



Empirical evaluation of link deletion methods for limiting information diffusion on social media

Shiori Furukawa¹ · Sho Tsugawa²

Received: 1 February 2022 / Revised: 24 October 2022 / Accepted: 25 October 2022 / Published online: 18 November 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

Although beneficial information abounds on social media, the dissemination of harmful information such as the so-called fake news has become a serious issue. Therefore, many researchers have devoted considerable effort to limiting the diffusion of harmful information. A promising approach to limiting diffusion of such information is link deletion methods in social networks. Link deletion methods have been shown to be effective in reducing the size of information diffusion cascades generated by synthetic models on a given social network. In this study, we evaluate the effectiveness of link deletion methods by using actual logs of retweet cascades, rather than by using synthetic diffusion models. Our results show that even after deleting 10–50% of links from a social network, the size of cascades after link deletion is estimated to be only 50% the original size under the optimistic estimation, which suggests that the effectiveness of the link deletion strategy for suppressing information diffusion is limited. Moreover, our results also show that there is a considerable number of cascades with many seed users, which renders link deletion methods inefficient.

Keywords Social network · Social media · Information diffusion · Link deletion

1 Introduction

An abundance of beneficial information is available on social media (Fan et al. 2013). Social media users can disseminate information posted by other users through functionalities such as retweeting on Twitter and sharing on Facebook (Doerr et al. 2012). Through such word-of-mouth information diffusion, information about important breaking news and newly created useful content is disseminated quickly and widely on social media. Because social media users can access such useful information, social media is considered to be an important platform for both those who produce information and those who consume it.

In contrast, the dissemination of harmful information, including the so-called fake news, misinformation, and hate speech, has become a serious issue (Kimura et al. 2008; Sharma et al. 2019; Mathew et al. 2019). For example, during the coronavirus disease 2019 (COVID-19) outbreak, there was widespread dissemination of misleading and false information on social media, resulting in people believing in false theories about the disease and the virus that causes it (Jolley and Paterson 2020). In addition, it has been reported that an increasing number of people are relying on misleading information on social media and as a result are hesitant to get vaccinated without first seeking professional advice (van der Linden et al. 2020).

Many researchers have devoted substantial effort to limiting the diffusion of harmful information on social media (Pérez-Rosas et al. 2018; Shah and Zaman 2011; Leskovec et al. 2009; Sharma et al. 2019). Pérez-Rosas et al. (2018) proposed a method to automatically detect fake news on social media. Shah and Zaman (2011) proposed a method to estimate the users who are the source of rumors on social media. Leskovec et al. (2009) proposed a framework for tracking the spread of fake news.

Methods to block information diffusion between (a small number of) specific users have been actively studied

✉ Shiori Furukawa
s.furu@mibel.cs.tsukuba.ac.jp

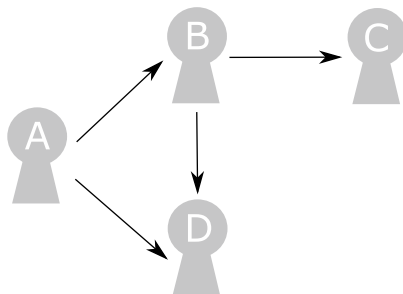
Sho Tsugawa
s-tugawa@cs.tsukuba.ac.jp

¹ Graduate School of Engineering Information and Systems,
University of Tsukuba, 1-1-1 Tennodai, Tsukuba,
Ibaraki 305-8573, Japan

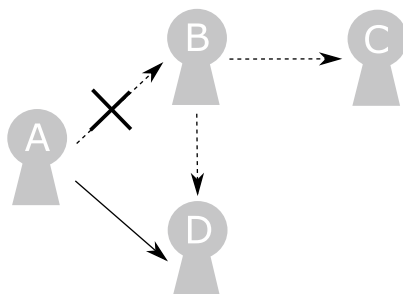
² Faculty of Engineering, Information and Systems, University
of Tsukuba, Tsukuba, Japan

as promising approaches to limiting harmful information diffusion (Yan et al. 2019; Kimura et al. 2008; Tong et al. 2012; Yao et al. 2015). Blocking one user's information from spreading to other users can reduce the final size of information diffusion cascades. Figure 1 shows an example of blocking information diffusion between users. As shown in Fig. 1, by blocking information diffusion between only users A and B, the number of users who finally receive the information will decrease from 4 to 2. However, if information diffusion is blocked too heavily, there is a risk of blocking diffusion of useful information. Therefore, it is desirable to control the diffusion of harmful information by blocking the diffusion of information between as few users as possible.

The problem with limiting the size of information diffusion cascades by blocking information diffusion between a limited number of users is formulated as a link deletion problem in social networks (Kimura et al. 2008). A social network is defined as a graph $G = (V, E)$, in which social media users are represented by nodes and relationships among them are represented by links (Domingos and Richardson 2001). Information posted by nodes in a social network spreads through links between nodes. Given a social network G and an integer k , the link deletion problem aims to determine the combination of k links that should



(a) Information flows before blocking information diffusion. Information posted by user A spreads to user A's followers, namely, users B and D. The information is reposted by user B, then the information spreads to users C and D.



(b) Information flows after blocking information diffusion. When information diffusion from user A to user B is blocked, user B cannot repost the information. As a result, user C cannot receive information from user B.

