Engineering and Applied Science Research 2021;48(2):221-229 Review Article

EASR

Engineering and Applied Science Research

https://www.tci-thaijo.org/index.php/easr/index

Published by the Faculty of Engineering, Khon Kaen University, Thailand

Influence maximization algorithm: Review on current approaches and limitations

Kristo Radion Purba*, David Asirvatham and Raja Kumar Murugesan

School of Computing and IT, Taylor's University, Malaysia

Received 10 February 2020
Revised 29 July 2020
Accepted 11 August 2020

Abstract

Influencing customers through social media is a new form of marketing. Recently, there were studies on the Influence Maximization (IM) problem, which aimed to identify influencers that can spread influence to a wider audience. The complex social media network requires efficient IM algorithms, in which small improvements will lead to a performance boost. In this research, recent articles on IM were reviewed. This review aims to identify the current approaches, enhancements, factors, diffusion models, and objectives of IM. In typical IM formulation, a social network is represented as a graph with nodes (user) and edges (relation). There are graph-based and non-graph-based IM approaches. Graph-based IM approaches include greedy and heuristic algorithms. The objectives of IM studies were optimizations on large or complex networks, on unknown networks, using bandit, using relation impacts, or general optimization. IM algorithms were continuously getting better. However, there are aspects that are still improvable, i.e. pre-calculation, thresholds estimation, seeds selection, integration of neural networks, and more importantly, real-life validation methods. This study will help in identifying possible improvements based on current IM limitations. Effective IM methods will help business users to identify influencers more accurately.

Keywords: Influence maximization, Social media, Graph theory, Algorithm design

1. Introduction

The recent growth of social media, especially Facebook and Instagram [1], caused the rise of influencers. An influencer is a user that is more central than others [2]. Influencer marketing has benefits compared to traditional marketing [3], i.e. trustworthy [4], and easily raises the need for a product [5]. Recently, there were studies on influence maximization (IM), intending to identify seeds set [6] that can reach a maximum number of users with minimum resources [7]. An IM algorithm uses simulation to diffuse influence [8].

This study aims to categorize the approaches and enhancements on IM, incorporated factors, diffusion models, and the objective of the IM algorithms. By doing the review, the limitations of recent IM algorithms can be identified and will help in identifying the possibilities for future work. The following questions were addressed, i.e. (R1) What are the categories of IM algorithms? (R2) What are the approaches, enhancements, incorporated factors, diffusion models, and objectives of IM algorithms? (R3) What are the limitations of current IM algorithms? This research is presented in the following sections, i.e. IM categorization, discussion, and limitations.

2. IM Categorization

In recent studies, various combination of approach, enhancements, factors, and diffusion model has been used to meet certain objectives. The list of categorizations is presented in Table 1. IM is a NP-Hard problem [9], which usually solved using randomizations. There are two approaches of IM, i.e. graph-based and non-graph based. Recent studies usually come with enhancements in specific areas in the common IM algorithm. Commonly, graph-based is divided into greedy and heuristic [10], and hybrid which is the combination of both. Hybrid method such as TIM and TIM$^+$ [11] used greedy for nodes selection, and heuristic approach prior to parameters estimation. To make IM more realistic, factors based on users and relations were incorporated.

3. Graph-Based IM approach

On graph-based IM, a diffusion model is used to simulate the spread of influence. The widely used models are Independent Cascade (IC) and Linear Threshold (LT) models [9]. A social media is represented as a directed graph G (V, E), where V is node (user), and E is an edge (relationship, i.e. indegree/follower, and outdegree/following). Besides IC and LT, there are other models which add various social network properties in the simulation. Below is the summary of the diffusion models:

- Independent Cascade (IC) [9]: At each time step t, each active node will be given a single chance to activate each of its followers, with a probability of $P_{u,v}$, where u is the source node, and v is the target node. If a node remains inactive in this time step, it can't be further activated in the future time steps. The $P_{u,v}$ can be set to random or an arbitrary number.
- Linear Threshold (LT) [9]: In this model, an edge has an influence weight, and each node has a threshold chosen uniformly at random. An inactive node will be activated if the sum of its incoming edges' weight $W_{u,v}$ exceeds its threshold T_u. This threshold is set to a uniform random number across all of the nodes.

*Corresponding author. Tel.: +6016 529 2630
Email address: kristoradionpurba@sd.taylors.edu.my
doi: 10.14456/easr.2021.24

Table 1 Categorization of Research on Influence Maximization

Approach	Algorithm Enhancements	Incorporated Factor	Objective
• Graph-based: Greedy • Graph-based: Heuristic • Graph-based: Hybrid (greedy and heuristic) • Non-graph-based	• Improving seeds selection • Improving node probing, or information propagation • Improve diffusion • Target-oriented • Use on multiple social networks	• Assortativity • Budget • Incentive • Inducement • Multiple networks • Novelty decay • Product • Sentiment • Time/period • Topic • User statistics • Community structure	• Optimization on large or complex networks • Optimization using bandit • Optimization on unknown network • Optimization using relation impact • Greedy improvement • Heuristic improvement

- Weighted cascade (WC) [9]: Oftentimes, IC is referred as WC, and in fact, most studies were using WC. The WC (and also LT) model uses an inverse proportion of the number of followees as the edge weight, instead of random in IC model.
- Game-based diffusion [10]: This model adds community aspects such as sociality, randomness, individual and group effect [12].
- Susceptible-Infected-Recovered (SIR) [13]: In this model, nodes can transition from S (susceptible) to I (Infected), and also vice versa (I to S), unlike only S to I in IC/LT models.
- Rumor model [14]: This model adds a probability μ which can reduce the influence of the spreader. This is because a spreader can become more active upon receiving positive responses and vice versa.

