Homophily - A Driving Factor for Hate Speech on Twitter

Seema Nagar, [1] **Sameer Gupta,** [2] **C.S. Bahushruth,** [3] **Ferdous Ahmed Barbhuiya,** [1] **Kuntal Dey** [4]
[1] Indian Institute of Information Technology, Guwahati
[2] National Institute of Technology, Kurukshetra
[3] Manipal University, Jaipur
[4] Accenture Technology Labs, Bangalore
seema.nagar@iiitg.ac.in, sameer.lego@gmail.com, bahushruth.bahushruth@gmail.com, ferdous@iiitg.ac.in,
kuntal.dey@accenture.com

Abstract

In this paper, we hypothesize that homophily is a powerful driver in the generation of hate speech on social media platforms. Prior work has shown the role of homophily in information diffusion, sustenance of online guilds, contagion in product adoption, the emergence of topics, and life-cycle on social networks. We note that familiarity computation techniques used in the literature for solving this problem, such as the presence of an edge, and the number of mutual friends between a pair of users, have ample scope of improvement. We address the limitations of existing familiarity computation methods by employing a variational graph auto-encoder to capture a user's position in the network and utilizing this representation to compute familiarity. We empirically demonstrate the presence of homophily on a dataset from Twitter. Further, we investigate how homophily varies with different hateful forms, such as hate manifesting in topics of gender, race, ethnicity, politics and nationalism. Our results indicate higher homophily in users associating with topics of racism, sexism and nationalism.

Introduction

Hate speech is prevalent on social media platforms. Often, hate speech campaigns on social media platforms have incited real-life violence(Ribeiro et al. 2017; Mathew et al. 2018). Facebook has been blamed for instigating anti-Muslim mob violence in Sri Lanka as well as for playing a leading role in the possible genocide of the Rohingya community in Myanmar by spreading hate speech. Consequently, studying hate speech spread is a fundamental problem with practical implications.

Homophily on social networks was first proposed by (McPherson, Smith-Lovin, and Cook 2001), using the assortative mixing hypothesis. Homophily is defined as the tendency of like-minded (similar) people to connect/befriend (familiar). A key observation about homophily is that it structures the ego-networks of individuals and impacts their communication behaviour. Therefore, many works have studied the role of homophily in information diffusion and dissemination(Aral, Muchnik, and Sundararajan 2009; De Choudhury et al. 2010; Halberstam and Knight 2016; Starbird and Palen 2012). (Dey et al. 2018) detect evidence of homophily in the topic formation and life cycle. Very recently, (Xu and Zhou 2020), utilize hashtag based homophily for re-marketing, a practical application with immense effect.

In this paper, we investigate the presence of homophily in hate speech on social media platforms, a new and unexplored facet of hate speech. Familiarity and similarity are the two essential factors in homophily. These two factors are readily available on social media platforms. The feature to become friends or follow other users is leveraged to compute familiarity. While the content produced, profile information or any other meta-data information present is utilized to calculate similarity. The purpose of familiarity computation is to capture familiarity between a pair of users. Existing approaches have used familiarity metrics such edge exists or not, the number of mutual friends or common neighbours, part of the same community/component or not (Dey et al. 2018).

However, the existing familiarity metrics fail to capture familiarity between a pair of users from the overall position in the social network. Moreover, familiarity metrics is applicable in a situation that has to be established empirically. This is a time consuming, manual, and expensive process. Therefore, we propose using variational graph autoencoder (VGAE) based embeddings of users in the social network to compute familiarity. VGAE based embeddings are shown to be powerful in capturing a vector representation of a user from the position in the graph. The embeddings have proven effective in many social network tasks such as link prediction, node clustering (Pan et al. 2018), node classification(Cai, Zheng, and Chang 2018), community detection(Agrawal, Arquam, and Singh 2020), citation recommendation(Jeong et al. 2019).