Fig. 1 Example of blocking information diffusion between users

be removed in order to minimize the size of information diffusion cascades on social network G . Many studies have proposed various link deletion methods to control the size of information diffusion cascades (Kimura et al. 2008; Tong et al. 2012; Khalil et al. 2014; Yan et al. 2019). Previous studies have used synthetic information diffusion models to evaluate the effectiveness of the link deletion methods (Tong et al. 2012; Yan et al. 2019; Zareie and Sakellariou 2022). For example, Tong et al. (2012) evaluated the effectiveness of link deletion methods in reducing the size of information diffusion using a popular synthetic diffusion model called the susceptible-infectious-susceptible (SIS) model.

In this study, we aim to reveal the benefits and limitations of link deletion in limiting the spread of information diffusion on *real* social media. Contrary to previous studies that used synthetic models, we utilize the actual log data to evaluate the effectiveness of link deletion methods. Synthetic information diffusion models cannot fully reproduce the characteristics of information diffusion on actual social media (Tripathy et al. 2010). Therefore, even if the link deletion method is effective in controlling the size of information diffusion generated by the synthetic information diffusion model, its effectiveness in controlling actual information diffusion on social media remains unclear. We examine the size of information diffusion cascades when it is possible to block information diffusion between a certain number of user pairs identified using link deletion methods. Because it is not easy to conduct an experiment that actually blocks information diffusion, we propose a method to estimate the size of retweet diffusion cascades after link deletion from a who-follows-whom network users and log data of information diffusion among them. Based on the proposed method, we examine the effectiveness of link deletion methods including NetMelt (Tong et al. 2012), a popular link deletion method, and other heuristics for limiting the size of information diffusion cascades on social media (Kashyap and Ambika 2019).

This paper is an extended version of our previous conference short paper (Furukawa and Tsugawa 2021). We have conducted experiments on two additional datasets, and two additional link deletion methods for comprehensive examination of the effectiveness of link deletion methods. Moreover, we propose a new method to estimate the size of diffusion cascade after link deletion. In our previous paper, we only evaluated the effectiveness of link deletion under the worst-case scenario. In this paper, we evaluate the effectiveness of link deletion both under the worst-case and more optimistic scenarios.

The contributions of the present study are summarized as follows.

- We use four large-scale datasets of cascades rather than synthetic cascade models to evaluate the effectiveness of

link deletion. We propose a method to estimate the size of a given cascade after link deletion. Using the proposed method, we examine the benefits and limitations of link deletion methods when applying them to actual social media.

- Our results suggest that there are limitations to link deletion methods in limiting the size of retweet cascades on social media. While existing studies (Tong et al. 2012; Yan et al. 2019) have demonstrated the effectiveness of link deletion methods in limiting the size of cascades generated by synthetic models, our results suggest that link deletion methods are not so effective in limiting the size of real diffusion cascades on social media both under the pessimistic and optimistic scenarios.
- We show that most of the cascades examined in this study have multiple sources, which makes limiting their sizes by link deletion difficult.

The remainder of this paper is organized as follows. In Sect. 2, we introduce related work. In Sect. 3, we propose a method to estimate the size of cascades when link deletion methods are applied. In Sect. 4, we explain the datasets and methodologies for the analyses. In Sect. 5, we present the results and discuss their implications. Finally, in Sect. 6, we summarize this study and discuss future work.

2 Related work

Many researchers have conducted research to detect harmful information on social media (Pérez-Rosas et al. 2018; Shah and Zaman 2011; Liu et al. 2020; Helmstetter and Paulheim 2018). Pérez-Rosas et al. (2018) proposed a method to automatically detect fake news on social media. They analyzed actual fake news texts and developed a model to predict whether a given news text is fake news by combining information on vocabulary, syntax, and semantics that are frequently used in actual fake news. Ghanem et al. (2020) proposed an approach to detect non-factual twitter accounts. Shah and Zaman (2011) proposed a method to estimate the source of rumors using maximum likelihood estimation. To control the spread of harmful information, both detecting harmful information and limiting its spread are required (Sharma et al. 2019). While the above mentioned studies focused on the former, the present study focuses on the latter.

To limit the spread of information diffusion on a social network, several studies have proposed methods to change the network topology of the given network so that the size of the spread of information diffusion cascades on the network are minimized (Khalil et al. 2014; Kimura et al. 2008). Some of these studies have proposed removing several nodes from the network (Wang et al. 2013; Alorainy et al. 2022). In

particular, previous studies have shown that removing high centrality nodes is an effective strategy (Kitsak et al. 2010). However, deleting high centrality nodes in the network means deleting influential users on social media. Therefore, such methods are considered to be costly (Zareie and Sakellariou 2021) and may even have a considerably negative effect because famous celebrities may be removed.

Deleting a certain number of links in a social network is another approach to limiting the spread of information diffusion (Tong et al. 2012; Yan et al. 2019; Zareie and Sakellariou 2022; Yao et al. 2015; Kanwar et al. 2022). Here, deleting a link between two nodes means blocking information diffusion between them. Tong et al. (2012) addressed the NetMelt problem of selecting links for deletion in order to minimize the spread of rumors. Because the largest eigenvalue of the adjacency matrix of a social network graph is positively correlated with the size of information diffusion on the network, they proposed an algorithm to remove k links so that the largest eigenvalue of the network is minimized. Yan et al. (2019) proposed a rumor spread minimization (RSM) problem and its algorithms. RSM algorithms aim to identify a set of links so that the size of the information diffusion cascade from given source nodes is minimized. In contrast, Zareie and Sakellariou (2022) proposed a source-ignorant method to identify a set of critical links for limiting the sizes of information diffusion cascades from arbitrary seed nodes. These studies used synthetic information diffusion models and demonstrated the potential of link deletion methods in limiting the size of information diffusion. Our study follows and extends these studies, evaluating the effectiveness of link deletion methods by using actual logs of retweet cascades on social media rather than using synthetic diffusion models.