Graph-based IM performance can be measured based on three criteria, i.e. influence spread, runtime, sometimes memory [7], or fairness benchmark [15]. Unfortunately, many variations of parameters and datasets were used in IM studies, which makes it hard to compare the performances. A study of performance comparison between 16 algorithms under IC model were studied by [16] using 100 datasets. The study found that methods based on adaptive degree centrality to be the best performer, however, with no large performance gap with other existing methods.

We categorize recent IM studies based on the main optimization. IM on large or complex networks address the dynamically growing or large networks, or various network properties. IM using bandit (or bandit-based IM) takes a different algorithmic approach in finding seeds set. Optimization on unknown networks is based on the idea of difficulties in getting social media data. IM using relation impacts focused on adding user-based properties to make IM more realistic. Lastly, general improvement means the enhancements of IM algorithms to improve theoretical performance. The IM studies are discussed in Sections 3.1 to 3.6.

3.1 Optimization on large or complex networks

Optimization on large networks focused on memory, accuracy, and stability. To make these enhancements, some studies improved existing algorithms. The research on IM on large networks is described in Table 2.

Runtime enhancements were done in various ways. One of the earliest efficient algorithms, called CELF (cost-efficient lazy forward) [17] keeps the marginal gain of each node in each iteration. CELF++ algorithm [18] improved the runtime of CELF by 35-55%, by reducing the number of marginal gain calculation. ASIM algorithm [19] further improved the runtime of CELF++ by assigning influence score to each node, which makes it 6-8 times faster than CELF++. Collective influence (CI) algorithm [20, 21] calculates influence (CI value) based on distance ℓ from each node, instead of globally.

TIM algorithm [21] enhanced accuracy by using lower bound θ estimation and picking seeds that cover a large number of RR (reverse reachable) sets. IMM algorithm [22] improved the TIM algorithm by adding martingale estimation. Accuracy enhancements can also be done by using new diffusion models [23, 24], or adding factors [25, 26]. IM based on real diffusion cascade [23, 24] used simulation from actual data, instead of Monte Carlo. Stop-and-Stare (SSA) algorithm [27] is one of the state-of-the-art algorithms, with remarkable runtime (typically under one minute) and influence spread similar or better than IMM. More recently, DISCO algorithm [28] employed deep learning technique, and has a similar influence spread with SSA, with a faster runtime. However, DISCO requires a training phase which takes up to three days of execution depending on the network size.

Stability enhancement or robust IM formulation was also common, where the goal of IM is to find maximum influence over the entire uncertain models. Methods for robust IM were branch-and-cut algorithm [29], IM in hyperparametric models [30]. In terms of adding factors, there were budget for competitive marketing [25, 26], and user, product, timing, item [31].

Complex network is another emerging area in IM, with a focus to increase the accuracy of seeds set identification to achieve the performance close to greedy algorithm [16]. Unlike trivial networks, modeling the social media connections into a graph will most likely produce complex networks, with some very high degree nodes and many low degree nodes [14]. The addition of centrality measures improved the IM performance in these studies, such as degree and closeness centrality measures [16], degree correlation [14], spreading influence related centrality (SIRC) based on multiple centrality indices [32]. The diffusion degree and maximum influence degree centrality measures [6] were based on the idea of reduced influence on increased distance. Overall, these centrality measures can improve many algorithms, even the less performant ones, by 2-5% [16].

3.2 Optimization using bandit

A bandit agent on IM learns from user activity changes, which aims to minimize regret. The bandit agent has a parameter of upper confidence bounds (UCB), which means a threshold of optimism in the uncertainty [33]. There is also cumulative regret function, in which less cumulative regret (loss of influenced nodes cumulated from every iteration) means more efficient. In contrast with a conventional IM approach, bandit-based IM is commonly benchmarked with the theoretical average regret value, instead of influence spread. There were various bandit-based IM improvements, especially for the parameters, as shown in Table 3.

The seeds selection process can be improved by adding factor [34], estimating influence diffusion [35], estimating activation

Table 2 Optimization on Large or Complex Network Research

Approach/Method	Ref, Year	Enhancement Area	Details
CELF	[17], 2007	Runtime	Marginal gain calculation
CELF++	[18], 2011	Runtime	Reducing calculation
ASIM	[19], 2015	Runtime, memory	Add influence score
Collective Influence	[20], 2015	Runtime	-
Improved Collective Influence	[21], 2016	Runtime	Remove redundancies in previous algorithm [19]
TIM	[11], 2014	Accuracy	Usage of RR-sets
IMM	[22], 2015	Accuracy	Martingale estimation
Stop-and-Stare (SSA)	[27], 2016	Accuracy, runtime	Overlap of RR-sets
Hyperparametric model	[30], 2019	Stability	-
Branch-and-cut	[29], 2019	Stability	-
Swarm intelligence based	[23], 2018	Diffusion model	Real diffusion cascade
Adapt to dynamic network	[24], 2018	Diffusion model	Real diffusion cascade
Competitive IM	[25], 2019	Deep reinforcement	-
Competitive IM	[26], 2019	Deep reinforcement	-
Continuous time diffusion	[31], 2017	Add factors	User, product, timing
IM with centrality measures	[6], 2014	Complex networks	Diffusion degree and maximum infl. degree
DISCO algorithm	[28], 2019	Accuracy, runtime	Deep learning and network embedding
Performance comparisons between algorithms/datasets	[16], 2019	Complex networks	-
IM with hybrid centrality	[32], 2019	Complex networks	SIRC
IM on correlated networks	[14], 2020	Complex networks, diffusion model, add factor	Rumor diff. model, community structure factor