We use VGAE to encode a user and then use various distance-based metrics such as Cosine, Euclidean, and Manhattan distance to compute familiarity between a pair of users. We use tweets posted by the users to calculate similarity. A user's tweets are concatenated to get a user document. The user document is encoded using an existing text encoder. The text embeddings are then used to compute similarity for a pair of users. We hypothesize that the newly proposed familiarity metrics are effective against the two existing (edge exists or not, mutual friends count). We ask ourselves a question, " *Is homophily exhibited by the users* *generating hateful content?"*.

Hate speech has multiple forms present, such as hate against race, religion, ethnicity, gender, among others. We believe that the multiple-forms of homophily might exhibit different homophilic behaviour; therefore, we ask ourselves another question, *"Is homophily more prominent for particular forms of hate?"*. We detect various hateful forms present in a corpus using topic modelling. We propose to use latent topic modelling to hateful-forms. We find that particular hateful-forms exhibit stronger homophilic behaviour.

In summary, we make the following contributions:

- We propose a novel way to compute familiarity using graph embedding technique
- We show homophily in hate speech on a dataset from Twitter
- We empirically demonstrate the effectiveness of the newly proposed metrics in establish familiarity against the existing metrics, using homophily as the benchmark of comparison
- We do a deep dive analysis of variations of homophily in different forms of hate

Related Work

A large body of works has tried to understand the homophily on social networks. (McPherson, Smith-Lovin, and Cook 2001) were the first to propose homophily in social networks. This paper defines homophily exists between two users if they follow each other because of shared common interests. Subsequently, (De Choudhury et al. 2010) establish that homophily plays a vital role in information diffusion and dissemination on social networks. Their work is built on a critical observation that homophily structures the ego-networks of individuals and impacts their communication behaviour. Another early work, Weng et al.(Weng et al. 2010) detects the existence of reciprocity that exists on Twitter. They show that users sharing common interests tend to follow each other, therefore, resulting in high reciprocity on Twitter due to homophily. They use similarity and familiarity together, to rank the influential users.

The role of homophily in information diffusion on social networks is further strengthened by (Halberstam and Knight 2016; Aral, Muchnik, and Sundararajan 2009; Starbird and Palen 2012). Specifically, (Halberstam and Knight 2016) investigate homophily in political information diffusion. They proved their hypotheses on two groups, conservative and liberal, and conclude "with homophily, members of the majority group have more network connections and are exposed to more information than the minority group". Additionally, they also show, "with homophily and a tendency to produce like-minded information, groups are disproportionately exposed to like-minded information, and the information reaches like-minded individuals more quickly than it reaches individuals of opposing ideologies." Another work by (Aral, Muchnik, and Sundararajan 2009) demonstrates homophily in contagion for product adoption on dynamic networks. They empirically establish on the Yahoo IM (instant messenger) system that peer influence is not sufficient to explain contagion in product adoption. They find that homophily is responsible for more than half of the contagion.

The role of homophily in sustaining online guilds, especially in the gaming world, is studied by (Ducheneaut et al. 2007). They find that homophily is an essential indicator of a guild's sustenance. Thus, homophily has been very well studied in the literature and is very important to explain that many social phenomena are happening in the virtual world.

Many papers have jointly studied familiarity and similarity for modelling solutions on Twitter. (Ying et al. 2010) show that semantic similarities in trajectories of users on mobile networks are a better indicator for user similarity compare to using the vanilla trajectory similarity. (Afrasiabi Rad and Benyoucef 2014) study communities formed over friendships on the YouTube social network. They observe that communities are formed from similar users on YouTube; however, they do not find large similarity values between friends in YouTube communities. Recently, topical homophily was proposed by (Dey et al. 2018), where they show the homophily is the driving factor in the emergence of topics and their life cycle.

However, the existing literature does not address homophily in hate speech.

Central Idea

The proposed approach has three main components:

- Define a novel metric for familiarity computation
- Utilizing the novel metrics in showing homophily in hate speech.
- Detecting hateful forms on social media platforms

In the first component, we use a graph variational autoencoder to encode a user's position in the social network. We hypothesize that these graph encodings can capture a user's position as well as their society from a network's perspective. The familiarity between a pair of users is computed using these graph encodings. We validate the newly proposed familiarity metrics by showing homophily in hate speech as part of the second component of the proposed approach. In the third component, we use latent topic modelling to detect multiple hateful forms present in hate speech. We hypothesize that individual hateful forms, differing in nature, might exhibit varied homophilic behaviours.

A Novel Familiarity Computation Approach

We aim to capture a user's unified view from a position in their social network. The existing familiarity metrics such as, whether an edge exists or not, mutual friends count, or if users are part of the same community or not, fail to incorporate the unified familiarity for a pair of users. We propose to use a graph encoder, to encode a user's position and then use these encoding vectors to compute familiarity between two users.

Graph Encoder

The graph encoder is capable of representing a user $u's$ latent position in a social network. It takes two inputs, 1) identity matrix and 2) the adjacency matrix, representing the social connectivity. It is trained in an unsupervised manner to learn a vector representation of the input graph. We use a well-established technique called variational graph auto-encoder (Kipf and Welling 2016) for our graph encoder. Variational graph auto-encoder (VGAE) constitutes of two parts, a) the encoder, to encode the graph and b) the decoder, to decode the graph. The encoder is consist of two layers of graph convolutional layers that encodes a graph into lower dimension latent representation Z. At the same time, the decoder reconstructs the adjacency matrix by optimizing cross-entropy loss between the original and reconstructed adjacency and KL-divergence of the approximate from the true posterior. Figure 1 illustrates the VGAE architecture used.

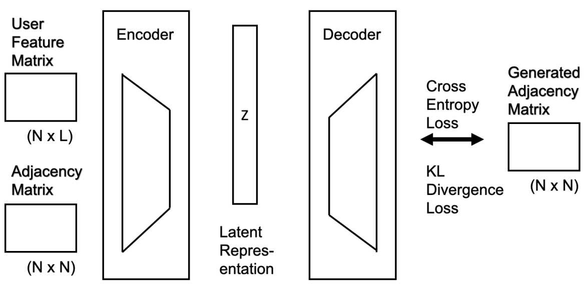

Figure 1: Variational Graph Auto Encoder

Variational Graph Auto Encoder (VGAE) The encoder takes a data point X and produces a distribution $q\phi(Z|X,A)$. The distribution is usually parameterized as a multivariate Gaussian. Therefore, the encoder predicts the means and standard deviation of the Gaussian distribution. The lower-dimensional embedding Z is sampled from this distribution. The decoder is a variational approximation, $p\theta(A|Z)$, which takes an embedding Z and produces the output $\hat{A}$.

Encoder or Inferencing Model

The encoder, called graph convolutional network (GCN), consists of two graph convolutional layers. GCN abstracts the social network graph into a latent representation through a convolutional network. The first layer takes adjacency matrix A and the identity matrix as inputs and generates the latent variable Z as output. The first GCN layer generates a lower-dimensional feature matrix $\bar{X}$, as shown in in the formula in Equation 1 where $\tilde{A} = D^{-1/2}AD^{-1/2}$ is a symmetrically normalized adjacency matrix. The second GCN layer generates μ and log of σ^2 as shown in Equation 2 and 3 .

$$\bar{X} = GCN(X, A) = ReLU(\tilde{A}XW_0) \tag{1}$$

$$\mu = GCN_\mu(X, A) = \tilde{A}\bar{X}W_1 \tag{2}$$

$$log\sigma^2 = GCN_\sigma(X, A) = \tilde{A}\bar{X}W_1 \tag{3}$$

The two layers of GCN can be combined as in Equation 4 and Z can be calculated using Equation 5, where $\epsilon \in \mathcal{N}(0, 1)$, wherein $\mathcal{N}$ represents a Gaussian (normal) distribution.

$$GCN(X, A) = \tilde{A}.ReLu(\tilde{A}XW_0)W_1 \tag{4}$$

$$Z = \mu + \sigma * \epsilon \tag{5}$$

In summary, encoder can be written as in Equation 6.

$$q(z_i|X, A) = N(z_i|\mu_i, diag(\sigma_i^2)) \tag{6}$$

Decoder

The generative model (decoder) is defined by an inner product between latent variable Z and its transpose Z^T. The output by the decoder (Equation 8) is reconstructed as adjacency matrix $\hat{A}$ as shown in Equation 7, where $\sigma()$ is the logistic sigmoid function. In summary, decoder can be represented as in Equation 8.

$$\hat{A} = \sigma(ZZ^T) \tag{7}$$

$$p(A_{ij} = 1|z_i, z_j) = \sigma(z_i^T z_j) \tag{8}$$

Loss in VGAE

The loss function for variational graph autoencoder has two parts. The first part of the loss function measures how well the network reconstructs the data, and it is called the variational lower bound. We model the loss as binary cross-entropy between the input and the output. The second part of the loss function is KL-divergence between $q(Z|X, A)$ and $p(Z)$, where$p(Z) = \mathcal{N}(0, 1)$. It measures how closely our $q(Z|X, A)$ matches with $p(Z)$. Formally, the loss can be written as in Equation 9.

$$L = E_{q(Z|X,A}[logp(A|Z)] - KL[q(Z|X, A)||p(Z)] \tag{9}$$

Computing Familiarity

Familiarity for a pair is computed, using the distance between a pair of users in the graph embedding space. We propose to use existing distance measures to compute familiarity. Distance is inversely proportional to familiarity, lesser the distance, higher the familiarity. Specifically, we experiment with, a) Cosine distance, b) Manhattan distance and c) Euclidean distance to compute familiarity. Formally, the metrics between two users u_1 and u_2 are defined in the equations 10 to 12.

Let $\bar{u_1}$ and $\bar{u_2}$ represent graph embedding of the users u_1 and u_2 respectively.

$$CosineSimiliarity(u1, u2) = \frac{(\bar{u_1} \cdot (\bar{u_2})}{||\bar{u_1}|| \cdot ||\bar{u_2}||} \tag{10}$$

$$ManhattanDistance(u1, u2) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p} \tag{11}$$

$$EuclideanDistance(u1, u2) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{12}$$

Hateful Forms Detection using Topic Modelling

We propose to use topical modelling to detect the hateful forms. We believe the latent topics present are capable of encompassing the hateful forms. Specially, we use LDA (Blei, Ng, and Jordan 2003) to detect the latent topics present in tweets.

Due to tweets being short in length and large in number, scaling of LDA to detect topics where every tweet is treated as one document is very challenging. Therefore, we create one document per user by concatenating all his posts, which includes tweets, retweets and quotes. Let a use u_i have made posts P , where $P = p_1, p_2, ..., p_N$. Then, the document d_i for user u_i is created by concatenating all the tweets in one document. Therefore, we have:

$$d_i = \cup_{(\forall p_j \in P)} p_j \tag{13}$$

Let $D = (\forall i \in 1..n) d_i$ be the corpus of documents. We further investigate D to detect latent topics present in it using LDA based techniques. We explore two variants of sampling for LDA, a) variational Bayes sampling method and b) Gibbs Sampling. Let the set of latent topics is T, where $T = t_1, t_2, ..., t_n$. The latent topic modelling produces a vector of topic affinity scores v_{d_i} for each document d_i. Let a_i is the affinity scores with respect to topic t_i, then we have topic affinity vector as follows:

$$T_{d_i} =< a_1, a_2, a_3, ..., a_T > \tag{14}$$

Experiments

Experiments Overview

The purpose of the experiments is to investigate the following research questions (RQs) as discussed in :

- RQ1: Is homophily exhibited by the users generating hateful content?
- RQ2: How effective is the newly proposed familiarity metric in comparison to the existing?
- RQ3: Is homophily more prominent for particular forms of hate?

Experiments Settings

Experimental dataset

We use hate speech dataset provided by (Ribeiro et al. 2017). This dataset contains 200 most recent tweets of $100,386$ users, totaling to around $19M$ tweets. It also contains a retweet induced graph of the users. The retweet-induced graph is a directed graph $G = (V, E)$ where each node $u \in V$ represents a user in Twitter, and each edge $(u1, u2) \in E$ represents a retweet in the network, where the user $u1$ has retweeted user $u2$. Furthermore, every tweet is categorized as an original tweet, retweet, or quote (retweet with a comment). Out of the $100,386$ users, labels (hateful or normal) are available for $4,972$ users, out of which 544 users are labeled as hateful and the rest as normal. The dataset does not have labels for the tweet content, therefore we manually annotate the tweets as hateful or not. Annotating $19M$ tweets is a very expensive and time-consuming process, hence we annotate a subset of tweets. We manually annotated $30,720$ tweets, where the users were picked based on their degree in the retweet network. The filtered retweet network has $7,71,401$ edges and $18,642$ nodes. The percentage of hateful tweets is 27.5 in the total annotated tweets. We employ pre-processing techniques suitable to tweets. We remove links, convert emoticons to text, and finally remove non ascii characters.

Parameter setting We use the source code provided by[1] to implement Variational Graph Auto-Encoder(VGAE). The adjacency matrix is constructed by using the retweet graph as an undirected graph. For training the VGAE, learning rate is set t0 0.01, number of epochs are 120, the batch size is the whole graph and Adam optimizer is used. The encoder portion of the trained VGAE is used as graph encoder for further experimentation.

To compute similarity between two users, we use universal sentence encoder (USE)[2] to encode a user's tweets. We create a document for a user by concatenating all his tweets and then use USE to encode the document. We use cosine similarity of the documents to get similarity score for a pair of users. Familiarity for a pair is computed using two existing and three newly proposed metrics. The two existing metrics are: edge exists or not and number of mutual friends between a pair of users. The three different novel ways of computing familiarity we implement are cosine similarity, Euclidean and Manhattan distance on the user vectors obtained from the graph encoder. We use scipy[3] to compute the similarity and familiarity metrics.

We create a document for each user by combining all the hateful posts (tweets, retweets and quotes). We create a corpus of all the documents of the users. We explored LDA based topic modelling using MALLET [4]. We perform grid search where we vary alpha between 0.1 and 0.01 and the number of topics from 6 to 12. The number of iterations for each run is 500. We look at the coherence scores (Röder, Both, and Hinneburg 2015) and visualization of topics in terms of overlap using pyLDAvis[5]. We find that the best performing topic model, which has both a high coherence score and the least number of overlapping topics, is when $\alpha = 0.1$ and number of topics equal to 8. We observe that $\alpha = 0.01$ gives us a higher coherence score for the same number of topics as compared to $\alpha = 0.1$. From this, we infer that a document in our corpus contains just a few topics.

Experiments Results

To answer RQ1 and RQ2, we plot similarity against familiarity for the three types of familiarity metrics in Figure 2. We vary the percentage of hateful tweets for a user and observe the corresponding homophily. We see that as the percentage of hateful tweets increases, similarity also increases. To measure the efficacy of our approach, we compare the Pearson correlation coefficient r for the three approaches.

[1]https://github.com/rusty1s/pytorch_geometric

[2]https://tfhub.dev/google/universal-sentence-encoder/4

[3]https://www.scipy.org/

[4]http://mallet.cs.umass.edu/

[5]https://pypi.org/project/pyLDAvis/

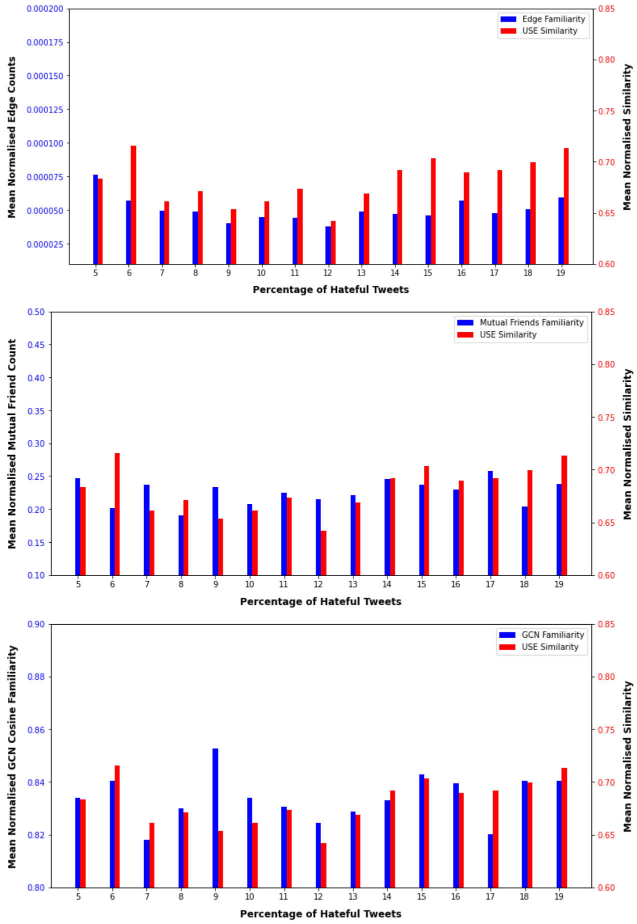

Figure 2: Variation in similarity and familiarity as hatefulness increases

We calculate that the value of r for our approach is 0.6, as compared to 0.5 for edge based familiarity; both indicating a moderate positive relationship. The r value for the mutual friends based familiarity is 0.2, which indicates a weak positive relationship. For the Euclidean and Manhattan distance metrics, as described in Equation 12 and Equation 11, the values of r are below 0.1, though greater than 0. Hence, we do not showcase these results and use the Cosine Similarity metric instead, as described in Equation 10.

To answer RQ3, we create a user base for each hateful form (topic). We pick users whose affinity score is above a certain threshold. We also rank the different hashtags used by users by frequency. This is shown in Table 1.

We observe that many users show higher values of association with specific topics (Topics - $0, 5, 7$), as compared to the rest. Therefore, we decide to have a dynamic threshold for topic affinity. To compute these affinity thresholds, we select users in such a way that there are a reasonable number (at least 10% of total users) of representative users. For each topic, we plot the average familiarity, using cosine similarity of graph embeddings, and average similarity in 3. The similarity and familiarity values are normalized by dividing

Table 1: Top Hashtags for the Hateful Topics

Topic	Hashtags
0	#maga, #trump, #realdonaldtrump, #trumptrain
1	#impeachtrump, #trump, #trumprussia, #jfkfiles
2	#bitch, #metoo, #harvey, #lockherup
3	#gobills, #pelicans, #mlscupplayoffs
4	#london, #fakenews, #cancer, #queen
5	#tormentedkashmir, #kashmirsuffering, #pakistan
6	#brexit, #crime, #terrorism, #illegal
7	#nigga, #bitch, #bitches, #somalia, #nigger

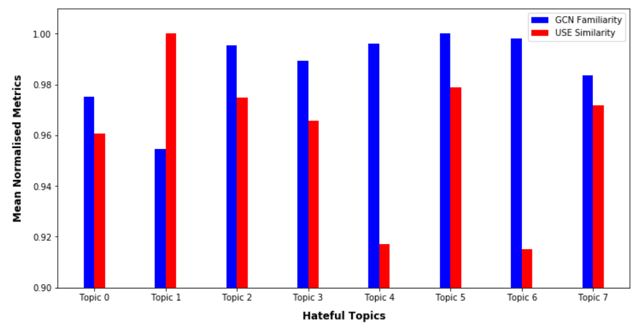

Figure 3: Homophily for the different hateful forms

by the maximum values, respectively. We observe that topics 2, 5 and 7 exhibit stronger homophily, as compared to others for both the communities. These topics can be broadly categorized into hate manifesting as sexism, nationalism and racism.

Conclusion

In this paper, we study homophily in hate speech on online social media platforms. We propose a novel way of computing familiarity using graph embeddings generated by variational graph auto-encoders. We empirically demonstrate that homophily is a significant driver in hate speech generation and also establish the effectiveness of graph embeddings in computing familiarity. Further, we observe the variation of homophily in different forms of hate. We report that there exists stronger homophily in users associated with certain hateful forms, as compared with others. These are racism, and xenophobia (nationalism). Therefore, we observe that homophily does play a role in the manifestation of hate speech. Further, we emphasize that the facet of homophily should be utilized for enhancing further research in this area.

References

Afrasiabi Rad, A.; and Benyoucef, M. 2014. Similarity and ties in social networks a study of the youtube social network. *Journal of Information Systems Applied Research* 7(4): 14.

Agrawal, R.; Arquam, M.; and Singh, A. 2020. Community detection in Networks using Graph Embedding. *Procedia Computer Science* 173: 372–381.

Aral, S.; Muchnik, L.; and Sundararajan, A. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences* 106(51): 21544–21549.

Blei, D.; Ng, A.; and Jordan, M. 2003. Latent dirichlet allocation. *the Journal of machine Learning research* 3: 993–1022.

Cai, H.; Zheng, V. W.; and Chang, K. C.-C. 2018. A comprehensive survey of graph embedding: Problems, techniques, and applications. *IEEE Transactions on Knowledge and Data Engineering* 30(9): 1616–1637.

De Choudhury, M.; Sundaram, H.; John, A.; Seligmann, D. D.; and Kelliher, A. 2010. "Birds of a Feather": Does User Homophily Impact Information Diffusion in Social Media? *arXiv preprint arXiv:1006.1702* .

Dey, K.; Shrivastava, R.; Kaushik, S.; and Garg, K. 2018. Assessing Topical Homophily on Twitter. In *International Conference on Complex Networks and their Applications*, 367–376. Springer.

Ducheneaut, N.; Yee, N.; Nickell, E.; and Moore, R. J. 2007. The life and death of online gaming communities: a look at guilds in world of warcraft. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 839–848.

Halberstam, Y.; and Knight, B. 2016. Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. *Journal of public economics* 143: 73–88.

Jeong, C.; Jang, S.; Shin, H.; Park, E.; and Choi, S. 2019. A context-aware citation recommendation model with BERT and graph convolutional networks. *arXiv preprint arXiv:1903.06464* .

Kipf, T. N.; and Welling, M. 2016. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308* .

Mathew, B.; Kumar, N.; Goyal, P.; Mukherjee, A.; et al. 2018. Analyzing the hate and counter speech accounts on Twitter. *arXiv preprint arXiv:1812.02712* .

McPherson, M.; Smith-Lovin, L.; and Cook, J. M. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27(1): 415–444.

Pan, S.; Hu, R.; Long, G.; Jiang, J.; Yao, L.; and Zhang, C. 2018. Adversarially regularized graph autoencoder for graph embedding. *arXiv preprint arXiv:1802.04407* .

Ribeiro, M.; Calais, P.; dos Santos, Y.; Almeida, V.; and Meira Jr, W. 2017. "Like Sheep Among Wolves": Characterizing Hateful Users on Twitter.

Röder, M.; Both, A.; and Hinneburg, A. 2015. Exploring the space of topic coherence measures. In *Proceedings of the eighth ACM international conference on Web search and data mining*, 399–408.

Starbird, K.; and Palen, L. 2012. (How) will the revolution be retweeted? Information diffusion and the 2011 Egyptian uprising. In *Proceedings of the acm 2012 conference on computer supported cooperative work*, 7–16.

Weng, J.; Lim, E.-P.; Jiang, J.; and He, Q. 2010. Twitterrank: finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on Web search and data mining*, 261–270.

Xu, S.; and Zhou, A. 2020. Hashtag homophily in twitter network: Examining a controversial cause-related marketing campaign. *Computers in Human Behavior* 102: 87–96.

Ying, J. J.-C.; Lu, E. H.-C.; Lee, W.-C.; Weng, T.-C.; and Tseng, V. S. 2010. Mining user similarity from semantic trajectories. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, 19–26.