Gatekeeping Twitter: message diffusion in political hashtags

Marco Toledo Bastos
LSE/University of São Paulo

Rafael Luis Galdini Raimundo
University of São Paulo

Rodrigo Travitzki
University of São Paulo

Abstract
This article explores the structure of gatekeeping in Twitter by means of a statistical analysis of the political hashtags #FreeIran, #FreeVenezuela and #Jan25, each of which reached the top position in Twitter Trending Topics. We performed a statistical correlation analysis on nine arrays of the dataset to evaluate if message replication in Twitter political hashtags was correlated with network topology. Our results suggest an alternative scenario to the hegemonic view regarding gatekeeping in Twitter political hashtags. Instead of depending on hubs that act as gatekeepers, we found that the intense activity of individuals with relatively few connections is capable of generating highly replicated messages that contributed to Trending Topics without relying on the activity of user hubs. The results support the thesis of social consensus through the influence of committed minorities, which states that a prevailing majority opinion in a population can be rapidly reversed by a small fraction of randomly distributed committed agents.

Corresponding author:
Marco Toledo Bastos, Rua Sta. Rosa Jr. 39, Ap.26, Butanta, São Paulo, 05579010, Brazil
Email: herrcafe@gmail.com

Gatekeeping in digital networks

The concept of gatekeeping was developed in applied mass communication research to address the topological bottleneck through which information was filtered. David Manning White (1950: 386) developed the concept in 1950 to describe Mr. Gates, a newspaper editor whose decisions were so subjective that he rejected one-third of incoming news items because of his personal feelings about the content. The concept was validated by Paul Snider, who repeated the study 17 years later and found that Mr. Gates was still rejecting about 90% of wire service copy (Snider, 1967: 402). These two studies gave shape to the concept of gatekeeping, a term originally coined by Kurt Lewin (1947) to describe a wife's role as the person who decides which food ends up on the family's dinner table. The gatekeeper was the person who chooses what should pass through each gate, of which there are several in any process. White (1964) seized upon Lewin's comments and applied them to journalism, using the word 'gatekeeping' to describe the decisions which determine what information will go forward and what will not.

Shoemaker defined gatekeeping as the process by which the billions of messages that are available in the world are winnowed down and transformed into the hundreds of messages that reach a given person on a given day (Shoemaker, 1991: 1). Defined as a process of selection, gatekeeping is consistently connected to the seminal description offered by Lewin (1947), who applied the concept to both interpersonal and mass communication (Shoemaker, 1991). On a very basic level, Lewin (1947) conducted experiments regarding group decisions and compared the process of being admitted to a channel to the process of accessing a 'gate'. Shoemaker (1991: 2) defined a gate as an in or out decision point, still in accordance with Lewin's (1951: 186) definition of gatekeeper as the controlling point between the channel and its external environment. This topological description of channels and gateways rendered gatekeeping a key concept for mass media research, first applied to editors and audiences in broadcasting media, and more recently to hubs and nodes in digital networks.

Nonetheless, the concept of gatekeeping is historically connected and methodologically tied to the broadcasting society. The overall design of the concept relied on highly centered networks that prevented the emergence of mechanisms for sharing information. Users who wanted to share content with other users had to deal with the high costs of production and distribution (as in print), or the use of scarce and expensive resources such as the electromagnetic spectrum (as in broadcasting). The concept of gatekeeping was sufficient to describe the control–communication infrastructure based upon sender–receiver roles and source–destination directions. But when digital networks superseded centralized wired networks, it became possible to communicate with millions of users at little or no cost. The previous sender-to-receiver role became increasingly obsolete, as the gated could also act as the source of information. Even information filtered by gatekeepers could be later redistributed or changed as it moved through the gateways. Hence, the traditional notion of source–destination was no longer a meaningful way to describe information control in information networks.

