Linking Friends in Social Networks using HashTag Attributes

Olga Gerasimova[1*][0000-0002-3598-7701] and Viktoriia Syomochkina[1]

HSE University, Moscow, Russia
ogerasimova@hse.ru

Abstract. Social networks are an integral part of modern life. They allow us to communicate online and exchange all kinds of information. In this paper, we consider the social network Instagram and its hashtags as a key tool for finding relevant information and new friends. The aim of our work is an empirical analysis of hashtags for posts in Instagram with certain locations. We obtain database of users of the Instagram network and collect a dataset of posts for three Far Eastern cities. Then, we build a friendship graph, for which we solve the link prediction problem. We show that both, structural and attributive graph information, such as hashtags, is important to achieve best quality.

Keywords: Social network · Link prediction · Hashtag.

1 Introduction

In our research, we study hashtags, their popularity and benefit of using them for finding new friends. We have chosen Instagram[1] as one of the most popular hashtag-sharing platforms. This social network is an American service launched in 2010 for uploading photos and videos, which can be organized with hashtags and geotags. To browse content by different tags, users can follow information enthralling them, check current trends or find their needs.

1.1 Hashtags

Concept of hashtags was created on Twitter[2] in 2007. A goal was to retrieve information as efficient as possible. To evolve in 2011, Instagram also began using hashtags to help users find specific posts, popular peoples, advertised services, and so on. The next step of Instagram progress was a way to follow not only personal pages, but also hashtags, which allows to show relevant topics in user's feeds. Nowadays, almost each Instagram post contains multiple hashtags, see Figure 1.

* The study has been funded within the framework of the HSE University Basic Research Program and the Russian Academic Excellence Project '5-100'.

[1] https://www.instagram.com/

[2] https://twitter.com

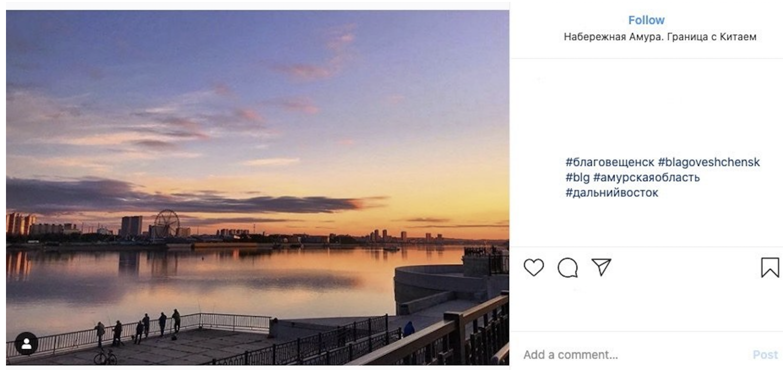

Fig. 1. Example of an Instagram post with hashtags.

Moreover, Instagram stimulates peoples to create specific hashtags, rather than using generic words, to diversify tendencies and draw attention to posts. As we said, Instagram users can create "trends" based on hashtags. The trends usually highlight a specific topic for posting information on. For example, hashtags can promote an election campaign, raise social issues or reflect food preferences.

1.2 Main contribution

Analyzing the conditions that cause people to subscribe to each other on social networks is a well-known task for developing recommendation systems. It is clear that people like to communicate with those who have similar interests. Hashtags can be considered as one of the forms of interests description. By hashtags, it is easy to search for posts with desired information, and, hence, for people who made these posts.

So, we are going to use hashtags for a link prediction task on friendship graphs based on data scraping from the Instagram.

The social network is represented by the graph $G(V, E)$, where V is a set of nodes corresponded people, and E is a set of edges - interactions between them. We obtained three friendship graphs for three Russian Far Eastern cities: Blagoveshchensk, Khabarovsk, Vladivostok. We are interested in comparison of users behavior in cities of different size, geographical, cultural and economical meanings in the same region.

