

RESEARCH

Open Access



Domain-relevance of influence: characterizing variations in online influence across multiple domains on social media

Bowen Shi¹, Ke Xu¹ and Jichang Zhao^{2*}

*Correspondence:
jichang@buaa.edu.cn

¹ State Key Laboratory of Software Development Environment, Beihang University, Beijing 10091, China

² School of Economics and Management, Beihang University, Beijing 10091, China

Abstract

Influentials play a key role in enhancing information diffusion on social media. However, how personal influence varies across multiple domains is rarely addressed. This study introduces a concept called Domain-Relevance of Influence to describe the relation between influence and domains, and establishes a methodological framework with a sample of 8,520,933 Weibo users to explore the cross-domain characteristics of influence. The results show that generalists with cross-domain attributes possess significantly higher influence than specialists in most domains, whereas in a single domain such as sports or technology, specialists and generalists can possess comparable influence. We further show that influence is positively associated with cross-domain capability in overall domains, but not necessarily in each single domain. This study contributes to better understanding of the influence variation across domains for influence enhancement, and provides a big data-based methodological basis for cross-domain communication research.

Keywords: Opinion leadership, Social media, Cross-domain, Generalist, Domain-relevance of influence, Specialist

Introduction

Many studies have highlighted that influence is crucial for information diffusion on social media [41, 43]. Initially, Katz and Lazarsfeld [22] indicated that personal influence was a powerful force in processing and interpreting information and, based on an analysis of leadership among 800 housewives, briefly described a phenomenon in which older women only dominated in public affairs, while one-third of opinion leaders were influential in more than one domain. With the booming of new media [6], mainstream media like Cable News Network show cross-domain influence on multiple aspects such as entertainment, politics and business, and many Hollywood stars have joined the #MeToo social movement, kicking off an online tidal wave [32]. As a result, cross-domain influence is becoming increasingly prevalent, thus elevating the importance of understanding influence across multiple domains.

We usually consider multidomain influence in terms of the old aphorism regarding the “Jack of all trades”, which was first mentioned in the book Greene’s Groat’s-worth

of Wit, written by Robert Greene [16]. Interestingly, the expression “Jack of all trades, master of none, but oftentimes better than master of one” and the shorter version “Jack of all trades, master of none” have two opposing meanings and emphasize the advantages of generalists and specialists, who are called polymorphic and monomorphic opinion leaders, respectively [33]. Put simply, the aforementioned contradiction has not been solved for hundreds of years and is the same for how influence varies across domains on social media. On the one hand, professional users may be influential in specific fields. For example, Usain Bolt is the trend leader in sports on Twitter. On the other hand, such generalists with cross-domain attributes can reach a wider audience. For example, the travel endorsements of Internet celebrities can significantly increase tourism GDP [44]. Against this backdrop, the question of how user influence varies across domains might essentially challenge seed targeting in marketing-like scenarios. Therefore, the comparison between specialists and generalists in multiple domains is a meaningful cross-domain issue.

However, previous studies only describe phenomena in one or several narrow domains based on small samples [22, 24], and cross-domain advantages have rarely involved personal influence [26, 46]. Generalists and specialists, which ones should be chosen as disseminators in the target domain? How does their influence vary across multiple domains? Is there a quantitative correlation between cross-domain capability and influence? These questions still need to be answered by existing studies, mainly due to the high costs of questionnaires or surveys. Because the existence of multiple domains necessitates that the dynamic changes of influence be tracked in each domain and traditional methods such as informant ratings are time-consuming and expensive, it is difficult to generate a suitable domain classification for complex contents [43].

Fortunately, the digital footprints that accumulate and aggregate on social media provide a more efficient but less costly platform for investigating this issue [4, 28]. In particular, Sina Weibo, China’s most popular Twitter-like service, had over 550 million monthly active users in March 2020 [40]. Furthermore, it has its own news topic classification and is the dominant social media in China for citizens’ public discussions [4, 14], which provides an ideal platform for cross-domain analysis. Due to the strict real-name registration regulations, each Weibo users must provide a real ID card and an available phone number, which further ensures sample authenticity. The question arises as to whether it is possible to develop a framework that can precisely measure the online influence of various groups to identify the best ones for specific domains.

Because existing theories and methodologies cannot provide suitable methods to explore these issues, this study proposes a concept called Domain-Relevance of Influence (DRI) to characterize variations in online influence across domains. Moreover, a novel quantifiable framework based on big data technologies is designed to demonstrate how personal influence varies across multiple domains and the analysis of 8.52 million users supports our contentions. To the best of our knowledge, this study is the first to explore the cross-domain characteristics of online influence. Finally, by explicitly mapping the influence on domains, this study enriches practical marketing strategies.

Theoretical background

Opinion leadership plays an important role in two-step flow of communication [25] and usually refers to popular individuals, celebrities, or organizations [18]. Although some scholars have criticized that model for oversimplifying the process of influence and proposed the multi-step flow model as an update [39], opinion leaders remain key elements in information cascades [18]. In addition, with the rise of social media, many studies have suggested that the communication process on social media is similar to the two-step flow of communication [8, 10], motivating this study to focus on opinion leaders. In diffusion theory, influentials are often called “innovators”, “early adopters” or “opinion leaders” [33]. Many scholars also use “elites” to represent this general concept [7, 36]. Given that “opinion leader” has evolved diverse concepts that are ambiguous and difficult to operationalize [5, 37], this study uses “elite” to broadly refer to users with actual influence.

Previous studies have mostly focused on the influence of opinion leaders but overlooked ordinary users [9, 12]. However, the influentials hypothesis, which posits that influentials trigger wide dissemination, has been questioned in recent years [38]. Zhang et al. [47] even pointed out that ordinary users possess greater odds of creating viral trends than do traditional opinion leaders. Especially on social media, everyone can reach the entire network through public content, which lowers the standards of elites to a certain extent [18]. For example, local bloggers with few followers can also have high influence in disasters [13]. Therefore, the elites in this study are not predetermined by socioeconomic status and ordinary users can be elites as long as they have actual influence.

Influence is a crucial skill used to transmit messages to a broader public [12] and influentials use their visible position in an extensive network to spread news [33]. Elite formation is a dynamic process where individual influence is constantly changing and can be quantified by interactive ability [25]. Moreover, some scholars summarized that the context of content influences communication structures and elites are influential in certain substantive areas but not others [20, 23]. Therefore, this study highlights that the influence of elites will vary across domains. Here, the influence is dynamic and quantifiable, and specifically refers to the individual communication effect in the spread of information, rather than the traditional static socioeconomic status.

On social media with diverse content, exploring how influence varies across multiple domains is crucial for information diffusion. However, a systematic study on the relation between influence and multiple domains on large-scale social media remains lacking, which motivates us to propose DRI to extend prior knowledge of advantages of specialization and cross-domain to social media and enrich our understanding of the variation in influence across multiple domains.