But gatekeeping remained an important concept for research about digital networks, and it continued to rely on the principle of a bottleneck of interconnections that determines the flow of information. However, the model of gatekeeping according to computing sciences was designed around a different set of operations (see Figure 1). In the
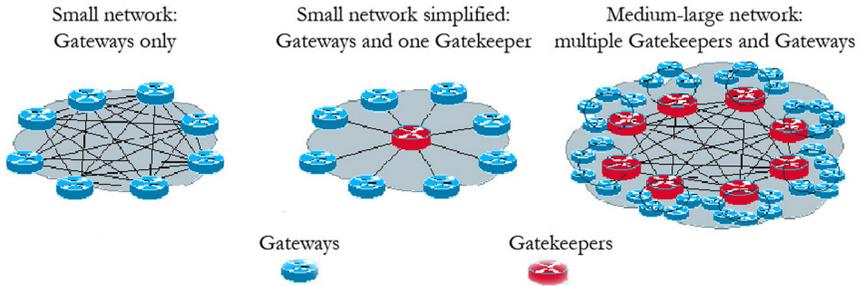
Figure 1. Understanding H.323 gatekeepers (Cisco, 2006)

computer sciences, gatekeepers are devices that manage domains and provide call control, supplying address information to terminals within the zone and to gatekeepers managing other zones. Gatekeepers control the gateways that, in telecommunication terms, are the devices interfacing two different zones and conveying the actual information. The concept of a gatekeeper as defined by mass media research better resembles what would be depicted in Figure 1 as a small network. As network scaling with gatekeepers grows into a medium-large network, the gatekeepers fade away and turn into a variant of another gateway. This overall design of the gatekeepers–gateways scaling is unknown in media research, as gatekeepers are described using topological features independent of network connectivity.

Media research describes gatekeepers only with reference to small networks. As network scaling grows to medium-large, gatekeepers turn themselves into gateways, as reporters, editors, witnesses, archives and readers add different elements to the same story and it is not always clear who is gatekeeping the final product. Gatekeeping is still a key mechanism in digital networks, only now it has been redesigned to incorporate a multitude of senders and receivers. Gatekeeping in digital networks occurs through a process in which a multitude of users pass information forward. Whereas gatekeeping in traditional media was performed by individual editors, in digital networks it embodies nodes taking part in the story and redesigning the process through which ideas and information are filtered for publication. What was before an internal decision-making process carried out by the media, which relayed to or withheld information from the public, is currently a decentralized process of following up on a story.

Gatekeeping on Twitter networks

Twitter is a free microblogging service that allows users to publish short messages, known as tweets, in a variety of ways. Users can post their tweets on the Twitter website or send SMS text-messages directly from their cell phones. Because Twitter enables real-time propagation of information to any number of users, the platform is an ideal environment for the dissemination of breaking news directly from the news source and/or from the geographical point of interest. Twitter network structure is reasonably simple. User *A* can follow *B* or any other user on the system, thus creating a social network in which user *A* receives the messages from the users he or she follows. Twitter being a large network

of users, it is the network of followers and followees that structures the channel through which messages are passed on to users. Twitter users with a large number of followers are identified as network hubs, which should account for the spreading of a message due to their specific positioning in the network. However, as we discuss further below, Twitter network topology is not correlated with the diffusion of messages. Gatekeeping might actually emerge as a function of the message frequency instead of the network structure, thus contradicting the description of gatekeepers as a bottleneck of interconnections that determines information flow.

We have investigated the connection between Twitter network structure and message diffusion. In order to address that, we analyzed the relationship between retweet network (RT), mention network (AT) and the followers' and followees' networks (FF). Retweets are posts that Twitter readers forward with full attribution to those who follow them, while the user's followers-and-followees network comprises a list of users who subscribe to one another's activity streams. Mentions are messages in which a specific Twitter user is mentioned using the at-sign (@), and even though these messages address specific receivers, they are also posted on the recipient's public page. Message diffusion within Twitter is heavily dependent on retweets, and because most retweets posted by a user are of tweets originally posted by someone the user follows (which can themselves be retweets), retweet activity reflects how the social network furthers the propagation of information. According to Suh et al. (2010), there is a very strong linear relationship between the number of followers and retweet rate. In other words, the larger the user's audience, the more likely that the tweet will be retweeted.

