Model for Twitter dynamics: Public attention and time series of tweeting

ARTICLE *in* PHYSICA A: STATISTICAL MECHANICS AND ITS APPLICATIONS · JUNE 2014

Impact Factor: 1.73 · DOI: 10.1016/j.physa.2014.02.034

5 AUTHORS, INCLUDING:

Ho Sung Kim
Sungshin Women's University
9 PUBLICATIONS **49** CITATIONS

Keumsook Lee
Sungshin Women's University
11 PUBLICATIONS **89** CITATIONS

MOO YOUNG CHOI
Seoul National University
230 PUBLICATIONS **3,376** CITATIONS

Physica A 404 (2014) 142–149

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Model for Twitter dynamics: Public attention and time series of tweeting

J. Ko [a], H.W. Kwon [b], H.S. Kim [c], K. Lee [d,*], M.Y. Choi [a,**]

[a] Department of Physics and Astronomy and Center for Theoretical Physics, Seoul National University, Seoul 151-747, Republic of Korea
[b] School of Physics, Korea Institute for Advanced Study, Seoul 130-722, Republic of Korea
[c] Department of Media Communication, Sungshin University, Seoul 136-742, Republic of Korea
[d] Department of Geography, Sungshin University, Seoul 136-742, Republic of Korea

HIGHLIGHTS

- We present a model for Twitter dynamics and extract the information-sharing tendency.
- A good measure for the general public attention is provided.
- The time evolution of tweeting, driven by articles on newspapers, is described well.
- Possibility of predicting the time evolution solely from external data is suggested.

ARTICLE INFO

Article history:
Received 18 October 2013
Received in revised form 11 December 2013
Available online 28 February 2014

Keywords:
Twitter dynamics
Mathematical model
Propensity to tweet or retweet
Information-sharing tendency
Mass media
Election

ABSTRACT

We present a simple mathematical model for the Twitter dynamics, and use the model to extract the information-sharing tendencies on two time scales, day and hour, about three contenders in the 2012 presidential election in South Korea. Comparison of the model results with actual data demonstrates that the information-sharing tendency on the day scale provides a good measure for the general public attention to the contenders, whereas the tendency on the hour scale reflects the daily cycle of twitter users. In addition, it is attempted to reproduce the time evolution of tweeting by taking the numbers of the articles on the online newspapers as the external driving to tweet, the validity of which is discussed.

1. Introduction

Sharing and transferring information, from useful knowledge to gossip, is an essential part of social life. In modern society, this needs are met by the social media, which allow people to share information more easily and rapidly [1]. Twitter, a microblogging service, is one of the popular social network services. In Twitter, a short text message, called a tweet, can be propagated by the forwarding function, called a retweet. The characteristics of Twitter and its applications are studied in various fields [1–8]. In recent years,a model for predicting how the number of tweets evolves with the propensity to tweet or retweet has been studied [9,10]. The propensity to tweet or retweet is related to the popularity of the topic [11]. The tendency to share information would be different according to the topics and varies with time. Here, through the use of social network services, the degrees of concern for particular topics are monitored virtually in real time.

* Corresponding author.
** Corresponding author. Tel.: +82 2 880 6615; fax: +82 2 884 3002.
E-mail addresses: kslee@sungshin.ac.kr (K. Lee), mychoi@snu.ac.kr (M.Y. Choi).

http://dx.doi.org/10.1016/j.physa.2014.02.034

One of the most interesting topics is the election event. In South Korea, the influence of Twitter on election has been addressed, especially after the Seoul mayor election in 2011 [12,13]. The information about candidates or their parties is very popular during the season of election [8]. Once the election is over, however, the interest tends to decrease sharply except some in the winner of the election. Influential candidates are mentioned extensively in the Twitter as well as in the mass media for political deliberation and campaign in the political arena. The case that tweets about a politician are retweeted many times should mean that he/she receives public attention much. In this manner, the information-sharing tendency, which may be interpreted as the propensity to retweet, reflects his or her popularity or influence, and the degree of public attention to a politician may be tracked by observing the information-sharing tendency.

While online news published by the media provides the main source of a new tweet, twitter users in general react to the news items with distinct degrees of influence [11,14]. For example, in case that the twitter users are biased to specific political positions, conservative or liberal, they would quote more articles from the corresponding media. Particularly, in South Korea, it is reported that the influence of liberal online media on Twitter is superior to those of conservative [15].

In this work, we present the mathematical model of tweeting, applied to the actual Twitter and online newspaper data during the 2012 election for the president of South Korea. The possibility of tracking the popular attention to candidates on different time scales is discussed. Predicting the time evolution of tweeting, we demonstrate the different influences of newspapers according to the political position. Specifically, it is revealed that the newspaper with a large circulation does not necessarily have large influence upon twitter users in South Korea.