3 Method of estimating size of diffusion cascades after link deletion

We propose a method to estimate the size of a given information diffusion cascade when deleting several links in the social network. Notations used in this paper are summarized in Table 1. Let T be a set of tweets or topics posted on social media during a certain period of time, and let L be a set of links to be deleted. Here, deleting link $(u, v) \in L$ means blocking information diffusion from user u to user v . We consider a method to estimate the size of each diffusion cascade regarding tweet or topic $t \in T$ after deleting links in L . Although each t can be either a specific tweet or a specific topic, from here on, we use tweet t rather than tweet or topic t for simplicity. We denote the set of users who post or repost tweet t as R_t , and the set of users involved in tweet set T as $V = \bigcup_{t \in T} R_t$. The time at which user $u \in V$ posts t or repost t is denoted as $\tau(u, t)$, and the social network that expresses

Table 1 Notations used in this paper

Notation	Description
$G = (V, E)$	Social network with node set V and link set E
T	Set of tweets or topics during a certain period
R_t	Set of users involved in tweet or topic t
$\tau(u, t)$	Time when user u posts a tweet or retweet about t
$(u, v) \in E$	Link representing user u following user v
$H_t = (R_t, E_t)$	Diffusion graph with node set R_t and link set E_t
E_t	Set of links in diffusion graph H_t
L	Set of links to be deleted
k	Number of links to be deleted
H'_t	Diffusion graph after link deletion
E'_t	Set of links after link deletion in diffusion graph H'_t

who-follows-whom relationships among users belonging to user set V is denoted as $G = (V, E)$. The link $(u, v) \in E$ represents user u following user v .

First, for each tweet $t \in T$, we construct a diffusion graph $H_t = (R_t, E_t)$ that represents the diffusion paths of tweet t . The set of nodes in H_t is equal to R_t , which is the set of users who tweeted or retweeted tweet t . The link $(u, v) \in E_t$ in H_t represents tweet t spreading from user u to user v . To obtain E_t is not obvious because each cascade does not contain information about the specific diffusion paths of the tweet. Here, we construct three diffusion graphs, one of which is non-tree diffusion graph (i.e., each node can have multiple parents), and the others are trees (i.e., each node has at most a single parent).

In the non-tree diffusion graph, we consider a tweet to have spread from user u to user v when user v is following user u and the timing of user u 's retweeted (or tweeted) tweet t is earlier than that of user v . In other words, if $(v, u) \in E$ and $\tau(u, t) < \tau(v, t)$, then $(u, v) \in E_t$.

While a user can be affected by multiple users in the non-tree diffusion graph, a user is affected by a only single user in the diffusion trees. We construct two-types of diffusion trees, which we call diffusion tree (first) and diffusion tree (last). In the diffusion tree (first), we consider a tweet to have spread from user u to user v when user v is following user u and the timing of user u 's retweeted tweet t is earlier than that of user v and the earliest among user v 's followees. In contrast, in the diffusion tree (last) we consider a tweet to have spread from user u to user v when user v is following user u and the timing of user u 's retweeted tweet t is earlier than that of user v and the latest among user v 's followees.

We call the node that does not have any incoming links in H_t the seed user of tweet t . The user who posted tweet t is the seed user. User u who retweeted tweet t is also a seed user if none of their followees retweeted tweet t before user u retweeted tweet t . We denote the set of seed users of tweet t as S_t .

Next, we use the link deletion method to select a set of links L ($|L| = k$) to be deleted and construct a diffusion graph after link deletion. By removing these links from diffusion graph H_t , we can estimate how the number of users who receive each tweet $t \in T$ changes. We remove the selected link set L from H_t . Namely, we construct $H'_t = (R_t, E'_t)$, where $E'_t = E_t \setminus L$. Deleting link (u, v) corresponds to blocking information propagation from user u to user v . Therefore, if user v receives tweet t from only user u , that is, if node v has no incoming link in H'_t , then deleting link (u, v) will prevent user v from receiving tweet t .

Finally, we use diffusion graph H'_t to find the number of users who receive tweet t after the link deletion. In graph H'_t , seed users can receive tweet t even after the link deletion. Users who follow the seed users can also receive this information. Furthermore, users who follow those users (i.e., the users following the seed users) can also receive the information. Therefore, we assume that all users within reach of the seed users in diffusion graph H'_t can receive tweet t . Then, we estimate the number of users who receive tweet t after link deletion as the number of nodes that are within the reach of the seed nodes $s \in S_t$ in diffusion graph H'_t .

Note that there is a only single diffusion path for a given node in the diffusion trees whereas there are multiple diffusion paths in the non-tree diffusion graph. Hence, when using diffusion trees, the estimated sizes of diffusion cascades after link deletion are smaller than those when using non-tree diffusion graphs. Using the non-tree diffusion graph and diffusion trees, we examine the effectiveness of link deletion both under the pessimistic and optimistic scenarios.