This is again the concept of gatekeeping based upon the principle that network topology determines message diffusion. Again, according to Suh et al. (2010), there is a close correlation between the tweet author's number of followers and the retweet rate. The more followers a Twitter user has, the more likely it is that their tweets will be retweeted. We tested this thesis with a dataset of Twitter political hashtags, which tend to be more persistent than other topics, thus assuring that once the message goes through the gate (once it goes viral), it remains in time and space. According to Romero et al. (2011), political hashtags are particularly persistent, with repeated exposures continuing to have large relative effects on adoption. The persistence has a significantly larger value than the average – in other words, successive exposures to a political hashtag have an unusually large effect relative to the peak. Once political hashtags become Trending Topics they tend to spread through the network in a way that corresponds with the complex contagion principle, which maintains that repeated exposures to an idea are of vital importance when the idea is controversial or contentious.

Previous studies (Bakshy et al., 2011; Huberman et al., 2009; Kwak et al., 2010) have shown that Twitter's topological features comprise a highly skewed distribution of followers and a low rate of reciprocated ties. Influence on Twitter was found to be connected to network topology, even though metrics such as the number of followers, page-rank, and number of retweets presented different results (Kwak et al., 2010; Wu et al., 2011). Bakshy et al. (2011) investigated the distribution of retweet cascades on Twitter and determined that although users with large follower counts and past success in triggering cascades were on average more likely to trigger large cascades in the future, these features were in general poor predictors of future cascade size. Jürgens et al. (2011)

analyzed political communication on Twitter and found out that users with a high number of message exchanges are in a position to exert strong, selective influence on the information passed within the network. Wu et al. (2011: 3) and Kwak et al. (2010: 10) made the point that Twitter does not conform to the usual characteristics of human social networks, which exhibit much higher reciprocity and far less skewed degree distributions, but instead better resembles a mixture of mass communication and face-to-face communication. Such topological features suggest that Twitter is a privileged system when it comes to analyzing gatekeeping in digital networks, as it better resembles an information-sharing network than a social network.

But Twitter challenges the traditional view on gatekeeping because users receive information from a plethora of distinct sources, most of which are not traditional media organizations. Even though media outlets are very active users on Twitter, only about 15% of tweets received by ordinary users are received directly from mass media channels (Wu et al., 2011: 5–6). Despite the fact that attention is fragmented and no longer centered upon media organizations, gatekeeping is a critical function, especially in view of the fact that less than 0.05% of users attract almost 50% of all attention in Twitter (Wu et al., 2011). If gatekeeping was previously identified with mass media channels, it is now shared among a number of unidentified elites who ensure that information flows have not become egalitarian. In order to further investigate that, we closely followed three political hashtags and mapped the social data related to each one. Hashtags are keywords or strings of text prefixed with a hash symbol (#) that can be clicked as a link to a global search of tweets using that same keyword. All hashtags in the dataset appeared in Twitter Trending Topics (TT), which is a feature that lists the most popular keywords being discussed on Twitter in real time. We performed a statistical analysis of retweet (RT), mention (AT) and the followers' and followees' network (FF). Our results show that gatekeeping in political hashtags is occurring through mechanisms that are not necessarily reliant on network topology. We expect the following analysis to shed some light on the ways gatekeeping is performed in Twitter networks.

#FreeIran, #FreeVenezuela and #Jan25

Political hashtags tend to exhibit particular network features, with a higher internal degree of nodes, a greater density of triangles and an overall larger number of nodes (Romero et al., 2011). An example of that is the political hashtags connected to the protests following the 2009 Iranian presidential election, clustered along with the hashtag #FreeIran. This pattern emerged again in the rising number of posts related to the socio-economic situation in Venezuela during the first days of February 2010, clustered along with the hashtag #FreeVenezuela. Both hashtags evolved in contexts of economic crisis, political uncertainty and threats to freedom of speech. We also analyzed the political hashtags related to the Arab Spring in Egypt, clustered along with the hashtag #Jan25, as a result of the wave of demonstrations and protests that took place in the Arab world in late 2010 and 2011. In the latter case, the dataset is significantly smaller due to the Egyptian networks being turned off by the government after the protests that erupted on 25 January.

The dataset spans approximately one and a half years of archiving processes. The #FreeIran dataset was archived over 575 days. The #FreeVenezuela dataset was collected over 373 days. The #Jan25 dataset is composed of 21 days of Twitter activity. Monitoring of the first two datasets continued over a longer time because the political agendas in Iran and Venezuela were not met, and therefore the hashtag and the archiving processes went on for several months. Because of the somewhat successful outcome of the Egyptian demonstrations, the #Jan25 hashtag fulfilled its purpose and was discontinued, having started on 21 January 2011 and then terminated on 10 February 2011. These three datasets together include 487,426 messages in total, including 353,518 unique messages and 133,908 redundant tweets (flooding). The datasets comprise a total of 23,702 messages that mention another user (AT-Network) and 212,982 messages that were echoed from another user (RT-Network). Twitter messages were retrieved from Twitter Streaming API, and the social graph was mapped by applying a one-step snowball sampling procedure based on users as seed nodes.

In order to analyze the gatekeeping function, we have focused on three types of networks within each hashtag. The first two networks refer to message diffusion and are the main channels for tweets to go viral. These are the network of mentions (AT-Network) and the network of retweets (RT-Network). The third network refers to Twitter's network of followers and followees (FF-Network), which we expected to play an important role in message replication. In order to evaluate whether message replication in Twitter political hashtags was correlated to the network topology, we performed a statistical correlation analysis on the following arrays of the dataset: retweet messages, mention messages, followers' connections, followees' connections, followers count, followees count and overall number of messages throughout the three datasets. This last feature accounts for the number of messages tweeted by single users regardless of their position in the network topology.

#FreeIran, #FreeVenezuela and #Jan25 are political hashtags that dominated Twitter's Trending Topics across the world through messages in up to 27 different languages. Eighty-three percent of the datasets from #FreeIran and #FreeVenezuela contain unique messages, while in #Jan25 the proportion is 55%. The same asymmetry was found in the number of retweets. While #FreeIran and #FreeVenezuela show numbers similar to each other (30% and 32%, respectively), #Jan25 shows a significantly higher number of retweets (60%) in the dataset. However, such differences tend to disappear in an in-depth analysis, which shows that of all unique messages in the three datasets, 8% of them were retweeted, thus implying that #Jan25 just had more redundant tweets (flood). The percentage of messages that include a mention (AT-Network) also varies little: 4.1% in #FreeIran, 5.7% in #FreeVenezuela and 3.8% in #Jan25. Mentions from a unique user to another unique user also show minor variations, being 2.7% in #FreeIran, 4.1% in #FreeVenezuela and 3.3% in #Jan25. However, network topology is not so symmetrical. #FreeIran, #FreeVenezuela and #Jan25 have, respectively, 19,599; 626,023; and 268,575 users interconnected as followers and followees. When we consider only users that are mutually followers of each other, homophily is considerably higher in #FreeIran (28%) and #FreeVenezuela (33%) than #Jan25 (16%).

In order to properly address gatekeeping in Twitter political hashtags, we focused on the data related to retweets. The action of retweeting is an emergent behavior in Twitter

that comprises the relaying of a tweet that has been written by another Twitter user. Retweeting makes it possible for a user to share a tweet written by another user with his or her followers. Retweeting can be understood as an instrument of information diffusion given that the original tweet is propagated to a new set of audiences, while the content of the message remains unaltered. This key mechanism for information diffusion in Twitter emerged as a simple way of passing on a message. The retweet rate is connected to the gatekeeping procedures, as it enables a message and the content it conveys to go viral in the Twitter mediascape. Suh et al. (2010) have shown that URLs and hashtags have strong relationships with retweetability, while suggesting that the number of followers and followees and the account creation date might also have an effect on retweetability. Moreover, Twitter literature found the amount of retweets to be the most effective technique for identifying intermediaries that transmit information to other users (Boyd et al., 2010; Cha et al., 2010; Kwak et al., 2010; Mendoza et al., 2010; Suh et al., 2010).