Table 3 Optimization Using Bandit Research

Approach/Method	Ref, Year	Enhancements				
		Cumulative regret	UCB	Seeds Selection	Influence degree	Scalability
Learn changes in network size	[35], 2018		v	v		
Probabilities estimation	[36], 2015	v	v	v		
Linear generalization	[37], 2017	v	v			
Factorization bandits	[38], 2019	v		v		
Two-phase (offline, online)	[39], 2015			v		
Turing estimator, diminishing influence	[33], 2017		v		v	
ERR-sets to pick seeds	[35], 2018			v		v
Influence checkpoints	[40], 2017				v	v
Partial feedback	[41], 2017			v		v
Calculation on network change	[42], 2017			v		v
Add incentives factor	[34], 2019			v		

Probabilities [36, 39]. Incentives [34] and assortativity [38] are among the added factors to replicate a real social network. Some bandit methods focused on network scalability [35, 40-42], which means new users are added during runtime. A limitation in these studies was the requirement to tune the parameters, i.e. ε in sampling phase [35], parameters k, β, N, L, U [40], delay α [41], and probability threshold η [42] that affect the tradeoff between performance and runtime.

3.3 Optimization on unknown network

Capturing data using either web scraping or API often produces incomplete information, which produces an "unknown" network. Recent studies predicted the missing information based on the behavior of influential and neighboring nodes. There were other common features from social media that are still not used in this area such as fake accounts [43] and hashtags [44], which are studied in other social network research.

The idea of unknown graph IM is to predict the expected degree (or influence) of nodes. There were two main approaches, i.e. heuristic [45-47], or neural network [48]. In the heuristic approaches, predictions were based on neighbors [45], querying sequences of nodes [46], friendship paradox [47]. The neural network approach [48] used the behavior of nodes for diffusion training.

The heuristic approach, since it relies on randomizations, often introduces inconsistencies. The IMUG algorithm [45] optimized probing efficiency by spreading seeds to the highest expected degree nodes. However, it required a lot of iterations in the greedy part. The ARISEN [46], in contrast, required less iteration but needs tuning of q (activation probability), K (number of seeds), and query. The CHANGE algorithm [47] improved ARISEN by removing the query-based method and leverage a friendship paradox (a randomly chosen neighbor of a node tends to have a higher degree than the node itself). However, it was based on a specific youth network, which might not be suitable for others.

3.4 Optimization using relation impacts

IM using relation impacts is relatable to the real world. Research on IM with relation impact can be described as shown in Table 4. In analyzing relation impact, recent research used several user factors, i.e. engagement, sentiment [49, 50], freeloaders [51], targeted ads [52]. Engagement in social media was described by different actions, i.e. user interaction [53], conversation content and reply [54], assortativity, the joint

Table 4 IM using Relation Impact Research

Approach/Method	Ref, Year	Enhancement Area	Details
Influence scoring on Twitter	[53]	Add Factors	User interaction
Influence scoring on Healthcare	[54]	Add Factors	Reply immediacy
Empirically motivated IM	[55]	Add Factors	Assortativity, influence, susceptibility
IM with balanced index (BI) & group performance index (GPI)	[56]	Add Factors	Resistance, influence on second neighborhoods
Evidential opinion-based IM	[50]	Add Factors	Opinion
OSIM and EaSyIM	[57]	Add Factors	Opinion
Negative-Aware IM	[51]	Add Factors	Freeloader
Personalized IM	[52]	Add Factors	Targeted ads
Modified PageRank algorithm	[58]	Add Factors	Topic
Topic-based Social Influence Measurement	[59]	Add Factors	Topic
Topic-aware IM (TIM)	[60]	Add Factors	Topic
Signed Voter model IM (SVIM)	[49]	Term-based, add factor	Short and long-term IM, friend and foe
IM with novelty decay (IMND)	[61]	Term-based, add factors	Long-term IM, novelty decay factor

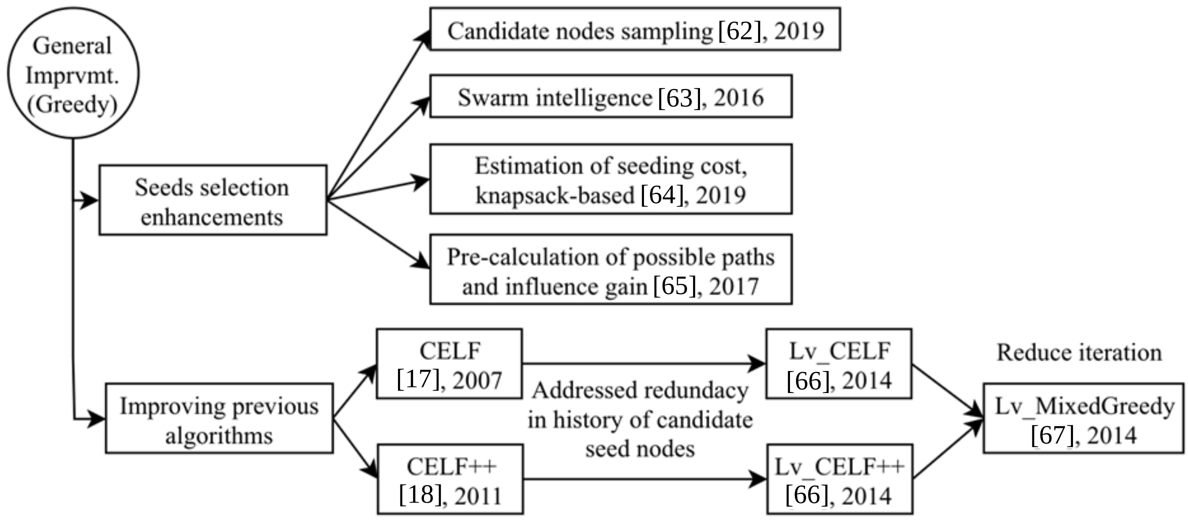

Figure 1 General Improvement on Greedy Algorithms

distribution of influence and susceptibility [55], resistance to influence and influence on second neighborhoods [56]. These factors are used to decide whether to activate nodes or not during the IM simulation.

