contains all data related to users. In addition, we created a special folder to store all images and videos linked by posts.

In more detail, the file `posts.csv` has the following fields: *(i)* `url`, which indicates the web address of the post; *(ii)* `parent_post`, which denotes the web address of the parent post, if the original post is a comment; it is empty otherwise; *(iii)* `user`, which indicates the username of the user who created the post; *(iv)* `caption`, which denotes any text or caption associated with the post; *(v)* `image_video`, which indicates the name of the visual content file in the associated folder, if the post includes images or videos; it is empty otherwise; *(vi)* `time`, which denotes the timestamp when the post was made; *(vii)* `likes`, which denotes the number of likes received by the post; it is set to 0 if the post received no like.

Instead, the file `users.csv` has the following fields: *(i)* `url`, which indicates the web address of the user's profile; *(ii)* `username`, which denotes the username of the user; *(iii)* `display_name`, which indicates the user's display name; *(iv)* `bio`, which indicates biographical information, if it is present in the user's profile; it is empty otherwise; *(v)* `bio_url`, which denotes web links if they are present in the user's profile; it is empty otherwise; *(vi)* `followers`, which indicates the number of followers of the user; it is set to 0 if the user has no follower.

Table 1 shows some basic statistics of the network $\mathcal{N}$ associated with our Threads dataset.

Statistic	*Value*
Number of nodes	26,248
Number of arcs	39,771
Number of isolated nodes	770
Density	$5.773 \cdot 10^{-5}$
Average Clustering Coefficient	0.003197
Average indegree of nodes (excluding isolated ones)	1.56
Average outdegree of nodes (excluding isolated ones)	1.56

Table 1
Some basic statistics of the network $\mathcal{N}$

3.2. Investigating status and value assortativity on Threads

The first analysis on status assortativity involved the computation of degree assortativity (see Equation 2.2). By performing such a computation we obtained that $r_D = -0.0303$. This result shows that, as far as this form of assortativity is concerned, users in Threads are neither assortative nor disassortative. This, in turn, implies that high-degree users tend to establish connections with high-degree, medium-degree and low-degree ones, and vice versa. This behavior differs from what happens to other social networks where researchers generally find the existence of degree assortativity among users. It may be motivated by the fact that Threads is still a new network and, therefore, there has not been time for backbones of influencers, or other strong typologies of interactions among them, to form. At the moment, therefore, the influencers' interest in spreading their content directly to everyone, without superstructures such as backbones with other influencers, prevails. Of course, we cannot exclude that as time goes on, influencers' backbone or other superstructures involving this type of users will also form in Threads.

The second analysis on status assortativity involved the computation of weighted degree assortativity (see Equation 2.3). By performing that computation, we obtained that $r_{WD} = 0.0157$. This result differs from the previous one only in the sign; indeed, in both cases the value of the assortativity coefficient is very close to 0. This implies that, again, no significant assortativity or disassortativity relationships are evident among the network users. Recall that weighted degree assortativity differs from degree assortativity in that it also takes into account the weight of nodes, which, in our model, is given by the number of followers of the corresponding users. Consequently, this form of assortativity focuses even more on influencers than the previous form; here, influencers are evaluated based on not only the number of connections they have but also the number of people who follow them. The fact that we again obtain a substantially null assortativity value is a further confirmation that in Threads there still does not seem to exist any superstructure (e.g., backbones) through which influencers support each other.

After evaluating status assortativity on the Threads dataset, we moved on to consider value assortativity on it. The first task to do was the extraction of the set of topics related to user interactions. For this purpose, we used OpenAI's GPT-3.5 model[3]. More specifically, for each post or comment of the dataset, we used the `gpt-3.5-turbo` to extract the topic that most represented it. Proceeding in this way, we identified 846 different topics. Figure 1 shows the distribution of the 100 most frequent ones. Note that there are 3,727 arcs whose posts and comments feature an "Uncategorized" topic. This is mainly due to the fact that these comments or posts consist only of emojis or single words like "Yes" or "No". We did not consider these comments or posts in the computation of value assortativity.