We compared similarity-based and several machine learning methods for link prediction. As model features, we considered both models with only hashtags information and combined data (structural features and hashtags). We obtained that adding hashtags to structural features increases accuracy of prediction.

1.3 Organization

The structure of the paper is as follows. In Section 2, we consider works in related fields. Next, in Section 3, we describe methods of our datasets collection. Section 4 contains empirical analysis on hashtags. In Section 5, we concentrate on link prediction task on Instagram graphs using structural features and hashtag attributes. Finally, in Section 6 we make the conclusion and discuss future work.

2 Related work

We made an overview of related papers in several directions. Firstly, we considered hashtag-based works. Secondly, we focused on researches related with Instagram. Finally, we mentioned papers about link prediction and the approaches that use social interests and network structure to recommend social relations.

2.1 Researches on hashtags

In our research, we concentrate our attention on hashtags, which can be used as a powerful tool for different tasks.

There are many works about certain hashtags, when their usage creates a new tendency, supports a protest or motivates to do something.

In [24, 15, 23, 26], some events became a motivation for investigations. The work [24] is inspired by a hostage incident in Sydney in December 2014 and is devoted to solidarity towards Muslims in Australia, in which author explored the nature of Twitter networks based on the hashtag #*illridewithyou*. Authors of [15] analyzed an online movement by women being sexually harassed and sharing their stories with the hashtag #*metoo*. Another research focused on elections to the constituent assembly in Venezuela and classified tweets with the #*Maduro* hashtag. The authors identified the explicit and implicit topics related to this important political event [23]. In [26], content analysis of tweet tagged with #*rescue* in the 2017 North Kyushu Heavy Rain disaster was investigated. Using only hashtags on Twitter, authors of [11] predicted the election results in India.

There are researches focusing on emotions' analysis. For example in [30], an image classifier to define Non-Suicidal Self-Injury (NSSI) or non-NSSI images was developed for posts, which contain a hashtag #*selfharm*. Another case is devoted to explore how users conceptualize trust for understanding ideas that help people to negotiate trust. The researchers sourced the Twitter data for the hashtag #*trust* to analyze the vocabulary used alongside the given hashtag.

It is also interesting to explore data in the entertainment sectors (movie, music, food, fitness and etc.). Authors of [13] chose to analyze the Twitter hashtag #*avengersendgame* in framework of sentiment analysis task. Their approach aimed to classify the posts with a given hashtag into positive and negative type. In [32], the aim of the study is to analyze users' emotional states in terms of their musical preferences. Authors collected datasets with #*nowplaying* tweets and retrieved affective context hashtags included in the same tweets. As for food analysis, in [33] there is a study on gender differences in terms of using hashtags for photos tagged with #*Malaysianfood* or in [20] the hashtags #*fitness*, #*brag* and #*humblebrag* are examined for exploring self-presentation strategies.

So, analyzing hashtags is a hot topic that promotes to appear many new researches related with them.

Different recommender systems based on hashtags' semantic analysis were constructed [9]. This research focuses on the classification of semantic words using a user's hashtag data and co-occurrence hashtag information. Understanding the meaning of a hashtag is one of the ways to learn latent semantic expressions of words. In [27], it was proposed a semi-supervised sentiment hashtag embedding model, which preserves both semantic and sentiment hashtags distribution.

2.2 Researches on Instagram

There are a lot of various researches around Instagram, because it is a enormous data source for a wide range of tasks.

In [22], Müngen et al. have found most effective Instagram posts, instead of users, focused on fuse motif analysis. Authors of [4] suggested method of matching of Instagram images and topics obtained aid topic modelling of hashtags. In [8], another way based on the HITS algorithm and the principles of collective intelligence for matching of Instagram hashtags and corresponding visual content of the images was presented. This method allowed to obtain noise-free training datasets for content-based image retrieval.

Also hashtags-based method for specific community detection problem was proposed in [5].