The proposed concept and methodological framework

Domain-relevance of influence

The characteristics of polymorphic and monomorphic opinion leaders in diffusion theory provide a direct theoretical starting point for our study [33]. Only one-third of women exert cross-domain influence in four domains (fashions, movies, public affairs

and consumer products), making it a rare attribute [22]. As a classic cross-domain issue, the comparison between specialists and generalists provides a breakthrough point. Intuitively, generalists with various skills should reach a wide audience, and cross-domain has been proved to generate high-quality content through recombinant innovation [26]. However, cognitive psychologists have long believed that human beings have a limited capacity for information processing [45]. Specifically, a person can only focus on a limited number of objects and the human cognitive system can use the resource-limiting strategy to handle information overload [30]. Yet some elites on social media, especially the media, can hire teams with collective knowledge and effort to overcome this limitation. Therefore, in overall domains, generalists who dominate multiple domains are more likely to be influential and their influence may be related to quantifiable cross-domain levels, which is defined as cross-domain capability later in this study. Here, overall domains refers to considering the information of all domains as a whole.

Regarding the single domain, on the one hand, some previous studies have shown that specialization promotes the effect of communication in a single domain [24, 29], but their experiments did not validate every single domain from a complete domain classification system. On the other hand, generalists familiar with other domains can reach a wider audience. Especially on social media, audiences interested in tweets from other domains of generalists can easily access and spread information from the current domain to increase the influence of generalists in this single domain. More importantly, instant and first-hand information in some domains is particularly crucial for opinion leadership and the interactions among some domains will be more frequent than others [46], indicating that their inherent characteristics and relations may differ. Thus, the advantages of specialization and cross-domain may not be unitary across each single domain and are related to the characteristics of the current domain. For example, in some specialized domains such as product launches or sports events, some professional media who purchase copyright will exclusively attract interested audiences, while in some highly interactive domains, generalists can more easily attract audiences from other domains and be more influential. Similarly, the correlation between influence and cross-domain capability may be unstable across each single domain. Consequently, in a single domain, the variation of influence is related to the domain-specific characteristics.

Combining the above theories with the characteristics of social media, we propose DRI to preliminarily characterize how influence varies across multiple domains on social media as follows: (a) personal influence has a domain scope, and varies with different domains; (b) cross-domain is a rare attribute of influence, and influence is positively associated with cross-domain capability in overall domains, while in a single domain, the variation of influence is related to the characteristics of the specific domain. In some specialized domains that value exclusive first-hand information, specialists will demonstrate dominant influence against those generalists or two sides will be evenly matched; (c) the characteristics and relations of various domains are different, and the cross-domain patterns reflect preferences across user groups and domain combinations. These propositions emphasize that influence depends on the set of domains and that the relationships between domains are complex. Therefore, when choosing early adopters in diffusion of innovations, it is necessary not only to observe their overall influence but also

to consider the characteristics of the target domain. Note that the influence in this study refers to the quantifiable individual communication effect in information cascades.

To test the propositions, three challenges in studying influence change on large-scale social networks with wide domains must be resolved. First, traditional methods are difficult to measure large populations' influence from comprehensive perspectives [5]. Second, different demographics can lead to distinctive domain preferences [48]. It is challenging to develop a suitable taxonomy for group users. Finally, determining the appropriate number of domains is critical in multi-domain research.

Reviewing the existing literature exposes the need for an available methodology to overcome the above three challenges, and our methodological framework abstracts them into three key elements: domain classification, user groups and influence measures. The following section offers details about our framework.

Domain classification

Domain, also called as field, broadly refers to various disciplines, content topics, platforms, organizations and countries [26, 46]. In the context of social media, the domain specifically represents the various topics of texts, i.e., tweets or posts. As the length of a tweet has historically been limited to 140 characters, machine learning technology and word embedding model have been applied to short text classification [14, 31]. To find a suitable number of topics, Li et al. [27] used the topic system predefined by journalists to train a classification model on Twitter. Inspired by this, a complete domain classification system with consistent and well-accepted categories specific to Weibo can be derived from Weibo itself. Weibo features a news taxonomy system from which the number and descriptions of domains can be determined accordingly [14]. Here, Weibo texts are split into seven domains such as entertainment, military and so on, which integrally reflect Weibo news topics.

Influence measures

In the context of the social media boom, many metrics of influence have been extensively presented based on statistical information and topological structure. For example, the ratio of the number of followees to followers reflects the user's broadcasting ability [17]. Meanwhile, many measures, such as in-degree, betweenness, and eigenvector centralities, as well as random walk methods [1], have been proposed for ranking users. The CI represents a comprehensive index of influence in information dissemination [28]. However, relatively static indicators such as followers cannot change across domains, and Al-Garadi et al. [2] noted that centrality methods have high computational complexity and that there are substantial differences between random walk methods and information spread in reality. Therefore, we select CI and in-degree to measure the dynamic changes of domain-relevant influence. For the large size of networks, the complexity of CI and the in-degree is $O(N \log N)$ and $O(1)$, respectively, implying an efficient calculation. In the network of information diffusion, in-degree is equal to the number of neighbors of the target user in one-hop subnet and reflects the reputation and influence recognized locally [11, 37], which is the basic criterion for selecting elites, while CI places greater emphasis on bridging giant components at the global level [28]. In short, they reflect the

communication effects in information cascades from different perspectives, which is in line with the concept of influence in this study.

User groups

Users in different demographic categories, e.g., professions [48], might exert different innate influence across multiple domains. For example, Spanish authorities actively interact with citizens regarding local issues [19], while enterprises are busy with online marketing. Fortunately, Weibo has a rigorous verification mechanism of verifying user demographics. After real-name certification, verified users must provide handwritten documentary evidence and a work certificate with the official seal. Here, users are split into five groups: celebrity, enterprise, etc.

By splitting users into five groups, content into seven domains and influence into two indicators, we establish an accurate and comprehensive mapping between personal influence and domains to test the DRI.

Study design

Specialist vs generalist

Traditionally, specialists were considered to be highly skilled in a narrowly defined domain, whereas generalists had knowledge and reputation in multiple domains [42]. Some scholars have theoretically emphasized that specialization is a recognized social accomplishment via communication others rather than a resource held by individuals [35]. Generally speaking, individuals need to invest considerable time and effort to gain the domain-specific recognition, but social media accounts can greatly reduce the time cost to develop a reputation in a certain domain by hiring spokesmen and professional teams. Moreover, the previous classification of generalists and specialists is subjective, and the levels of expertise in multiple domains lack a unified assessment standard. On social media, although everyone can post content on diverse topics, it is important for evaluating specialization that the value of the content is recognized by local topical networks. Therefore, this study defines specialists as users who are not only active in the domain, but also have sufficient regional reputation, i.e. a certain regional influence in this domain, while generalists need to possess sufficient reputation in multiple domains. Specifically, we use in-degree, which reflects the regional reputation, to select elites in each domain-oriented network. Specialists and generalists refer to being identified as elites in only one domain and in two or more domains, respectively.

As mentioned earlier, the comparison between specialists and generalists is a classic question [24, 29]. However, the existing empirical evidence for such comparison remains ambiguous and irrelevant to social media. For example, consumers trust professional websites more during shopping [24], whereas Leahey and Moody [26] suggested that a multidisciplinary article will usually receive more citations. According to DRI, different domains are expected to influence this comparison on social media. Thus, given the prior inconsistent findings, we explore this question from the perspectives of overall domains and single domain, respectively. In overall domains, social media accounts can use the collective knowledge and effort of the hired team to ensure the quality and quantity of multi-domain tweets to reach a wider audience. Considering the advantages of generalists in overall domains, we hypothesize following:

H1. In overall domains, generalists are more likely to be influential than specialists.