Relation impact factors were proven to improve the performance of IM, such as, 21.7% influence propagation improvement [55], 70% of influence spread improvement [51], some improvements on performance and time complexity [56]. There were also studies using topic analysis, such as topic influence score based on Twitter message [58, 59], and topic learnt from connections data [60]. Some enhancements were used, such as skipping users with insignificant topic influence [68] by estimating threshold of influence spread.

There were also term-based IM studies, i.e. long-term IM (e.g. political campaign) and short-term IM (e.g. vote in election day). Examples of long-term IM algorithms are SVIM-L [49] and IM with novelty decay (repeated exposure that causes diminishing influence) [61]. An example of short-term influence is SVIM-S [49]. These "terms" are used to skip nodes with small influence in later iterations. The short-term IM can be suitable for identifying viral hashtags, which can occur at any time on a social network.

3.5 General improvements on greedy algorithm

Commonly, greedy approaches have a major weakness, i.e. requires a lot of iterations [69]. It is derived from the first Greedy IM approach [9], which simulates each node to find the best influence spreader at every *k* (number of seeds). To mitigate this drawback, there were enhancements on seeds selection, the addition of factors, and improvement to existing algorithms, as shown in Figure 1.

Optimization on seeds selection was done using methods such as sampling [62], swarm intelligence [63], knapsack [64], pre-calculation [65]. A combination of heuristic and greedy was often done to achieve better results. The MATI algorithm [65] used greedy pre-calculation of possible paths and influence gain, and heuristic for approximations. Another example is the combination of candidate nodes sampling using a heuristic, and verification using greedy [62].

There were improvements to existing algorithms, such as Lv_CELF and Lv_CELF++ [66], that improved CELF [17] and CELF++ [18]. It reduced the number of candidate seed nodes kept in table Q to reduce runtime. Further improvement was done by the Lv_MixedGreedy algorithm [67]. It used three strategies to reduce iterations, i.e. random live edge selection that replaces Monte Carlo, threshold θ, and strongly connected components.

3.6 General improvements on heuristic algorithm

The goal of a heuristic approach is to achieve better stability. Improvements in this area include seeds selection, runtime improvements, target-based, the addition of factors, an improvement on diffusion models, analysis on multiple social networks, as shown in Figure 2.

There were various approaches used to enhance seeds selection, such as Chernoff bounds sampling [70], swarm optimization [71], influence of neighbors [72], group search [73], neural network (NN) [28, 74]. Some specific methods in this area include the combination of SelectTopK, RankedReplace

