COEVOLVE: A Joint Point Process Model for Information Diffusion and Network Co-evolution

Mehrdad Farajtabar*
Georgia Tech
mehrdad@gatech.edu

Manuel Gomez-Rodriguez
MPI for Software Systems
manuelgr@mpi-sws.org

Yichen Wang
Georgia Tech
yichen.wang@gatech.edu

Shuang Li
Georgia Tech
sli370@gatech.edu

Hongyuan Zha
Georgia Tech
zha@cc.gatech.edu

Le Song
Georgia Tech, Ant Financial
lsong@cc.gatech.edu

ABSTRACT

Information diffusion in online social networks is affected by the underlying network topology, but it also has the power to change it. Online users are constantly creating new links when they are exposed to new information sources, and in turn these links are alternating the way information spreads. However, these two highly intertwined stochastic processes—information diffusion and network evolution—have been typically studied *separately*, ignoring their co-evolutionary dynamics.

In this work, we propose a temporal point process model, Coevolve, for such joint dynamics, allowing the intensity of one process to be modulated by that of the other. The model allows us to efficiently simulate interleaved diffusion and network events, and generate traces obeying common diffusion and network patterns observed in real-world networks. Moreover, we develop a convex optimization framework to learn the parameters of the model from historical diffusion and network evolution traces. Experiments in both synthetic data and real data gathered from Twitter show that our model provides a good fit to the data as well as more accurate predictions than alternatives.

CCS CONCEPTS

• **Information systems → Social networks**; • **Networks → Online social networks**; **Topology analysis and generation**; *Network dynamics*;

KEYWORDS

Social Networks; Network Structure; Information Diffusion; Coevolutionary Dynamics; Point Processes; Hawkes processes; Survival Process

ACM Reference Format:
Mehrdad Farajtabar, Manuel Gomez-Rodriguez, Yichen Wang, Shuang Li, Hongyuan Zha, and Le Song. 2018. COEVOLVE: A Joint Point Process Model for Information Diffusion and Network Co-evolution . In *The 2018 Web Conference Companion, April 23–27, 2018, Lyon, France.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3184558.3186236

*Corresponding author

ACM ISBN 978-1-4503-5640-4/18/04..
https://doi.org/10.1145/3184558.3186236

1 INTRODUCTION

Online social networks such as Facebook, Twitter or Weibo have become large information networks where people share, discuss and search for information of personal interest as well as breaking news [19]. In this context, users often forward to their *followers* information they are exposed to via their *followees*, triggering the emergence of information *cascades* that travel through the network [4], and constantly create new links to information sources, triggering changes in the network itself over time. Importantly, recent empirical studies with Twitter data have shown that both information diffusion and network evolution are coupled and network changes are often triggered by information diffusion [1, 25, 32].

While there have been many recent works on modeling information diffusion [4, 5, 12, 13] and network evolution [3, 20, 21], most of them study these two stochastic processes separately, ignoring the influence one may have on the other over time. Therefore, to better understand information diffusion and network evolution, there is an urgent need for joint probabilistic models of the two processes, which are largely inexistent to date.

In this work, we propose a probabilistic generative model, Coevolve, for the joint dynamics of information diffusion and network evolution. Our model is based on the framework of temporal point processes, which explicitly characterizes the continuous time interval between events, and it consists of two interwoven and interdependent components:

I. **Information diffusion process.** We design an "identity revealing" multivariate Hawkes process [24] to capture the mutual excitation behavior of retweeting events, where the intensity of such events in a user is boosted by previous events from her time-varying set of followees. Although Hawkes processes have been used for information diffusion before [7, 10, 17, 23], the key innovation of our approach is to explicitly model the excitation due to a particular source node, hence revealing the identity of the source. Such design reflects the reality that information sources are explicitly acknowledged, and it also allows a particular information source to acquire new links in a rate according to her "informativeness".

II. **Network evolution process.** We model link creation as an "information driven" survival process, and couple the intensity of this process with retweeting events. Although survival processes have been used for link creation before [16, 30], the key innovation in our model is to incorporate retweeting events as the

* Extended version of this work appeared in (Farajtabar et al. [8]).
◇ Code is available at https://github.com/farajtabar/Coevolution.

