

# SCIENTIFIC REPORTS



OPEN

## Uncovering and Predicting the Dynamic Process of Collective Attention with Survival Theory

Peng Bao<sup>1</sup> & Xiaoxia Zhang<sup>2</sup>

The subject of collective attention is in the center of this era of information explosion. It is thus of great interest to understand the fundamental mechanism underlying attention in large populations within a complex evolving system. Moreover, an ability to predict the dynamic process of collective attention for individual items has important implications in an array of areas. In this report, we propose a generative probabilistic model using a self-excited Hawkes process with survival theory to model and predict the process through which individual items gain their attentions. This model explicitly captures three key ingredients: the intrinsic attractiveness of an item, characterizing its inherent competitiveness against other items; a reinforcement mechanism based on sum of each previous attention triggers; and a power-law temporal relaxation function, corresponding to the aging in the ability to attract new attentions. Experiments on two population-scale datasets demonstrate that this model consistently outperforms the state-of-the-art methods.

The subject of collective attention is central to an information era from knowledge database to online media, where millions of people are inundated with the explosive growth of user generated items<sup>1–3</sup>. In the heart of collective attention lies a competing process through which a few items become popular while most fade with time<sup>4–7</sup>. For example, papers increase their visibility by competing for citations from new papers<sup>8–10</sup>, tweets or Hashtags in Twitter become more popular as being re-tweeted<sup>11,12</sup>, videos on YouTube or stories on Digg gain their popularity by striving for views or votes<sup>13,14</sup>. Therefore, to understand the process underlying attention in large groups and predict the dynamic process of collection attention for individual items within a dynamically evolving system not only probes our understanding of complex systems, but also has important implications in a wide range of domains, including viral marketing, traffic control, public opinion monitoring, etc.<sup>15–18</sup>. However, to predict the dynamics of collective attention is challenging since numerous factors can affect the attention gathered by online content. Moreover, attention is very asymmetric and broadly-distributed<sup>19,20</sup>. Early studies devote to characterizing the distribution of the collective attention over an aggregation of user generated items<sup>21–23</sup> and making prediction on the final scale of attentions by exploiting temporal correlations<sup>24–26</sup>.

In recent year, there has been heightened research interest regarding the predictive modeling of the dynamics of collective attention for online content<sup>27,28</sup>. In general, current models fall into two main paradigms, each with known strengths and limitations. One focuses on making predictions by exploring relevant factors and applying standard regression or classification methods<sup>29–31</sup>. These models reveal many effective factors for prediction. However, there are still numerous factors to be investigated and they lack predictive power for the dynamics of collective attention for individual items. The other line of enquiry, in contrast, treats the dynamics as time series, making predictions by fitting these time series into certain class of functions<sup>32,33</sup>. Despite their initial success in certain domains, these models are deterministic and ignore the underlying arrival process of attentions. Recently, more sophisticated models have been proposed to simulate the dynamics of attentions for individual items, treating the diffusion process as a reinforced Poisson process<sup>34,35</sup> or a double stochastic process<sup>36</sup>. However, these models usually assume an aggregate stochastic process without distinguishing the triggering effects of different attentions in the diffusion-and-reaction process. Therefore, we still lack an effective method to uncover and predict the dynamics of collective attention.

In this report, we propose a generative probabilistic model using a self-excited Hawkes process with survival theory to model and predict the dynamic process through which individual items gain their attentions. This

<sup>1</sup>School of Software Engineering, Beijing Jiaotong University, Beijing, China. <sup>2</sup>School of Economics and Management, Tsinghua University, Beijing, China. Correspondence and requests for materials should be addressed to P.B. (email: [baopeng@bjtu.edu.cn](mailto:baopeng@bjtu.edu.cn))

model explicitly captures three key ingredients simultaneously: the intrinsic attractiveness of an item, characterizing its inherent competitiveness against other items; a reinforcement mechanism based on sum of each previous attention triggers, documenting the well-known “rich-get-richer” phenomenon; and a power-law temporal relaxation function, corresponding to the aging in the ability to attract new attentions. We validate the proposed model by applying it on two different types of population-scale datasets, of which one is a citation dataset, the other one a micro-blogging dataset. Experimental results demonstrate that our proposed model consistently outperforms the state-of-the-art methods on two datasets.

## Material and Methods

**Data description.** We use two population-scale datasets for this study, as follows.

- **APS:** It comprises the papers published in all the journals in American Physical Society from 1893 to 2009, consisting of 245,365 authors, 463,344 papers, and 4,692,026 citations. For each paper, the dataset includes title, DOI, PACS code, date of publication (day, month, year), names and affiliations of every author, a list of the previous papers cited, and so on (<http://journals.aps.org/datasets>).
- **WEIBO:** It is collected from the most popular micro-blogging service in China, namely Sina Weibo, which has more than 300 million registered users and generates about 100 million messages per day. Here we only use the messages that were originally posted between July 1, 2011 and July 31, 2011. There are 2.6 million messages. For each message, we collect its forwardings between July 1, 2011 and August 31, 2011 (<http://www.wise2012.cs.ucy.ac.cy/challenge.html>).

See Supplementary Section S1 for details.