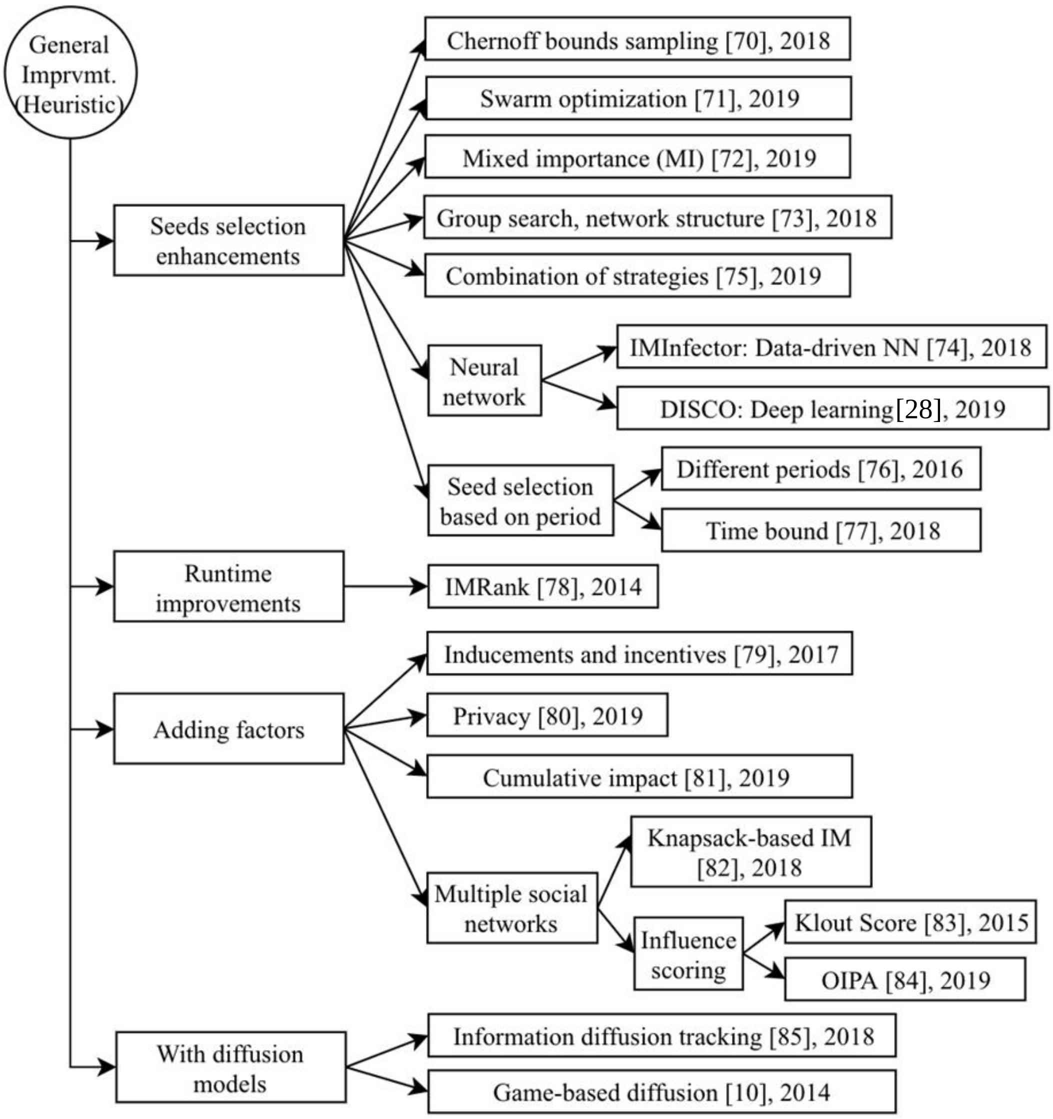

Figure 2 General Improvement on Heuristic Algorithms

and greedy [75]. Recently, studies that include neural networks in IM became more prominent. It was used for diffusion cascade training [74], or estimating the influence of nodes [28]. Seeds selection can be done in a single time period, different time periods [76], or specific time-bound [77]. Runtime improvement was done by keeping previous iterations results, such as in IMRank [78].

The addition of factors in information propagation was also done on heuristic algorithms, such as inducements and incentives [79], privacy protection [80], cumulative impact [81]. Cumulative impact calculation was used to adapt to the fact that users need to be impacted multiple times before buying a product.

There were studies that analyzed users on multiple social networks, such as in [82]. However, the limitation here was the performance degrade as the number of overlapping users increase. Another research [83, 84] focused on influence scoring on users on multiple social networks. Some studies proposed new diffusion models to adapt to users' behavior. Game-based diffusion [10] adapts users' rational and irrational decisions that can be caused by friends' influence. Information diffusion tracking [85] captured direct and indirect influences.

4. Non-Graph based IM approach

Viral marketing is the key problem to be solved using IM [86], so ideally, the graph approach is more suitable. The impact of the relations in the network can be acquired by simulating influence diffusion. Research also showed that influence effect is chaining, e.g. friends' influence [87], connection density [88]. Despite the limitations, there were non-graph approaches for IM, such as user statistics based [83, 89], machine learning [90], evolutionary algorithms [91], and swarm intelligence [23]. Some limitations include no measurement of influence spread [89] [83], local optimum [23].

Non graph-based IM methods are more focused on finding top ranked influencers individually, instead of a set of influencers, which doesn't require diffusion models. More practical metrics were used, such as engagement [89], post reactions [83], users similarity [90]. Compared to influence spread metric, which is theoretical, these metrics can be more useful especially in a small campaign. The IM with evolutionary algorithm [91] was one of the state-of-the-art non-graph IM, with performance almost similar to greedy, but with a much faster runtime. However, recent graph-based IM algorithms didn't suffer slow runtimes anymore, such as SSA [27] which typically runs under a minute on a billion-scale network.

5. IM Algorithms limitations

In this section, limitations or possible improvements on existing research are presented in Table 5. On IM research, continuous improvements were always made by researchers.

Table 5 Analysis of Current IM Limitations

Category	Identified Problem	Limitations
IM Optimization on Large or Complex Network	Existing solutions may not scale as the network become larger, or lacked real-life social network properties	Pre-calculations or keeping previous results use more RAM. Validation based on ground truth data still has to be improved.
IM Optimization Using Bandit	Conventional IM approach and diffusion models did not incorporate node-level and edge-level feedbacks	Bandit-based IM is often more algorithmically complicated than conventional IM approach
IM Optimization on Unknown Graph	Existing solutions relied on a full graph which is difficult to obtain	Relied on thresholds, stability on different cases was untested
IM Optimization Using Relation Impacts	Existing solutions only analyzed influence diffusion based only on following/follower, that may not be realistic to the real world	Some other factors can still be added or improved, i.e. sentiment, a chain of topic similarity, fake account. Validation based on ground truth data still has to be improved.
General Improvements on Greedy Approach	Existing solutions can't adapt well on different scenarios, existing solutions have calculation redundancies	Solutions that used pre-calculations or keeping previous results may add more RAM usage
General Improvements on Heuristic Approach	Existing solutions can't adapt well on different scenarios, existing solutions have calculation redundancies	Relied on approximations. Pre-calculation can be added to improve, but it may degrade performance.
Non-graph IM (machine learning, statistics, AI)	Existing solutions were not fast enough, and mostly relied on a graph which is difficult to obtain	Subject to local optimum, or cannot adapt well with various properties on social network

Another common issue in IM and other network science studies is network noise, which can be caused by opinion polarity, emergence of technology [92], and virality caused by only a few early adopters [93]. Simulating IM without noise assumes a perfect connection between users. Network noise can disrupt the sub modularity of IM [94]. In recent studies, noise was generated by using Ising model [92] and shuffling edges [93]. Overall, various experiments of noise parameters in these studies showed that noise can greatly affect the performance of IM. Even though noise is inspired by real-world scenario, validation using ground truth data still lacked in these studies.