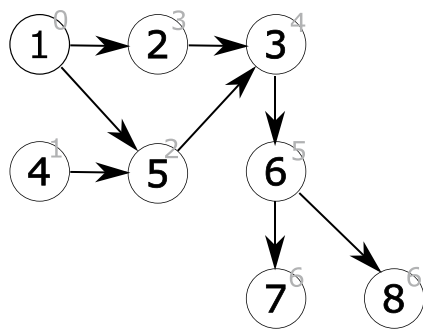
An example of non-tree diffusion graph $H_t = (R_t, E_t)$ of tweet t is shown in Fig. 2a. We consider the case where the set of links to be deleted is $L = \{(1, 5), (3, 6)\}$. Graph H'_t after deleting these links from H_t is shown in Fig. 2b.

4 Dataset and methodology

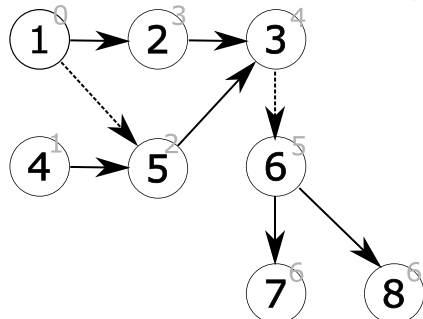
4.1 Datasets

To evaluate the effectiveness of link deletion in limiting the size of cascades, we use four datasets containing both cascades and social networks representing who-follows-whom relationships among users involved in the cascades. A statistical summary of the datasets is shown in Table 2. The datasets are explained in detail as follows.

1. Ordinary dataset (Tsugawa 2019) This dataset contains 10,000 tweets and their retweets posted during November 19–25, 2018. The 10,000 tweets were randomly selected from a collection of English tweets with one or more retweets posted during the above period. Note that there were no specific events during the period. In total,



(a) Example of non-tree diffusion graph $H_t = (R_t, E_t)$ of tweet t . The set of users who have tweeted or retweeted tweet t is $R_t = \{1, 2, 3, 4, 5, 6, 7, 8\}$, and the set of links is $E_t = \{(1, 2), (1, 5), (2, 3), (4, 5), (5, 3), (3, 6), (6, 7), (6, 8)\}$. The seed node set that does not have any incoming links is $S_t = \{1, 4\}$.



(b) Graph H'_t after deleting links $L = \{(1, 5), (3, 6)\}$ from H_t . In H'_t , node 6 has no incoming links, and therefore user 6 will not receive tweet t that was originally tweeted by users 1 and 4. Moreover, users 7 and 8 who received tweet t from user 6 also will not receive tweet t . The nodes within the reach of the seed node set $S_t = \{1, 4\}$ are $\{1, 2, 3, 4, 5\}$, which is the set of users who will receive tweet t after the link deletion.

Fig. 2 Examples of original diffusion graph H_t and the diffusion graph after link deletion H'_t . Numbers on the nodes indicate the time at which the user posted a tweet or retweet regarding t . Link (1, 2) indicates that user 2 is following user 1, and the time at which user 1 posted a tweet or retweet regarding t was earlier than the time at which user 2 did

114,249 users tweeted or retweeted these tweets, and these users are target user set V . This dataset also contains the social network of who-follows-whom relationships among the target users. The social network is used to construct the diffusion graph and to determine links to be deleted. To examine the effectiveness of link deletion in limiting the size of retweet cascades, tweets that have a large cascade are preferable for analysis. There-

fore, among the 10,000 tweets, we extracted those that were retweeted 100 times or more, which resulted in 90 original tweets. These 90 tweets that were retweeted 100 times or more are the set of tweets T that we analyzed.

2. Higgs dataset (Domenico et al. 2013) This dataset contains tweets and retweets about the discovery of the Higgs boson in 2012. In this dataset, the discovery of the Higgs boson is regarded as topic t . That is, all sets of tweets and retweets in the dataset are regarded as a single cascade of information diffusion. Therefore, the number of topics to be analyzed $T = 1$ in this dataset. The social network representing who-follows-whom relationships among users who tweet and retweet is available in this dataset and is used to construct the diffusion graph and to determine links to be deleted.
3. URL dataset (Hodas and Lerman 2014) This dataset contains tweets containing URLs and retweets during October 2010. Each URL is considered as a topic propagated among users. As with the ordinary dataset, we extract tweets that were retweeted 100 times or more.
4. Douban dataset (Zhong et al. 2012) This dataset contains data from a Chinese social website. This website allows users to post their book reading status and also allows them to follow the status of other users. Each book is considered as a topic propagated among users. As with the ordinary dataset, we extract cascades whose size were 100 or more.

4.2 Link deletion methods

In this study, we focus on the effectiveness of three popular link deletion methods, NetMelt (Tong et al. 2012), Betweenness (Kashyap and Ambika 2019), and Edge-Degree (Kashyap and Ambika 2019) in limiting the size of actual diffusion cascades. We also use a random baseline. The details of the link deletion methods are explained as follows.

1. NetMelt (Tong et al. 2012) We apply the NetMelt to the social network G representing who-follows-whom relationships among the target users and obtain $|L| = k$ links to be deleted.

Table 2 Statistics of the datasets used in the analysis

	Ordinary	Higgs	Douban	URL
Number of target users V	118,162	456,626	25,306	12,627
Number of links E	3,130,963	14,855,842	758,310	619,262
Number of target tweets or topics T	90	1	439	253
Average cascade size	830.4	228,556	272.6	246.5

2. Betweenness (Kashyap and Ambika 2019) Betweenness selects $|L| = k$ links to be deleted in descending order of link betweenness of the links in social network G .
3. Edge Degree (Kashyap and Ambika 2019) The edge degree of link (u, v) is defined as the product of in-degree of node u and out-degree of node v (Kashyap and Ambika 2019). This method selects $|L| = k$ links to be deleted in descending order of edge degree of the links in social network G .
4. Random Random selects $|L| = k$ links to be deleted randomly from all links in social network G .