Regarding the single domain, since most prior studies only focused on one domain [24, 29], we advance this line of research by exploring how the comparison varies in each domain from a complete news taxonomy system. DRI shows that the variations in influence of a single domain are related to domain-specific characteristics, so the comparison of each single domain may be inconsistent and need to be explored separately. Hence, we ask a research question:

RQ1. How does the comparison of influence between specialists and generalists vary in each single domain?

Cross-domain capability and influence

Generalists have different levels of cross-domain attributes. Yet prior research has largely neglected the level of cross-domain, such as how to define and quantify it. Here, we define cross-domain capability as the ability to have a certain proven reputation and regional influence in multiple domains, and measure it by the number of domains in which target users are identified as elites. It is worth noting that many followers and tweets do not necessarily bring enough reputations. For instance, ETtoday News has received popular recognition in international and military, but not enough feedback in other domains (see Additional file 1: Table S1 for detailed attribute variations of some users across domains). Surprisingly, it is not as influential as the Chinese Air Force official with fewer followers and tweets in military, or even the individual blogger (Lawyer-Kevin in New York) in both overall domains and each single domain. Interestingly, the Shanghai Morning Post, which has a certain reputation in six domains, did not get enough recognition for its tweets on sports. In contrast, some specialized enterprises with first-hand information on sports or technology (such as Super Sports Media Inc., PPSports, NBA, Huawei, Smartisan, etc.) occupied leading positions in the corresponding single domain. Therefore, cross-domain capability as an attribute of regional influence is not a simple spillover of the influence of elites with massive fan basis, but depends on their sufficient efforts and content recognition in the domain.

On the whole, the ability to possess reputations in multiple domains is more likely to reach a wider audience, and social media accounts can use fine divisions of labor and collaboration to overcome the information overload of cognitive system to improve cross-domain capability [30]. As a result, cross-domain capability may be connected to influence enhancement. However, previous studies have rarely addressed the correlation between cross-domain capability and personal influence. The research on journal impact suggests that citing interdisciplinary disciplines can increase the journal's impact factor under certain conditions [46]. DRI reminds that the variation of influence is related to domain-specific characteristics, so we explore this issue from the perspectives of overall domains and single domain, respectively. Therefore, we expect the following relation in overall domains:

H2. In overall domains, influence is positively associated with cross-domain capability.

Lots of information dissemination and influencer marketing are targeted at domain-specific events and products [13, 32, 34, 44], and within a single target domain, influence maximization is crucial for most communicators and marketers to maximize information diffusion. Taken overall, cross-domain is a viable means to expand your

target audience to enhance influence. However, in each single domain, is user influence within that domain always positively correlated with cross-domain capability? That is, whether the audience interested in the content published by the target user in other domains will eventually continue to recognize his content of the current domain, thereby spreading information to increase user influence of the current domain. Unfortunately, there is little research on the correlation between cross-domain capability and single-domain influence due to the lack of metrics to measure the dynamic influence across each domain. Even the above impact factor is an indicator for overall disciplines and cannot vary across disciplines. In our methodological framework, the influence indicators based on the networks of each domain can vary across each single domain, which makes it possible to explore this issue. In addition, DRI implies that this correlation may vary in different single domains. To provide novel insights on the relation between influence and cross-domain capability in each single domain, we ask:

RQ2. How does the relationship between influence and cross-domain capability vary in each single domain?

Cross-domain pattern

Just as the frequencies of interaction between various disciplines differ [46], the degree of closeness between various domains will also change. Katz and Lazarsfeld [22] demonstrated some common cross-domain patterns, but there is no empirical evidence for successful cross-domain patterns on social media. Frequent cross-domain patterns show the interaction preference between domains, which provides valuable experience for elites who want to effectively use cross-domain behavior to expand their influence. Moreover, DRI also suggests that cross-domain patterns can differ according to user demographics such as status. Thus, for a target user group, which domains are the shortcuts for him to gain cross-domain capability? To enrich the marketing strategies, we ask:

RQ3. Which domains interact more frequently in cross-domain patterns?

RQ4. What are the differences in the cross-domain patterns of various user groups?

Methods

Data collection

Over 140,000,000 tweets from 10 October 2016 to 10 January 2017 were collected using the Weibo application program interface (API), and we identified 8,520,933 users by their unique IDs. For each tweet, the fields in the JavaScript Object Notation (JSON) file contain the following information: text, retweet status and user demographics such as gender, verified types and so on. Each user's posting activity and retweet count accumulate by the retweet status.

In particular, verified users are called "Big Vs" on Weibo [9], and a red or blue "V" next to their account names. The description and examples of the main types are shown in Additional file 1: Table S2 and thus we can divide users into five main groups to examine at the microscopic level. Note that in the following, the term "all" refers to overall domains considering all tweets as a whole, regardless of their domain labels.

Domain classifier

The text topic can represent the content domain on Weibo. This study adopted a previously well-developed Naive Bayesian classifier to perform domain classification [14]. The classifier is trained on more than 410,000 manually labeled tweets, and its seven topics fit well with the news taxonomy provided by Weibo. If none of the topics has a significant advantage in probability, the text will be classified into the “unknown” category. Note that tweets labeled “unknown” by the classifier will be omitted in our later analysis due to the lack of confidence in determining their domains.

Here, the labeled dataset is divided into ten equal partitions for cross-validation. In each iteration, a piece is selected as the test set, while the others are all used as the train set. After averaging all measurements, the model performance is shown in Additional file 1: Table S3. The F-measure and precision of the classifier in the ten-fold cross-validation experiment on the test set are greater than 84%. Compared with other machine learning technologies, unique incremental training mechanism of Bayesian analysis can also solve the problem of new words in to-do tasks.

Selection of elites

As a primary form of online interaction, retweeting does not necessarily mean agreeing with views, but it indicates that the retweeter pays attention to the message and acknowledges the informational value of retweetee [11, 18]. Also noteworthy, the information on retweets and corresponding authors extracted from the “retweet status” field of each tweet was used to create a network. The retweet network can be represented by a directed weighted graph in which the nodes represent users (those without edges are omitted), the edges are the sets of relationships, and the weight of an edge is the total number of retweets between user pairs (in sampling period). Accordingly, we built eight networks using the whole and separate retweet data from the seven domains. Their degree distributions show the power-law trend with long tails (see Additional file 1: Fig. S1), indicating that a minority of users have many connections. Additional file 1: Fig. S2 shows a sampled snapshot of the military retweet network.

In every domain’s retweet network, in-degree is often used to measure users’ regional influence and identify elites [11, 18, 37]. Due to the uneven network size across domains, the top k is used to select elites instead of probability to handle imbalanced domain problem that the number of elites is skewed towards large domain networks. To guarantee the reliability of the retweeting relationships, only pairs with at least two retweets are connected to establish the network. This study defines elites who are top- k in only one domain as specialists and in more than one domain as generalists. Considering that cross-domain is a rare attribute, the value of k should not be too large. Additionally, Additional file 1: Figs. S3 and S4 show the changes in influence scores of top- k users and the proportion of reachable nodes, respectively. To ensure sufficient individual influence and collective communication effects, we set k to 200. The proportional distribution of specialists and generalists in the ranking is shown in Additional file 1: Fig. S5, where “all” indicates the condition of overall domains and has a total of 990 unique users from a collection of 200 elites in each domain. Note that in the ranking of overall domains, the calculation of influence is based on the network composed of all retweets. Compared with other domains, specialists in sports and technology comprise a larger proportion

among the top 60 (the three bars on the far left). In addition, the result of $k=100$ shown in Additional file 1: Fig. S6 is similar to that of $k=200$.