6. Discussion

In this research, recent studies on influence maximization (IM) on social media were reviewed. Graph-based IM was first coined by Kempe, et.al. [9], which usually consist of 2 ideas, i.e. (1) Technique to choose most influential nodes (seeds set) on the network that can spread influence faster to a wider audience, (2) Technique to diffuse information. Most of the recent studies were the combination of greedy and heuristic techniques. Greedy approach is mostly used for nodes selection, which has a high computational cost but also high theoretical warranty. Heuristic approach aims to increase efficiency, such as calculation before parameters estimation [11], node ranking [75], estimating node's degree [45], etc.

There were also non-graph techniques, which is fast and simple, and can be as accurate as graph techniques [91]. However, with the emergence of fast graph IM techniques [27, 28], runtime of IM algorithm is not an issue anymore. Graph-based IM is more robust and scalable, which can adapt to various network properties such as noise [93], dynamic network [24], term-based IM [49], decaying influence [61], multiple social networks analysis [82] [84], direct and indirect influence [85], etc.

7. Conclusion

Recent studies in Influence Maximization (IM) can be categorized based on the main optimizations, i.e. on large/complex network, using bandit, on unknown network, using relation impacts, and general optimization. There were plenty of enhancements in IM algorithms in recent research, i.e. pre-calculations, thresholds estimation, information diffusion technique, seeds selection. However, parameters in IM, i.e. propagation probability (for IC model), threshold (in LT model), as well as other algorithm-specific parameters need to be tuned to get an optimal result. Sometimes, it's necessary to find a balance between influence spread and runtime by tuning parameters manually. Only a few algorithms did an automated tuning process, by using learning.

Most of IM studies measured performance based on the theoretical result, i.e. the influence spread. Studies under the category of general greedy and heuristic optimizations were mostly focused on the algorithmic enhancements to improve theoretical results. Research under IM on large/complex network and optimization using bandit studied various network properties, such as noise, epidemic model, competitive IM, to mimic information diffusion in real-life as much as possible. Studies in IM optimization using relation impacts have added multiple relation aspects such as engagement, resistance, susceptibility.

Despite the addition of relation impacts and network properties, real-life validation is still a missing factor. A recent study used retweet chains as the validation method [74], however, retweet (or repost) data is not natively available or widely-used on other social networks, such as Instagram, YouTube, Facebook. Open questions to be addressed in future research are "how much is the correlation between the activated users (from the diffusion models) with the actual users influenced in real-life?" and "how to deal with inactive or spammy users in real-life?".

8. References

[1] Statista. Most famous social network sites worldwide as of July 2018, ranked by number of active users (in millions) [Internet]. 2018. [cited 2018 Oct 3]. Available: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/.

[2] Tanase D, Garcia D, Garas A, Schweitzer F. Emotions and activity profiles of influential users in product reviews communities. Front Phys. 2015;3(87):1-12.

[3] Odell P. The power of influencers: how brands can harness word-of-mouth and peer recommendations to connect with consumers. Norwalk: Chief Marketer; 2015.

[4] Raudeliūnienė J, Davidavičienė V, Tvaronavičienė M, Jonuška L. Evaluation of advertising campaigns on social media networks. Sustainability. 2018;10(4):973.

[5] Akar E, Yüksel HF, Z. Bulut A. The impact of social influence on the decision-making process of sports consumers on facebook. J Internet Appl Manag. 2015;6(2):5-27.

[6] Pal SK, Kundu S, Murthy CA. Centrality measures, upper bound, and influence maximization in large scale directed social networks. Fundam Inform. 2014;130(3):317-42.

[7] Arora A, Galhotra S, Ranu S. Debunking the myths of influence maximization. Proceedings of the 2017 ACM International Conference on Management of Data (SIGMOD'17); 2017 May 14-19; Chicago, USA. New York: ACM; 2017. p. 651-66.

[8] Ko YY, Cho KJ, Kim SW. Efficient and effective influence maximization in social networks: a hybrid-approach. Inform Sci. 2018;465:144-61.

[9] Kempe D, Kleinberg J, Tardos É. Maximizing the spread of influence through a social network. Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining; 2003 Aug 24-27; Washington, USA. New York: ACM; 2003. p. 137-46.

[10] Ok J, Jin Y, Choi J , Shin J, Yi Y. Influence maximization over strategic diffusion in social networks. 48th Annual Conference on Information Sciences and Systems (CISS); 2014 Mar 19-21; Princeton, USA. USA: IEEE; 2014. p. 1-5.

[11] Tang Y, Xiao X, Shi Y. Influence maximization: near-optimal time complexity meets practical efficiency. SIGMOD'14; 2014 Jun 22-27; Snowbird, USA. New York: ACM; 2014. p. 75-86.

[12] Li M, Wang X, Gao K, Zhang S. A survey on information diffusion in online social networks: models and methods. Information. 2017;8(4):118.

[13] Pastor-Satorras R, Castellano C, Mieghem PV, Vespignani A. Epidemic processes in complex networks. Rev Mod Phys. 2015;87(3):925.

[14] Vega-Oliveros DA, Costa LdF, Rodrigues FA. Influence maximization by rumor spreading on correlated networks through community identification. Comm Nonlinear Sci Numer Simulat. 2020;83:105094.

[15] Stoica AA, Chaintreau A. Fairness in social influence maximization. Proceedings of the 2019 World Wide Web Conference (WWW '19 Companion); 2019 May 13-14; San Francisco, USA. New York: ACM; 2019. p. 569-74.