There are various important studies on brand mentioning practice of influencers [31], in which authors suggested a neural network-based model for post classification according to (non-)sponsorship parameter.

Among political related researches, it is clear there are also investigation in Instagram. For example in [3], Zahra Aminolroaya and Ali Katanforoush proposed novel ideas of hashtag diffusion for Iranian communities in Instagram during the last legislative election in Iran.

In travel segment, a model-based location recommender system for designing a location's profile helps to recommend locations based on user preferences[21].

To sum up, we can identify the following task that are solved on Instagram data: automatic image annotation [4, 8], topic modelling [4], community detection [5, 3], user/post/image classification [31, 30], information diffusion modeling [3], sentiment analysis [13], recommender system [21], and others.

We focus on the structure of the user network that connect Instagramm users and their content as a core goal of our research.

2.3 Link prediction

In network analysis, one of the main tasks is a link prediction [18], the objective of which is to predict the pairs of nodes that will be connected by the link or not in the next state of the network. Being important task, the link prediction has become significant in various fields. Hence, many different methodologies to solve it have been suggested such as similarity-based [2], maximum likelihood models [7], probabilistic models [17], or based on deep learning [6].

There are applications of link prediction in most domains from classical task of prediction social relationships [34], web linking [1] to developing various recommender systems [28, 16].

Systematic surveys on link prediction methodologies were described in [14, 19, 29, 12, 25].

3 Data collection

3.1 Methods

To perform the link prediction task, we collected data from the Instagram social network based on the geotags of chosen locations.

After detailed study on extracting data from Instagram, the following solutions were proposed: (1) the Instagram API, (2) searching for posts by hashtags of given cities through a web page, (3) creating a bot.

As for the first method, there is a problem of missing important attribute information in the Instagram API, in particular, the location id attribute. This parameter has been removed since October 2019. The significant drawback of the second way is that it misses many posts without city hashtags, but tagged with the given locations. As a result, the third method was chosen, because it allows to solve our task on obtaining data.

To automatically collect posts, a data retrieval bot was written. The following technologies were chosen to implement the bot: Java, spring, springjpa, MySQL. The MySQL DBMS was used to store the database. Based on the dumped database, we obtained graphs in the form of edges lists and sets of users' hashtags.

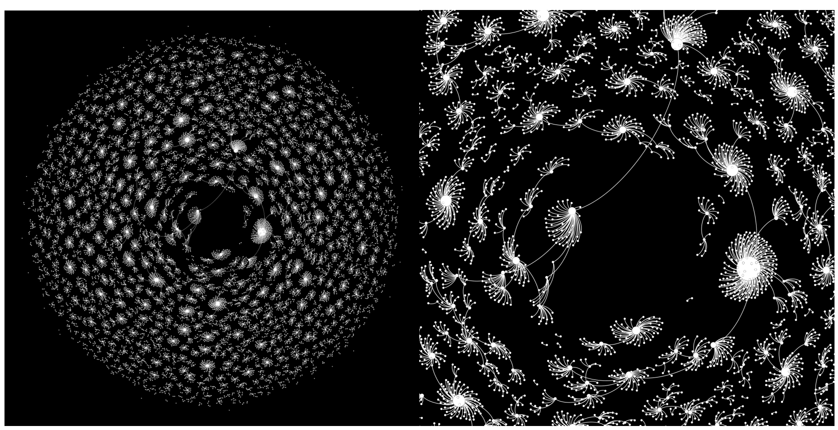

Fig. 2. Visualisation of the Blagoveshchensk graph.

3.2 Datasets

Graphs. We built three friendship graphs based on geotags for the following locations: Blagoveshchensk, Khabarovsk, Vladivostok. The Blagoveshchensk graph is illustrated via the Gephi[3] visualization platform, see the Figure 2.