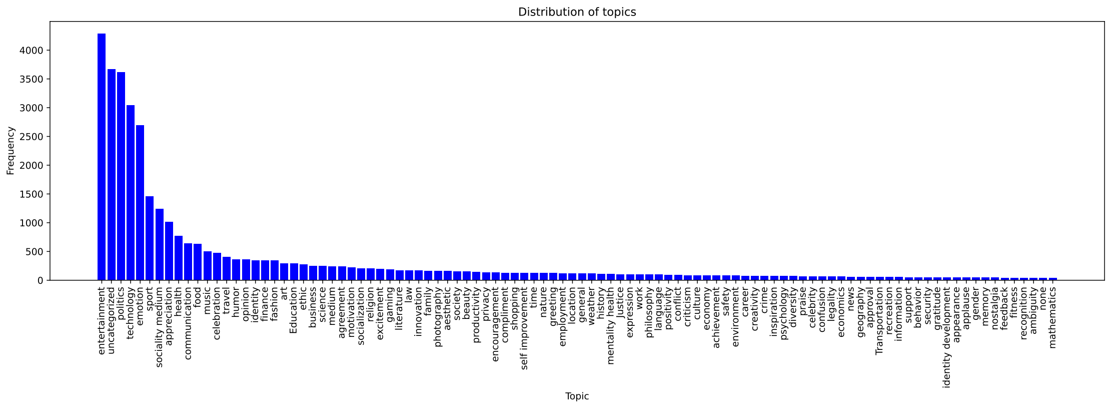

Figure 1: Distribution of the 100 most frequent topics in our Threads dataset

At this point, we computed the value assortativity coefficient r_V by applying Equation 2.4. At the end of this task, we obtained a value of r_V equal to 0.103. This value reveals a slight tendency of Threads users to communicate with other users who posted similar content, and thus showed common interests. This result suggests the existence within Threads of communities or clusters of users who are interesed in the same topics and like to interact with each other. The positive, but not extremely high, assortativity value can be explained by considering that a user is generally interested in multiple topics and therefore tends to interact with other users within multiple communities each of which could be only minimally intersecting with another.

3.3. Comparison with other social platforms

Having computed status and value assortativity in Threads, in Table 2 we compare our results with those obtained by other researchers in the past for other social platforms.

[3] www.openai.com

Source	*Investigated network*	*Assortativity Type*	*Assortativity based on*	*Obtained result*
This paper	Threads	Status	Degree of nodes	Unassortative (i.e., neutral), $r = -0.0303$
This paper	Threads	Status	Number of followers	Unassortative (i.e., neutral), $r = 0.0157$
This paper	Threads	Value	Posts and Comments	Assortative, $r = 0.103$
[1]	Several complex networks	Status	Degree of nodes	It depends on networks
[9]	Several complex networks	Status	Node degrees, links and weights	It depends on networks
[10]	Several complex networks	Status	Network configuration	It depends on networks
[11]	Arxiv:condmat	Status	Degree of nodes	Assortative, $r \in [0.14, 0.35]$ depending on parameters
[12]	Several online social networks	Status	Node degrees	It depends on networks
[13]	Several social networks	Status	Social status	Assortative, depending on parameters
[14]	X	Value	Happiness	Assortative, r constant (about 0.4)
[15]	X	Value	Subjective Well-Being	Assortative, $r = [0.1, 0.4]$
[16]	X	Value	Suicide	Assortative
[4]	Reddit	Status	Degree centrality	Assortative for some centrality intervals
[4]	Reddit	Status	Eigenvector centrality	Assortative for some centrality intervals
[4]	Reddit	Value	Co-posting activity	Assortative
[17]	Facebook	Status	Age	Assortative, $r = 0.226$
[18]	Facebook	Status	Membership of common platforms	Assortative, $r = 0.074$
[19]	Facebook	Status	Based on bridge users	Assortative, $r = 0.894$
[19]	X	Status	Based on bridge users	Assortative, $r = 0.675$
[20]	X	Status	Network structure	Assortative, $r = 0.996$
[21]	X	Status	Degree assortativity	Assortative, $r = 0.414$
[21]	X	Value	Shared interests and language	Assortative, r positive for all attributes considered.
[22]	X	Value	Music Category	Assortative, $r = 0.737$
[22]	X	Value	Cinema Category	Assortative, $r = 0.811$
[22]	X	Value	Sport Category	Assortative, $r = 0.776$

Table 2
An overview of investigations on assortativity

4. Conclusions

In this paper, we proposed an in-depth study of the concept of assortativity proposed by Newman [1]. First, we introduced the concept of status assortativity, which takes into account the structure of the network, and that of content assortativity, which takes into account the content exchanged between nodes, if it exists. We then applied our definitions of assortativity to Threads. In order to perform our analysis, we needed a dataset of Threads storing the network users, their interactions, and the content they exchanged. Unfortunately, we could not find an existing dataset that was suitable for our purposes. Therefore, we had to build a new dataset from scratch.