**The model.** We now introduce the proposed generative probabilistic model from the perspective of individual items. Supposing that there are a set of time moments  $\{t_i\}$  ( $1 \leq i \leq N$ ) which denote the occurrence time of each attention for individual item  $d$  during observed time period  $[0, T]$ . Here,  $N$  is the total number of attentions. Without loss of generality, we have  $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_i \leq \dots \leq t_N \leq T$ . In this report, we model its dynamic process of attentions using a *self-excited Hawkes process*<sup>37</sup>, incorporating three key ingredients simultaneously: (1) attractiveness of an item, characterizing its inherent competitiveness against other items; (2) a reinforcement mechanism based on sum of previous attention triggers, capturing the well-known “richer-get-richer” phenomenon; (3) a general temporal relaxation function corresponding to the aging effect, characterizing time-dependent attractiveness of individual items. Taken these three factors together, for an individual item  $d$ , we model its dynamics of attentions characterized by the rate function  $\lambda(t)$  as

$$\lambda(t) = \mu + \sum_{0 < t_i < t} \varphi(t - t_i), \quad (1)$$

where  $\mu$  is the intrinsic attractiveness of the item,  $\varphi(\tau)$  is the relaxation function that characterizes the temporal inhomogeneity due to the aging effect. The explicit form of  $\varphi(\tau)$  will be investigated in the following section. Our model generalizes the reinforcement function as the sum of each previous attention triggers with time decaying, instead of the total count of attentions<sup>34,35</sup>.

The length of time interval between two consecutive attentions follows a self-excited Hawkes process. Therefore, given that the  $(i - 1)$ -th attention arrives at  $t_{i-1}$ , the probability that the  $i$ -th attention arrives at  $t_i$  follows

$$p(t_i | t_{i-1}) = e^{-\int_{t_{i-1}}^{t_i} \lambda(t) dt} \lambda(t_i), \quad (2)$$

which is the product of the survival and hazard functions. Specifically, the survival function  $e^{-\int_{t_{i-1}}^{t_i} \lambda(t) dt}$  captures the probability that no attention arrives in the interval  $(t_{i-1}, t_i)$ , and the hazard function  $\lambda(t_i)$  captures the instantaneous rate of the  $i$ -th attention arrives at  $t_i$ . Similarly, because there is no attention arrives between  $t_N$  and  $T$ , the probability can be written as

$$p(T | t_N) = e^{-\int_{t_N}^T \lambda(t) dt}. \quad (3)$$

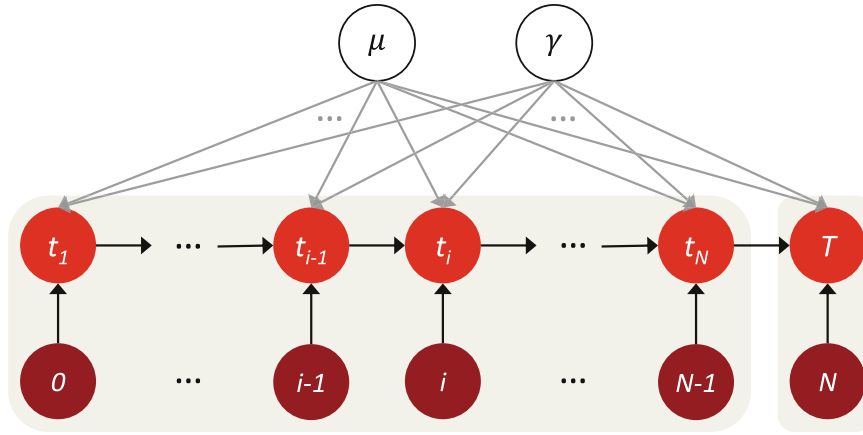
Assuming that attentions during different time intervals are statistically independent, by incorporating equation (2) and (3), the likelihood of observing the dynamics  $\{t_i\}$  during time interval  $[0, T]$  follows

$$L = p(T | t_N) \prod_{i=1}^N p(t_i | t_{i-1}). \quad (4)$$

For clarity, we illustrate the proposed model in the graphical representation, as shown in Fig. 1.

## Results

**Empirical validation of power-law temporal relaxation function.** The temporal relaxation function  $\varphi(\tau)$  can be measured directly from the real data. As shown in the rate function in equation (1), the temporal dynamics of an item is controlled by three forces, which are difficult to separate from each other. Hence to determine the specific form of temporal relaxation function, we need to control the other factors, isolating the temporal decay. To achieve this we should group items with the same attractiveness and cumulative attentions, and look at the time when they receive next attention. However, we do not know the attractiveness beforehand. Therefore,



**Figure 1.** Graphical representation of the generative probabilistic model. Here  $\{t_i\}$  ( $1 \leq i \leq N$ ) is a set of time moments which denote the occurrence time of each attention for an individual item during observed time period  $[0, T]$ .  $\mu$  and  $\gamma$  represent the parameters in the model. This generative probabilistic model explicitly models the arrival process of attentions.

by aggregating different items, we will measure a superposition of different temporal relaxation functions. We therefore select papers published between 1950 and 1970 in the APS dataset with fixed cumulative citations  $N_c$  and track the moment when their citations changed from  $N_c$  to  $N_c + 1$ . We denote  $\Delta t$  the time interval between two consecutive attentions and measure  $\Delta t$  in years, i.e. years passed when  $N_c \rightarrow N_c + 1$  took place. Here  $P(\Delta t|N_c)$  is the probability that a paper gets cited after time  $\Delta t$  elapsed with fixed cumulative citations  $N_c$ , capturing a paper's attractiveness to the research community. In Fig. 2a and b, we present the distribution of  $P(\Delta t|N_c)$  for fixed  $N_c = 10$  and  $N_c = 20$ . We find that  $P(\Delta t|N_c)$  roughly follows a power-law distribution with an exponent 2.11 for  $N_c = 10$  and an exponent 2.03 for  $N_c = 20$  respectively, indicating that collective attention is allocated in a rather asymmetric way, with a burst of rapidly arriving attentions followed by long periods of no attention.