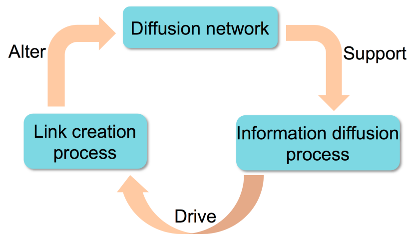


Figure 1: Illustration of how information diffusion and network structure processes interact

driving force for such processes. Since our model has captured the source identity of each retweeting event, new links will be targeted toward the information sources, with an intensity proportional to their degree of excitation and source's influence.

Point processes have been successfully applied to different phenomena in social networks like community detection [28], fake news mitigation [9], recommendation [15], message broadcasting [18], modeling crowd-generated data [27], and cascade modeling [33].

Our model (Fig. 1) is designed in such a way that it allows the two processes, information diffusion and network evolution, unfold simultaneously in the same time scale and excise bidirectional influence on each other. Moreover, we experimentally verify that our model can produce coevolutionary dynamics of information diffusion and network evolution, and generate retweet and link events that obey common information diffusion patterns (*e.g.*, cascade structure, size and depth), static network patterns (*e.g.*, node degree) and temporal network patterns (*e.g.*, shrinking diameter) described in the related literature [11, 21, 22]. Finally, we show that, by modeling the coevolutionary dynamics, our model provide significantly more accurate link and diffusion event predictions than alternatives in large scale Twitter data [1].

2 MODEL

We will model the generation of two types of events: tweet/retweet events, e^r, and link creation events, e^l. We represent the events as

$$e^r \text{ or } e^l := (\underset{\substack{\uparrow \\ \text{destination}}}{u}, \overset{\substack{\text{source} \\ \downarrow}}{s}, \underset{\substack{\uparrow \\ \text{time}}}{t}).$$

For retweet event, the triplet means that the destination node u retweets at time t a tweet originally posted by source node s. This event can happen when u is retweeting a message by another node u' where the original information source s is acknowledged. Given a list of retweet events $\{e^r_1 = (u_1, s_1, t_1), \ldots\}$ up to time t, the history $\mathcal{H}^r_{us}(t)$ of retweets by u due to source s is $\mathcal{H}^r_{us}(t) = \{e^r_i = (u_i, s_i, t_i) | u_i = u \text{ and } s_i = s\}$.

For link creation event, the triplet means that destination node u creates at time t a link to source node s, *i.e.*, from time t on, node u starts following node s. We restrict ourselves to the case where each (directed) link is created only once. We denote the link creation history as $\mathcal{H}^l(t)$.

Then, given m users, we will use two sets of counting processes to record the generated events:

- **Counting processes for retweets** are denoted as a matrix $\boldsymbol{N}(t)$ of size $m \times m$ for each fixed time t. The (u, s)-th entry in the matrix, $N_{us}(t) \in \{0\} \cup \mathbb{Z}^+$, counts the number of retweets of u due to source s up to t. These counting processes are "identity revealing".
- **Survival processes for links** are denoted also as a matrix $\boldsymbol{A}(t)$ of size $m \times m$ for each fixed time point t. The (u, s)-th entry in the matrix, $A_{us}(t) \in \{0, 1\}$, indicates whether u is directly following s, *i.e.*, $A_{us}(t) = 1$ means the directed link has been created before t. There is only 1 event for an instantiation of a survival process.

An important way to characterize point processes is via the conditional intensity function $\lambda^*(t)$. Formally, it is the conditional probability of observing an event in a small window $[t, t + dt)$ given history $\mathcal{H}(t)$, *i.e.*, $\lambda^*(t)dt := \mathbb{P}\{\text{event in } [t, t + dt) | \mathcal{H}(t)\} = \mathbb{E}[dN(t) | \mathcal{H}(t)]$, where we usually assume $dN(t) \in \{0, 1\}$ for the small window dt.

The interwoven information diffusion and network evolution processes are characterized as $\mathbb{E}[d\boldsymbol{N}(t) \,|\, \mathcal{H}^r(t) \cup \mathcal{H}^l(t)] = \boldsymbol{\Gamma}^*(t)\, dt$ and $\mathbb{E}[d\boldsymbol{A}(t) \,|\, \mathcal{H}^r(t) \cup \mathcal{H}^l(t)] = \boldsymbol{\Lambda}^*(t)\, dt$.