2. Model

Tweets are categorized into new tweets and retweets, i.e., newly posted tweets and transferred tweets the original of which have already been posted by someone else. The total number N_t of tweets at time t is given by the sum

$$N_t(t) = N_n(t) + N_r(t), \tag{1}$$

where $N_n(t)$ and $N_r(t)$ are the numbers of new tweets and retweets at time t, respectively.

Suppose that new tweets are driven by external stimuli $E(t)$ such as the mass media [9]. Among the total tweets at time τ, some will be retweeted at some time or later while others never will. The ratio $r(\tau)$ of the former to the total tweets at time τ, which we call 'the information-sharing tendency', suggests how much the tweets are worth retweeting, thus providing a measure of the value to share or transfer the information [10]. With the information-sharing tendency, we write $N_n(t)$ and $N_r(t)$ in the form:

$$N_n(t) = \mathcal{N} \sum_{\tau=0}^{t} P_n(t-\tau)E(\tau), \tag{2}$$

$$N_r(t) = \sum_{\tau=0}^{t} P_r(t-\tau)r(\tau)N_t(\tau), \tag{3}$$

where $\mathcal{N}$ is the number of twitter users, and $P_n(t-\tau)$ and $P_r(t-\tau)$ represent the probabilities of a new tweet and of a retweet, respectively, at time t, given that the twitter user was driven at time τ (by an external stimulus and by a tweet, respectively). Since the information flow from the environment to the Twitter space is very fast, tweeting old news should not be meaningful. It is thus expected that $P_n(t-\tau)$ is appreciable only for short time duration $t-\tau$. Specifically, we assume that twitter users are influenced by online news for no longer than two days including the very day of the news, i.e., $P_n(t-\tau) = 0$ unless $t-\tau = 0$ or 1 in units of the day. Unlike new tweets, somewhat old tweets may also be retweeted. In fact it was reported that 55% of the retweeted tweets are retweeted in one hour and 75% are in 24 h. Further, the cumulative distribution of the time taken for a tweet to be retweeted was observed to follow the logarithmic form in general, albeit with a slight deviation from the logarithmic form for times shorter than 24 h [2]. Disregarding the slight deviation for simplicity, we adopt the logarithmic cumulative function at all times. The empirical results for P_r, together with the assumed form of P_n are summarized as follows:

$$P_n^{\mathrm{d}}(t-\tau) = q\delta_{t,\tau} + (1-q)\delta_{t,\tau+1}, \tag{4}$$

$$P_r^{\mathrm{d}}(t-\tau) = 0.75\delta_{t,\tau} + 0.16\log_{30}[1+(t-\tau)^{-1}](1-\delta_{t,\tau}) \tag{5}$$

$$P_r^{\mathrm{h}}(t-\tau) = 0.55\delta_{t,\tau} + 0.20\log_{24}[1+(t-\tau)^{-1}](1-\delta_{t,\tau}) \tag{6}$$

for $t \geq \tau$, where q describes the probability that the twitter user tweets on the very day of the driving (and $1-q$ the probability on the next day). Note that the superscript 'd' means that the associated probability function P has a time variable in units of the day and the superscript 'h' in units of the hour. Accordingly, Eq. (6) holds only for $0 \leq t-\tau < 24$ h.

In Ref. [10], variations of the tendency to tweet with the day of the week was taken into consideration by introducing the character function, multiplied on the right-hand side of Eq. (2). Empirically, the numbers of tweets on weekends or holidays reduce approximately to half of the numbers on weekdays, leading to the factor 1/2 [9]. In contrast, there is no character function here because the numbers of articles, which are introduced as the external stimuli, include already such variations. Indeed the measured ratio of the numbers of articles on weekends or holidays to those on weekdays takes the value 0.44 on average, for the data analyzed in this paper. Accordingly, we do not need to introduce the character function of the empirical nature.

Substituting Eqs. (4) and (5) into Eqs. (2) and (3), we obtain

$$N_n(t) = \mathcal{N}\left[qE(t) + (1-q)E(t-1)\right], \tag{7}$$

$$N_r(t) = 0.75r(t)N_t(t) + \sum_{\tau=0}^{t-1} 0.16 r(\tau)N_t(\tau)\log_{30}\left(1 + \frac{1}{t-\tau}\right) \tag{8}$$

which, upon substitution into Eq. (1), lead to the dynamics of the number of total tweets:

$$N_t(t) = \frac{4}{4-3r(t)}\left[q\mathcal{N}[E(t) + \gamma E(t-1)] + 0.16\sum_{\tau=0}^{t-1} r(\tau)N_t(\tau)\log_{30}\left(1 + \frac{1}{t-\tau}\right)\right] \tag{9}$$

