Influence Maximization in Online Social Networks

Cigdem Aslay*
ISI Foundation
Turin, Italy
cigdem.aslay@isi.it

Laks V.S. Lakshmanan
University of British Columbia
Vancouver, B.C., Canada
laks@cs.ubc.ca

Wei Lu
Rupert Labs
Mountain View, CA
welu@cs.ubc.ca

Xiaokui Xiao
Nanyang Technological University
Singapore
xkxiao@ntu.edu.sg

ACM Reference Format:
Cigdem Aslay, Laks V.S. Lakshmanan, Wei Lu, and Xiaokui Xiao. 2018. Influence Maximization in Online Social Networks. In *WSDM 2018: WSDM 2018: The Eleventh ACM International Conference on Web Search and Data Mining , February 5–9, 2018, Marina Del Rey, CA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3159652.3162007

MOTIVATION

Viral marketing, a popular concept in the business literature, has recently attracted a lot of attention also in computer science, due to its high application potential and computational challenges.The idea of viral marketing is simple yet appealing: by targeting the most influential users in a social network (*e.g.*, by giving them free or price-discounted samples), one can exploit the power of the network effect through word-of-mouth, thus delivering the marketing message to a large portion of the network analogous to the spread of a virus.

Influence maximization is the key algorithmic problem behind viral marketing. The problem, as originally defined by Kempe *et al.* [32], is as follows: given (*i*) a directed social network, (*ii*) a set of weights associated with edges, representing strengths or probabilities of influence among users, (*iii*) a stochastic influence propagation model that governs how a certain behavior would diffuse among users, and (*iv*) a cardinality constraint k, aim is to identify a set of k nodes, called the "seed set", that can be targeted to maximize the expected number of influenced nodes. Kempe *et al.* studied influence maximization as a discrete optimization problem, obtaining provable approximation guarantees under several social influence propagation models. Following this seminal work, research on the dynamics of social influence propagation and influence maximization took off in several dimensions.

In this tutorial we cover major algorithmic and theoretical developments and issues arising in this field. A good chunk of this research has been done in the data mining and databases communities. While related tutorials [2, 10, 23, 24, 35, 55] appeared in VLDB'11, KDD'11, KDD'12, WSDM'13, WWW'15, and KDD'15, our tutorial showcase recent advances in the field not covered by the previous tutorials. A tutorial like the one that we propose can allow interested researchers and practitioners to gain up-to-date knowledge on the recent theoretical and algorithmic developments and seize the opportunity to contribute to the advancement of this fast-paced field.

*Main contact.

WSDM 2018, February 5–9, 2018, Marina Del Rey, CA, USA

ACM ISBN 978-1-4503-5581-0/18/02.
https://doi.org/10.1145/3159652.3162007

TUTORIAL OUTLINE

Part I: Influence maximization problem (70%)
We start the tutorial by introducing social influence propagation models and the standard influence maximization problem as defined by Kempe *et al.* We then extensively survey the algorithmic efforts for devising scalable and efficient influence maximization algorithms. We also discuss recently defined context-aware influence maximization problems, as well as closely related alternative optimization problems. The detailed outline of Part I is as follows:

- Influence maximization problem (15%)
  - Discrete- [32, 33, 43] and continuous-time [21, 51] models
  - Hardness & greedy approximation [15, 17, 32, 36]
- Improving efficiency and scalability (20%)
  - Ranking and score based heuristics [15, 17, 22, 27, 31, 47]
  - Sketches, reverse influence sampling [8, 19, 30, 46, 57, 58]
- Context-aware influence maximization (25%)
  - Topic- [3, 11, 13, 38] and time-aware [14, 39]
  - Location-aware [37, 53, 54, 61]
  - Competitive and comparative [7, 9, 40, 41, 50]
  - Dynamic influence maximization [18, 48]
- Other optimization objectives (10%)
  - Business-oriented optimization: adoption, profit, recommendation, etc [6, 42]
  - Robustness of influence maximization [12, 28, 49]

Part II: Modeling and learning social influence (15%)
In this part of the tutorial we first review the body of research that aims to learn influence weights from past propagation traces [25, 45, 52]. We then cover recent efforts that aim to directly learn the influence function [20, 26, 29, 44], thus, enabling the estimation of

the expected influence of a given set of nodes directly from given propagation traces. We also review a recent line of work which leverages active learning and multi-armed bandits to tackle "on-line" influence maximization where influence probabilities are not known and no past propagation data is available [16, 34, 59, 60].