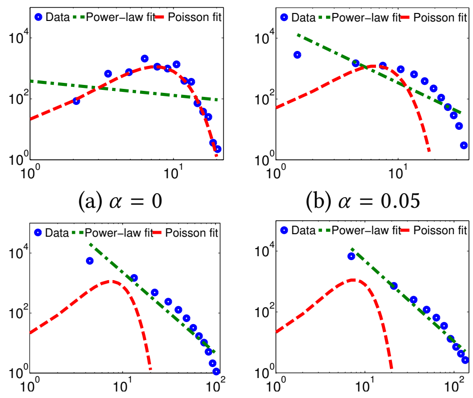
where $\boldsymbol{\Gamma}^*(t) = (\ \gamma^*_{us}(t)\ )_{u,s \in [m]}$ and $\boldsymbol{\Lambda}^*(t) = (\ \lambda^*_{us}(t)\ )_{u,s \in [m]}$. We model the intensities, $\boldsymbol{\Gamma}^*(t)$, for retweeting events and $\boldsymbol{\Lambda}^*(t)$, for link creation as $\gamma^*_{us}(t) = \beta_s \sum_{v \in \mathcal{F}_u(t)} \kappa_{\omega_1}(t) \star (A_{uv}(t)\ dN_{vs}(t))$ and $\lambda^*_{us}(t) = (1 - A_{us}(t))(\mu_u + \alpha_u\, \kappa_{\omega_2}(t) \star dN_{us}(t)$. Here $f(t) \star dN(t) := \int_0^t f(t - \tau)\, dN(\tau) = \sum_{t_i \in \mathcal{H}(t)} f(t - t_i)$, and we have $\mathcal{F}_u(t) := \{v \in [m] : A_{uv}(t) = 1\}$ is the current set of followees of u, $\gamma_{uu} = \eta_u$, and $\lambda_{uu} = 0$. The term $\eta_u \geqslant 0$ is the intensity of original tweets by a user u on his own initiative, becoming the source of a cascade, and the term $\beta_s \sum_{v \in \mathcal{F}_u(t)} \kappa_\omega(t) \star (A_{uv}(t)\ dN_{vs}(t))$ models the propagation of peer influence over the network, where the triggering kernel $\kappa_{\omega_1}(t)$ models the decay of peer influence over time. The term $1 - A_{us}(t)$ effectively ensures a link is created only once, and after that, the corresponding intensity is set to zero, the term $\mu_u \geqslant 0$ denotes a baseline intensity, which models when a node u decides to follow a source s spontaneously at her own initiative, and the term $\alpha_u \kappa_{\omega_2}(t) \star dN_{us}(t)$ corresponds to the retweets of node u due to tweets originally published by source s, where the triggering kernel $\kappa_{\omega_2}(t)$ models the decay of interests over time. Here, the higher the corresponding retweet intensity, the more likely u will find information by source s useful and will create a *direct* link to s. Please refer to [8] for details of the model, its efficient simulation and estimation procedure.

3 EXPERIMENTAL RESULTS

3.1 Synthetic Data

In this section, we simulate the evolution of a 8,000-node network as well as the propagation of information using our simulation algorithm [8]. For simulated networks we set the exogenous intensities of the link and diffusion events to $\mu_u = \mu = 4 \times 10^{-6}$ and $\eta_u = \eta = 1.5$ respectively, and the triggering kernel parameter to $\omega_1 = \omega_2 = 1$. The parameter μ determines the independent growth of the network—roughly speaking, the expected number of links each user establishes spontaneously before time T is μT.

Degree Distribution. Empirical studies have shown that the degree distribution of online social networks and microblogging sites follow a power law [3, 19], and argued that it is a consequence of the rich get richer phenomena. The degree distribution of a network is a power law if the expected number of nodes m_d with degree d is given by $m_d \propto d^{-\gamma}$, where $\gamma > 0$. Intuitively, the higher the values



(a) $\alpha = 0$ (b) $\alpha = 0.05$

(c) $\alpha = 0.1$ (d) $\alpha = 0.2$

Figure 2: Degree distributions when network sparsity level reaches 0.001 for different α's and fixed $\beta = 0.1$. The degree distributions spans from random to scale-free networks illustrating the flexibility to model real networks.