Using these methods, we determine $|L| = k$ links to be deleted and estimate the size of cascade $t \in T$ after link deletion.

5 Results and discussion

First, we analyze how the total size of the cascades in the four datasets changes by link deletion. Figure 3 shows the estimated total size of retweet cascades against the number of deleted links in each dataset when using non-tree diffusion graph. From Fig. 3, we can see that although the total cascade size decreases by link deletion, the decrease is not large relative to the number of deleted links. For example, in the case of ordinary data, when 1.5 million links are deleted from the social network using NetMelt, the total cascade size is about 50% that of the original cascade. 1.5 million links are equivalent to approximately 50% of the links in the original social network. It is practically difficult to block information diffusion between such a large number of user pairs. Therefore, this result suggests that the effects of link

deletion methods to limit the size of tweet diffusion cascades are limited. Looking at the difference among link deletion methods, the NetMelt and the Betweenness are generally effective. In particular, except for the URL dataset, the NetMelt achieves the best results in the three of four datasets.

We next examine the effectiveness of link deletion under more optimistic scenarios by using diffusion trees. Figures 4 and 5 compare the total size of cascades when using non-tree diffusion graph, diffusion tree (first), and diffusion tree (last). Figure 4 shows the results for the NetMelt, and Fig. 5 shows the results for the Betweenness. These results show that although the estimated sizes of cascades when using the diffusion trees are smaller than those when using the non-tree diffusion graph, the estimated sizes of cascades are still large. Even under the best case for each dataset, 10%–50%

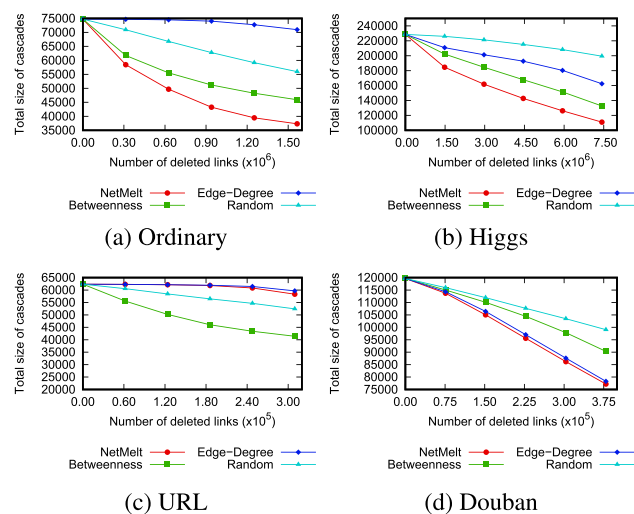


Fig. 3 Total size of cascades after link deletion versus the number of deleted links (non-tree diffusion graph)

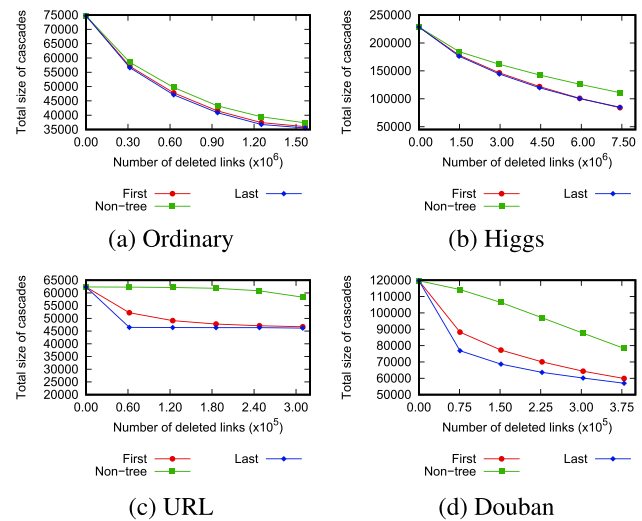


Fig. 4 Comparison among different diffusion graphs (NetMelt)

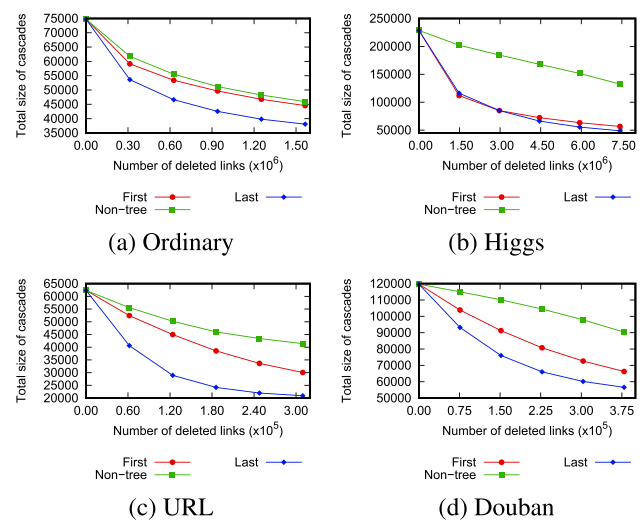


Fig. 5 Comparison among different diffusion graphs (Betweenness)

links are necessary to be deleted for limiting the estimated diffusion sizes to be under 50% of the original sizes. These results again suggest the limitations of link deletion methods for limiting the sizes of real diffusion cascades. It is suggested that even under the optimistic estimation, many links should be deleted to reduce the cascade sizes.