with $\gamma \equiv (1-q)/q$ [10]. From the Twitter data, we can extract successively the information-sharing tendency:

$$r^{\mathrm{d}}(t) = \frac{4}{3}\left[f_r^{\mathrm{d}}(t) - 0.16\sum_{\tau=0}^{t-1} r(\tau)\frac{N_t(\tau)}{N_t(t)}\log_{30}\left(1 + \frac{1}{t-\tau}\right)\right], \tag{10}$$

$$r^{\mathrm{h}}(t) = \frac{20}{11}\left[f_r^{\mathrm{h}}(t) - 0.22\sum_{\tau=0}^{t-1} r(\tau)\frac{N_t(\tau)}{N_t(t)}\log_{24}\left(1 + \frac{1}{t-\tau}\right)\right], \tag{11}$$

where $f_r^{\mathrm{i}}(t) \equiv N_r(t)/N_t(t)$ ($i = \mathrm{d}, \mathrm{h}$) are the percentages of retweets among the total tweets on the corresponding time scale (i.e., day, hour) [10].

3. Public attention to candidates

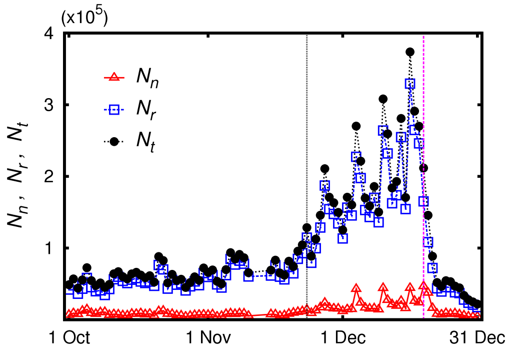
There were three strong contenders in the presidential election held on 19 December, 2012, in South Korea: "Park, Geun-Hye", "Moon, Jae-In", and "Ahn, Cheol-Soo". Two opposition contenders, Moon and Ahn, tried to unify opposition candidates. The negotiation did not progress smoothly; eventually, Ahn withdrew on 23 November, just before the due date for candidate registration. After that day, the presidential election race consisted of the competition between the two candidates. It then turned out that 75.8% of the electorate participated in voting: Park got 51.6% of the total valid votes and Moon 48.0% of the votes.

We have obtained the numbers of total tweets and of retweets addressing the names of contenders "Park", "Moon", and "Ahn" in Korean, from 1 October to 31 December, through the use of the keyword search program. Then the numbers of new tweets are obtained by subtracting the numbers of retweets from those of the total tweets. Fig. 1 shows the time-series of the numbers of total tweets, new tweets and retweets for each contender. The vertical black dotted line and the magenta dashed line indicate the day of Ahn's withdrawal and the election day, respectively. It is observed in Fig. 1(a) and (b) that the numbers of total tweets for Park and Moon exhibit four peaks after Ahn's withdrawal. The first peak on 27 November reflects that it was the first day of electioneering. The second (4 December), third (10 December), and fourth (16 December) peaks, arise from the TV debates between candidates. On the other hand, the number of tweets for Ahn reached a peak on his withdrawal day, 23 November, as shown in Fig. 1(c).

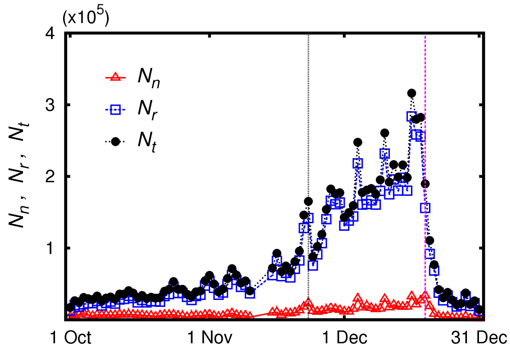
Fig. 2 shows how the information-sharing tendency r^{d} and the retweet percentage f_r^{d} varied with time t on the day scale. Recall that $r^{\mathrm{d}}(t)$ measures the fraction of tweets on day t to be retweeted at some time or later whereas $f_r^{\mathrm{d}}(t)$ describes merely the percentage of retweets among the total tweets on day t. Despite large variations of the total tweets (see Fig. 1), the fraction $f_r^{\mathrm{d}}(t)$ stays more or less constant. It is thus inferred that the mere retweet fraction on a given day does not serve as a good measure of the degree of public concerns about contenders. In contrast, the information-sharing tendency $r^{\mathrm{d}}(t)$ displays characteristic behavior on the days of big events. In particular, both $r^{\mathrm{d}}(t)$ for Park and $r^{\mathrm{d}}(t)$ for Moon show dramatic decrease right after the election day, as shown in Fig. 2(a) and (b), respectively, while they persist essentially as constants until the election day. On the other hand, the information-sharing tendency for Ahn in Fig. 2(c) exhibits a deep valley right after the day of his withdrawal. After that, the tendency $r^{\mathrm{d}}(t)$ gets recovered to some extent and decreases rapidly after the election, like that for other candidates. In fact, after the announcement of withdrawal, Ahn kept silent for a few days and then made a gesture of support of the other opposition candidate, Moon. Public attention to Ahn must have followed this trend and the change of $r^{\mathrm{d}}(t)$ apparently reflects this.