In addition, it is similar to measuring  $P(\Delta t)$  from empirical data in the WEIBO dataset. We roughly consider messages which are posted in a fixed time period and receive the same number of forwardings in one hour after being posted as having the same attractiveness. By selecting messages with same number of forwardings, the reinforcement is also controlled. Therefore we select messages posted between 10 am and 12 am with fixed cumulative number of forwardings  $N_f$  in the first hour after being posted, and track the moment when their number of forwardings changed from  $N_f$  to  $N_f + 1$ . Note that we measure  $\Delta t$  in minutes to track the time interval due to the granularity of time scale<sup>26,34</sup>. As shown in Fig. 2c and d,  $P(\Delta t|N_f)$  also displays a power-law distribution with an exponent 1.37 for  $N_f = 10$  and an exponent 0.88 for  $N_f = 20$  respectively.

The result reflects the emergence of bursty human behaviors<sup>4</sup>, exhibiting the temporal nature of collective attention. Meanwhile, although the dynamic behaviors on both datasets obey the power-law temporal scaling, the power exponents are quite different. Therefore, we need to assign an item-specific exponent to capture the inhomogeneous aging effect among individual items. Hence in this report, we model the aging effect by adopting a power-law temporal relaxation function for individual items as follows

$$\varphi(\tau) \propto \tau^{-\gamma}. \tag{5}$$

Note that  $\gamma$  is one of the parameters of our proposed model, which characterizes the item-specific aging effect and can be estimated by maximum likelihood estimation methods in the section below.

**Parameter estimation and prediction.** By substituting the power-law temporal relaxation function in equation (5) into the general rate function in equation (1), we can get the specific form of rate function for dynamics  $\{t_k\}$  as

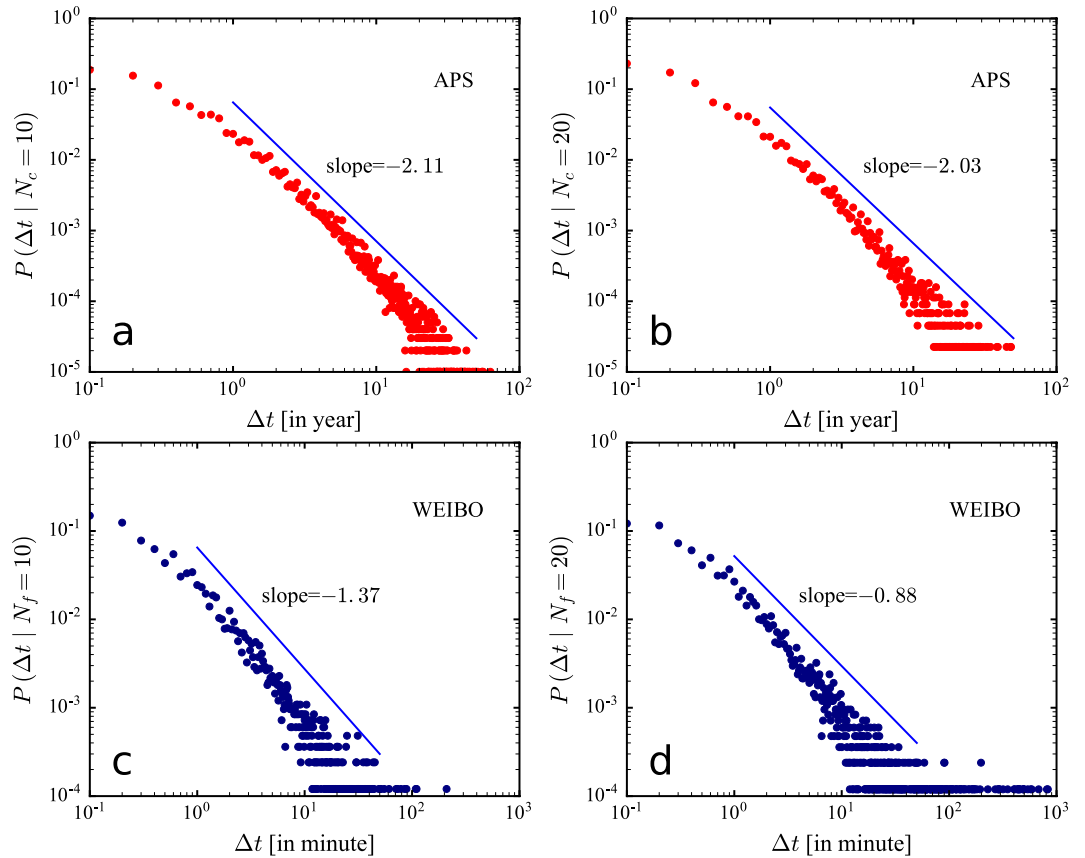
$$\lambda(t) = \mu + \sum_{0 < t_i < t} (t - t_i)^{-\gamma}. \tag{6}$$