We collected posts with the given locations. Next, we identified users who created these posts and built friendship graphs. We meant that two users are

[3] https://gephi.org

friends if they follow each other. We removed users with posts not containing any hashtags. Table 1 gives a summary statistics of all the three datasets.

Hahstags preprocessing. All hashtags were converted to lowercase. Also, we removed all the symbols except letters and numerals.

Table 1. Summary statistics of datasets.

	Blagoveshchensk	Khabarovsk	Vladivostok
Num. of users	15 262	36 807	51 430
Num. of links	14 119	34 118	47 523
Num. of hashtags	440 872	1 173 889	1 504 696
Num. of posts		2 695 509	

4 Empirical analysis

Top hashtags. We analyzed most popular hashtags in our three cities and presented our results in Table 2. We generalized and defined several topics of popular hashtags. Hashtags were translated from Russian except the Instagram class and were ordered by popularity from top to bottom within categories.

Table 2. Most popular hashtags in three datasets.

Topic	**Blagoveshchensk**	**Khabarovsk**	**Vladivostok**
Instagram	#instagood #repost #followme	#instakhv #instagood #repost	#instagood #repost #followme
General	#family #love #beautiful	#love #family #beautiful	#sea #love #beautiful
Locations	#blg #blagoveshchensk #blaga	#khv #khabarovsk #khv27	#vladivostok #vl #vdk
Seasons	#summer #spring #winter	#spring #summer #autumn	#summer #spring #autumn
Celebrations	#newyear #9may #birthday	#newyear #birthday #8march	#newyear #birthday #8march
Services	#manicureblagoveshchensk #nailsblagoveshchensk #gelpolishblagoveshchensk	#manicurekhabarovsk #manicure #photographerkhabarovsk	#photographervladivostok #manicure #manicurevladivostok
Quarantine	#stayhome #selfisolation #stayinghome	#stayhome #quarantine #selfisolation	#quarantine #stayhome #selfisolation

Because of the idea for data collection, many popular hashtags contain names of the cities. The current quarantine situation related with COVID-19 also produces many relevant tags. Also, general hashtags about seasons, celebrations or popular services are retrieved. We have noted several interesting patterns. While Vladivostok is a seaport, #*sea* is the most popular in this city. Different seasons preferences are identified. Among these cities, the celebration #*9may* is more often mentioned in Blagoveshchensk, possibly due to the fact that Blagoveshchensk is a border town. In addition, Vladivostok is also a tourism center, that is why there are more Instagram posts advertising #*photographer* services than #*manicure*, unlike other cities.

Distributions. We plot the distribution of the number of hashtags in each post for the united dataset of posts for three locations in Figure 3 (left). There are around $\sim 50\%$ posts included up to 10 hashtags. The upper limit for the number of hashtags per Instagram post is 30. So, we can see an increase of around 30, because there are many sponsored ad posts with the maximum number of tags. However, there are a few posts with exceeding limit of the number of hashtags.

Figure 3 (right) illustrates distributions of the number of users according to the number of shared hashtags for three graphs. It can be seen that most people use quite small set of hashtags. In the next Figure 4, we plots the distributions of the number of times hashtag was shared in all posts, i.e. we analyze the frequency of using hashtags. There are a lot of specific hashtags, which are popular among narrow circle of users. Also, there are general widespread hashtags, but their number is much smaller. It is clear and logical that both distributions look like the power law.