We next analyze how the link deletion affects the size of each cascade. Because the Higgs dataset contains only a single cascade, the other three datasets are used in the following analyses. Figures 6 and 7 show the relation between the original cascade size and the estimated cascade size after link deletion for each tweet when using the NetMelt and the Betweenness, respectively. The number of deleted links is 50% of the original links, and the diffusion tree (last) is used as the diffusion graph. These results show that the cascade size of some tweets is substantially decreased by link deletion. However, we should note that this result is obtained when 50% of links are deleted. Considering the fact that a large number of links are deleted, this result also suggests that the effect of link deletion on information diffusion is limited.

Finally, we examine the reason why the link deletion does not substantially reduce the size of diffusion cascades. Figure 8 shows the relationship between the original cascade size and the number of seed users of each tweet. For the sake of visibility, we only show the tweets with less than 1000 retweets. These figures show that many tweets have a large ratio of seed users relative to the total cascade size. A seed user is a user who retweets a tweet without being influenced by their followers. Twitter users may obtain information about tweets from sources other

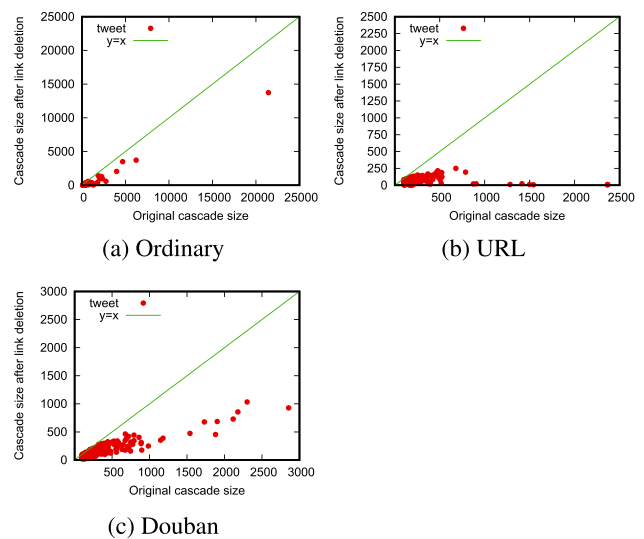


Fig. 7 Relation between the original cascade size and the cascade size after link deletion for each tweet (link deletion method: Betweenness, diffusion graph: diffusion tree (last), number of deleted links $|L| = 0.5|E|$)

than their followers. For example, if a tweet is *trending* on Twitter, or if it is introduced on another web page, Twitter users may obtain information about the tweet not from their social relationships on Twitter. Link deletion only limits information diffusion through social relationships and cannot prevent seed users from obtaining information from other sources. If a tweet spreads through only social relationships among users, link deletion should be effective in limiting its cascade size. However, if many users obtain information about a tweet from sources other than their social

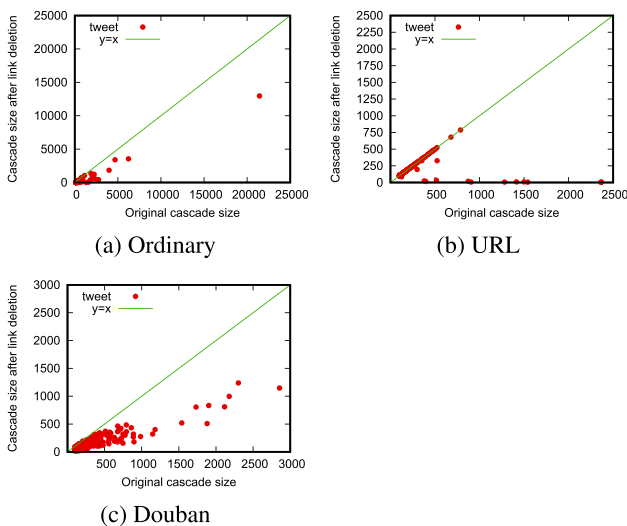


Fig. 6 Relation between the original cascade size and the cascade size after link deletion for each tweet (link deletion method: NetMelt, diffusion graph: diffusion tree (last), number of deleted links $|L| = 0.5|E|$)

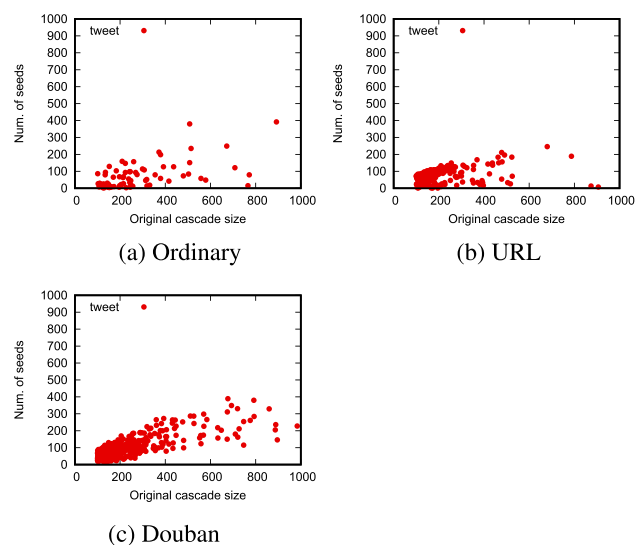


Fig. 8 Relation between the original cascade size and the number of seed nodes for each tweet (tweets with less than 1000 retweets)