Next, by substituting equation (6) into the likelihood function in equation (4) and taking logarithm, we can get the log-likelihood for the dynamics  $\{t_k\}$  up to  $T$  as

$$\ell = \ln \left[ p(T|t_N) \prod_{i=1}^N p(t_i|t_{i-1}) \right] = \sum_{i=1}^N \ln \lambda(t_i) - \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \lambda(t) dt - \int_{t_N}^T \lambda(t) dt = \frac{1}{1-\gamma} X - \mu T, \tag{7}$$

where  $X = \sum_{i=1}^N ((1-\gamma) \ln(\mu + \sum_{t_j < t_i} (t_i - t_j)^{-\gamma}) - (T - t_i)^{1-\gamma})$

Then we utilize maximum likelihood estimation methods to estimate the parameters in the proposed model. For parameter  $\{\mu, \gamma\}$ , the optimal values can be found by maximizing the log-likelihood in equation (7) using the gradient ascent method. See Supplementary Section S2 for details.



**Figure 2.** Empirical validation of temporal relaxation function. **(a)**  $P(\Delta t)$  when the number of citations  $N_c$  change from 10 to 11 in the APS dataset. **(b)**  $P(\Delta t)$  when the number of citations  $N_c$  change from 20 to 21 in the APS dataset. **(c)**  $P(\Delta t)$  when the number of forwardings  $N_f$  change from 10 to 11 in the WEIBO dataset. **(d)**  $P(\Delta t)$  when the number of forwardings  $N_f$  change from 20 to 21 in the WEIBO dataset. We find that the temporal relaxation function roughly follows a power-law distribution on two datasets. Note that a smaller power exponent indicates a slower decaying speed of the attractiveness of items.

Here, we denote the optimal values for parameters  $\{\mu, \gamma\}$  as  $\{\mu^*, \gamma^*\}$ . With the obtained parameters, the model can be used to predict the expected number of attentions gathered by item  $d$  up to any given time  $t$ , which is denoted as  $c(t)$ . Incorporating with the rate function in equation (6), for  $t > T$ , we treat the prediction task as the following differential equation

$$\frac{dc(t)}{dt} = \mu + \sum_{0 < t_i < t} (t - t_i)^{-\gamma} \tag{8}$$

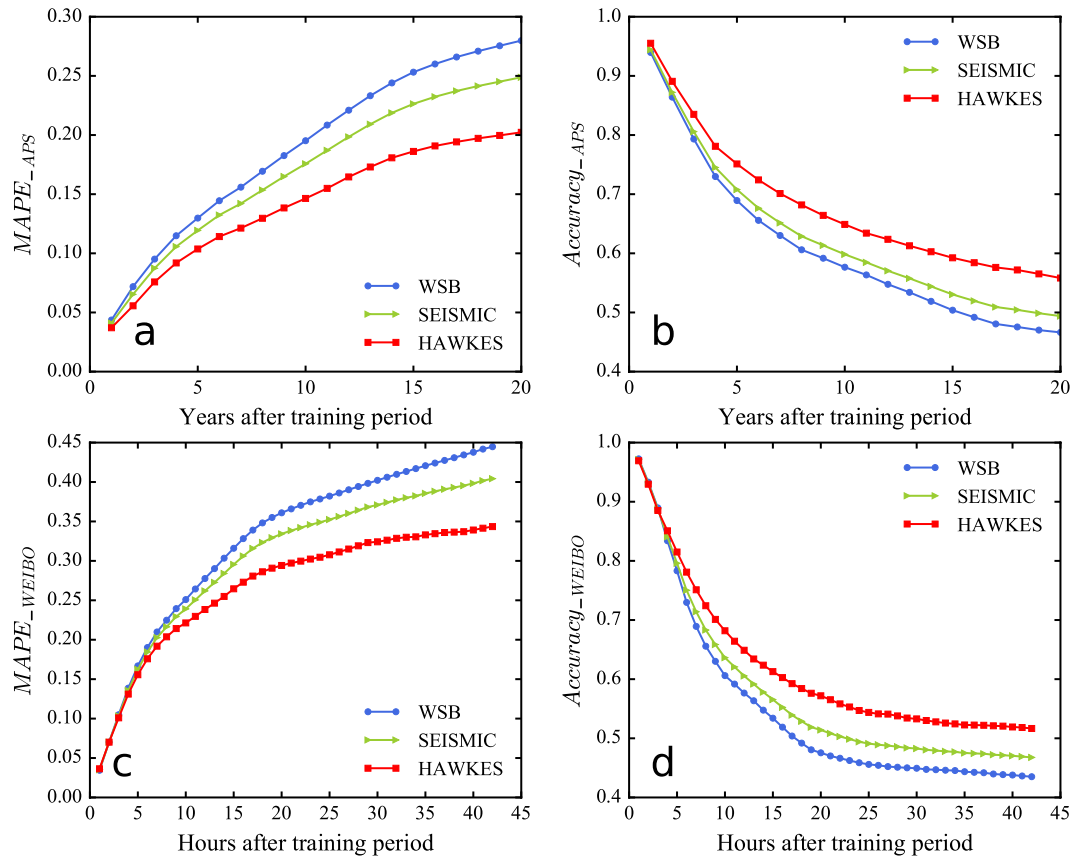
with the boundary condition  $c(T) = N$ . Solving this differential equation, we obtain the prediction function

$$c(t) = N + \mu^*(t - T) + \sum_{0 < t_i < t} \frac{1}{1 - \gamma^*} ((t - t_i)^{1-\gamma^*} - (T - t_i)^{1-\gamma^*}). \tag{9}$$