Additional file 1: Table S4 reflects the differences of various elites, as shown by the distribution of 930 elites in five groups across domains (the other user groups, such as school, are ignored due to the small proportions). Sports and technology have the most enterprise elites. Moreover, the proportion of specialists across user groups is shown in Additional file 1: Fig. S7. Enterprises tend to be the most professional, while media and government users have a high proportion of generalists.

Measures

As mentioned earlier, this study highlights the dynamic variation of influence. Through eight networks, we measure user influence in each specific domain as well as overall domains from the two perspectives of in-degree and CI. Considering that the influence value is affected by the network scale, we use the ranking value of in-degree and CI in the experiment. Lower rank values indicate greater influence. The degree of cross-domain capability is measured by calculating how many times the user is selected as elite in all domains. In other words, the more domains in which a user is an elite, the greater that user's cross-domain capability is. Because many factors might affect influence, this study employs multiple regression models to explore the correlation between influence and cross-domain capability. Boster et al. [5] emphasized that persuasive people often advance their opinions frequently, so posting more will bring more opportunities to be noticed. Meanwhile, Weibo posts are automatically broadcast to the author's followers, hence, the number of followers reflects the odds that content will be disseminated and is a factor affecting influence. Moreover, gender is an important demographic in understanding human behaviors [15], and the verified system is a prominent feature of Weibo. To sum up, one dependent variable (influence) and five groups of variables are included in the regression models: (1) cross-domain capability, (2) activity, (3) number of followers, (4) gender, (5) verified type. Here, the number of tweets within the sampling period is used to measure activity, and some numerical independent variables are processed by logarithm to limit the large range of variation for better effects. The gender and verified type are dummy variables, and their values are 1 when they belong to this group and 0 otherwise. The Means and standard deviations of independent variables across user groups are shown in Additional file 1: Table S5.

Results

Specialist vs generalist

H1 and RQ1 examined the comparison of influence between specialists and generalists across domains. As a classical cross-domain question, we compare the two sides in terms of CI and in-degree. As Additional file 1: Fig. S5 shows, generalists in most domains occupy a larger proportion of the top positions, especially the top 60. However, the proportion of generalists across each interval in sports and technology is almost always less than 50%, implying that there are more top-ranked specialists in both domains and the comparison will differ within different single domains. Figure 1 visually presents the comparison of in-degree rank, and the Z-test results are shown in Table 1. The results show that the influence of generalists is significantly greater than that of specialists in

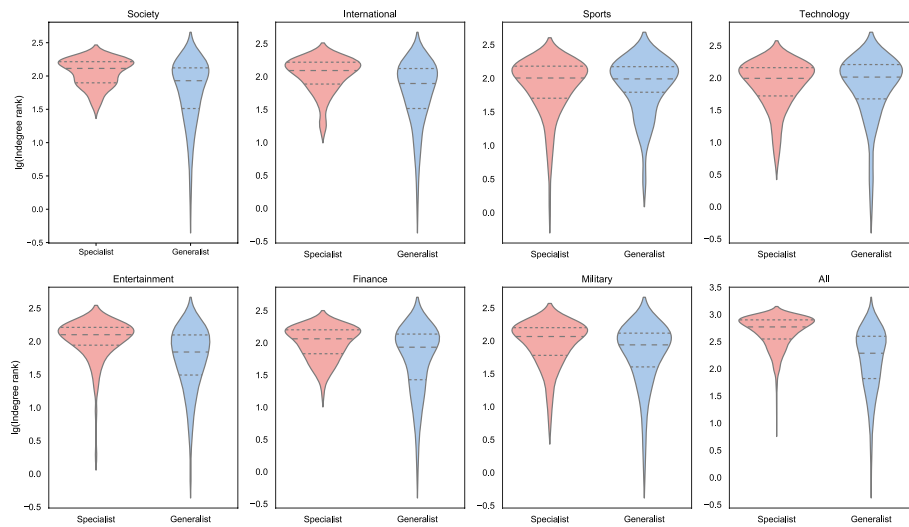


Fig. 1 Comparison of influence (in-degree rank) between specialists and generalists across domains. Lower value is better in in-degree rank. All refers to considering all domains as a whole

Table 1 Z-test of the influence (in-degree rank) of specialists and generalists across domains

Domain	Z-value	P-value
Society	4.410	0.000***
International	4.699	0.000***
Sports	-0.098	0.922
Technology	-0.259	0.796
Entertainment	5.293	0.000***
Finance	3.255	0.001**
Military	2.701	0.006**
All	16.284	0.000***

Level of significance: ***p < 0.001; **p < 0.01; *p < 0.05. All refers to considering all domains as a whole

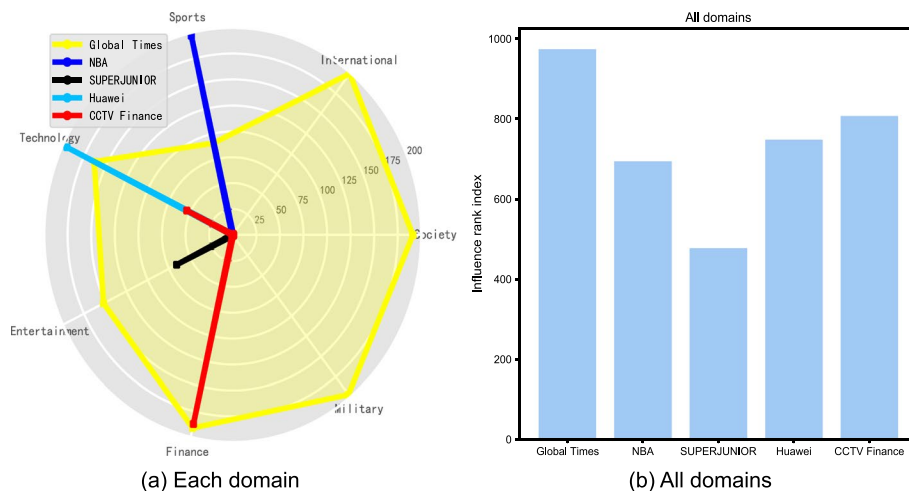


Fig. 2 Variations in the influence rank index of some users across domains. All refers to considering all domains as a whole. In each single domain, we calculate the influence rank index by subtracting the user ranking from the maximum ranking value of the selected elites for better visualization. Consequently, those with a higher rank index are more influential

overall domains, providing solid support to H1. In addition, Fig. 2 shows some examples, indicating that generalists are more likely to be influential in overall domains.