[16] Erkol Ş, Castellano C, Radicchi F. Systematic comparison between methods for the detection of infuential spreaders in complex networks. Sci Rep. 2019;9(1):1-11.

[17] Leskovec J, Krause A, Guestrin C, Faloutsos C, VanBriesen J, Glance N. Cost-effective outbreak detection in networks. Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2017 Aug 12-15; San Jose, USA. New York: ACM; 2007. p. 420-9.

[18] Goyal A, Lu W, Lakshmanan LV. CELF++: Optimizing the greedy algorithm for influence maximization in social networks. Proceedings of the 20th international conference companion on World Wide Web; 2011 Mar 28 – Apr 1; Hyderabad India. New York: ACM; 2011. p. 47-8.

[19] Galhotra S, Arora A, Virinchi S, Roy S. ASIM: A scalable algorithm for influence maximization under the independent cascade model. Proceedings of the 24th International Conference on World Wide Web; 2015 May 18-22; Florence, Italy. New York: ACM; 2015. p. 35-6.

[20] Morone F, Makse HA. Influence maximization in complex networks through optimal percolation. Nature. 2015;524:65-8.

[21] Morone F, Min B, Bo L, Mari R, Makse H. Collective influence algorithm to find influencers via optimal percolation in massively large social media. Sci Rep. 2016;6:30062.

[22] Tang Y, Shi Y, Xiao X. Influence maximization in near-linear time: a martingale approach. Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data; 2015 May 31 - Jun 4; Melbourne, Australia. New York: ACM; 2015. pp. 1539-54.

[23] Şimşek A, Kara R. Using swarm intelligence algorithms to detect influential individuals for influence maximization in social networks. Expert Syst Appl. 2018;114:224-36.

[24] Panagopoulos G, Malliaros FD, Vazirgiannis M. DiffuGreedy: an influence maximization algorithm based on diffusion cascades. In: Aiello L, Cherifi C, Cherifi H, Lambiotte R, Lió P, Rocha L, editors. Complex Networks and Their Applications VII. Cham: Springer; 2018. p. 392-404.

[25] Pham CV, Duong HV, Hoang HX, Thai MT. Competitive influence maximization within time and budget constraints in online social networks: an algorithmic approach. Appl Sci. 2019;9(11):2274.

[26] Chung TY, Ali K, Wang CY. Deep reinforcement learning-based approach to tackle competitive influence maximization. Proceedings of 2019 ACM SIGKDD conference (KDD19 (MLG Workshop)) (MLG'19); 2019 Aug 5; Anchorage, USA. New York: ACM; 2019. p. 1-8.

[27] Nguyen HT, Thai MT, Dinh TN. Stop-and-Stare: Optimal sampling algorithms for viral marketing in billion-scale networks. Proceedings of the 2016 International Conference on Management of Data; 2016 Jun 26 - Jul 1; San Francisco, USA. New York: ACM; 2016. p. 695-710.

[28] Li H, Xu M, Bhowmick SS, Sun C, Jiang Z, Cui J. DISCO: influence maximization meets network embedding and deep learning. arXiv:1906.07378. 2019:1-14.

[29] Nannicini G, Sartor G, Traversi E, Calvo RW. An exact algorithm for robust influence maximization. Math Program. 2020;183:419-53.

[30] Kalimeris D, Kaplun G, Singer Y. Robust influence maximization for hyperparametric models. Proceedings of the 36th International Conference on Machine Learning; 2019 Jun 10-15; Long Beach, USA. PMLR; 2019. p. 3192-200.

[31] Du N, Liang Y, Balcan MF, Gomez-Rodriguez M, Zha H, Song L. Scalable influence maximization for multiple products in continuous-time diffusion networks. J Mach Learn Res. 2017;18:1-45.

[32] Chen X, Tan M, Zhao J, Yang T, Wu D, Zhao R. Identifying influential nodes in complex networks based on a spreading influence related centrality. Phys Stat Mech Appl. 2019;536:122481.

[33] Lagrée P, Cappé O, Cautis B, Maniu S. Algorithms for online influencer marketing. arXiv:1702.05354. 2017:1-29.

[34] Mothilal RK, Yadav A, Sharma A. Optimizing peer referrals for public awareness using contextual bandits. Proceedings of the Conference on Computing & Sustainable Societies; 2019 Jul 3-5; Accra, Ghana. New York: ACM; 2019. p. 74-85.

[35] Wu X, Fu L, Meng J, Wang X. Evolving influence maximization. arXiv:1804.00802. 2018:1-23.

[36] Vaswani S, Lakshmanan LV, Schmidt M. Influence maximization with bandits. arXiv:1503.00024. 2015:1-12.

[37] Wen Z, Kveton B, Valko M, Vaswani S. Online influence maximization under independent cascade model with semi-