**Experiment results.** To compare the predictive power of our proposed model against other models, we introduce two widely-used models that have been used or can be used to model and predict the dynamics of collective attention: the WSB model<sup>34</sup> and the SEISMIC model<sup>36</sup>. See Supplementary Section S3 for details.

In order to validate the prediction performance of all the prediction models, we utilize two evaluation metrics: Mean Absolute Percentage Error (*MAPE*) and *Accuracy*. Let  $c_d(t)$  be the observed number of attentions for an item  $d$  up to time  $t$ , and  $\hat{c}_d(t)$  be the predicted value.

- *MAPE* measures the average deviation between the predicted and empirical number of attentions over an aggregation of items. For a dataset of  $D$  items, the *MAPE* is defined as



**Figure 3.** Prediction comparison for different models. **(a)** MAPE of model prediction with respect to the 1st to 20th year after the training period in the APS dataset. **(b)** Accuracy of model prediction with respect to the 1st to 20th year after the training period in the APS dataset. **(c)** MAPE of model prediction with respect to the 1st to 42nd hour after the training period in the WEIBO dataset. **(d)** Accuracy of model prediction with respect to the 1st to 42nd hour after the training period in the WEIBO dataset. We find that the proposed model (HAWKES) consistently outperforms the state-of-the-art methods (WSB and SEISMIC), exhibiting lower error **(a,c)** and higher accuracy **(b,d)**.

$$MAPE = \frac{1}{D} \sum_{d=1}^D \left| \frac{\hat{c}_d(t) - c_d(t)}{c_d(t)} \right|.$$

- Accuracy measures the fraction of items, correctly predicted under a given error tolerance  $\varepsilon$ . Specifically, the Accuracy of prediction over  $D$  items is defined as

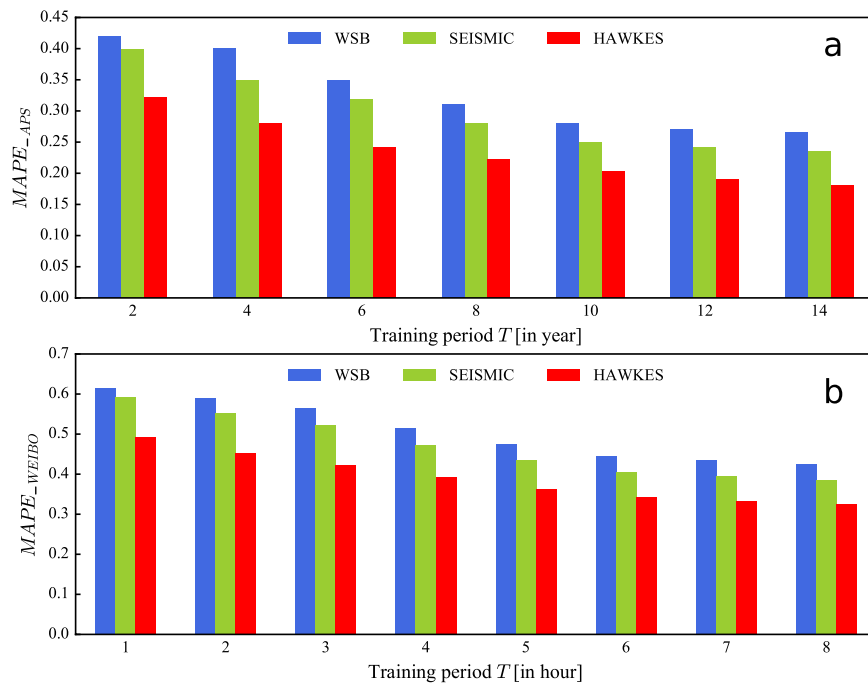
$$Accuracy = \frac{1}{D} \sum_{d=1}^D \mathbb{I} \left[ \left| \frac{\hat{c}_d(t) - c_d(t)}{c_d(t)} \right| \leq \varepsilon \right].$$

where  $\mathbb{I}[X]$  is an indicator function which return 1 if the statement  $X$  is true and 0 otherwise. In this report, the threshold  $\varepsilon$  is set as 0.1.

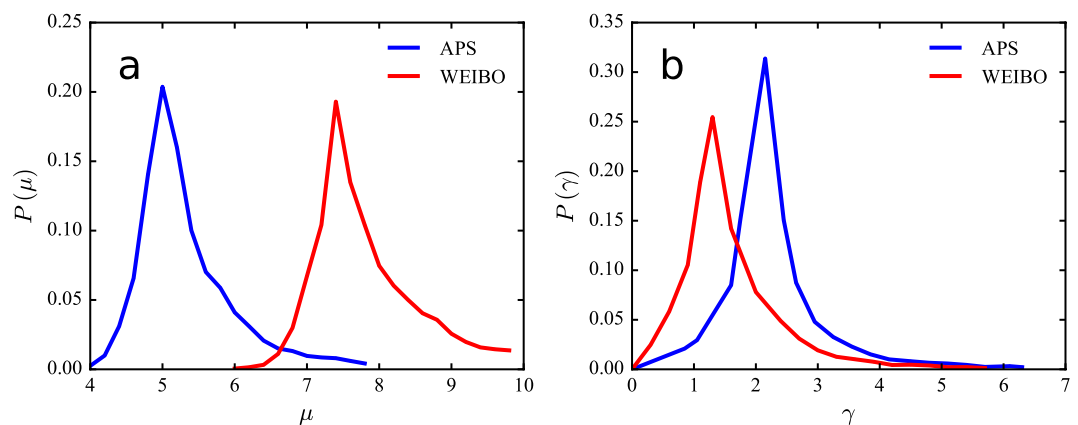
Therefore, for the APS dataset, we set the training period  $T$  as 10 years and then predict the number of citations for each paper from the 1st to 20th year after the training period. Similarly, for the WEIBO dataset, the training period is 6 hours and we predict the number of forwardings for each message from the 1st to 42nd hour after the training period.

Figure 3 shows the comparison results of these models with respect to different prediction time on two datasets. We find that the proposed model consistently outperforms the state-of-the-art methods, exhibiting lower error (Fig. 3a and c) and higher accuracy (Fig. 3b and d).

Furthermore, we carry out extensive experiments on two datasets to examine the prediction performance of different models when the training period varies. To be specific, we apply these models on the APS dataset with the training period varying from 2 to 14 years. And we fix the prediction time  $t$  to be 20 years after publication. For the WEIBO dataset, we change the training period from 1 to 8 hour. Since most messages in the dataset stop receiving more forwarding after being posted for 48 hours<sup>25</sup>, we fix the prediction time  $t$  to be 48 hours to check



**Figure 4.** Prediction comparison for different models with the variation of training period. **(a)** *MAPE* of model prediction with respect to different length of training period in the APS dataset. **(b)** *MAPE* of model prediction with respect to different length of training period in the WEIBO dataset. We find that the proposed model (HAWKES) performs the best on the entire range of training period on two datasets. It also shows that the *MAPE* decreases as the training period increases.



**Figure 5.** Distribution of model parameters in two datasets. **(a)** Distribution of learned intrinsic attractiveness  $\mu$  of items in two datasets. It shows that most values of  $\mu$  lie around 5 in the APS dataset and 7.5 in the WEIBO dataset respectively, indicating that the average intrinsic attractiveness of messages on micro-blogging network is higher than that of papers in citation network. **(b)** Distribution of learned power-law exponent  $\gamma$  in the temporal relaxation function in two datasets. We find that most values of  $\gamma$  lie around 1.30 in the APS dataset and shift to around 2.15 in the WEIBO dataset, indicating that the average decaying speed of the attractiveness of messages in micro-blogging network is slower than that of papers in citation network.

the ability for different models in predicting the final number of received attentions. We use *MAPE* to measure the prediction performance.

Are shown in Fig. 4, for all models, the *MAPE* decreases as the training period increases on two datasets, indicating that increasing the training period can improve the prediction performance for all the models. More importantly, we find that the proposed model performs the best on the entire range of training period on two datasets, indicating the effectiveness of the proposed model. In addition, we can also see that the rate at which *MAPE* declines slows down quickly. This means the marginal gain for performance improvement diminishes with the increasing of the training period.

**Analysis of model parameters.** In our model, there are in total two parameters  $\{\mu, \gamma\}$  and they are derived from the model learning process. Here we investigate the characteristics of the learned parameters.

Figure 5a illustrates the distribution of intrinsic attractiveness parameter  $\mu$  of items in two datasets. We observe that most values of  $\mu$  lie around 5 in the APS dataset and 7.5 in the WEIBO dataset respectively, indicating that the average intrinsic attractiveness of messages on micro-blogging network is higher than that of papers in citation network. Moreover, for the exponent of power-law temporal relaxation function, parameter  $\gamma$ , as shown in Fig. 5b, most values of  $\gamma$  lie around 1.30 in the APS dataset. Nevertheless, most values of  $\gamma$  shift to around 2.15 in the WEIBO dataset. Note that a smaller  $\gamma$  indicates a slower decaying speed of the attractiveness of items. This means that the average decaying speed of the attractiveness of messages in micro-blogging network is slower than that of papers in citation network. One possible explanation for these findings is that micro-blogging system, a typical type of social media for sharing and spreading information, can help messages improve their visibility and prolong their lifespan through a variety of features<sup>38, 39</sup>.

## Discussion

In this report, we propose a general framework to model and predict the dynamic process of collective attention. Our main contribution are three-folds: (1) We proposed a generative probabilistic framework and employed a self-excited Hawkes process to captures the triggering effect of each attention, distinguishing itself from the existing deterministic approaches; (2) We investigated three key ingredients for the dynamics of collective attention and combined them into the proposed model: the intrinsic attractiveness of an item, a reinforcement mechanism corresponding to the “rich-get-richer” effect, and a power-law temporal relaxation function explaining the aging effect in attracting new attentions; (3) We validated the proposed model by applying it on two population-scale datasets. Experimental results demonstrate that the proposed model consistently outperforms the state-of-the-art methods. We hope that this study will provide us richer understanding of the fundamental mechanism of information diffusion and shed light on the collective attention of online human behavior, paving ways towards better management of online content.

The proposed model is flexible, being able to incorporate exogenous information to improve its accuracy. To show this, we consider the inhomogeneous influence between individual attentions. Note that we employ the PageRank score as the influence of a paper in the APS dataset and the logarithmic of the number of a user’s followers in the followship network to represent its influence in the WEIBO dataset (See Supplementary Section S4). We find that when we incorporate the inhomogeneous influence between individuals, the accuracy increases. Therefore, if exogenous information is available, our method can absorb that, improving its predictive power.

There are still a few limitations on the proposed method. Although the overall performance is very well, it does not hold for some abnormal dynamic processes with specific patterns (by using machines, zombie followers, etc.). In addition, maximum likelihood parameter estimation suffers from the over-fitting problem for small sample size. Both of these are very interesting and we will try to solve them in our future work.

A long list of extensions can be conducted based on our findings. Examples include thorough investigation of the effect of the choice of temporal relaxation function, deep exploration on the interplay between the dynamics of collective attention and the structural characteristics of the networks spanned by early adopters, i.e., the users who view or forward the item in the early stage of dissemination. Moreover, it is also an interesting research topic to analyze the effect of the inhomogeneous influence among individuals. In addition, one is also encouraged to enrich the proposed model by incorporating more factors such as, different network behavior for particular types of content extracted from the item itself. More broadly, one is also encouraged to investigate the potential connection between the theoretical approach applied in this paper and the revolution occurring in physics with an increasing interest for renewal processes and ergodicity breaking<sup>40–42</sup>.

## References

- Lazer, D. *et al.* Computation social science. *Science* **323**, 721–723, doi:10.1126/science.1167742–723 (2009).
- Barabási, A. L. The network takeover. *Nat. Phys.* **8**, 14–16, doi:10.1038/nphys2188 (2012).
- Muchnik, L., Aral, S. & Taylor, S. J. Social influence bias: a randomized experiment. *Science* **341**, 647–651, doi:10.1126/science.1240466 (2013).
- Barabási, A. L. The origin of bursts and heavy tails in human dynamics. *Nature* **435**, 207–211, doi:10.1038/nature03459 (2005).
- Petersen, A. M., Tenenbaum, J. N., Havlin, S., Stanley, H. E. & Perc, M. Languages cool as they expand: Allometric scaling and the decreasing need for new words. *Sci. Rep.* **2**, 943, doi:10.1038/srep00943 (2012).
- Perc, M. Evolution of the most common English words and phrases over the centuries. *J. R. Soc. Interface* **9**, 3323–3328, doi:10.1098/rsif.2012.0491 (2012).
- Perc, M. The Matthew effect in empirical data. *J. R. Soc. Interface* **11**, 20140378–20140378, doi:10.1098/rsif.2014.0378 (2014).
- Perc, M. Self-organization of progress across the century of physics. *Sci. Rep.* **3**, 1720, doi:10.1038/srep01720 (2013).
- Kuhn, T., Perc, M. & Helbing, D. Inheritance patterns in citation networks reveal scientific memes. *Phys. Rev. X* **4**, 041036 (2014).
- Sinatra, R., Wang, D., Deville, P., Song, C. & Barabási, A. L. Quantifying the evolution of individual scientific impact. *Science* **354**, 596, doi:10.1126/science.aaf5239 (2016).
- Romero, D. M., Meeder, B. & Kleinberg, J. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the 20th International Conference on World Wide Web*, 695–704 (2011).
- Cheng, T. & Wicks, T. Event detection using Twitter: a spatio-temporal approach. *PLoS ONE* **9**, e97807, doi:10.1371/journal.pone.0097807 (2014).
- Lerman, K. & Hogg, T. Using a model of social dynamics to predict popularity of news. In *Proceedings of the 19th International Conference on World Wide Web*, 621–630 (2010).
- Pinto, H., Almeida, J. M. & Gonçalves, M. A. Using early view patterns to predict the popularity of youtube videos. In *Proceedings of the 6th ACM International Conference on Web Search and Data Mining*, 365–374 (2013).
- Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200–3203, doi:10.1103/PhysRevLett.86.3200 (2001).
- Leskovec, J., Adamic, L. A. & Huberman, B. A. The dynamics of viral marketing. *ACM Transactions on the Web* **1**, 5–es, doi:10.1145/1232722 (2007).