Concerning RQ1, in a single domain, the comparison is dependent on the corresponding domain. Specifically, in the domains of society and entertainment, the influence of generalists is significantly greater than that of specialists, while in sports and technology, the difference is not statistically significant (see Fig. 1 and Table 1), and specialists can also exert high influence. According to Additional file 1: Fig. S7 and Table S4, more professional enterprise accounts may be the reason for the specialization in sports and technology. For instance, Huawei is more influential than Nubia in technology (see Additional file 1: Fig. S8), though the latter has gained cross-domain capability by sponsoring China Open and even signing Cristiano Ronaldo as the brand ambassador. In fact, while paying more attention to the product itself, Huawei also invited some sports stars to boost its brand, but this failed to bring enough influence in sports. Similarly, many sports accounts such as Hupu Soccer only report soccer news, which reflects the advantages of specialization. Another possible explanation is that events in sports and technology, such as matches and product launches, are scheduled in advance. From the audience perspective, interested users will automatically look for experts with first-hand information [37], which is consistent with Uses and Gratifications Theory [3]. The results of CI are similar to those of in-degree, as shown in Additional file 1: Fig. S9 and Table S6. Therefore, the variation of influence across domains should be seriously considered when targeting domain-oriented marketing seeds.

Relationships between cross-domain capability and influence

In testing H2, intuitively, influence may increase with cross-domain capability. As stated earlier, activity and followers are also key factors affecting influence. Moreover, according to DRI, various user groups have preferences for domains. Therefore, Table 2 presents the multiple regression results of 10 independent variables such as cross-domain ability, activity and followers to predict influence across domains. In overall domains, the model fits well, and its coefficient of determination is 0.56. The cross-domain capability is significantly positively associated with influence ($p < 0.001$) when all other variables were controlled, which confirms H2. In fact, this phenomenon can explain why users who post across domains will have more opportunities to attract potential followers from diverse domains. Moreover, detailed statistics show that there are only six elites across all seven domains and they are mainly media users such as People's Daily, indicating that teamwork gives them an innate advantage in cross-domain capability.

To address RQ2, models based on various domains further confirm that the cross-domain capability is significantly positively associated with influence in most domains, but not in sports. In most situations, cross-domain is a good attempt, but the models for sports and technology have the lowest absolute value of coefficients of cross-domain capability and determination. The results based on CI are shown in Additional file 1: Table S7, where the coefficients in sports and technology are nonsignificant. In these two domains, the significance of most variables is unstable across two influence indicators. In contrast to other user types with significant coefficients, enterprise is most likely to get top CI and in-degree rank in sports and women with many followers are more likely to be influential in technology, indicating that they are important variables in

Table 2 Multiple regression to predict influence (in-degree rank) based on cross-domain capability across domains

	Society		International		Sports		Technology		Entertainment		Finance		Military		All	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Intercept	1.67***	0.15	1.30***	0.11	1.86***	0.16	1.66***	0.17	1.71***	0.15	1.43***	0.12	1.69***	0.11	2.37***	0.04
Cross-domain capability	-0.14***	0.02	-0.12***	0.02	0.02	0.02	-0.05*	0.02	-0.13***	0.02	-0.12***	0.02	-0.09***	0.02	-0.21***	0.01
Ig(Activity)	-0.03	0.05	-0.05	0.05	-0.22***	0.05	-0.10*	0.05	-0.15**	0.05	-0.12**	0.04	-0.30***	0.04	-0.09***	0.02
Ig(Followers)	-0.08	0.04	0.00	0.03	-0.16**	0.05	-0.09*	0.04	-0.04	0.04	-0.01	0.04	-0.05	0.03	-0.15***	0.01
Ordinary (dummy)	0.30***	0.07	0.31***	0.05	0.48***	0.06	0.31**	0.09	0.23**	0.09	0.34***	0.06	0.39***	0.05	0.41***	0.02
Celebrity (dummy)	0.28***	0.06	0.31***	0.05	0.47***	0.08	0.25**	0.07	0.22**	0.08	0.37***	0.05	0.34***	0.05	0.36***	0.02
Government (dummy)	0.42***	0.12	0.11	0.10	0.00	0.00	0.52*	0.24	0.89**	0.30	0.20*	0.09	0.27**	0.08	-0.61***	0.05
Enterprise (dummy)	0.34*	0.14	0.38***	0.11	0.37***	0.08	0.11	0.08	0.09	0.14	0.25**	0.08	0.39***	0.09	0.48***	0.03
Media (dummy)	0.33***	0.08	0.21**	0.07	0.54***	0.09	0.47***	0.10	0.27**	0.10	0.26***	0.07	0.30***	0.07	0.51***	0.03
Male (dummy)	0.85***	0.08	0.70***	0.06	0.91***	0.09	0.87***	0.09	0.84***	0.08	0.68***	0.07	0.90***	0.06	1.25***	0.02
Female (dummy)	0.82***	0.08	0.60***	0.06	0.94***	0.09	0.79***	0.09	0.87***	0.08	0.74***	0.07	0.80***	0.06	1.12***	0.03
Observations	199		194		164		196		194		192		192		930	
R ²	0.36		0.30		0.22		0.12		0.30		0.34		0.44		0.56	
Adjusted R ²	0.33		0.27		0.19		0.08		0.27		0.31		0.42		0.56	
F-Statistic	13.16***		9.84***		6.30***		3.16**		9.73***		11.95***		17.82***		146.40***	

Level of significance: ***p<0.001; **p<0.01; *p<0.05. Lower value is better in in-degree rank. All refers to considering all domains as a whole

predicting influence in these single domains. Taken together, independent of indicators, the results suggest that the relationship between influence of single domain and cross-domain capability depends on the specific domain. From this perspective, cross-domain behavior does not always enhance the influence of a single domain, and targeting the right domains can be an option for boosting the probability of an elite serving as a seed.

Preference in cross-domain patterns

Regarding RQ3, Fig. 3 shows the mapping between frequent patterns and domains, in which the number of users in each pattern must be greater than five. It should be noted that the most frequent pattern is (society, entertainment), which provides goal for the elites involved in these domains. Compared with sports and technology, other domains have a higher level of inter-domain interactions, which proves the existence of interaction preference among domains. Hence, most elites need to follow an empirical pattern that suits them, and top influencers should consider increasing their number of domains to further boost influence. As for sports and technology, elites in these domains should concentrate on improving their professionalism to strengthen influence.

RQ4 asked about the differences in the cross-domain patterns of various user groups. Similar to the overlap of leadership in different areas [21], Fig. 4 demonstrates the overlapping of elites across domain pairs for each group. Specifically, the elites of a certain group are further divided into partitions that occur simultaneously in any pair of domains, and larger intersections reflect a higher concentration of elites in the corresponding domain pairs. The distribution of the intersection sizes for each group is extremely unbalanced, which is consistent with the proposition that various elites have heterogeneous preferences for cross-domain patterns. For instance, celebrities prefer to discuss entertainment and society topics, while government elites focus on domain pairs among the international, finance and military domains. From a point of practical