bandit feedback. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; 2017 Dec 4-9; Long Beach, USA. New York: Curran Associates Inc; 2017. p. 1-11.
[38] Wu Q, Li Z, Wang H, Chen W, Wang H. Factorization bandits for online influence maximization. KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; 2019 Jul 4-8; Anchorage, USA. New York: ACM; 2019. p. 636-46.
[39] Lei S, Maniu S, Mo L, Cheng R, Senellart P. Online influence maximization. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2015 Aug 10-13; Sydney, Australia. New York: ACM; 2015. p. 645-54.
[40] Wang Y, Fan Q, Li Y, Tan KL. Real-Time influence maximization on dynamic social streams. PVLDB. 2017;10(7):805-16.
[41] Yuan J, Tang S. No Time to observe: adaptive influence maximization with partial feedback. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17); 2017 Aug 19-25; Melbourne, Australia. p. 3908-14.
[42] Wang Y, Zhu J, Ming Q. Incremental influence maximization for dynamic social networks. In: Zou B, Han Q, Sun G, Jing W, Peng X, Lu, editors. Data Science. Singapore: Springer; 2017. p. 13-27.
[43] Mohammadrezaei M, Shiri ME, Rahmani AM. Identifying fake accounts on social networks based on graph analysis and classification algorithms. Secur Comm Network. 2018;2018:5923156.
[44] Martın EG, Lavesson N, Doroud M. Hashtags and followers : an experimental study of the online social network Twitter. J Market Manag. 2013;31(1):221-43.
[45] Mihara S, Tsugawa S, Ohsaki H. Influence maximization problem for unknown social networks. 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM); 2015 Aug 25-28; Paris, France. USA: IEEE; 2015. p. 1539-46.
[46] Wilder B, Immorlica N, Rice E, Tambe M. Maximizing influence in an unknown social network. The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18); 2018 Feb 2-7; New Orleans, USA. USA: AAAI Publications; 2018. p. 1-8.
[47] Wilder B, Onasch-Vera L, Hudson J, Luna J, Wilson N, Petering R, et al. End-to-End influence maximization in the field. AAMAS '18: Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems; 2018 Jul 10-15; Stockholm, Sweden. Richland: IFAAMAS; 2018. p. 1414-22.
[48] Yan B, Song K, Liu J, Meng F, Liu Y, Su H. On the maximization of influence over an unknown social network. AAMAS '19: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems; 2019 May 13-17; Montreal, Canada. Richland: IFAAMAS; 2019. p. 2279-81.
[49] Li Y, Chen W, Wang Y, Zhang ZL. Influence diffusion dynamics and influence maximization in social networks with friend and foe relationship. Proceedings of the sixth ACM international conference on Web search and data mining; 2013 Feb 4-8; Rome, Italy. New York: ACM; 2013. p. 657-66.
[50] Jendoubi S, Martin A. Evidential positive opinion influence measures for viral marketing. Knowl Inf Syst. 2020;62:1037-62.
[51] Chen Y, Li H, Qu Q. Negative-Aware influence maximization on social networks. The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18); 2018 Feb 2-7; New Orleans, USA. USA: AAAI Publications; 2018. p. 1-2.
[52] Guo J, Zhang P, Zhou C, Cao Y, Guo L. Personalized influence maximization on social networks. Proceedings of the 22nd ACM international conference on Information & Knowledge Management; 2013 Oct 27 – Nov 1; San Francisco, USA. New York: ACM; 2013. p. 199-208.
[53] Razis G, Anagnostopoulos I. InfluenceTracker: Rating the impact of a twitter account. In: Iliadis L, Maglogiannis I, Papadopoulos H, Sioutas S, Makris C, editors. Artificial Intelligence Applications and Innovations. Berlin: Springer; 2014. p. 184-95.
[54] Tang X, Yang CC. Ranking user influence in healthcare social media. ACM Trans Intell Syst Technol. 2012;3(4):73-96.
[55] Aral S, Dhillon PS. Social influence maximization under empirical influence models. Nat Hum Behav. 2018;2:375-82.
[56] Karampourniotis PD, Szymanski BK, Korniss G. Influence maximization for fixed heterogeneous thresholds. Sci Rep. 2019;9:5573.
[57] Galhotra S, Arora A, Roy S. Holistic influence maximization: combining scalability and efficiency with opinion-aware models. Proceedings of the 2016 International Conference on Management of Data; 2016 Jun 26 - Jul 1; San Francisco, USA. New York: ACM; 2016. p. 743-58.
[58] Cano AE, Mazumdar S, Ciravegna F. Social influence analysis in microblogging platforms – a topic-sensitive based approach. Semant Web. 2014;5(5):357-72.
[59] Hamzehei A, Jiang S, Koutra D, Wong RK, Chen F. TSIM: Topic-based social influence measurement for social networks. Proceedings of the 14th Australasian Data Mining Conference; 2016 Dec 6-8; Canberra, Australia. New York: ACM; 2016. p. 1-9.
[60] Aslay C, Barbieri N, Bonchi F, Baeza-Yates R. Online Topic-aware influence maximization queries. The 17th International Conference on Extending Database Technology (EDBT); 2014 Mar 24-28; Athens, Greece. p. 295-306.
[61] Feng S, Chen X, Cong G, Zeng Y, Chee YM, Xiang Y. Influence maximization with novelty decay in social networks. Twenty-Eighth AAAI Conference on Artificial Intelligence; 2014 Jul 27-31; Québec, Canada. USA: AAAI; 2014. p. 37-43.
[62] Wang F, Zhu Z, Liu P, Wang P. Influence maximization in social network considering memory effect and social reinforcement effect. Future internet. 2019;11(4):95.
[63] Sankar CP, Asharaf S, Kumar KS. Learning from bees: an approach for influence maximization on viral campaigns. PLOS One. 2016;11(12):e0168125.
[64] Talukder A, Alam MGR, Tran NH, Niyato D, Hong CS. Knapsack-based reverse influence maximization for target marketing in social networks. IEEE Access. 2019;7:44182-98.
[65] Rossi M. Graph mining for influence maximization in social networks. France: Université Paris-Saclay; 2017.
[66] Lv J, Guo J, Yang Z, Zhang W, Jocshi A. Improved algorithms of CELF and CELF++ for influence maximization. J Eng Sci Tech Rev. 2014;7(3):32-8.
[67] Lv J, Guo J, