Temporal Modeling of Information Diffusion using MASEP: Multi-Actor Self-Exciting Processes

Bowen Zhang

The Chinese University of Hong Kong
zb016@ie.cuhk.edu.hk

Wing Cheong Lau

The Chinese University of Hong Kong
wclau@ie.cuhk.edu.hk

ABSTRACT

Predicting the popularity of a discussion topic in an online social network (OSN) or the responses to an online fund-raising campaign is a practical challenge of immense value. Previous work tries to predict the popularity of an online campaign by modeling information diffusion as a homogeneous temporal point process within a network of a single-type of actors. However, real-world information propagation often involved multiple types of actors. In particular, there are the so-called opinion leaders, e.g. online celebrities or influential OSN users with a huge number of followers, who can create a great impact on the visibility and thus the final popularity of an event by simply mentioning it in their tweets or postings. In this paper, we propose MASEP, a Multi-actor Self-exciting Process, to model and predict the popularity of different online campaigns involving multiple types of actors. MASEP combines a self-exciting branching process with a periodical decay process to capture the dynamics and interdependent relationship between opinion leaders and ordinary users during an online campaign. A closed-form expression is derived for the temporal campaign popularity under the MASEP model. Based on this closed-form expression, we can efficiently perform regression against the empirical activity measurements of an online campaign during its early stage to estimate the parameters of the corresponding MASEP model. The final popularity of the campaign can then be predicted. To demonstrate the efficacy of the MASEP-based approach, we apply it to predict the popularity of three types of online campaigns from different large-scale real-world datasets, namely, the total number of posts in retweeting cascades, the overall count of individual hashtags in posting streams, and the final number of sponsors for crowd-funding campaigns. In particular, using the initial 30% of each campaign data trace for training, our approach can achieve absolute prediction error (APE) of 13.25%, 15.7%, and 36.9% respectively for datasets of 3 different types of campaigns. This corresponds to a 26.1% to 63.2% reduction in prediction error when comparing to state-of-the-art approaches including SEISMIC, SpikeM, and STRM.

CCS CONCEPTS

• Information systems → Data mining;

KEYWORDS

Popularity prediction; Multi-actor network; Temporal point process

This paper is published under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW '18 Companion, April 23–27, 2018, Lyon, France

© 2018 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC BY 4.0 License.

ACM ISBN 978-1-4503-5640-4/18/04.

<https://doi.org/10.1145/3184558.3191634>

ACM Reference Format:

Bowen Zhang and Wing Cheong Lau. 2018. Temporal Modeling of Information Diffusion using MASEP: Multi-Actor Self-Exciting Processes. In *WWW '18 Companion: The 2018 Web Conference Companion, April 23–27, 2018, Lyon, France*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3184558.3191634>

1 INTRODUCTION

OSN services, such as Twitter, Weibo, and Facebook, allow their users to share posts in the form of tweets, images, and videos. In OSN, there are at least two types of actors, namely, the so-called opinion leaders such as online celebrities or news agencies, and the ordinary users of the OSN service [6]. On one hand, opinion leaders, serving as trend-setters, often have much larger influence on the propagation of information than ordinary users. On the other hand, an ordinary user would help the further dissemination of information through his/ her social circles after exposed to the posts of opinion leaders or other ordinary users he/ she follows. In this way, the information of an online event can be spread.

A key challenge of immense practical value is to predict the number of posts related to an online event as time progresses. Accurate predictions of the popularity of online events can enable effective content aggregation, data management and ultimately improve the user experience of the OSN service [14]. Most existing prediction schemes use different types of homogeneous temporal point process to model the dynamics of information propagation among ordinary users within an OSN and neglect the impact of the highly influential opinion leaders. As such, previous work fails to provide accurate predictions on the popularity of online events which involved multiple types of actors. In this paper, We propose the **Multi-Actor Self-Exciting Process (MASEP)** by combining a self-exciting branching process with a periodical decay process to model the dynamics and the interdependent relationship between opinion leaders and ordinary users. In summary, the technical contributions of this paper include:

- A qualitative analysis on the empirical temporal behavior of 3 different types of online campaigns using large-scale real-world datasets.
- The proposed MASEP model which uses two heterogeneous and intertwined temporal point processes to capture the dynamics and interactions of opinion leaders and ordinary users during an online information propagation event.
- Derivation of a closed-form expression for the temporal popularity of a MASEP-modeled online event. The resultant expression enables fast parameter estimation and accurate prediction of the final popularity of the online event.

This work is supported in part by an RGC-CUHK Direct Grant, Project number 4055076.

Figure 1: The underlying diffusion topologies of different information propagation models.

- Extensive evaluation/ validation of the proposed prediction scheme based on large-scale empirical datasets.

The rest of this paper is organized as follows: Section 2 introduces the related work, and Section 3 discusses the underlying diffusion topologies of existing temporal models. Section 4 talks about the temporal behavior of the popularity of an online campaign. Section 5 introduces the MASEP model. Section 6 shows our experimental evaluation based on three real-world datasets. We conclude the paper in Section 7.

2 RELATED WORK

Recently, point-process-based methods have been introduced to predict the popularity of online events. Previous work [3, 4, 8, 11–14] focuses on the spreading of information in a network of ordinary users and models the resultant dynamics as a homogeneous temporal point process. Different classical point processes including the Poisson process, the Hawkes process, and the epidemiology process, have been used to model the social phenomena that modulate the information diffusion. Additional effects like aging [14], ‘rich-get-richer’ phenomenon, or so-called reinforcement [12], and the periodic nature of activities [4, 8] have also been considered. Among the existing types of approaches, the two most successful (and popular) ones are: i) self-exciting-process-based methods (SEISMIC [14], TiDeH [4], and STRM [11], etc.) which try to capture the bursty and cascading effect of information diffusion in OSNs; and ii) Susceptible-Infected model (SI model)-based schemes such as SpikeM [8] which aim to characterize the rise-and-fall pattern of information diffusion. Other methods have also been proposed to account for the multiple components of OSN activities. For instance, Coevolve [2] and SLANT [1] use two multi-variate Hawkes point processes to model information diffusion under evolving network structures/user opinions. HIP [10] combines a Hawkes process with an impulse response function: the former is for the modeling of single-actor-type information propagation while the latter aims to characterize continuous stimulation from external sources/ media.

MASEP differs from the above work in two major ways. First, most of the prior work [3, 4, 8, 11–14] deals with information diffusion within networks consisting of a single-type of (ordinary) users. MASEP extends the model to cover networks with multiple actor types and captures the interdependent dynamics among them. Second, MASEP performs online event popularity prediction solely based on activities observed on the target OSN without relying on additional measurements or data of the external simulating sources.

3 DIFFUSION TOPOLOGY OF ONLINE EVENTS

The diffusion topology of an online event is a graph which visualizes the properties of the temporal behavior over fixed and discrete time intervals, like each hour or each day, of an information diffusion process. Figure 1(a) and 1(b) depict the diffusion topology of the self-exciting model and the SI model, respectively. The diffusion topology of a self-exciting model forms a tree. Each node represents a user in the network. An online campaign starts from the seed user corresponding to the root of the tree and spreads to other users in the network. The child node can only have access to the information after the parent node gets infected, i.e. resharing the post or participating in the online campaign. Consider a discrete time series $\{r(t)\}$ over N observing intervals where $r(t)$ (for $t = 1, 2, \dots, N$) represents the number of nodes which are newly infected during the t -th interval. Let τ be the number of time intervals between the arrival of a post in a node’s timeline and the resultant (triggered) response performed by the node, e.g. a retweet/ reshare of the post, or becoming a sponsor of a crowd-funding campaign. Let $P(\tau)$ be the corresponding probability mass function. Common choices of $P(\tau)$ include the power-law function [9, 14] and the exponential function [4, 11]. In other words,

$$P(\tau) = \lambda(\tau + c)^{-(1+\theta)} \quad \text{or} \quad P(\tau) = \lambda e^{-\theta\tau}$$

where λ , θ , and c are the shape parameters whose value may depend on various factors, including the quality of the online event content, the social network structure, and the influence of infected users, etc. Based on the diffusion topology of the self-exciting model, we can estimate the future values of $r(t)$. Denote such estimate by $\hat{r}(t)$, which is given by:

$$\hat{r}(t) = \lambda_0(t) + \alpha \sum_{n \in I(t)} P(t - t_n), \quad (1)$$

where $\lambda_0(t)$ is the number of converted nodes due to external simulation at time interval t . α is the scaling factor. $I(t)$ is the set of users who already got infected before time t . In the sequel, we will also refer $\hat{r}(t)$ as the predicted popularity of the online campaign/ event at time interval t .

Alternatively, researchers have also proposed the SI model to characterize the information diffusion process of an online event. Under this model, the diffusion topology is assumed to be a clique, i.e., there is an edge between each pair of users in the network and all users can infect each other. $P(\tau)$ is also defined as in the case of the self-exciting model discussed above. Under the SI model, the

Figure 2: Examples of four information diffusion patterns.

predicted popularity of an online campaign/ event at time interval t can be expressed as:

$$\hat{r}(t) = \alpha(S - \sum_{i=1}^{t-1} \hat{r}(i)) \sum_{i=1}^{t-1} \hat{r}(i)P(t-i), \quad (2)$$

where α is the scaling factor. S is the total number of nodes in the clique. Under this model, $\hat{r}(t)$ shows a rising-and-falling pattern.

4 TEMPORAL BEHAVIOR OF THE POPULARITY OF A CAMPAIGN

Figure 2 depicts 4 common temporal patterns of the popularity of online events observed from two large-scale real-world datasets, namely, the retweet cascades recorded by the TWEET-1Mo dataset [14] and the hashtag streams of the Hashtag dataset in [13]. For the TWEET-1Mo dataset, most of the retweet cascades follow Pattern C2 which shows very limit fluctuations after the first peak. Moreover, around 30% of the retweets of a given tweet arrive within first five minutes of the initial one. In the Hashtag dataset, 17.3% of the hashtag streams follow Pattern C1 while 13.5% follow Pattern C2. 40.3% of the hashtag streams follow Pattern C3 and 28.9% belong to Pattern C4. Notice that both Pattern C3 and C4 contain approximately periodic peaks. For Pattern C3, the amplitudes of peaks decrease monotonically as time progresses. In contrast, Pattern C4 contains peaks of non-monotonic, fluctuating amplitude.

In our experiments, we find that the self-exciting model and the SI model can provide good fit for empirical traces following Pattern C1 and C2. However, both of these models fail to generate good matches for empirical traces of Pattern C3 and C4. As such, popularity predictions using the self-exciting and SI models perform poorly for online campaigns which follow the temporal patterns of C3 and C4. In the following sections, we will demonstrate that, by taking the effects of opinion leaders into account using a Multi-Actor Self-Exciting Process (MASEP), we can accurately model information diffusion processes that exhibit temporal popularity patterns of C3 and C4.

5 THE MASEP-BASED MODEL FOR ONLINE CAMPAIGN POPULARITY

In this section, we first define MASEP by considering a new diffusion topology. We then derive a closed-form expression for the temporal popularity of an online campaign modeled by MASEP.