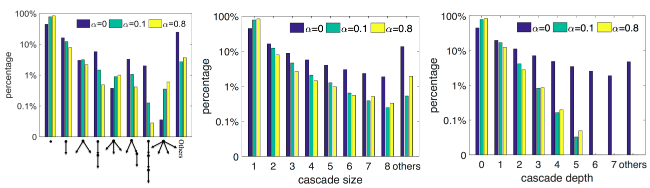


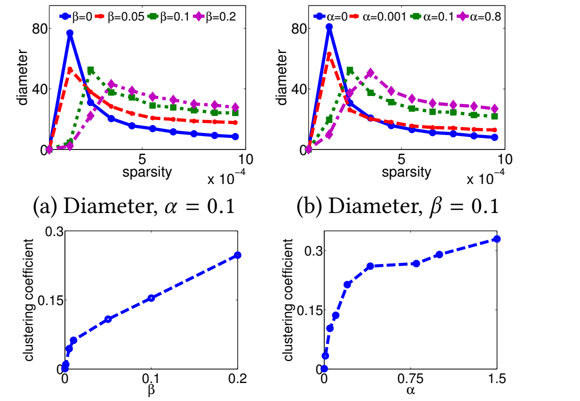
Figure 3: Distribution of cascade structure, size and depth for different values of α and fixed $\beta = 0.2$

of the parameters α and β, the closer the resulting degree distribution follows a power-law. This is because the network grows more locally. Interestingly, the lower their values, the closer the distribution to an Erdos-Renyi random graph [6], because, the edges are added almost uniformly and independently without influence from the local structure. Figure 2 confirms this intuition by showing the degree distribution for different values of α.

Cascade Patterns. Our model can produce the most commonly occurring cascades structures as well as heavy-tailed cascade size and depth distributions, as observed in historical Twitter data [11]. Figure 3 summarizes the results and provide empirical evidence that the higher the α (β) value, the shallower and wider the cascades.

Small (shrinking) Diameter. There is empirical evidence that the diameter of online social networks and microblogging sites exhibit relatively small diameter and shrinks (or flattens) as the network grows [2, 3, 22]. Figures 4(a-b) show the diameter on the largest connected component (LCC) against the sparsity of the network over time for different values of α and β. Although at the beginning, there is a short increase in the diameter due to the merge of small connected components, the diameter decreases as the network evolves. Moreover, larger values of α or β lead to higher levels of local growth and, as a consequence, slower shrinkage.

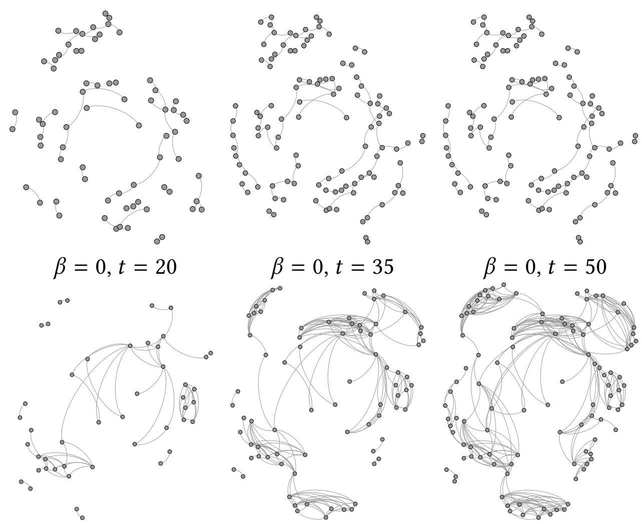
Clustering Coefficient. Triadic closure [14, 20, 26] has been often presented as a plausible link creation mechanism. However, different social networks and microblogging sites present different levels of triadic closure [29]. Importantly, our method is able to generate



(a) Diameter, $\alpha = 0.1$ (b) Diameter, $\beta = 0.1$

(c) CC, $\alpha = 0.1$ (d) CC, $\beta = 0.1$

Figure 4: Diameter and clustering coefficient for network sparsity 0.001; (a,b) The diameter against sparsity over time for fixed $\alpha = 0.1$, and for fixed $\beta = 0.1$ respectively; (c,d) The clustering coefficient (CC) against β and α, respectively.