Retweet activity can also shed light on aspects of gatekeeping in Twitter. As previous studies have shown (Wu et al., 2011), elite users act as hubs in the network and they tend to generate more retweets than ordinary users. The difference is significant enough to qualify these users as new gatekeepers (Jürgens et al., 2011). Although gatekeeping as an effect of network topology might hold true for other sets of hashtags – and our data does indicate that, particularly for Twitter-idioms hashtags – in political hashtags our results show a different scenario. Political hashtags show a pattern that contrasts with the traditional concept of gatekeepers because gatekeeping is not based upon network topology. In fact, not only were the strongest statistically significant correlations with retweet rates found in the mention network (AT-Network) rather than the followers-and-followees network (FF-Network), but the strongest correlations were also primarily with the frequency of messages from ordinary users that were not active hubs. We measured this particular activity by the overall number of messages posted by users and the relation of these messages to the retweet rate. In Table 1 this activity is shown in the field 'Number of tweets per user'. A correlation coefficient of 0 demonstrates that the variables are independent; a correlation coefficient of 1 means that the two numbers are perfectly correlated.

#FreeIran and #FreeVenezuela show a statistically significant correlation between retweet rates and the number of tweets per user ($r = 0.83$ and $r = 0.73$, respectively, $p < 0.001$). As shown in Table 1, these are the forces that drove retweet activity in these hashtags, followed by messages that mentioned another user ($r = 0.62$ and $r = 0.33$, respectively, $p < 0.001$). That means the growing number of retweets is not correlated to network properties such as a high number of followers and followees, but mostly to the assiduous activity of ordinary users who were not necessarily hubs or elite users themselves. In #Jan25 the strongest correlation of retweet rates ($r = 0.83$, $p < 0.001$) is with messages that mentioned another user (AT-Network), which suggests that messages were retweeted at the same time that they mentioned another user. #Jan25 data shows that network topology played an important role in retweetability rates, as retweets have a significant correlation with the followers-and-followees network ($r = 0.40$, $p < 0.001$). Retweet rates and the number of tweets per user is only the third statistically significant correlation in #Jan25, as retweeted messages show a correlation coefficient of 0.28 ($p < 0.001$) with the number of tweets per user in the dataset. It is interesting to note that

Table 1. Pearson's correlation coefficients between retweets and network attributes in #FreeIran, #FreeVenezuela and #Jan25. Only significant correlations are given (*p* < 0.001).

	#FreeIran	#FreeVenezuela	#Jan25
	retweets	retweets	retweets
AT-messages sent	0.62	0.33	**0.83**
AT-messages received	0.22	0.33	0.13
FF-connection with followers	0.08	0.07	0.09
FF-connection with followees	0.15	0.08	0.40
User's number of followers	***	***	***
User's number of followees	***	0.03	***
Twitter account creation date	***	***	0.04
User's number of messages	0.09	0.07	0.06
Number of tweets per user	**0.83**	**0.73**	0.28

nowhere in the three hashtags was network topology the most important factor in determining message retweetability. Contrary to what traditional gatekeeping theories would suggest, the number of connections and the strategic position of users within these networks played a minor role during the process in which the hashtags #FreeIran, #FreeVenezuela and #Jan25 were going viral.

Based on these findings, we investigated additional political hashtags in order to test the hypothesis that highly active, non-hub users play a central role during the emergence of Twitter political hashtags. We investigated the hashtag #SOSNatal, which refers to the protests of Brazilian students against policies adopted by the local government. The dataset contains 6318 tweets and shows a statistically significant correlation between retweet rates and the frequency of tweets per user ($r = 0.68$, $p < 0.001$). Statistically significant correlations between retweet rates and the number of tweets per user were again found in the datasets of political hashtags #NaoFoiAcidente, comprising 15,124 tweets, and #AmandaGurgel, comprising 2356 tweets ($r = 0.67$ and 0.41, respectively, $p < 0.001$). We also tested this hypothesis by examining Twitter-idioms hashtags (Romero et al., 2011). Contrary to the results of our analysis of Twitter political hashtags, no statistically significant correlation between retweet rates and the frequency of tweets per user was found in the Twitter-idioms hashtags #WordsThatCanStartaWar, #ThingsTheDevilInvented and #IllPunchuInTheFaceIf ($r = 0.01$, $r = 0.08$ and $r = 0.02$, respectively, $p < 0.001$). These Twitter-idioms datasets contain, respectively, 102,278, 114,173 and 48,789 messages. In these Twitter-idioms hashtags the strongest statistically significant correlations with retweet rates were found with the followers-and-followees network (FF-Network) and mention network (AT-Network) instead of with the frequency of tweets per user, hence being compatible with the understanding of gatekeeping designed around network topology.