Table 1: Table of symbols

Symbol	Description
$X(t)$	The popularity of ordinary users at time interval t
$Y(t)$	The popularity of opinion leaders at time interval t
ω_1	The aging rate of infectiousness
ω_2	The aging rate of opinion leaders' interest
ω_0	The frequency of opinion leaders' activity
θ	The phase shift of periodicity
α	The scaling factor
γ	The intensity of opinion leaders' activity
β	The weight of opinion leaders' influence
p_0	The strength of periodicity
$r(t)$	The empirical popularity at time interval t
$\hat{r}(t)$	The prediction popularity at time interval t
R_∞	The empirical final popularity
$\hat{R}_\infty$	The prediction final popularity

5.1 Diffusion Topology of MASEP

Figure 1(c) shows the diffusion topology of MASEP which combines a tree with external stimulating nodes. The two types of nodes correspond to the ordinary users and the opinion leaders, respectively. They possess different properties and dynamics. The MASEP-based diffusion topology preserves the properties of ordinary users in the self-exciting model while taking the impact of the opinion leaders into account. As we will demonstrate in the Section 6, MASEP can also capture other commonly observed properties of online campaigns including the aging effect, the reinforcement effect, and the periodicity of campaign popularity.

5.2 The MASEP Model

Let $X(t)$ and $Y(t)$ be the number of posts published by the ordinary users and the opinion leaders at time interval t , respectively. Based on the diffusion topology of MASEP discussed above, we have:

$$X(t) = \alpha \sum_{i=1}^{t-1} e^{-\omega_1(t-i)} [X(i) + \beta Y(i)]. \quad (3)$$

$$Y(t) = \gamma [1 - p_0 \cos(\omega_0 t + \theta)] e^{-\omega_2 t}. \quad (4)$$

$\{X(t)\}$ and $\{Y(t)\}$ correspond to a self-exciting branching process and a periodical decay process, respectively. The popularity of the corresponding online campaign can then be estimated by:

$$\hat{r}(t) = X(t) + Y(t). \quad (5)$$

Refer to Table 1 for the definition of other parameters in Eqs.(3-5).

Parameter Estimation: MASEP has a set of eight model parameters: $\Omega = \{\alpha, \beta, \theta, \gamma, p_0, \omega_0, \omega_1, \omega_2\}$. By comparing the empirically observed popularity values, $r(t)$'s, as defined in Section 3, with their predicted values $\hat{r}(t)$ under the MASEP model, we can determine Ω^* , the optimal set of model parameters, for a given online campaign. Specifically, we obtain Ω^* by solving the following minimization problem using the Levenberg-Marquardt (LM) method [5]:

$$\Omega^* = \arg(\minimize_{\Omega} \{ \sum_{t=1}^N [r(t) - \hat{r}(t)]^2 \}). \quad (6)$$

Figure 3: Comparison of MASEP and SpikeM in fitting complex information diffusion pattern. In this example, we use the first 70 hours temporal reshare trace to train parameter.

5.3 Closed-form Expression for $\hat{r}(t)$

During each iteration of the minimization process in Eq. (6), we need to evaluate $\sum_{i=1}^t X(i)$. A naive approach based on the recurrent equation of $X(t)$ in Eq. (3) would incur a computational cost of $O(t^2)$. To speed up model parameter estimation for MASEP, we have derived a closed-form expression of $X(t)$ by rewriting Eq. (3):

$$\begin{aligned} X(t) &= \alpha b[X(t-1) + \beta Y(t-1)] + \alpha b \sum_{i=1}^{t-2} b^{(t-1-i)}[X(i) + \beta Y(i)] \\ &= \alpha b[X(t-1) + \beta Y(t-1)] + bX(t-1) \\ &= (b + \alpha b)^{(t-1)}X(1) + \alpha b\beta \sum_{i=1}^{t-1} (b + \alpha b)^{(t-1-i)}Y(i). \end{aligned} \quad (7)$$

Substitute $Y(\cdot)$ and $\cos \omega = \frac{1}{2}(e^{j\omega} + e^{-j\omega})$, where j is the imaginary unit, into Eq. (7), we have:

$$X(t) = (b + \alpha b)^{t-1}[X(1) + \alpha b\beta Y(t)]. \quad (8)$$

where

$$\begin{aligned} b &= e^{-\omega_1}, \quad c = e^{-\omega_2}, \quad F = \frac{c}{b + \alpha b}, \\ H(t) &= \frac{F - F^t}{1 - F} + \frac{[\cos(\omega_0 t + \theta)F^t + \cos(\theta)F^2]p_0}{1 + F^2 - 2\cos(\omega_0)F} \\ &\quad - \frac{[\cos(\omega_0 + \theta)F + \cos(\omega_0(t-1) + \theta)F^{(t+1)}]p_0}{1 + F^2 - 2\cos(\omega_0)F}. \end{aligned}$$

Eq. (8) provides a closed-form expression for $X(t)$ without the recursive summation term. The computational complexity of $\sum_{i=1}^t X(i)$ is reduced from $O(t^2)$ to $O(t)$, which results in an order of magnitude reduction in the time of MASEP model parameter estimation. For instance, when running on the same machine, the average parameter estimation time for traces in the Hashtag dataset reduces from 4.93 second to 0.672 second by using Eq. (8) instead of Eq. (3).

5.4 Closed-form Expression for $\hat{R}_\infty$

In many applications, we are interested in how far a post can spread across the network and need to calculate the final popularity: $R_\infty = \sum_{t=1}^\infty r(t)$. Therefore, we have also derived a closed-form expression

Table 2: Comparing the Final Popularity Prediction Error of MASEP, SpikeM, SEISMIC, and STRM (using the initial 30% (time-wise) of the data trace of each campaign for training)

Methods	Three Real-world Validation Datasets		
	TWEET-1Mo	Hashtag	Crowdfunding
MASEP	13.25	15.7	36.9
SpikeM	25.65	35.1	46.7
SEISMIC	38.07	51.6	68.3
STRM	32.12	40.3	63.6

for R_∞ under the MASEP model. Consider the Z-transform of $(X(\cdot) + Y(\cdot))$ given by $\mathcal{X}(z) + \mathcal{Y}(z) = \sum_{i=0}^\infty (X(i) + Y(i))z^i$. Note that if $z = 1$, $\mathcal{X}(1) + \mathcal{Y}(1) = \sum_{i=0}^\infty (X(i) + Y(i)) = \hat{R}_\infty$, which yields:

$$\hat{R}_\infty = \left(\frac{\alpha b \beta}{1 - (1 + \alpha)b} + 1 \right) \cdot \left[\frac{\gamma}{1 - c} - \frac{\gamma \lambda p_0 (\cos(\theta) - c \cdot \cos(\omega_0 - \theta))}{1 - 2c \cdot \cos \omega_0 + c^2} \right]. \quad (9)$$

In the experiments, we use the empirical total number of posts of an online campaign in the dataset as an approximation of the real final popularity.

6 EXPERIMENTS

In this section, we evaluate the popularity prediction performance of MASEP using three real-world datasets:

TWEET-1Mo [14]: This dataset is the complete set of over 3.2 billion tweets on Twitter from October 7 to November 7, 2011. For each tweet, the dataset includes the tweet id, the posting time, and the number of followers.

Hashtag Dataset [13]: This dataset has 1000 hashtag streams selected from more than 580 million Twitter posts covering an 8-month period from June 2009 to February 2010. For each hashtag stream, the dataset includes the start time, the length of time interval, and the number of posts in each time interval.

Crowdfunding Dataset [7]: This dataset contains the crowdfunding projects launched after November 2012 and completed before April 2013 from Kickstarter. For each project, the dataset records the proposers, the duration, the fundraising goal, the length of time interval, and the number of backers in each time interval.

6.1 Temporal Evolution Pattern Prediction

To validate the effectiveness of MASEP for predicting $r(t)$, we plot the prediction information diffusion pattern using Eq. (8) and Eq. (4) and evaluate the fitting accuracy between prediction and empirical information diffusion pattern by calculating the normalized root mean square error (NRMSE):

$$\frac{1}{\max_{1 \leq i \leq T} r(i) - \min_{1 \leq i \leq T} r(i)} \sqrt{\frac{\sum_{i=1}^T (\hat{r}(i) - r(i))^2}{T}}.$$

We choose SpikeM as the baseline as SEISMIC and STRM are both point-estimators designed to predict the final popularity of a campaign (as opposed to the generative models like SpikeM and MASEP which can readily predict the entire time-series of a campaign). Figure 3 depicts a representative empirical time series of the popularity of a campaign as well as its counterparts predicted by the MASEP

Figure 4: Effect of training dataset size on popularity prediction error (APE and NRMSE) of the Hashtag dataset. The bottom figures are the performance variance of each method. For each data bar, the meaning of each line from top to bottom is the 95% percentile, the 75% percentile, the median, the 25% percentile, and the 5% percentile.

and SpikeM model. Figure 4(a) shows the effect of training dataset size on prediction performance of four information diffusion patterns on the Hashtag dataset.

Both MASEP and SpikeM take the aging effect of infectiousness into account, so they can capture the heavy tail pattern of C1 and C2. Since MASEP addresses the reinforcement effect and periodicity of the opinion leaders, MASEP has an advantage over SpikeM in fitting the latter peaks more accurately, e.g. the fitting result after 60 hours in Figure 3. While SpikeM uses a data-driven and limited estimation of the scale of the underlying clique, the fitting power of SpikeM will drop dramatically once the prediction popularity exceed the estimated scale of the underlying clique. Moreover, we find that there is an "incubation period" for some online content, like the first 30 hours in Figure 2(d). Via removing the incubation period, MASEP reduces NRMSE by 8%. Figure 5 validates the robustness of the MASEP-based popularity predictor for online campaigns of

different sizes (in terms of the total number of postings or sponsors). We classify the hashtag streams on the Hashtag dataset into four non-overlap groups based on their final number of hashtag reshares. The robustness of MASEP is larger than SpikeM with 20% training dataset size. A careful scrutiny reveals that, SpikeM has a tendency to down-estimate the count of reshares because of the long incubation period in the hash streams with the final reshares between 1000 and 5000.

6.2 Final Popularity Prediction

Table 2 summarizes the final popularity prediction results of MASEP, SpikeM, SEISMIC, and STRM on three real-world datasets using the initial 30% of a campaign data trace for training. The performance metric is the absolute percentage error (APE):

$$\frac{|\hat{R}_{\infty} - R_{\infty}|}{R_{\infty}}.$$

Figure 5: Robustness of Popularity Estimator w.r.t. the size of online campaigns

Figure 4(b) shows the effect of training dataset size on final popularity prediction performance on the Hashtag dataset. We also apply MASEP to the scenarios involving multiple OSN services, like the crowd-funding campaigns. Figure 6 depicts the results on the Crowdfunding dataset. We do not include the results of SEISMIC and STRM because of their poor prediction performance.

Both SESIMIC and STRM are generative methods based on point process, and require to observe the reshare time of each infected user. So, their prediction error of the final popularity with discretized training data is larger than the other fitting-based methods. For campaigns of Pattern C1 and C2, SpikeM already provides good prediction results using the initial 10% to 20% of the resharing trace as training data. MASEP can achieve the best performance in four patterns with more than 30% training data. For Pattern C3 and C4, MASEP’s performance is superior to SpikeM as long as more than 10% of the initial portion of a campaign trace is used for training.

Notice that, unlike retweeting cascades or hashtag streams, a crowd-funding campaign always has a fixed deadline, which often results in a final surge of sponsors right before the end of the campaign [7]. Such last-minute bursty arrival behavior directly contradicts the aging effect assumption made by all of the models considered in this paper. This explains the relatively poor popularity prediction performance for the Crowdfunding dataset when comparing to the other two datasets.

In the previous section, we claim that the contributions of the opinion leaders are larger in Pattern C3 and C4. To validate this claim, we calculate the ratio between the number of posts published by the opinion leaders and the ordinary user: $\sum_1^T Y(i) / \sum_i^T X(i)$. The average ratio of each pattern on the Hashtag dataset: 0.43%, 0.39%, 1.72%, and 2.73%. The average ratio of the Tweet-1Mo dataset, the Hashtag dataset, and the Crowdfunding dataset: 0.37%, 1.13%, and 9.89%. The first group of results verify our claim. As for the second group of results, because most of the tweet cascades on the Tweet-1Mo dataset are spontaneous activities of the ordinary users, the ratio is small. Since there are intentional promotions of crowd-funding campaigns in OSN, the fraction of opinion leaders of crowd-funding campaigns is larger than other online events.

7 CONCLUSIONS AND FUTURE WORK

In this paper, we propose MASEP to model and predict the popularity of different online campaigns involving multiple types of

Figure 6: Effect of training dataset size on popularity prediction error (APE) for the Crowdfunding dataset.

actors. MASEP combines a self-exciting branching process and a periodical decay process to model the dynamics and interdependent relationship between two major types of actors in OSN: the opinion leaders and the ordinary users. MASEP also captures the reinforcement effect, the aging effect, and the periodicity of temporal popularity of online campaigns. We discover four information diffusion patterns and validate the prediction power of MASEP on each pattern. A closed-form expression is derived for the MASEP-modeled popularity and yields an order of magnitude improvement in training speed. For the prediction of the final popularity of online campaigns in 3 real-world datasets, MASEP successfully reduces the absolute prediction errors by 26.1% to 63.2% when comparing to other state-of-the-art approaches.

REFERENCES

- [1] Abir De, Isabel Valera, Niloy Ganguly, Sourangshu Bhattacharya, and Manuel Gomez Rodriguez. 2016. Learning and forecasting opinion dynamics in social networks. In *NIPS*.
- [2] Mehrdad Farajtabar, Yichen Wang, Manuel Gomez Rodriguez, Shuang Li, Hongyuan Zha, and Le Song. 2015. Coevolve: A joint point process model for information diffusion and network co-evolution. In *NIPS*.
- [3] Manuel Gomez-Rodriguez, Jure Leskovec, and Bernhard Schölkopf. 2013. Modeling information propagation with survival theory. In *ICML*.
- [4] Ryota Kobayashi and Renaud Lambiotte. 2016. TiDeH: Time-Dependent Hawkes Process for Predicting Retweet Dynamics. In *ICWSM*.
- [5] Kenneth Levenberg. 1944. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics* 2, 2 (1944), 164–168.
- [6] Feng Li and Timon C Du. 2011. Who is talking? An ontology-based opinion leader identification framework for word-of-mouth marketing in online social blogs. *Decision Support Systems* 51, 1 (2011).
- [7] Chun-Ta Lu, Sihong Xie, Xiangnan Kong, and Philip S Yu. 2014. Inferring the impacts of social media on crowdfunding. In *WSDM*.
- [8] Yasuko Matsubara, Yasushi Sakurai, B Aditya Prakash, Lei Li, and Christos Faloutsos. 2012. Rise and fall patterns of information diffusion: model and implications. In *KDD*.
- [9] Swapnil Mishra, Marian-Andrei Rizoio, and Lexing Xie. 2016. Feature driven and point process approaches for popularity prediction. In *CIKM*.
- [10] Marian-Andrei Rizoio, Lexing Xie, Scott Sanner, Manuel Cebrian, Honglin Yu, and Pascal Van Hentenryck. 2017. Expecting to be HIP: Hawkes Intensity Processes for Social Media Popularity. In *WWW*.
- [11] Bidisha Samanta, Abir De, and Niloy Ganguly. 2017. Strm: A sister tweet reinforcement process for modeling hashtag popularity. In *INFOCOM*.
- [12] Hua-Wei Shen, Dashun Wang, Chaoming Song, and Albert-László Barabási. 2014. Modeling and Predicting Popularity Dynamics via Reinforced Poisson Processes. In *AAAI*.
- [13] Jaewon Yang and Jure Leskovec. 2010. Modeling information diffusion in implicit networks. In *ICDM*.
- [14] Qingyuan Zhao, Murat A Erdogdu, Hera Y He, Anand Rajaraman, and Jure Leskovec. 2015. Seismic: A self-exciting point process model for predicting tweet popularity. In *KDD*.