Part III: Broader optimization objectives (15%)
The area of computational advertising has attracted a lot of interest during the last decade. However, with the advent of social advertising, the standard interest-driven allocation of ads to users has become inadequate as it fails to leverage the potential of social influence. In comparison to computational advertising, social advertising is still in its early stage. In this part of the tutorial, we review the recent initial efforts that aim to bridge the gap between viral marketing and social advertising [1, 4, 5, 56].

REFERENCES

[1] Zeinab Abbassi, Aditya Bhaskara, and Vishal Misra. 2015. Optimizing Display Advertising in Online Social Networks. In *WWW*. 1–11.
[2] Divyakant Agrawal, Ceren Budak, and Amr El Abbadi. 2011. Information diffusion in social networks: Observing and influencing societal interests. *PVLDB* 4, 12 (2011), 1–5.
[3] Cigdem Aslay, Nicola Barbieri, Francesco Bonchi, and Ricardo A Baeza-Yates. 2014. Online Topic-aware Influence Maximization Queries.. In *EDBT*. 295–306.
[4] Cigdem Aslay, Francesco Bonchi, Laks VS Lakshmanan, and Wei Lu. 2017. Revenue maximization in incentivized social advertising. *PVLDB* 10, 11 (2017), 1238–1249.
[5] Cigdem Aslay, Wei Lu, Francesco Bonchi, Amit Goyal, and Laks VS Lakshmanan. 2015. Viral marketing meets social advertising: Ad allocation with minimum regret. *PVLDB* 8, 7 (2015), 814–825.
[6] Smriti Bhagat, Amit Goyal, and Laks VS Lakshmanan. 2012. Maximizing product adoption in social networks. In *WSDM*. 603–612.
[7] Bharathi et al., 2007. Competitive influence maximization in social networks. In *International Workshop on Web and Internet Economics*. Springer, 306–311.
[8] Christian Borgs, Michael Brautbar, Jennifer T Chayes, and Brendan Lucier. 2014. Maximizing Social Influence in Nearly Optimal Time.. In *SODA*. SIAM, 946–957.
[9] Allan Borodin, Yuval Filmus, and Joel Oren. 2010. Threshold models for competitive influence in social networks. In *WINE*. Springer, 539–550.
[10] Carlos Castillo, Wei Chen, and Laks VS Lakshmanan. 2012. Information and influence spread in social networks. In *KDD*. ACM.
[11] Shuo Chen, Ju Fan, Guoliang Li, Jianhua Feng, Kian-lee Tan, and Jinhui Tang. 2015. Online topic-aware influence maximization. *PVLDB* 8, 6 (2015), 666–677.
[12] Wei Chen, Tian Lin, Zihan Tan, Mingfei Zhao, and Xuren Zhou. 2016. Robust Influence Maximization. In *KDD (KDD '16)*. ACM, 795–804.
[13] Wei Chen, Tian Lin, and Cheng Yang. 2015. Real-time Topic-aware Influence Maximization Using Preprocessing. In *CSoNet*. Springer, 1–13.
[14] Wei Chen, Wei Lu, and Ning Zhang. 2012. Time-critical influence maximization in social networks with time-delayed diffusion process.. In *AAAI*, Vol. 2012. 1–5.
[15] Wei Chen, Chi Wang, and Yajun Wang. 2010. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *KDD*. 1029–1038.
[16] Wei Chen, Yajun Wang, and Yang Yuan. 2013. Combinatorial Multi-Armed Bandit: General Framework and Applications. In *ICML*. 151–159.
[17] Wei Chen, Yifei Yuan, and Li Zhang. 2010. Scalable Influence Maximization in Social Networks under the Linear Threshold Model. In *ICDM*. 88–97.
[18] Xiaodong Chen, Guojie Song, Xinran He, and Kunqing Xie. 2015. On influential nodes tracking in dynamic social networks. In *SDM*. SIAM, 613–621.
[19] Cohen et al., 2014. Sketch-based Influence Maximization and Computation: Scaling up with Guarantees. In *CIKM*.
[20] Nan Du, Yingyu Liang, Maria Balcan, and Le Song. 2014. Influence function learning in information diffusion networks. In *ICML*. 2016–2024.
[21] Nan Du, Le Song, Manuel Gomez-Rodriguez, and Hongyuan Zha. 2013. Scalable influence estimation in continuous-time diffusion networks. In *NIPS*. 3147–3155.
[22] Sainyam Galhotra, Akhil Arora, and Shourya Roy. 2016. Holistic influence maximization: Combining scalability and efficiency with opinion-aware models. In *SIGMOD*. ACM, 743–758.
[23] M. Gomez-Rodriguez and Le Song. 2015. Diffusion in Social and Information Metworks: Research Problems; Probabilistic Models & Machine Learning Methods. In *WWW*. ACM, 1527–1528.
[24] M. Gomez Rodriguez and Le Song. 2015. Diffusion in social and information networks: research problems, probabilistic models and machine learning methods. In *KDD*. ACM, 2315–2316.
[25] Amit Goyal, Francesco Bonchi, and Laks V. S. Lakshmanan. 2010. Learning influence probabilities in social networks. In *WSDM*. 241–250.
[26] Amit Goyal, Francesco Bonchi, and Laks V. S. Lakshmanan. 2011. A Data-Based Approach to Social Influence Maximization. *PVLDB* 5, 1 (2011), 73–84.
[27] Amit Goyal, Wei Lu, and Laks V. S. Lakshmanan. 2011. SIMPATH: An Efficient Algorithm for Influence Maximization under the Linear Threshold Model. In *ICDM*. 211–220.
[28] Xinran He and David Kempe. 2016. Robust Influence Maximization. In *KDD*. ACM, 885–894.
[29] Xinran He, Ke Xu, David Kempe, and Yan Liu. 2016. Learning Influence Functions from Incomplete Observations. In *NIPS*. 2065–2073.
[30] Keke Huang, Sibo Wang, Glenn Bevilacqua, Xiaokui Xiao, and Laks VS Lakshmanan. 2017. Revisiting the Stop-and-Stare Algorithms for Influence Maximization. *PVLDB* 10, 9 (2017).
[31] Kyomin Jung, Wooram Heo, and Wei Chen. 2012. Irie: Scalable and robust influence maximization in social networks. In *ICDM*. IEEE, 918–923.
[32] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *KDD*. ACM, 137–146.
[33] David Kempe, Jon M. Kleinberg, and Éva Tardos. 2005. Influential Nodes in a Diffusion Model for Social Networks. In *ICALP*. 1127–1138.
[34] Siyu Lei, Silviu Maniu, Luyi Mo, Reynold Cheng, and Pierre Senellart. 2015. Online influence maximization. In *KDD*. ACM, 645–654.
[35] Jure Leskovec. 2011. Social media analytics: tracking, modeling and predicting the flow of information through networks. In *WWW*. ACM, 277–278.
[36] Leskovec et al., 2007. Cost-effective outbreak detection in networks. In *KDD*. ACM, 420–429.
[37] Guoliang Li, Shuo Chen, Jianhua Feng, Kian-lee Tan, and Wen-syan Li. 2014. Efficient location-aware influence maximization. In *SIGMOD*. ACM, 87–98.
[38] Yuchen Li, Dongxiang Zhang, and Kian-Lee Tan. 2015. Real-time targeted influence maximization for online advertisements. *PVLDB* 8, 10 (2015), 1070–1081.
[39] Bo Liu, Gao Cong, Dong Xu, and Yifeng Zeng. 2012. Time constrained influence maximization in social networks. In *ICDM*. IEEE, 439–448.
[40] Lu et al., 2013. The bang for the buck: fair competitive viral marketing from the host perspective. In *KDD*. ACM, 928–936.
[41] Wei Lu, Wei Chen, and Laks VS Lakshmanan. 2015. From competition to complementarity: comparative influence diffusion and maximization. *PVLDB* 9, 2 (2015), 60–71.
[42] Wei Lu and Laks VS Lakshmanan. 2012. Profit maximization over social networks. In *ICDM*. IEEE, 479–488.
[43] Elchanan Mossel and Sebastien Roch. 2007. On the submodularity of influence in social networks. In *STOC*. 128–134.
[44] Harikrishna Narasimhan, David C Parkes, and Yaron Singer. 2015. Learnability of influence in networks. In *NIPS*. 3186–3194.
[45] Praneeth Netrapalli and Sujay Sanghavi. 2012. Learning the graph of epidemic cascades. In *ACM SIGMETRICS*, Vol. 40. ACM, 211–222.
[46] Hung T. Nguyen, My T. Thai, and Thang N. Dinh. 2016. Stop-and-Stare: Optimal Sampling Algorithms for Viral Marketing in Billion-scale Networks. In *SIGMOD*.
[47] Ohsaka et al., 2014. Fast and Accurate Influence Maximization on Large Networks with Pruned Monte-Carlo Simulations.. In *AAAI*. 138–144.
[48] Ohsaka et al., 2016. Dynamic influence analysis in evolving networks. *PVLDB* 9, 12 (2016), 1077–1088.
[49] Naoto Ohsaka and Yuichi Yoshida. 2017. Portfolio Optimization for Influence Spread. In *WWW*. ACM, 977–985.
[50] Pathak et al., 2010. A generalized linear threshold model for multiple cascades. In *ICDM*. IEEE, 965–970.
[51] Gomez-Rodriguez et al., 2014. Uncovering the structure and temporal dynamics of information propagation. *Network Science* 2, 01 (2014), 26–65.
[52] Saito et al., 2008. Prediction of Information Diffusion Probabilities for Independent Cascade Model. In *KES*.
[53] Saleem et al., 2017. Location Influence in Location-based Social Networks. In *WSDM*. ACM, 621–630.
[54] Chonggang Song, Wynne Hsu, and Mong Li Lee. 2016. Targeted Influence Maximization in Social Networks. In *CIKM*. ACM, 1683–1692.
[55] Jimeng Sun and Jie Tang. 2013. Models and algorithms for social influence analysis. In *WSDM*. ACM, 775–776.
[56] Shaojie Tang and Jing Yuan. 2016. Optimizing Ad Allocation in Social Advertising. In *CIKM*. ACM, 1383–1392.
[57] Youze Tang, Yanchen Shi, and Xiaokui Xiao. 2015. Influence maximization in near-linear time: A martingale approach. In *SIGMOD*. ACM, 1539–1554.
[58] Youze Tang, Xiaokui Xiao, and Yanchen Shi. 2014. Influence Maximization: Near-Optimal Time Complexity Meets Practical Efficiency. *SIGMOD* (2014).
[59] Vaswani et al., 2017. Model-Independent Online Learning for Influence Maximization. In *ICML*. 3530–3539.
[60] Wen et al., 2017. Online Influence Maximization under Independent Cascade Model with Semi-Bandit Feedback. In *NIPS*. 3025–3035.
[61] Zhang et al., 2012. Evaluating geo-social influence in location-based social networks. In *CIKM*. ACM, 1442–1451.