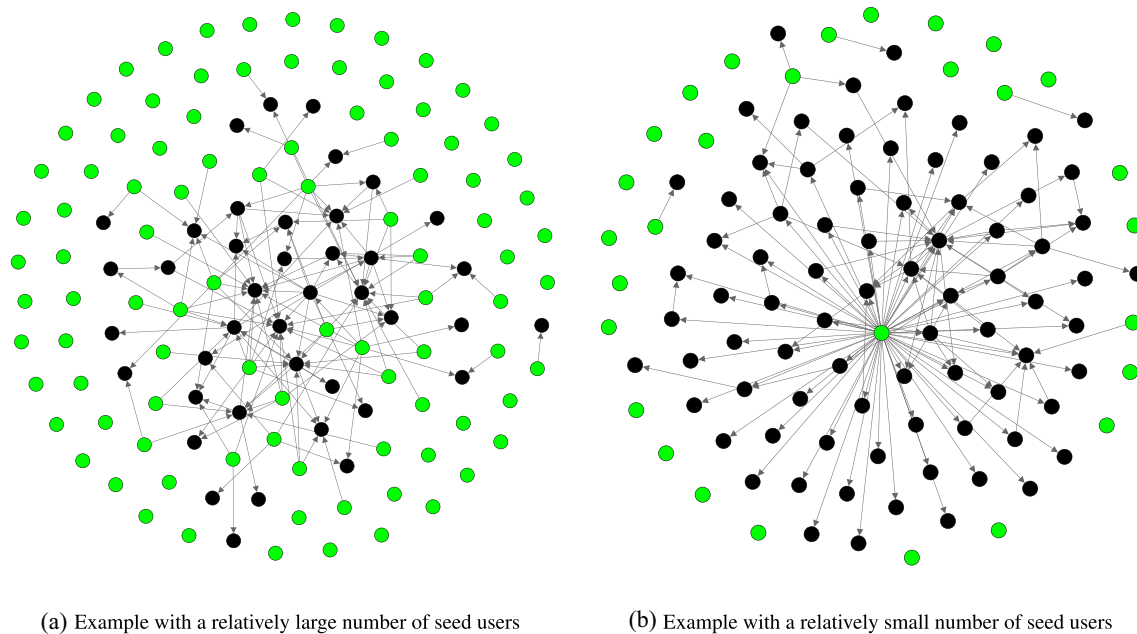


Fig. 9 Examples of non-tree diffusion graph (ordinary dataset): Light green nodes represent seed users

relationships, link deletion may not be effective in limiting its cascade size.

From the examples of diffusion graphs shown in Fig. 9, we can also confirm that limiting the cascade sizes by removing a small number of links is difficult. Seed users are represented as light green nodes in the figure. If seed users will be isolated from other users in the diffusion graph, the cascade size can be reduced. However, in the example of Fig. 9a, many seed users involve in the tweet diffusion, and isolating these seed users requires many links to be deleted. Also for the example of Fig. 9b, a seed user shown in the center of the figure is connected to many other users, and isolating the seed user requires many links to be deleted.

In summary, our results suggest that link deletion methods are limited in efficiently limiting the size of diffusion cascades. NetMelt has been shown to be effective in limiting the size of synthetic cascades in (Tong et al. 2012) and requires many links (e.g., 50% of links) to be deleted to limit the size of real tweet diffusion cascades. In practice, it is difficult to block information diffusion between such a large number of user pairs on social media, which suggests that link deletion on social media is limited in practice. Our results also suggest that the existence of many seed users is one of the reasons for the inefficiency of link deletion methods in actual social media. We show that the information cascades analyzed in this paper have many seed users who may receive the information, not from their social relationships, but from other sources. Link deletion methods cannot prevent seed users from receiving information, so other strategies (e.g., advertising counter campaigns (Budak et al.

2011; Erd et al. Erd et al.)) are suggested as necessary methods for limiting the size of diffusion cascades with many seed users. Moreover, we recognize that other link deletion methods that delete links by using the information about the seed nodes are available in the literature (Yan et al. 2019), and their effectiveness should be evaluated using actual diffusion cascade logs in future research.

6 Conclusion and future work

In this study, we used actual logs of retweet cascades and evaluated the effectiveness of link deletion methods to suppress the spread of information. We proposed a method to estimate the size of a given retweet cascade after link deletion and examined the potential and limitations of link deletion methods when applying them to actual social media. Our results suggest that link deletion methods are limited in limiting the size of retweet cascades on social media. We showed that most of the retweet cascades have multiple sources, which makes limiting their size by link deletion difficult.

In future work, we plan to evaluate the effectiveness of link deletion methods that incorporate the information of seed nodes (Yan et al. 2019). In addition, validating the proposed estimation method of cascade size is also an important future work.

Acknowledgments This work was partly supported by JSPS KAKENHI Grant Number 19K11917.