In general, the total tweet number N_t in Eq. (9) varies due to variations in the information inflow E as well as in the information-sharing tendency r. Note here that information sharing is sometimes not correlated with the information inflow. For instance, the information inflow (or the number of new tweets) increases since mid-November (see Fig. 1), whereas the information-sharing tendency does not change considerably until the election day (see Fig. 2). Therefore, tracking just the total tweet number does not inform how the information-sharing tendency varies.

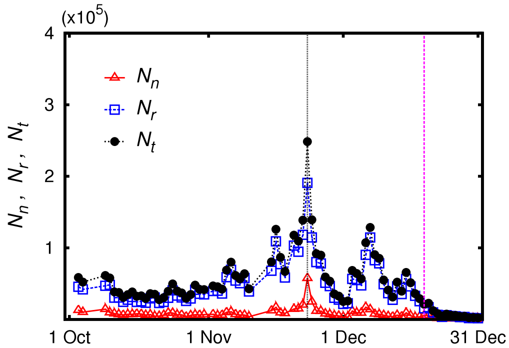
The information-sharing tendency on the day scale stays more or less constant in the election season; in contrast, that on the hour scale shows characteristic variations during a day, reflecting the daily cycle in the background, still exposing



(a) Park.



(b) Moon.



(c) Ahn.

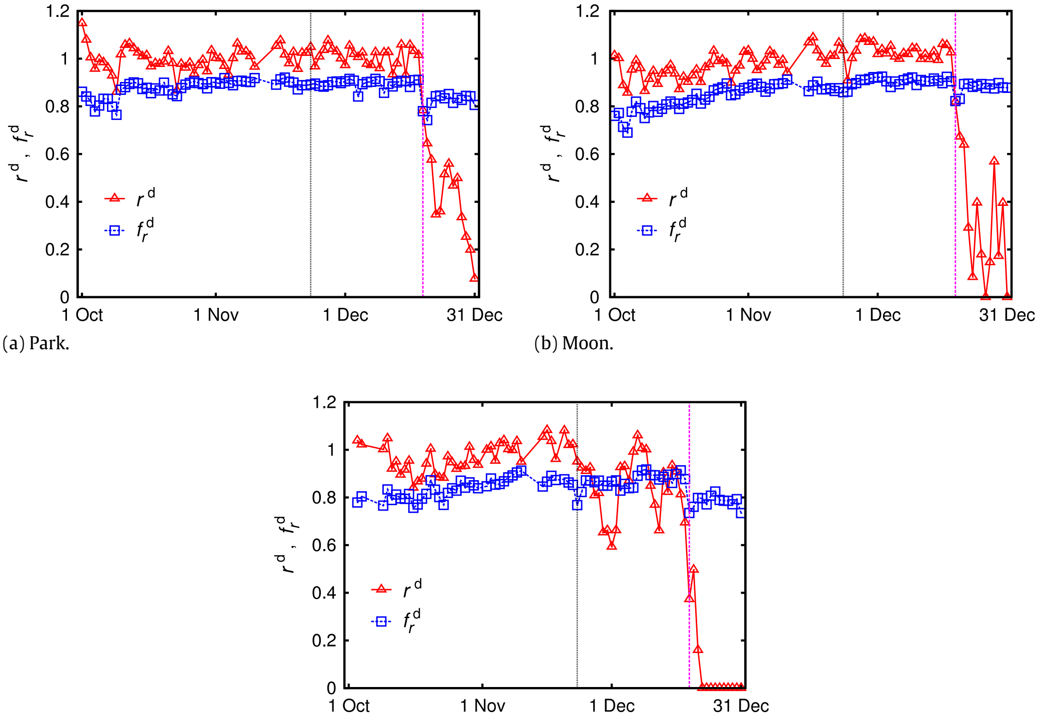
Fig. 1. Daily changes of the numbers of new tweets (red triangles), retweets (blue squares), and total tweets (black circles), searched by keywords (a) "Park, Geun-Hye", (b) "Moon, Jae-In", and (c) "Ahn, Cheol-Soo" in Korean, from 1 October to 31 December in the year 2012. The vertical black dotted line indicates the day 23 November, when Ahn, Cheol-Soo withdrew, whereas the vertical magenta dashed line indicates the election day, 19 December. Lines connecting data points are merely guides to the eye. The scales on the vertical axes are given in units of 10^5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