The correlation between the variables does not imply that one causes the other. However, the correlation between variables is a necessary condition for linear causation in the absence of any third variable. The case studies provide two significant conclusions about gatekeeping in Twitter political hashtags. The first of these is the non-existence of

a statistically significant correlation between retweet rate and the networks of followers and followees. The second indicates a statistically significant correlation between retweets and the number of tweets posted per user, thus suggesting that gatekeeping in Twitter political hashtags is not reliant on the network topology, but rather on the features and properties of the message itself.

Forms of digital gatekeeping

In spite of the variety of areas to which the concept of gatekeeping has been applied, it has consistently remained a function of the network topology. While studying Twitter political hashtags, we expected to confirm previous research that claims network topology is of high importance for message replication (Suh et al., 2010). However, the results presented herein show that the bottleneck created by hubs towards nodes, which topologically configure gatekeeping, is not of great importance to Twitter political hashtags. Results also show that a large number of followers (the followers count) or followees (the followees count) are of minor importance to message replication, while user activity and mention network seem to be the prevailing forces in message diffusion. In these cases, it is not the hubs which account for major message replication, but a few users whose activity is so intense that they end up accounting for the spreading of a message. Therefore, gatekeeping emerges in digital networks in a form not necessarily dependent on network topology, and the results suggest an alternative scenario to the understanding of gatekeeping in mass media research. Instead of depending on hubs that act as gatekeepers, political hashtags seemingly rely on highly active users with relatively few connections generating highly replicated messages.

This approach to gatekeeping is only possible because social networks undermined the validity of the dichotomy of mass versus interpersonal communication. Furthermore, social networks made it possible to compile large datasets regarding the massive circulation of information in media (Wu et al., 2011). Although the concept of gatekeeping is historically connected to the effects of mass communication, digital media is reshaping the concept by transforming a mass audience into smaller and atomized audiences. The research of Wu et al. (2011) pointed to this change, showing that the top ten most-followed users in Twitter were not corporations or media organizations but individuals who communicate directly with their followers via messages which were usually written by themselves or publicists. They thus bypass the intermediation of what are usually understood as gatekeepers. Still, according to the authors, Twitter has revealed a new class of individuals who often become more prominent than traditional public figures, such as entertainers or official gatekeepers. The function of such individuals is neither of broadcasting nor narrowcasting, but rather a form of directed-casting.

Even though the research of Wu et al. (2011) on Twitter network was broadly consistent with the original conception of the two-step flow, our results seem to partially contradict that. The two-step flow of communication theory emphasized that ideas flow from mass media to opinion leaders and from them to a wider audience. This general layout, first introduced by Katz and Lazarsfeld (1955), contrasted with the hypodermic needle model of direct effect from mass media, as it introduced the agency of opinion

leaders in the two-step flow model. This theoretical framework also supported gatekeeping theory by separating the opinion leaders from the audience, an interpretation that is not consistent with our results. Nonetheless, Twitter literature supports the thesis that gatekeepers have become increasingly atomized and fragmented, as users receive and pass on information without the mediation of media outlets. Traditional gatekeepers such as newspaper and media corporations are well represented in Twitter and their coverage is widely distributed throughout the environment, but only a small portion of tweets received by ordinary users comes from media outlets.