References

- Alorainy W, Burnap P, Liu H, Williams M, Giommoni L (2022) Disrupting networks of hate: characterising hateful networks and removing critical nodes. *Social Netw Anal Mining* 12(1):27
- Budak C, Agrawal D, El Abbadi A (2011) Limiting the spread of misinformation in social networks. In: *Proceedings of the 20th International Conference on World Wide Web (WWW'10)*, pp 665–674
- Doerr Benjamin, Fouz Mahmoud, Friedrich Tobias (2012) Why rumors spread so quickly in social networks. *Commun ACM* 55(6):70–75
- Domenico D, Manlio Lima, Antonio Paul M, Musolesi Mirco (2013) The anatomy of a scientific rumor. *Sci Rep* 3:2980
- Domingos P, Richardson M (2001) Mining the network value of customers. In: *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'01)*, p 57–66
- Erd FC, Vignatti AL, da Silva MV (2020) Blocking the spread of misinformation in a network under distinct cost models. In: *Proceedings of the 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'20)*, pp 232–236
- Fan L, Lu Z, Wu W, Thuraisingham B, Ma H, Bi Y (2013) Least cost rumor blocking in social networks. In: *Proceedings of the 33rd IEEE International Conference on Distributed Computing Systems (ICDCS'13)*, pp 540–549
- Furukawa S, Tsugawa S (2021) Limitations of link deletion for suppressing real information diffusion on social media. In: *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'21)*, p 111–115
- Ghanem B, Ponzetto SP, Rosso P (2020) Factweet: profiling fake news twitter accounts. *Proc Stat Language Speech Process* 2020:35–45
- Helmstetter S, Paulheim H (2018) Weakly supervised learning for fake news detection on Twitter. In: *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'18)*, pp 274–277
- Hodas NO, Lerman K (2014) The simple rules of social contagion. *Sci Rep* 4(1):1–7
- Jolley D, Paterson JL (2020) Pylons ablaze: examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *British J Social Psychol* 59(3):628–640
- Kanwar K, Kaushal S, Kumar H, Gupta G, Khari M (2022) BC_{DCN} : a new edge centrality measure to identify and rank critical edges pertaining to SIR diffusion in complex networks. *Social Netw Anal Mining* 12(1):49
- Kashyap G, Ambika G (2019) Link deletion in directed complex networks. *Phys A: Stat Mech Appl* 514:631–643
- Khalil EB, Dilkina B, Song L (2014) Scalable diffusion-aware optimization of network topology. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'14)*, pp 1226–1235
- Kimura M, Saito K, Motoda H (2008) Minimizing the spread of contamination by blocking links in a network. In: *Proceedings of the 23rd National Conference on Artificial Intelligence (AAAI'08)*, pp 1175–1180
- Kitsak Maksim, Gallos KL, Havlin Shlomo, Liljeros Fredrik, Muchnik Lev, Stanley Eugene H, Makse AH (2010) Identification of influential spreaders in complex networks. *Nature Phys* 6(11):888–893
- Leskovec J, Backstrom L, Kleinberg J (2009) Meme-tracking and the dynamics of the news cycle. In: *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'09)*, pp 497–506
- Liu Yang W, Brook YF (2020) FNEED: a deep network for fake news early detection on social media. *ACM Trans Inf Syst* 38(3):1–33
- Mathew B, Dutt R, Goyal P, Mukherjee A (2019) Spread of hate speech in online social media. In: *Proceedings of the 10th ACM Conference on Web Science (WebSci'19)*, pp 173–182
- Pérez-Rosas V, Kleinberg B, Lefevre A, Mihalcea R (2018) Automatic detection of fake news. In: *Proceedings of the 27th International Conference on Computational Linguistics (COLING'18)*, pp 3391–3401
- Shah D, Zaman T (2011) Rumors in a network: who's the culprit? *IEEE Trans Inf Theor* 57(8):5163–5181
- Sharma K, Qian F, Jiang H, Ruchansky N, Zhang M, Liu Y (2019) Combating fake news: a survey on identification and mitigation techniques. *ACM Trans Intell Syst Technol* 10(3):1–42
- Tong H, Prakash BA, Eliassi-Rad T, Faloutsos M, Faloutsos C (2012) Gelling, and melting, large graphs by edge manipulation. In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM'12)*, pp 245–254
- Tripathy RM, Bagchi A, Mehta S (2010) A study of rumor control strategies on social networks. In: *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM'10)*, p 1817–1820
- Tsugawa S (2019) Empirical analysis of the relation between community structure and cascading retweet diffusion. *Proc Int AAAI Conf Web Social Med* 13:493–504
- van der Linden S, Roozenbeek J, Compton J (2020) Inoculating against fake news about COVID-19. *Front Psychol* 11:2928
- Wang S, Zhao X, Chen Y, Li Z, Zhang K, Xia J (2013) Negative influence minimizing by blocking nodes in social networks. In: *Proceedings of Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence*, pp 134–136
- Yan R, Li Y, Wu W, Li D, Wang Y (2019) Rumor blocking through online link deletion on social networks. *ACM Trans Knowl Discov Data* 13(2):1–26
- Yao Q, Zhou C, Xiang L, Cao Y, Guo L (2015) Minimizing the negative influence by blocking links in social networks. In: *Proceedings of Trustworthy Computing and Services*, pp 65–73
- Zareie A, Sakellariou R (2021) Minimizing the spread of misinformation in online social networks: a survey. *J Netw Comput Appl* 186:103094
- Zareie A, Sakellariou R (2022) Rumour spread minimization in social networks: a source-ignorant approach. *Online Social Netw Med* 29:100206
- Zhong E, Fan W, Wang J, Xiao L, Li Y (2012) Comsoc: Adaptive transfer of user behaviors over composite social network. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'12)*, pp 696–704

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.