occurrence of big events. Fig. 3 shows the information-sharing tendency of three contenders (a) on the Ahns withdrawal day and (b) on the election day. In Fig. 3(a), the information-sharing tendencies are small during bedtime and reach the peaks at 21:00 (indicated by the dotted line) when Ahn announced the withdrawal. The largest value at 1:00 might be due to the TV debate for the unification of opposition candidates. On the election day, the information-sharing tendencies for three contenders are vanishingly small from 1:00 to 10:00, but that for the keyword vote (in Korean) is consistently large until the vote except bedtime, as shown in Fig. 3(b). The election law in Korea specifies that the online campaign and debate about candidates are prohibited on the very election day, while the vote stimulation is allowed. The vertical dashed line indicates the time when the result of the exit poll was announced. The information sharing tendencies of the two candidates reached the peaks at that moment.

In the presidential election, many key issues are picked out and catch attention usually during the electioneering period. People would have a chat together and discuss these issues often; this would lead to a large number of opinions posted on newspapers. Intuitively, however, fundamental interests and concerns, which are not influenced by the transient events, should persist unchanged until the election day. The mere number of retweets, that of total tweets, or the fraction f_r^{d} is not adequate for characterizing such fundamental interests. Instead, the information-sharing tendency r^{d} fits in with describing fundamental interests, although r^{h} varies according to the daily cycle. It is expected that the information-sharing tendency may be used for tracing the genuine public attention to politicians or other issues.

4. External driving to tweet: dependence on the media

To verify the validity of the model described by Eqs. (7) and (9), we apply the model to the real data. The numbers of articles are obtained from keyword search on the Korean portal site, and employed as the external driving $E(t)$. For each contender, we consider three kinds of numbers of articles: numbers of articles on the total online newspapers and on the two most representative Korean newspapers, "Chosunilbo" which is a conservative newspaper and "Hankyoreh" which is a liberal one. Taking these three kinds of $E(t)$ for each contender, we obtain the simulated time-series of tweets. Here $E(t)$ needs to be normalized and we choose the normalization in such a way that the sum of the simulated data for $N_t(t)$ during

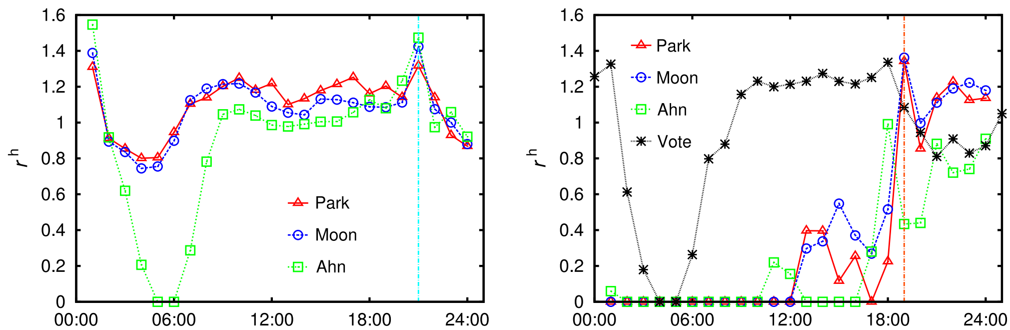


(a) Park.

(b) Moon.

(c) Ahn.

Fig. 2. Daily changes of the information-sharing tendency r^d (red triangles) and the retweet percentage f_r^d (blue squares) for three contenders. The vertical black dotted and magenta dashed lines indicate the day of Ahn's withdrawal and the election day, respectively. Lines connecting data points are merely guides to the eye. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) Ahn's withdrawal day.

(b) Election day.

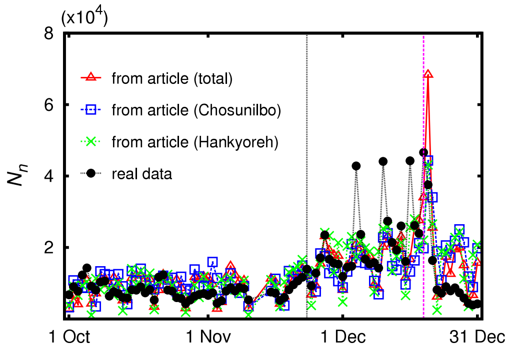
Fig. 3. Hourly changes of the information-sharing tendency r^h for three contenders and keyword 'vote'. The vertical black dotted line in (a) and the magenta dashed line in (b) indicate the time of announcement of Ahn's withdrawal and that of the result of the exit poll, respectively. Lines connecting data points are merely guides to the eye. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Normalization constant for external driving $E(t)$.