Despite the concept of gatekeeping being traditionally linked to a bottleneck of interconnections, our results support the hypothesis that gatekeeping in digital networks is shaped by a number of actions and channeling routines that are reliant not only upon network topology, but also upon message fitness. Unfortunately, the role of messages is usually handled as a secondary consideration in original models of gatekeeping (Roberts, 2005: 12). We found out that this is especially true for political messages, which are not neutral and which follow a planned agenda often immune to persuasion. As we have shown throughout this article, these messages enable a somewhat self-reliant gatekeeping mechanism. Our results also support the thesis of social consensus through the influence of committed minorities. According to Xie et al. (2011), a prevailing majority opinion in a population can be rapidly reversed by a small fraction of randomly distributed committed agents who consistently proselytize the opposing opinion and are immune to influence. As a result, the tipping point at which Twitter messages gain traction and go viral could be further investigated by looking at additional data regarding the relative fitness of Twitter political messages.

References

Bakshy E, Hofman J, Mason W et al. (2011) Everyone's an influencer: quantifying influence on Twitter. In: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining.* Hong Kong, China: ACM, pp. 65–74.

Boyd D, Golder S and Lotan G. (2010) Tweet, tweet, retweet: conversational aspects of retweeting on Twitter. Paper presented at Hawaii International Conference on System Sciences, Honolulu, 5–8 January.

Cha M, Haddadi H, Benevenuto F et al. (2010) Measuring user influence in Twitter: the million follower fallacy. In: *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media.* Washington, DC: Association for the Advancement of Artificial Intelligence.

Cisco (2006) Understanding H.323 gatekeepers. In: Cisco SystemsDocument ID 5244. Available at: http://www.cisco.com/en/US/tech/tk1077/technologies_tech_note09186a00800c5e0d. shtml (accessed November 2012).

Huberman B, Romero D and Wu F (2009) Social networks that matter: Twitter under the microscope. *First Monday* 14(1).

Jürgens P, Jungherr A and Schoen H (2011) Small worlds with a difference: new gatekeepers and the filtering of political information on Twitter. In: *Proceedings of the ACM WebSci '11*. New York: ACM.

Katz E and Lazarsfeld P (1955) *Personal Influence*. New York: Free Press.

Kwak H, Lee C, Park H et al. (2010) What is Twitter, a social network or a news media? In: *Proceedings of the ACM WebSci '11*. New York: ACM, pp. 591–600.

Lewin K (1947) Frontiers in group dynamics. *Human Relations* 1.

Lewin K (1951) *Field Theory in Social Science: Selected Theoretical Papers*. New York: Harper.

Mendoza M, Poblete B and Castillo C (2010) Twitter under crisis: can we trust what we RT? In: *1st Workshop on Social Media Analytics (SOMA '10).* Washington, DC: ACM.

Roberts C (2005) Gatekeeping theory: an evolution. Paper presented at Communication Theory and Methodology Division, Association for Education in Journalism and Mass Communication, San Antonio, TX, Augst.

Romero D, Meeder B and Kleinberg J (2011) Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter. In: *Proceedings of the 20th ACM International World Wide Web Conference*, Hyderabad, India, 28 March–1 April.

Shoemaker P (1991) *Gatekeeping*. Newbury Park, CA: Sage.

Snider PB (1967) Gates revisited – 1966 version of 1949 case study. *Journalism Quarterly* 44: 419–427.

Suh B, Hong L, Pirolli P et al. (2010) Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. In: *Proceedings of the IEEE Second International Conference on Social Computing (SocialCom)*, Minneapolis, MN, August.

White DM (1950) The 'gate keeper': a case study in the selection of news. *Journalism Quarterly* 27: 383–390.

White DM (1964) *People, Society and Mass Communications*, London: Collier-Macmillan.

Wu S, Hofman JM, Mason WA et al. (2011) Who says what to whom on Twitter. In: *Proceedings of the 20th ACM International World Wide Web Conference*, Hyderabad, India, 28 March–1 April.

Xie J, Sreenivasan S, Korniss G et al. (2011) Social consensus through the influence of committed minorities. *Physical Review E* 84.