$\beta = 0, t = 20$ $\beta = 0, t = 35$ $\beta = 0, t = 50$

$\beta = 0.8, t = 20$ $\beta = 0.8, t = 35$ $\beta = 0.8, t = 50$

Figure 5: Evolution of a network with values $\beta = 0, 0.8$

networks with different levels of triadic closure, as shown by Figure 4(c-d), where we plot the clustering coefficient [31], which is proportional to the frequency of triadic closure.

Network Visualization. Figure 5 visualizes several snapshots of the largest connected component (LCC) of two 300-node networks for two particular realizations of our model, under two different values of β. In both cases, we used $\mu = 2 \times 10^{-4}$, $\alpha = 1$, and $\eta = 1.5$. The three left graphs correspond to $\beta = 0$ and represent one end of the spectrum, *i.e.*, Erdos-Renyi random network. Here, the network evolves uniformly. The right three graphs correspond to $\beta = 0.8$ and represent the other end, *i.e.*, scale-free networks. Here, the network evolves locally, and clusters emerge naturally as a consequence of the local growth.

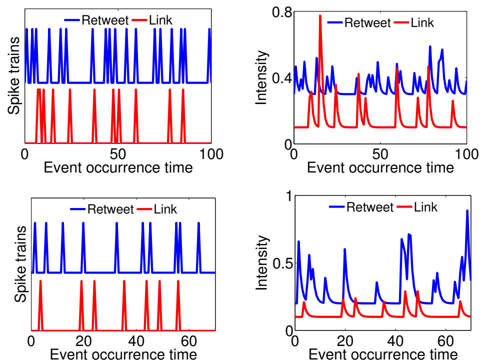


Figure 6: Link and retweet behavior of 2 typical users in the real data. The left column shows the spike trains of link and retweet events for them and the right column shows the estimated link and retweet intensities for the same two users.

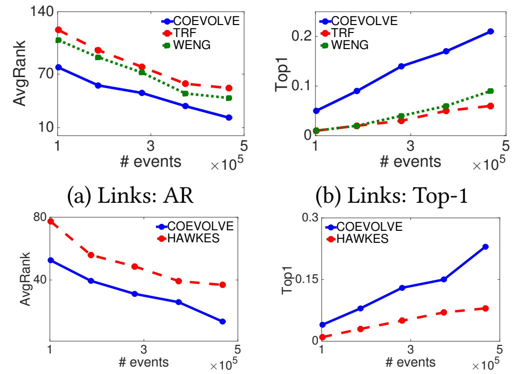
3.2 Real Data

We use a dataset that contains both link events as well as tweets and retweets from millions of Twitter users [1]. In particular, the dataset contains data from three sets of users in 20 days; nearly 8 million tweet, retweet, and link events by more than 6.5 million users. The first set of users (8,779 users) are source nodes s, for whom all their tweet times were collected. The second set of users (77,200 users) are the followers of the first set of users, for whom all their retweet times (and source identities) were collected. The third set of users (6,546,650 users) are the users that start following at least one user in the first set during the recording period.

In our experiments, we focus on all events (and users) during a 10-day period (Sep 21 2012 - 30 Sep 2012) and used the information before Sep 21 to construct the initial social network (original links between users). We model the co-evolution in the second 10-day period using our framework. More specifically, in the coevolution modeling, we have 5,567 users in the first layer who post 221,201 tweets. In the second layer, 101,465 retweets are generated by the whole 77,200 users in that interval. And, in the third layer, we have 198,518 users who create 219,134 links to 1978 users (out of 5567) in the first layer. Finally, we split events into a training set (covering 85% of events) and a test set (covering the remaining 15%) according to time, *i.e.*, all events in the training set occur earlier than those in the test set. Then, we use our model estimation procedure [8] to fit the parameters from an increasing proportion of training data.

Retweet and Link Coevolution. Figure 6 visualizes the retweet and link events, aggregated across different targets, and the corresponding intensities given by our trained model for four source nodes, picked at random. Here, it is clear that retweets (of his posts) and link creations (to him) are clustered in time and often follow each other, and our fitted model intensities successfully track them.

Link Prediction. We use our model to predict the identity of the source for each test link event, given the historical (link and retweet) events before the time of the prediction, and compare its performance with TRF [1] and WENG [32], which are state of the art methods. We evaluate the performance by computing the probability of all potential links using different methods, and then compute (i) the average rank of all true (test) events (AvgRank) and, (ii) the success probability (SP) that the true (test) events rank among the top-1 potential events at each test time (Top-1). We summarize the results in Figure 7(a-b), where we consider an increasing number of training retweet/tweet events. Our model outperforms TRF and WENG consistently. For example, for $8 \cdot 10^4$ training events, our model achieves a SP 2.5x times larger than TRF and WENG.



(a) Links: AR (b) Links: Top-1

(c) Activity: AR (d) Activity: Top-1

Figure 7: Prediction performance in the Twitter dataset by means of average rank (AR) and success probability that the true (test) events rank among the top-1 events (Top-1).

Activity Prediction. We use our model to predict the identity of the node that generates each test diffusion event, given the historical events before the time of the prediction, and compare its performance with a baseline consisting of a Hawkes process without network evolution. For the Hawkes baseline, we take a snapshot of the network right before the prediction time, and use all historical retweeting events to fit the model. Here, we evaluate the performance the via the same two measures as in the link prediction task and summarize the results in Figure 7(c-d) against an increasing number of training events. The results show that, by modeling the co-evolutionary dynamics, our model performs significantly better than the baseline.

4 CONCLUSION

In this work, we proposed a joint continuous-time model of information diffusion and network evolution, which can capture the coevolutionary dynamics, can mimic the most common static and temporal network patterns observed in real-world networks and information diffusion data, and can predict the network evolution and information diffusion more accurately than previous state-of-the-arts. Using point processes to model intertwined events in information and social networks opens up many interesting venues for future. Our current model is just a show-case of a rich set of possibilities offered by a point process framework, which have been rarely explored before in large scale social network modeling. For example, a large and diverse range of point processes can also be used instead in the framework and augment the current model without changing the efficiency of simulation and the convexity of parameter estimation. Developing an efficient mechanism to account for heterogeneity in time resolution would improve the model's ability to predict. Finally, we may augment the framework to allow time-varying parameters. The simulation would not be affected and the estimation of time-varying interaction can still be carried out via a convex optimization problem [34].

REFERENCES

[1] D. Antoniades and C. Dovrolis. Co-evolutionary dynamics in social networks: A case study of twitter. *arXiv preprint arXiv:1309.6001*, 2013.
[2] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna. Four degrees of separation. In *Proceedings of the 4th Annual ACM Web Science Conference*, pages 33–42, 2012.
[3] D. Chakrabarti, Y. Zhan, and C. Faloutsos. R-mat: A recursive model for graph mining. *Computer Science Department*, page 541, 2004.
[4] J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec. Can cascades be predicted? In *Proceedings of the 23rd international conference on World wide web*, pages 925–936, 2014.
[5] N. Du, L. Song, M. Gomez-Rodriguez, and H. Zha. Scalable influence estimation in continuous-time diffusion networks. In *Advances in Neural Information Processing Systems 26*, 2013.
[6] P. Erdos and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.
[7] M. Farajtabar, N. Du, M. Gomez-Rodriguez, I. Valera, H. Zha, and L. Song. Shaping social activity by incentivizing users. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.
[8] M. Farajtabar, Y. Wang, M. Gomez-Rodriguez, S. Li, and H. Zha. Coevolve: A joint point process model for information diffusion and network evolution. *Journal of Machine Learning Research*, 18:1–49, 2017.
[9] M. Farajtabar, J. Yang, X. Ye, H. Xu, R. Trivedi, E. Khalil, S. Li, L. Song, and H. Zha. Fake news mitigation via point process based intervention. *arXiv preprint arXiv:1703.07823*, 2017.
[10] M. Farajtabar, X. Ye, S. Harati, L. Song, and H. Zha. Multistage campaigning in social networks. In *Advances in Neural Information Processing Systems*, pages 4718–4726, 2016.
[11] S. Goel, D. J. Watts, and D. G. Goldstein. The structure of online diffusion networks. In *Proceedings of the 13th ACM conference on electronic commerce*, pages 623–638, 2012.
[12] M. Gomez-Rodriguez, D. Balduzzi, and B. Schölkopf. Uncovering the temporal dynamics of diffusion networks. In *Proceedings of the International Conference on Machine Learning*, 2011.
[13] M. Gomez-Rodriguez, J. Leskovec, and A. Krause. Inferring networks of diffusion and influence. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1019–1028. ACM, 2010.
[14] M. Granovetter. The strength of weak ties. *American journal of sociology*, pages 1360–1380, 1973.
[15] S. A. Hosseini, K. Alizadeh, A. Khodadadi, A. Arabzadeh, M. Farajtabar, H. Zha, and H. R. Rabiee. Recurrent poisson factorization for temporal recommendation. *arXiv preprint arXiv:1703.01442*, 2017.
[16] D. Hunter, P. Smyth, D. Q. Vu, and A. U. Asuncion. Dynamic egocentric models for citation networks. In *Proceedings of the 28th International Conference on Machine Learning*, pages 857–864, 2011.
[17] T. Iwata, A. Shah, and Z. Ghahramani. Discovering latent influence in online social activities via shared cascade poisson processes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 266–274. ACM, 2013.
[18] M. R. Karimi, E. Tavakoli, M. Farajtabar, L. Song, and M. Gomez-Rodriguez. Smart broadcasting: Do you want to be seen? *arXiv preprint arXiv:1605.06855*, 2016.
[19] H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, pages 591–600, New York, NY, USA, 2010. ACM.
[20] J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins. Microscopic evolution of social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 462–470. ACM, 2008.
[21] J. Leskovec, D. Chakrabarti, J. Kleinberg, C. Faloutsos, and Z. Ghahramani. Kronecker graphs: An approach to modeling networks. *Journal of Machine Learning Research*, 11(Feb):985–1042, 2010.
[22] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 177–187. ACM, 2005.
[23] S. W. Linderman and R. P. Adams. Discovering latent network structure in point process data. In *International Conference on Machine Learning (ICML)*, 2014.
[24] T. J. Liniger. *Multivariate Hawkes Processes*. PhD thesis, Swiss Federal Institute of Technology Zurich, 2009.
[25] S. A. Myers and J. Leskovec. The bursty dynamics of the twitter information network. In *23rd International Conference on the World Wide Web*, pages 913–924, 2014.
[26] D. M. Romero and J. Kleinberg. The directed closure process in hybrid social-information networks, with an analysis of link formation on twitter. In *ICWSM*, 2010.
[27] B. Tabibian, I. Valera, M. Farajtabar, L. Song, B. Schölkopf, and M. Gomez-Rodriguez. Distilling information reliability and source trustworthiness from digital traces. *arXiv preprint arXiv:1610.07472*, 2016.
[28] L. Tran, M. Farajtabar, L. Song, and H. Zha. Netcodec: Community detection from individual activities. In *SDM*, 2015.
[29] J. Ugander, L. Backstrom, and J. Kleinberg. Subgraph frequencies: Mapping the empirical and extremal geography of large graph collections. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1307–1318. International World Wide Web Conferences Steering Committee, 2013.
[30] D. Q. Vu, D. Hunter, P. Smyth, and A. U. Asuncion. Continuous-time regression models for longitudinal networks. In *Advances in Neural Information Processing Systems*, pages 2492–2500, 2011.
[31] D. J. Watts and S. H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(6684):440–442, June 1998.
[32] L. Weng, J. Ratkiewicz, N. Perra, B. Gonçalves, C. Castillo, F. Bonchi, R. Schifanella, F. Menczer, and A. Flammini. The role of information diffusion in the evolution of social networks. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 356–364. ACM, 2013.
[33] A. Zarezade, A. Khodadadi, M. Farajtabar, H. R. Rabiee, and H. Zha. Correlated cascades: Compete or cooperate. *arXiv preprint arXiv:1510.00936*, 2015.
[34] K. Zhou, H. Zha, and L. Song. Learning triggering kernels for multi-dimensional hawkes processes. In *International Conference on Machine Learning (ICML)*, 2013.