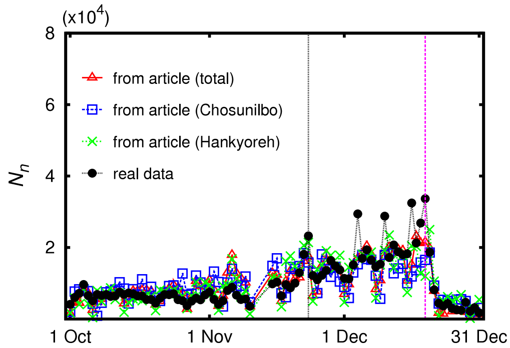
	Units: 10^{-6}		
	Park	Moon	Ahn
All (online)	2.29	1.83	2.02
Chosunilbo	181	138	143
Hankyoreh	175	157	156

the period equals the corresponding sum of real data. The resulting normalization constants, which are optimal in respect of the mean difference between the model results and the actual data, are listed in Table 1.

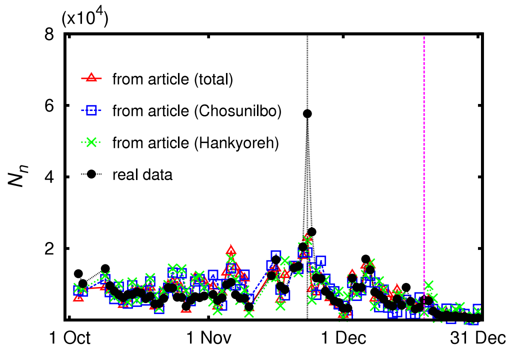
In Fig. 4 we present the numbers of new tweets, obtained from Eq. (7), together with the real data. Except the peaks on the days of the presidential TV debates, the model results exhibit behaviors indeed parallel with the ups and downs of the



(a) Park.



(b) Moon.



(c) Ahn.

Fig. 4. Numbers $N_n(t)$ of new tweets, obtained from the model with the numbers of articles on all the online press (red triangles), on "Chosunilbo" (blue squares), and on "Hankyoreh" (green crosses), chosen as $E(t)$, in comparison with actual data (black circles) for three contenders. The parameter values are $q = 0.9$ and $\mathcal{N} = 4 \times 10^6$, which is roughly the number of Korean users of Twitter. The vertical black dotted and magenta dashed lines indicate the day of Ahn's withdrawal and the election day, respectively. Lines connecting data points are merely guides to the eye. The scales on the vertical axes are given in units of 10^4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Root-mean-square deviations in the number of total tweets between the model results and actual data. The upper two rows correspond to the null model in Eq. (12) and our model in Eq. (9), respectively, with all (online) newspapers chosen as $E(t)$; the lower two rows to our model with "Chosunilbo" and "Hankyoreh" as $E(t)$, respectively.

	Park	Moon	Ahn
Null model (all)	68 851	45 563	23 979
Our model (all)	27 407	19 688	19 642
Our model (Chosunilbo)	34 008	26 671	19 570
Our model (Hankyoreh)	25 188	20 729	19 674

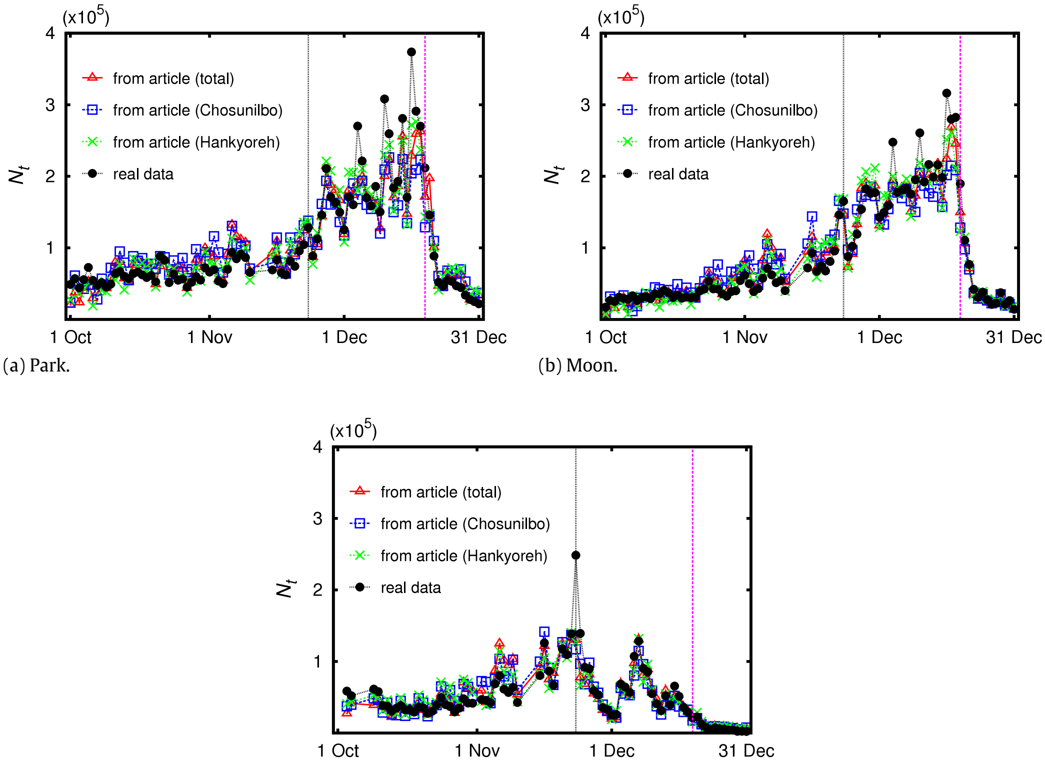
actual data. Obviously, this agreement does not prove that new tweets are driven only by the external media, for the cause and effect is not clear. Nevertheless, it is evident that these two numbers are highly related and the numbers of articles can be used as a measure of propensity to tweet a new one.