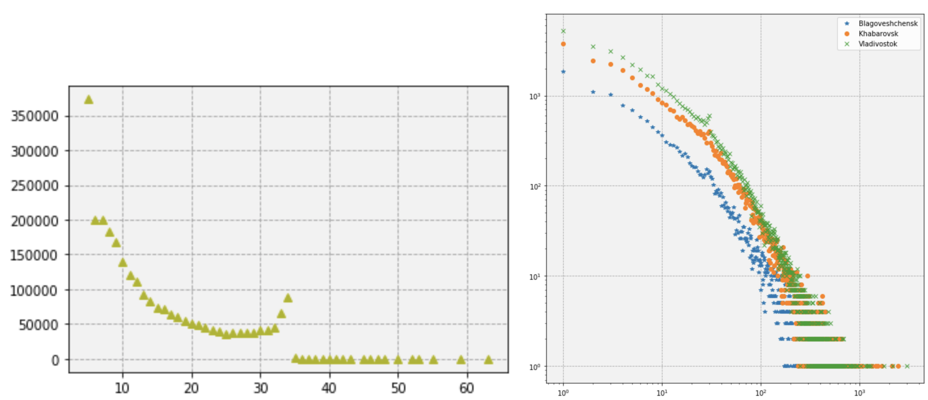

Fig. 3. Left: distribution of the number of hashtags in each post in all three datasets, x-axis corresponds to the number of hashtags and y-axis – number of posts. Right: distributions of the number of times a hashtag is used (loglog scale), x-axis corresponds to the number of users and y-axis – number of hashtags.

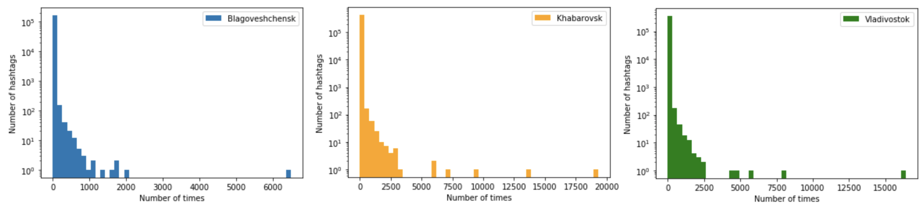

Fig. 4. Distributions of the number of times a hashtag is used (loglog scale).

5 Experiments

We consider the link prediction as a binary classification task: we predict unit for edges, which appear in the next state of the network, and zero for non-existent edges.

Training settings and metrics. To receive examples for the negative class, we applied the negative sampling strategy. We randomly choose non-existent edges such that the number of them would be approximately the same, as the number of existent edges for a balanced sample. Moreover, we aggregated results over five negative samplings. We split our data into training and test sets as 70% and 30%, respectively.

We use standard classification performance metrics for evaluating quality such as Accuracy (Acc.), F1-score (micro, macro), Log-Loss, ROC-AUC. The smaller Log-Loss is better, whereas the greater other metric is better.

Hashtag vectorization. The bag of words method was used to vectorize information about users' hashtags. We had more than 160, 350 and 424 thousand different hashtags for Blagoveshchensk, Khabarovsk, Vladivostok locations, respectively. According to the chosen vectorization method, the the vector dimension describing the user's hashtags should be equal to the number of different hashtags. It is reasonable to reduce the dimension choosing hashtags that were shared more than once, in other words, by more than one user. The resulting dimensions are approximately 40, 93 and 112 thousand hashtags, respectively. We can see that there are a lot of hashtags used by only one user in our datasets.

Features. In our problem statement, each user is characterized by a set of hashtags and connections with other users. So, we used two types of features: binary vectors describing used hashtags and binary vectors corresponding to a set of friends (rows from the graph adjacency matrix).

5.1 Similarity-based models

As a basic model for link prediction, we used Cosine Similarity (CS) and Jaccard Index (JI) metrics.

$$CS(v_1, v_2) = \frac{v_1 \cdot v_2}{\|v_1\| \|v_2\|} \qquad JI(v_1, v_2) = \frac{|v_1 \cap v_2|}{|v_1 \cup v_2|}$$

These metrics were calculated for the test set containing existent edges and non-connected pairs of nodes from negative sampling set. The threshold for prediction was obtained on training step. We used a grid search from 0 to 1 with a step of 0.0001. We predicted existence of edge if metric value was greater than corresponding threshold chosen for each metric.

We made three experiments to calculate the similarity metrics with different feature spaces: vectors with links data only, vectorized hashtags only, and combined users' features as concatenated binary vectors with vectorized hashtags and information about friends. We got a statistically significant increase in accuracy by $\sim$ 2-4 % adding information about hashtags to graph information, as shown in Table 3. We can see that prediction accuracy using only hashtags is greater than using graph information, which support our hypothesis that users of Instagram are linked by their interests. However, this approach did not show sufficiently high accuracy, that is why, we also looked at more advanced techniques using machine learning models.

Table 3. Accuracy on test data for similarity-based models.

	Model	Blagoveshchensk	Khabarovsk	Vladivostok
JI	friends	0.513	0.5	0.515
	tags	0.537	**0.534**	0.54
	friends+tags	**0.538**	0.533	**0.542**
CS	friends	0.499	0.5	0.5
	tags	0.537	0.536	0.538
	friends+tags	**0.538**	**0.539**	**0.54**

5.2 Machine learning models

We considered machine learning (ML) models for binary classification task such as Logistic Regression (LogReg) and Extreme Gradient Boosting (XGB).

As model features, we decided to use node2vec network embeddings [10] with random walks parameters $p, q = (1, 1)$, dimension of the embedding $d = 64$, length of walks $l = 30$, and number of walks per node equaled $n = 200$. Also, we made experiments with combining node2vec embeddings and vectorized hashtags.

Edge functions. To receive edge feature, we applied specific component-wise functions to node features for source u and target v nodes of a given edge. This model was suggested in [10], in which four functions for such edge embeddings were presented:

Average:	$\frac{u+v}{2}$	WeightedL_1:	$\|u-v\|$
Hadamard:	$u \cdot v$	WeightedL_2:	$(u-v)^2$

In Table 4 and Table 5, we presented values of quality metrics on test data for some experiments on the giant connected component (GCC) of the Blagoveshchensk graph. All results are significantly better than for similarity-based models.

Table 4. Comparison of ML models on GCC of the Blagoveshchensk graph, part 1.

		Average					Hadamard				
		Acc.	F1-micro	F1-macro	Log-loss	ROC-AUC	Acc.	F1-micro	F1-macro	Log-loss	ROC-AUC
node2vec	LogReg	0.745	0.745	0.735	8.82	0.745	**0.964**	**0.964**	**0.964**	**1.26**	**0.964**
	XGB	0.953	0.953	0.953	1.64	0.953	0.953	0.953	0.953	1.64	0.953
node2vec+#	LogReg	0.803	0.803	0.799	6.8	0.803	**0.964**	**0.964**	**0.964**	**1.26**	**0.964**
	XGB	**0.956**	**0.956**	**0.956**	**1.51**	**0.956**	0.953	0.953	0.953	1.64	0.953

Table 5. Comparison of ML models on GCC of the Blagoveshchensk graph, part 2.

		WeightedL_1					WeightedL_2				
		Acc.	F1-micro	F1-macro	Log-loss	ROC-AUC	Acc.	F1-micro	F1-macro	Log-loss	ROC-AUC
node2vec	LogReg	0.839	0.839	0.837	5.55	0.839	0.836	0.836	0.834	5.67	0.836
	XGB	0.843	0.843	0.839	5.42	0.843	0.843	0.843	0.839	5.42	0.843
node2vec+#	LogReg	**0.964**	**0.964**	**0.964**	**1.26**	**0.964**	**0.887**	**0.887**	**0.886**	**3.9**	**0.887**
	XGB	0.956	0.956	0.956	1.51	0.956	0.858	0.858	0.855	4.91	0.858

For features with node2vec embeddings only, XGB showed better quality in terms of all metrics for Average, WeigtedL_1 and WeigtedL_2 edge functions than LogReg. However, the situation is opposite for the Hadamard edge function. If we add features with vectorized hashtags, we obtain better results for LogReg with all edge functions except Average. Generally, combined features lead to higher quality, which supports our claim and core idea of the paper.

6 Conclusion and discussion

In the framework of this paper, we made a survey of different works on hashtags to show the motivation for hashtag-based researches, to present hashtags as a powerful tool for describing the preferences of social media users, influencing public opinion and disseminating information. While we focused on the Instagram, we also decided to describe a wide range of research opportunities based on this social network.

For our study, we created a database based on Instagram posts tagging with chosen locations. Constructing Instagram friendship graphs, we solved link prediction task using different feature types. In addition, we made general empirical analysis of hashtags to be aware of their usage trends at Russian Far Eastern cities.

For the future work, we aim to study another methods of graph embedding, which can contribute to link prediction problem from both, hashtag similarity and structural information feature engineering. We are also going to look at additional features related to social profiles of Instagram users for our problem of friend recommendations.

References

1. Adafre, S.F., de Rijke, M.: Discovering missing links in wikipedia. In: Proceedings of the 3rd international workshop on Link discovery. pp. 90–97 (2005)
2. Adamic, L.A., Adar, E.: Friends and neighbors on the web. Social networks **25**(3), 211–230 (2003)
3. Aminolroaya, Z., Katanforoush, A.: How iranian instagram users act for parliament election campaign? a study based on followee network. In: 2017 3th International Conference on Web Research (ICWR). pp. 1–6. IEEE (2017)
4. Argyrou, A., Giannoulakis, S., Tsapatsoulis, N.: Topic modelling on instagram hashtags: An alternative way to automatic image annotation? In: 2018 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP). pp. 61–67. IEEE (2018)
5. Bejandi, S.A., Katanforoush, A.: How unseen communities of instagram users are revealed using the real-valued collocations of hashtags. In: 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI). pp. 0487–0491. IEEE (2017)
6. Berg, R.v.d., Kipf, T.N., Welling, M.: Graph convolutional matrix completion. arXiv preprint arXiv:1706.02263 (2017)
7. Clauset, A., Moore, C., Newman, M.E.: Hierarchical structure and the prediction of missing links in networks. Nature **453**(7191), 98 (2008)

8. Giannoulakis, S., Tsapatsoulis, N.: Filtering instagram hashtags through crowdtagging and the hits algorithm. IEEE Transactions on Computational Social Systems **6**(3), 592–603 (2019)
9. Gorrab, A., Kboubi, F., Ghezala, H.B., Le Grand, B.: New hashtags' weighting schemes for hashtag and user recommendation on twitter. In: 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA). pp. 564–570. IEEE (2017)
10. Grover, A., Leskovec, J.: Node2vec: Scalable feature learning for networks. In: Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 855–864. KDD '16, ACM, New York, NY, USA (2016)
11. Gupta, S., Singh, A.K., Buduru, A.B., Kumaraguru, P.: Hashtags are (not) judgemental: The untold story of lok sabha elections 2019. arXiv preprint arXiv:1909.07151 (2019)
12. Haghani, S., Keyvanpour, M.R.: A systemic analysis of link prediction in social network. Artificial Intelligence Review **52**(3), 1961–1995 (2019)
13. Handayanto, R.T., Setiyadi, D., Retnoningsih, E., et al.: Corpus usage for sentiment analysis of a hashtag twitter. In: 2019 Fourth International Conference on Informatics and Computing (ICIC). pp. 1–5. IEEE (2019)
14. Hasan, M.A., Zaki, M.J.: A Survey of Link Prediction in Social Networks, pp. 243–275. Springer US, Boston, MA (2011)
15. Hassan, N., Mandal, M.K., Bhuiyan, M., Moitra, A., Ahmed, S.I.: Can women break the glass ceiling?: An analysis of #metoo hashtagged posts on twitter. In: 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). pp. 653–656. IEEE (2019)
16. He, Q., Pei, J., Kifer, D., Mitra, P., Giles, L.: Context-aware citation recommendation. In: Proceedings of the 19th International Conference on World Wide Web. pp. 421–430. WWW '10, ACM, New York, NY, USA (2010)
17. Heckerman, D., Meek, C., Koller, D.: Probabilistic entity-relationship models, prms, and plate models. Introduction to statistical relational learning pp. 201–238 (2007)
18. Liben-Nowell, D., Kleinberg, J.: The link-prediction problem for social networks. Journal of the Association for Information Science and Technology **58**(7), 1019–1031 (2007)
19. Lü, L., Zhou, T.: Link prediction in complex networks: A survey. Physica A: statistical mechanics and its applications **390**(6), 1150–1170 (2011)
20. Matley, D.: "this is not a #humblebrag, this is just a #brag": The pragmatics of self-praise, hashtags and politeness in instagram posts. Discourse, context & media **22**, 30–38 (2018)
21. Memarzadeh, M., Kamandi, A.: Model-based location recommender system using geotagged photos on instagram. In: 2020 6th International Conference on Web Research (ICWR). pp. 203–208. IEEE (2020)
22. Müngen, A.A., Kaya, M.: Quad motif-based influence analyse of posts in instagram. In: 2017 2nd International Conference on Advanced Information and Communication Technologies (AICT). pp. 51–55. IEEE (2017)
23. Niklander, S.: Content analysis on social networks: Exploring the #maduro hashtag. In: 2017 International Conference on Computing Networking and Informatics (ICCNI). pp. 1–4. IEEE (2017)
24. Rathnayake, C., Suthers, D.D.: Networked solidarity: An exploratory network perspective on twitter activity related to #illridewithyou. In: 2016 49th Hawaii International Conference on System Sciences (HICSS). pp. 2058–2067. IEEE (2016)

25. Raut, P., Khandelwal, H., Vyas, G.: A comparative study of classification algorithms for link prediction. In: 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA). pp. 479–483 (2020)
26. Sato, S.: Analysis of tweets hashtagged "#rescue" in the 2017 north kyushu heavy rain disaster in japan. In: 2018 5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM). pp. 1–7. IEEE (2018)
27. Singh, L.G., Anil, A., Singh, S.R.: She: Sentiment hashtag embedding through multitask learning. IEEE Transactions on Computational Social Systems **7**(2), 417–424 (2020)
28. Su, Z., Zheng, X., Ai, J., Shang, L., Shen, Y.: Link prediction in recommender systems with confidence measures. Chaos: An Interdisciplinary Journal of Nonlinear Science **29**(8), 083133 (2019)
29. Wang, P., Xu, B., Wu, Y., Zhou, X.: Link prediction in social networks: the state-of-the-art. Science China Information Sciences **58**(1), 1–38 (Jan 2015)
30. Xian, L., Vickers, S.D., Giordano, A.L., Lee, J., Kim, I.K., Ramaswamy, L.: #self-harm on instagram: Quantitative analysis and classification of non-suicidal self-injury. In: 2019 IEEE First International Conference on Cognitive Machine Intelligence (CogMI). pp. 61–70. IEEE (2019)
31. Yang, X., Kim, S., Sun, Y.: How do influencers mention brands in social media? sponsorship prediction of instagram posts. In: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. pp. 101–104 (2019)
32. Zangerle, E., Chen, C.M., Tsai, M.F., Yang, Y.H.: Leveraging affective hashtags for ranking music recommendations. IEEE Transactions on Affective Computing (2018)
33. Zhang, Y., Baghirov, F., Hashim, H., Murphy, J.: Gender and instagram hashtags: A study of #malaysianfood. In: Conference on Information and Communication Technologies in Tourism (2016)
34. Zhang, Y.: Language in our time: an empirical analysis of hashtags. In: The World Wide Web Conference. pp. 2378–2389 (2019)