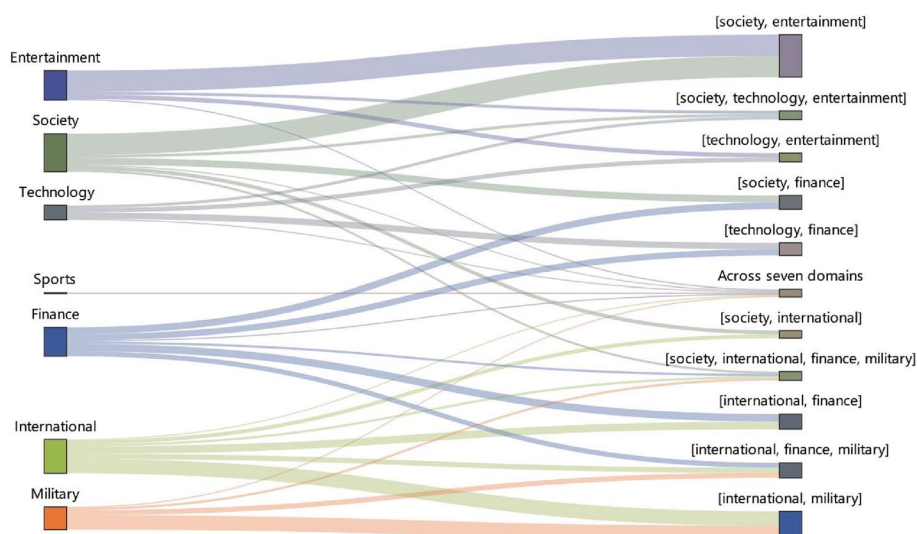


Fig. 3 Mapping between frequent cross-domain patterns and various domains. The length of pattern indicates its user size; the number of users in each pattern must be greater than five. The value of each pattern is equally assigned to the weights of all domains it contains

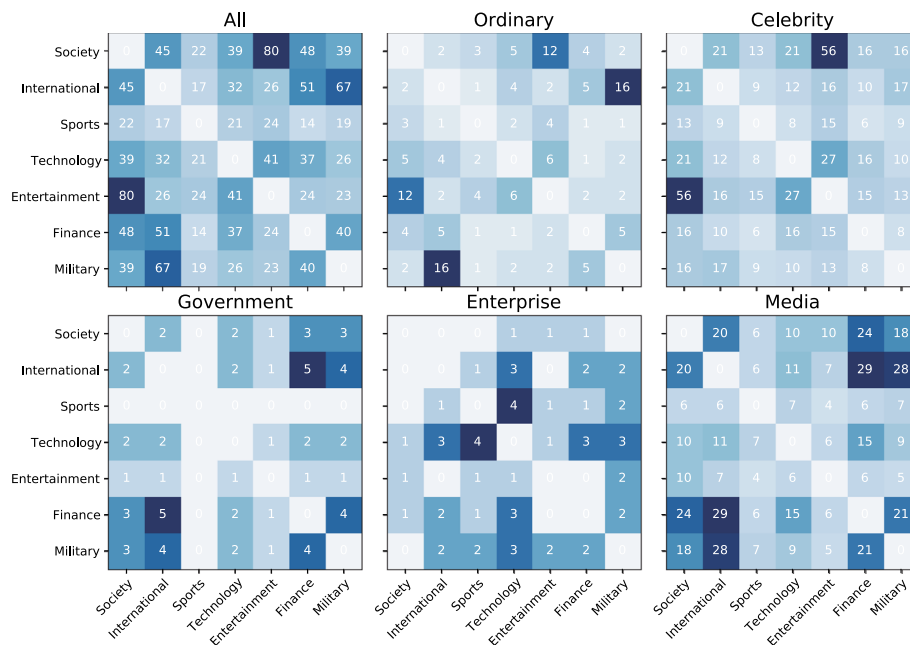


Fig. 4 Overlapping elites across domain pairs for each user group. All refers to considering all domains as a whole

inspiration, targeting a given user group can be effectively implemented based only on the domains in question.

Conclusion and future works

Influence varies across multiple domains, which is a prevalent phenomenon worth exploring in the context of social media. Considering that previous studies on influence have ignored the domain element [5, 12], we introduced DRI to characterize how influence varies across domains on social media and proposed a methodological framework based on machine learning and big-data analytics from social networks of 8,520,933 users to examine the hypotheses. In general, the results speak to the effectiveness of our framework in capturing the dynamics of online influence across multiple domains and provide empirical evidences for our perspectives. Specifically, influence is positively associated with cross-domain capability in overall domains, whereas in a single domain, influence variation depends on the domain-specific characteristics, so specialists can also possess comparable influence. These results fill an important gap in cross-domain research and provide practical implications for maximizing influence in multiple domains. The detailed theoretical and methodological contributions of this study are discussed below.

This study reconceptualized a series of key elements, such as influence and generalists, to fit the characteristics of social media, theorizing how online influence varies across domains. To increase influence, users face a trade-off in deciding whether to focus on one domain or invest time in multiple domains, which remains a controversial issue in prior literature. Some works only emphasized the advantages of specialization in a single domain [24, 29], while other interdisciplinary studies focused more on collaboration and recombinant innovation [26, 42, 46]. Our DRI extends this literature by comprehensively

characterizing variations in influence from both overall domains and single-domain perspectives and highlighting that the differences among domains can affect varying patterns.

Regarding the classical cross-domain question, the comparison between generalists with cross-domain attributes and specialists, most prior studies suggest that specialized channels or websites have better effects in communication [24, 29], but they only focus on one domain. In overall domains, independent of the influence measurements, our results reveal that high influence is in fact related to the cross-domain attributes and generalists are more likely to be influential. Furthermore, our experiments in each domain also demonstrate that in most single domains, the advantages of specialization no longer exist. In other words, the influence of generalists is significantly higher than that of specialists, apart from the domains of sports and technology, which further implies that social media has revolutionized the landscape of conventional communication systems [37]. On social media, the cognitive diversity of generalists often helps to integrate knowledge of diverse domains to reach a wider audience. In contrast, specialists can be particularly influential in sports and technology, which is in line with our proposition that influence within a single domain is related to the domain-specific characteristics. Specifically, many sports-focused media companies have purchased the exclusive right to broadcast sporting events, and technological enterprises monopolize the release of their new product information. According to the Uses and Gratifications Theory, audiences actively select the exclusive news sources they consume to gratify various psychological needs [3]. Thus, in these single domains that value exclusive first-hand information, specialists can also possess considerable influence.

Notably, in overall domains, influence enhancement is significantly positively related to the increase in cross-domain capability, i.e., posting in more domains can be a feasible path for generalists to increase their influence. We further prove the existence of top influencers across all seven domains by retrieving detailed ranking. Interestingly, most of them are media users such as the People's Daily and Global Times, which indicates that in the age of social media, team accounts can overcome cognitive limitations to gain cross-domain capability and are more likely to be influential in overall domains, while solo authors are difficult to become generalists across many domains. In addition, this result also proves that Chinese traditional newspaper media has successfully transformed on Weibo through cross-domain capability and remains dominant in overall domains, which can be explained by the group's professional expertise, commercial motivation and reputation accumulated in offline communication. However, within a single domain, high influence is not necessarily accompanied by high cross-domain capability. This study suggests that this correlation is not unitary across each single domain and related to the domain-specific characteristics. Unlike the case in which highly interdisciplinary journal articles are more frequently cited [46], in some specialized domains such as sports, there is no clear correlation between cross-domain capability and influence. This might have important implications for communication and marketing in multiple domains. On the one hand, in a specialized target domain, it is necessary to provide high-quality content timely to attract audiences, and reputation beyond the scope of target domain should not be weighed highly, on the other hand, elites still have to strive for reputation in

multiple domains to enhance the communication effects in overall domains, and the coefficients of cross-domain capability can offer practical clues for enhancing influence across domains.

The frequent and binary patterns of various users confirm the existence of cross-domain preferences. On the whole, elites with high cross-domain capability are uncommon, and there are preferential attachments between different domains. The collection of diverse patterns provides prior knowledge for target seeds in marketing-like scenarios. Based on the overlap of leadership [21], another finding is that the changing patterns of influence in any domain pairs are complex across user groups. In other words, being active in multiple but random domains might not enhance influence because the preferential attachment differs across user group. For example, the government group trivially influences trends in sports and entertainment, while enterprises have the highest influence in technology. Therefore, these intersection patterns of elites across domain pairs have practical implications for the cross-domain marketing of various users.

Methodologically, the study contributes to cross-domain research on social media by developing a framework with three key elements to capture the variations of influence across multiple domains. The schemes for dividing both users and content are fine-grained and derived from the official taxonomies with regard to both verified types and news topics. Moreover, Weibo's real-name regulation and rigorous verification process further guarantee the authenticity of anonymous users. Finally, our experiments, based on two influence indicators and big data from Weibo including over 140,000,000 tweets, provide powerful supporting evidence for propositions. Unlike traditional methods, machine learning and network analysis make our approach less expensive and easier to operate, which shed light on future research on cross-domain leadership.

To conclude, this study theoretically and methodologically links the influence and domains on social media and the results about cross-domain capability deserve particular emphasis. As a vital step, this study helps enrich our understanding of how influence varies across multiple domains, which prior research has not addressed, and suggests that domain-relevance can be used to maximize influence or identify influencers in multiple domains on social media. In short, most elites should follow a suitable cross-domain pattern according to the target domain and top generalists should improve cross-domain capability to boost influence.

We are entering the age of social sensing, in which every individual in society can be a sensor who voluntarily reports and disseminates sophisticated signals anytime and anywhere. The previous offline communication landscape has been inherently and profoundly reshaped and revolutionized, evolving into a new online version. As a preliminary foundation, DRI will be helpful for achieving a better balance between popularity and cost across multiple domains on social media. Nevertheless, the study is limited in that results here are static, not dynamic, and only focus on Weibo user behavior. One promising direction of future studies is to gain insight into the spatio-temporal variation of influence under more diverse social media platforms and domain classification systems with diverse cultural contexts to extend and refine the concept.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40537-023-00764-x>.

Additional file 1: Table S1. Variations in the attributes of some representative samples of users across domains. All refers to considering all domains as a whole. Lower rankings are omitted to highlight the user's cross-domain capability. **Table S2.** Description and examples of user groups. **Table S3.** The precision, recall and F-measure of the cross-validation. All refers to considering all domains as a whole. **Table S4.** Distribution of the number of users across user groups and domains. All refers to considering all domains as a whole. **Table S5.** Means and standard deviations of independent variables across user groups. All refers to considering all user groups as a whole. **Table S6.** Z-test of the influence (CI rank) of specialists and generalists across domains. Level of significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. All refers to considering all domains as a whole. **Table S7.** Multiple regression to predict influence (CI rank) based on cross-domain capability across domains. Level of significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Lower value is better in CI rank. All refers to considering all domains as a whole. **Figure S1.** Probability distribution of in-degree of each domain-oriented retweet network. All refers to considering all domains as a whole. **Figure S2.** Retweet network of various users in the military domain. The threshold of the edge weight is set to 10 for better visualization, and the size of the node is related to its number of retweeters. We color each node by its verified type, i.e., blue represents the media, green represents enterprises, red represents the government, orange represents celebrities and gray represents others. Note that the color of the edge is the same as that of the source node. **Figure S3.** Influence scores (in-degree) of the top-k influentials in each domain-oriented retweet network. **Figure S4.** Proportion of users reached by the top-k influentials in each domain-oriented retweet network. **Figure S5.** Proportional distribution of specialists and generalists in the influence ranking across domains. The threshold of each single domain is set to 200 and each ranking interval is 20. All refers to considering all domains as a whole. **Figure S6.** Proportional distribution of specialists and generalists in the influence ranking across domains. The threshold of each single domain is set to 100 and each ranking interval is 10. All refers to considering all domains as a whole. **Figure S7.** Percentages of specialists and generalists across user groups. **Figure S8.** Variations in the influence rank index of some enterprise users across domains. All refers to considering all domains as a whole. In each single domain, we calculate the influence rank index by subtracting the user ranking from the maximum ranking value of the selected elites for better visualization. Consequently, those with a higher rank index are more influential. **Figure S9.** Comparison of influence (CI rank) between specialists and generalists across domains. Lower value is better in CI rank. All refers to considering all domains as a whole.

Acknowledgements

Not applicable.

Author contributions

KX, JZ and BS conceived the study and designed the research; KX provided data samples and experimental environment; BS performed the data analysis and modeling studies; KX, JZ and BS participated in the discussion; JZ and BS wrote the paper. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (Grant No. 71871006) and the National Social Science Funds of China (Grant No. 18BKS180).

Availability of data and materials

The data sets used in this study are publicly available after careful anonymization and can be downloaded freely through <https://github.com/bowenshi/cross-domains>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors have read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 23 December 2022 Accepted: 8 May 2023

Published online: 17 May 2023

References

1. Ahajjam S, Badir H. Identification of influential spreaders in complex networks using HybridRank algorithm. *Sci Rep.* 2018;8(1):11932.
2. Al-garadi MA, Varathan KD, Ravana SD. Identification of influential spreaders in online social networks using interaction weighted K-core decomposition method. *Physica A.* 2017;468:278–88.
3. Blumler JG, Katz E. The uses of mass communications: current perspectives on gratifications research. *Sage Annu Rev Commun Res.* 1974;3:318.

4. Barnett GA, Xu WW, Chu J, Jiang K, Huh C, Park JY, Park HW. Measuring international relations in social media conversations. *Gov Inf Q*. 2017;34(1):37–44. <https://doi.org/10.1016/j.giq.2016.12.004>.
5. Boster FJ, Kotowski MR, Andrews KR, Serota K. Identifying influence: development and validation of the connectivity, persuasiveness, and maven scales. *J Commun*. 2011;61(1):178–96. <https://doi.org/10.1111/j.1460-2466.2010.01531.x>.
6. Brossard D, Scheufele DA. Science, new media, and the public. *Science*. 2013;339(6115):40–1.
7. Chong D, Druckman JN. A theory of framing and opinion formation in competitive elite environments. *J Commun*. 2007;57(1):99–118.
8. Choi S. The two-step flow of communication in Twitter-based public forums. *Soc Sci Comput Rev*. 2015;33(6):696–711. <https://doi.org/10.1177/0894439314556599>.
9. Chen J, She J. An analysis of verifications in microblogging social networks—Sina Weibo. In: 2012 32nd international conference on distributed computing systems workshops. IEEE; 2012. p. 147–54.
10. Dang-Xuan L, Stieglitz S, Wladarsch J, Neuberger C. An investigation of influentials and the role of sentiment in political communication on Twitter during election periods. *Inf Commun Soc*. 2013;16(5):795–825.
11. Doerfel ML, Taylor M. The story of collective action: the emergence of ideological leaders, collective action network leaders, and cross-sector network partners in civil society. *J Commun*. 2017;67(6):920–43.
12. Dubois E, Gaffney D. The multiple facets of influence: identifying political influentials and opinion leaders on Twitter. *Am Behav Sci*. 2014;58(10):1260–77. <https://doi.org/10.1177/0002764214527088>.
13. Fan C, Jiang Y, Mostafavi A. The role of local influential users in spread of situational crisis information. *J Comput-Mediat Commun*. 2016;26(2):108–27. <https://doi.org/10.1093/jcmc/zmaa020>.
14. Fan R, Zhao J, Xu K. Topic dynamics in Weibo: a comprehensive study. *Soc Netw Anal Min*. 2015;5(1):41. <https://doi.org/10.1007/s13278-015-0282-0>.
15. Gefen D, Straub DW. Gender differences in the perception and use of e-mail: an extension to the technology acceptance model. *MIS Q*. 1997;21:389–400.
16. Greene R. Greene's groats-worth of wit, 1592. *Scolar P*; 1969.
17. González-Bailón S, Borge-Holthoefer J, Moreno Y. Broadcasters and hidden influentials in online protest diffusion. *Am Behav Sci*. 2013;57(7):943–65.
18. Guo L, Rohde JA, Wu HD. Who is responsible for Twitter's echo chamber problem? Evidence from 2016 US election networks. *Inf Commun Soc*. 2020;23(2):234–51. <https://doi.org/10.1080/1369118x.2018.1499793>.
19. Haro-de-Rosario A, Saez-Martin A, del Carmen Caba-Perez M. Using social media to enhance citizen engagement with local government: Twitter or Facebook? *New Media Soc*. 2018;20(1):29–49. <https://doi.org/10.1177/146144816645652>.
20. Hilbert M, Vasquez J, Halpern D, Valenzuela S, Arriagada E. One step, two step, network step? Complementary perspectives on communication flows in twittered citizen protests. *Soc Sci Comput Rev*. 2017;35(4):444–61.
21. Katz E. Where are opinion leaders leading us? *Int J Commun*. 2015;9:1023.
22. Katz E, Lazarsfeld PF. Personal influence: the part played by people in the flow of mass communications. New York: Free Press; 1955.
23. Katz E. The two-step flow of communication: an up-to-date report on an hypothesis. *Public Opin Q*. 1957;21(1):61–78.
24. Koh YJ, Sundar SS. Heuristic versus systematic processing of specialist versus generalist sources in online media. *Hum Commun Res*. 2010;36(2):103–24.
25. Lazarsfeld PF, Berelson B, Gaudet H. The people's choice: how the voter makes up his mind in a presidential campaign. New York: Duell, Sloan, and Pearce; 1944.
26. Leahey E, Moody J. Sociological innovation through subfield integration. *Soc Curr*. 2014;1(3):228–56. <https://doi.org/10.1177/2329496514540131>.
27. Li Q, Shah S, Liu X, Nourbakhsh A, Fang R. Tweetsift: tweet topic classification based on entity knowledge base and topic enhanced word embedding. In: Proceedings of the 25th ACM international on conference on information and knowledge management. ACM; 2016. p. 2429–32.
28. Morone F, Makse HA. Influence maximization in complex networks through optimal percolation. *Nature*. 2015;524(7563):65. <https://doi.org/10.1038/nature14604>.
29. Nass C, Reeves B, Leshner G. Technology and roles: a tale of two TVs. *J Commun*. 1996;46(2):121–8.
30. Norman DA, Bobrow DG. On data-limited and resource-limited processes. *Cogn Psychol*. 1975;7(1):44–64. [https://doi.org/10.1016/0010-0285\(75\)90004-3](https://doi.org/10.1016/0010-0285(75)90004-3).
31. Peng S, Cao L, Zhou Y, Ouyang Z, Yang A, Li X, Jia W, Yu S. A survey on deep learning for textual emotion analysis in social networks. *Digit Commun Netw*. 2022;8(5):745–62.
32. Quan-Haase A, Mendes K, Ho D, Lake O, Nau C, Pieber D. Mapping #MeToo: a synthesis review of digital feminist research across social media platforms. *New Media Soc*. 2021. <https://doi.org/10.1177/1461444820984457>.
33. Rogers EM. Diffusion of innovations. Simon and Schuster; 2010.
34. Tian S, Tao W, Hong C, Tsai WHS. Meaning transfer in celebrity endorsement and co-branding: meaning valence, association type, and brand awareness. *Int J Advert*. 2022;41(6):1017–37.
35. Treem JW, Leonardi PM. Recognizing expertise: factors promoting congruity between individuals' perceptions of their own expertise and the perceptions of their coworkers. *Commun Res*. 2017;44(2):198–224.
36. Van Aelst P, Walgrave S. Information and arena: the dual function of the news media for political elites. *J Commun*. 2016;66(3):496–518.
37. Walter S, Brüggemann M. Opportunity makes opinion leaders: analyzing the role of first-hand information in opinion leadership in social media networks. *Inf Commun Soc*. 2020;23(2):267–87. <https://doi.org/10.1080/1369118x.2018.1500622>.
38. Watts DJ, Dodds PS. Influentials, networks, and public opinion formation. *J Consum Res*. 2007;34(4):441–58. <https://doi.org/10.1086/518527>.
39. Weimann G. On the importance of marginality: one more step into the two-step flow of communication. *Am Social Rev*. 1982;47(6):764–73.

40. Wang Y, Wu P, Liu X, Li S, Zhu T, Zhao N. Subjective well-being of Chinese Sina Weibo users in residential lockdown during the COVID-19 pandemic: machine learning analysis. *J Med Internet Res*. 2020;22(12): e24775.
41. Wang CJ, Zhu JJ. Jumping over the network threshold of information diffusion: testing the threshold hypothesis of social influence. *Internet Res*. 2021. <https://doi.org/10.1108/INTR-08-2019-0313>.
42. Woo D, Pierce CS, Treem JW. Specialists over generalists? Examining discursive closures and openings in expert collaborations. *Commun Monogr*. 2021. <https://doi.org/10.1080/03637751.2021.1950917>.
43. Xiong Y, Cheng Z, Liang E, Wu Y. Accumulation mechanism of opinion leaders' social interaction ties in virtual communities: empirical evidence from China. *Comput Hum Behav*. 2018;82:81–93. <https://doi.org/10.1016/j.chb.2018.01.005>.
44. Xu X, Pratt S. Social media influencers as endorsers to promote travel destinations: an application of self-congruence theory to the Chinese Generation Y. *J Travel Tour Mark*. 2018;35(7):958–72.
45. Zhu JH. Issue competition and attention distraction: A zero-sum theory of agenda-setting. *J Q*. 1992;69(4):825–36.
46. Zhu Y, Fu KW. The relationship between interdisciplinarity and journal impact factor in the field of communication during 1997–2016. *J Commun*. 2019;69(3):273–97. <https://doi.org/10.1093/joc/jqz012>.
47. Zhang L, Zhao J, Xu K. Who creates trends in online social media: the crowd or opinion leaders? *J Comput-Mediat Commun*. 2016;21(1):1–16. <https://doi.org/10.1111/jcc4.12145>.
48. Zhao Y, Wang G, Yu PS, Liu S, Zhang S. Inferring social roles and statuses in social networks. In: Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining. ACM; 2013. <https://doi.org/10.1145/2487575.2487597>.

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