Substituting the information-sharing tendency $r(t)$ and three kinds of external driving $E(t)$ in Eq. (9), we obtain the simulated time evolution of the number $N_t(t)$ of total tweets and present the results, together with the actual data, in Fig. 5. The overall good agreement indicates that Eq. (9) mimics the real data indeed well.

To test further the validity of this model, we also consider a null model in which the total tweets are driven only by the external media:

$$N_t(t) = \alpha[E(t) + \gamma E(t-1)], \tag{12}$$

where α is an appropriate normalization constant. Since the normalization constants are set to make the mean difference in $N_t(t)$ between the model results and actual data vanish for each model with all newspapers chosen as $E(t)$, the relevance of each model, Eq. (9) or Eq. (12), may be probed, e.g., by the root-mean-square (rms) deviations in $N_t(t)$, displayed in Table 2. Comparison of the data in the upper two rows manifests that the rms deviations in the second row are far smaller than those in the first row. It is thus concluded that our model described by Eq. (9) gives significantly better agreement with the actual data than the null model in Eq. (12).



(a) Park. (b) Moon. (c) Ahn.

Fig. 5. Number $N_t(t)$ of total tweets, obtained from the model with the numbers of articles on all the online press (red triangles), on "Chosunilbo" (blue squares), and on "Hankyoreh" (green crosses), chosen as $E(t)$, in comparison with actual data (black circles) for three contenders. The information-sharing tendency $r(t)$ has been taken from Fig. 2 and parameter values are $q = 0.9$ and $\mathcal{N} = 4 \times 10^6$. The vertical black dotted and magenta dashed lines indicate the day of Ahn's withdrawal and the election day, respectively. Lines connecting data points are merely guides to the eye. The scales on the vertical axes are given in units of 10^5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Circulations (of conventional newspapers) and followers (of Twitter accounts) of "Chosunilbo" and "Hankyoreh". Circulations were investigated in 2011 [16] and numbers of followers were in June 26, 2013.

	Units: 10^3	
	Circulations	Followers
Chosunilbo	1353	33
Hankyoreh	211	155

In addition, we may also probe the relevance of different choices for $E(t)$ in our model, i.e., all the press, "Chosunilbo", or "Hankyoreh", by means of the rms deviations. The rms deviations in $N_t(t)$ between the model results and actual data for the three choices of $E(t)$ are listed in Table 2 (see the lower three rows). In general, for Park and Moon, all newspapers or "Hankyoreh" as $E(t)$ is observed to give better results than "Chosunilbo" whereas the three choices of $E(t)$ give similar results for Ahn. The fact that the conservative newspaper, "Chosunilbo", gives worse results than the liberal newspaper "Hankyoreh" or all the press, can be interpreted in two ways: One is that the keynote of "Chosunilbo" is more or less dislocated from the press in general. The other interpretation is that all the online media react on average much more to the event that are beneficial to the liberal parties. "Chosunilbo" has the largest circulation in Korea but its influence on twitter users is questionable, as shown in the results. Table 3 exhibits the gaps between circulations and numbers of followers of Twitter accounts of "Chosunilbo" and "Hankyoreh". Even considering that the number of followers does not reflect the influence as reported [2], the diffusion power of mass media should be reconsidered on the Twitter.

5. Summary

We have presented a mathematical model for the propensity to tweet and retweet, applied to the 2012 South Korean presidential election. The information-sharing tendency, measuring the value to share or transfer the information, is extracted from the actual Twitter data in different time scales. It provides the possibility of monitoring fundamental public attention to the contenders virtually in real time, thus suggesting many potential applications to politics and other fields. On the other hand, the mere numbers of tweets and retweets, or their ratio do not function as good measures for the essential attention of twitter users since such numbers are heavily influenced by the exposure to the external media and their ratio hardly varies around the election.

We have carried out simulations of the model, to obtain the time evolution of tweeting. It has been demonstrated that the overall results are consistent with the actual data. The motivation to tweet a new one appears to be related with the in-flow information from the external media. The effects of the external media have been analyzed and the difference between the external media has been identified: for the two major candidates, Park and Moon, the liberal newspaper "Hankyoreh" gives better results than the conservative one "Chosunilbo", which reveals that the newspaper with a large circulation does not necessarily have large influence, at least upon twitter users.

Finally, we address a few shortcomings of our model and discuss how to improve the model. In this study, the influence of Twitter on the mass media has not been considered although in reality some news is likely to flow from Twitter to the mass media. Presumably, the direction of the information flow can be measured by tracing the numbers of tweets and articles in a higher time resolution. Moreover, while only the numbers of tweets including names of contenders have been considered in the model, the contents of tweets are also important for measuring the trend of public opinion about politicians or any other things. In this direction, our model may be extended to take into account the wording of tweets; probing other motivations to tweet or retweet can also refine the model. These issues are left for future study.

Acknowledgments

This work was supported by the National Research Foundation of Korea Grant NRF-2013S1A5A2A01017764 and by the Sungshin University Research Grant of 2013 (K.L.) and Grants 2012R1A2A4A01004419 and 2011-0012331 through the Basic Science Research Program (M.Y.C.).

References

[1] A. Java, X. Song, T. Finin, B. Tseng, Why we twitter: understanding microblogging usage and communities, in: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, ACM, pp. 56–65.
[2] H. Kwak, C. Lee, H. Park, S. Moon, What is twitter, a social network or a news media? in: Proceedings of the 19th International Conference on World Wide Web, ACM, pp. 591–600.
[3] B.A. Huberman, D.M. Romero, F. Wu, Social networks that matter: twitter under the microscope, First Monday 14 (1) (2009).
[4] M. Cha, H. Haddadi, F. Benevenuto, K.P. Gummadi, Measuring user influence in twitter: the million follower fallacy, in: 4th International AAAI Conference on Weblogs and Social Media, Volume 14, p. 8.
[5] B.J. Jansen, M. Zhang, K. Sobel, A. Chowdury, Twitter power: tweets as electronic word of mouth, J. Am. Soc. Inf. Sci. Technol. 60 (2009) 2169–2188.
[6] T. Sakaki, M. Okazaki, Y. Matsuo, Earthquake shakes twitter users: real-time event detection by social sensors, in: Proceedings of the 19th International Conference on World Wide Web, ACM, pp. 851–860.
[7] J. Bollen, H. Mao, X. Zeng, Twitter mood predicts the stock market, J. Comput. Sci. 2 (2011) 1–8.
[8] A. Tumasjan, T.O. Sprenger, P.G. Sandner, I.M. Welpe, Election forecasts with twitter how 140 characters reflect the political landscape, Soc. Sci. Comput. Rev. 29 (2011) 402–418.
[9] H.W. Kwon, M.Y. Choi, H.S. Kim, K. Lee, Dynamic characteristics of tweeting and tweet topics, J. Korean Phys. Soc. 60 (2012) 590–594.
[10] H.W. Kwon, H.S. Kim, K. Lee, M.Y. Choi, Information-sharing tendency on twitter and time evolution of tweeting, Europhys. Lett. EPL 101 (2013) 58004.
[11] R. Bandari, S. Asur, B.A. Huberman, The pulse of news in social media: forecasting popularity, in: Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media, AAAI, pp. 26–33.
[12] T.-j. Kim, Twitter proves mighty force in election again, Koreatimes, 27 October [Online], 2011. Available at: http://koreatimes.co.kr/www/news/nation/2011/10/113_97457.html (Last accessed: 10.10.13).
[13] M.-h. Jung, Will social media be game changer, Koreatimes, 20 November [Online], 2012. Available at: http://koreatimes.co.kr/www/news/nation/2012/11/113_125136.html (Last accessed: 10.10.13).
[14] C. Castillo, M. El-Haddad, J. Pfeffer, M. Stempeck, Characterizing the life cycle of online news stories using social media reactions, e-print 2013. arXiv:1304.3010.
[15] Y. Hwang, A comparative study of information diffusion between conservative and liberal media, Korean J. Journalism Commun. Stud. (in Korean) 56 (2012) 350–372 (in Korean).
[16] B.G. Kim, Declination of the circulation of the daily newspapers (in Korean), Yonhap News, 14 December [Online], 2012. Available at: http://www.yonhapnews.co.kr/bulletin/2012/12/14/0200000000AKR20121214164900017.HTML (Last accessed: 10.10.13).