17. Watts, D. J. & Dodds, P. S. Influentials, networks, and public opinion formation. *Journal of Consumer Research* **34**, 441–458, doi:10.1086/518527 (2007).
18. Song, C., Qu, Z., Blumm, N. & Barabási, A. L. Limits of predictability in human mobility. *Science* **327**, 1018–1021, doi:10.1126/science.1177170 (2010).
19. Salganik, M., Dodds, P. & Watts, D. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* **311**, 854–856, doi:10.1126/science.1121066 (2006).
20. Wang, C. & Huberman, B. A. How random are online social interaction? *Sci. Rep.* **2**, 633, doi:10.1038/srep00633 (2012).
21. Lü, L., Chen, D. B. & Zhou, T. The small world yields the most effective information spreading. *New J. Phys.* **13**, 123005, doi:10.1371/journal.pone.0077455 (2011).
22. Gleeson, J. P., Cellai, D., Onnela, J., Porter, M. A. & Reed-Tsochas, F. A simple generative model of collective online behavior. *Proc. Natl. Acad. Sci.* **111**, 10411–10415, doi:10.1073/pnas.1313895111 (2014).
23. Delvenne, J., Lambiotte, R. & Rocha, L. E. C. Diffusion on networked systems is a question of time or structure. *Nat. Commun.* **6**, 7366, doi:10.1038/ncomms8366 (2015).
24. Szabo, G. & Huberman, B. A. Predicting the popularity of online content. *Communications of the ACM* **53**, 80–88, doi:10.1145/1787234 (2010).
25. Bao, P., Shen, H. W., Huang, J. & Cheng, X. Q. Popularity prediction in microblogging network: a case study on sina weibo. In *Proceedings of the 22nd International Conference on World Wide Web*, 177–178 (2013).
26. Gao, S., Ma, J. & Chen, Z. Modeling and predicting retweeting dynamics on microblogging platforms. In *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*, 107–116 (2015).
27. Ratkiewicz, J., Fortunato, S., Flammini, A., Menczer, F. & Vespignani, A. Characterizing and modeling the dynamics of online popularity. *Phys. Rev. Lett.* **105**, 15870, doi:10.1103/PhysRevLett.105.158701 (2010).
28. Gomez-Rodriguez, M., Leskovec, J. & Schölkopf, B. Modeling information propagation with survival theory. In *Proceedings of the 30th International Conference on Machine Learning*, 666–674 (2013).
29. Ugander, J., Backstrom, L., Marlow, C. & Kleinberg, J. Structural diversity in social contagion. *Proc. Natl. Acad. Sci.* **109**, 5962–5966, doi:10.1073/pnas.1116502109 (2012).
30. Cheng, J., Adamic, L., Dow, A., Kleinberg, J. & Leskovec, J. Can cascades be predicted? In *Proceedings of the 23rd International Conference on World Wide Web*, 925–936 (2014).
31. Anderson, A., Huttenlocher, D., Kleinberg, J., Leskovec, J. & Tiwari, M. Global diffusion via cascading invitations: structure, growth, and homophily. In *Proceedings of the 24th International Conference on World Wide Web*, 66–76 (2015).
32. Crane, R. & Sornette, D. Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Natl. Acad. Sci.* **105**, 15649–15653, doi:10.1073/pnas.0803685105 (2008).
33. Matsubara, Y., Sakurai, Y., Prakash, B. A., Li, L. & Faloutsos, C. Rise and fall patterns of information diffusion: model and implications. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 6–14 (2012).
34. Wang, D., Song, C. & Barabási, A.-L. Quantifying long-term scientific impact. *Science* **342**, 127–132, doi:10.1126/science.1237825 (2013).
35. Shen, H. W., Wang, D., Song, C. & Barabási, A. L. Modeling and predicting popularity dynamics via reinforced poisson processes. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence* **345**, 291–297, doi:10.1126/science.1248961 (2014).
36. Zhao, Q., Erdogdu, M. A., He, H. Y., Rajaraman, A. & Leskovec, J. SEISMIC: a self-exciting point process model for predicting tweet popularity. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1513–1522 (2015).
37. Hawkes, A. G. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83–90, doi:10.1093/biomet/58.1.83 (1971).
38. Kwak, H., Lee, C., Park, H. & Moon, S. What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, 591–600 (2010).
39. Bakshy, E., Rosenn, I., Marlow, C. & Adamic, L. The role of social networks in information diffusion. In *Proceedings of the 21st International Conference on World Wide Web*, 519–528, doi:10.1145/2187836 (2012).
40. Allegrini, P., Bologna, M., Fronzoni, L., Grigolini, P. & Silvestri, L. Experimental quenching of harmonic stimuli: universality of linear response theory. *Phys. Rev. Lett.* **103**, 030602, doi:10.1103/PhysRevLett.103.030602 (2009).
41. Grigolini, P. Emergence of biological complexity: Criticality, renewal and memory. *Chaos, Solitons and Fractals* **81**, 575–588, doi:10.1016/j.chaos.2015.07.025 (2015).
42. Geneston, E., Tuladhar, R., Beig, M. T., Bologna, M. & Grigolini, P. Ergodicity breaking and localization. *Phys. Rev. E* **94**, 012136, doi:10.1103/PhysRevE.94.012136 (2016).

## Acknowledgements

This work was funded by the Fundamental Research Funds for the Central Universities under grant number 2015RC031 and the State Visiting Scholar Funds from the China Scholarship Council under grant number 201607095027. This work was also supported in part by National Natural Science Foundation of China under grant number 61370128.

## Author Contributions

P.B. designed research. P.B. and X.Z. performed experiments, P.B. and X.Z. wrote and reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-02826-6

**Competing Interests:** The authors declare that they have no competing interests.

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017