This paper should not be seen as an end point but as a starting point for further researches in this context. In fact, in the future, we can think of designing a framework that makes it easy to define new forms of status and value assortativity on Threads and other social networks as the need or opportunity arises. This framework could include a machine learning component to predict changes in assortativity based on trends that should gradually emerge in interactions of users and the content they exchange. At a later stage, we could extend this framework to consider other forms of network dynamics beyond assortativity. Another possible extension of our approach would be to consider not only textual posts, but also images and videos in the study of value assortativity. Finally, we would like to make an in-depth comparison between X and Threads with respect to different forms of value and status assortativity, since these two networks are direct competitors. This comparison may allow us to gain insight into the differences in user connections and behavior in the two networks.

References

[1] M. Newman, Assortative mixing in networks, Physical Review Letters 89 (2002) 208701. APS.

[2] A. A. Musawi, S. Roy, P. Ghosh, Examining indicators of complex network vulnerability across diverse attack scenarios, Scientific Reports 13 (2023) 18208. Nature Publishing Group UK London.

[3] L. Xue, P. Cui, Z. Di, Identifying the structure patterns to govern the performance of localization in regulating innovation diffusion, New Journal of Physics 25 (2023) 053009. IOP Publishing.

[4] F. Cauteruccio, E. Corradini, G. Terracina, F. Ursino, L. Virgili, Investigating Reddit to detect subreddit and author stereotypes and to evaluate author assortativity, Journal of Information Science 48 (2022) 783–810. SAGE Publications Sage UK: London, England.

[5] A. Al-Amery, J. Herrera-Diestra, Z. Du, Z. Ertem, Epidemic Surveillance Systems and Containment Strategies in Complex Networks, Available at SSRN 4354792 (2023).

[6] P. Lazarsfeld, R. Merton, Friendship as a social process: A substantive and methodological analysis, Freedom and control in modern society 18 (1954) 18–66.

[7] R. Merton, Social theory and social structure, 1968. Simon and Schuster.

[8] M. Tsvetovat, A. Kouznetsov, Social Network Analysis for Startups: Finding connections on the social web, Sebastopol, CA, USA, 2011. O'Reilly Media, Inc.

[9] R. Noldus, P. V. Mieghem, Assortativity in complex networks, Journal of Complex Networks 3 (2015) 507–542. Oxford University Press.

[10] S. Johnson, J. Torres, J. Marro, M. Munoz, Entropic origin of disassortativity in complex networks, Physical Review Letters 104 (2010) 108702. APS.

[11] M. Catanzaro, G. Caldarelli, L. Pietronero, Assortative model for social networks, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 70(3) (2004) 037101–037104. The American Physical Society.

[12] H. Hu, X. Wang, Disassortative mixing in online social networks, EPL (Europhysics Letters) 86 (2009) 18003. IOP Publishing.

[13] B. Zhou, X. Lu, P. Holme, Universal evolution patterns of degree assortativity in social networks, Social Networks 63 (2020) 47–55. Elsevier.

[14] C. Bliss, I. Kloumann, K. Harris, C. Danforth, P. Dodds, Twitter reciprocal reply networks exhibit assortativity with respect to happiness, Journal of Computational Science 3 (2012) 388–397. Elsevier.

[15] J. Bollen, B. Gonçalves, G. Ruan, H. Mao, Happiness is assortative in online social networks, Artificial life 17 (2011) 237–251. MIT Press.

[16] I. Cero, T. Witte, Assortativity of suicide-related posting on social media, American Psychologist 75 (2020) 365. American Psychological Association.

[17] J. Ugander, B. Karrer, L. Backstrom, C. Marlow, The anatomy of the facebook social graph, arXiv preprint arXiv:1111.4503 (2011).

[18] F. Buccafurri, G. Lax, S. Nicolazzo, A. Nocera, Comparing Twitter and Facebook user behavior: Privacy and other aspects, Computers in Human Behavior 52 (2015) 87–95. Elsevier.

[19] F. Buccafurri, G. Lax, A. Nocera, A new form of assortativity in online social networks,

International Journal of Human-Computer Studies 80 (2015) 56–65. Elsevier.
[20] V. Singh, A. Singh, D. Jain, V. Kumar, P. Verma, Patterns affecting structural properties of social networking site “Twitter”’, International Journal of Business Information Systems 27 (2018) 411–432. Inderscience Publishers (IEL).
[21] S. Šćepanović, I. Mishkovski, B. Gonçalves, T. Nguyen, P. Hui, Semantic homophily in online communication: Evidence from Twitter, Online Social Networks and Media 2 (2017) 1–18. Elsevier.
[22] F. Buccafurri, G. Lax, S. Nicolazzo, A. Nocera, Interest Assortativity in Twitter, in: Proc. of the 12th International Conference on Web Information Systems and Technologies (WEBIST 2016), Rome, Italy, 2016, pp. 239–246. "SCITEPRESS – Science and Technology Publications, Lda".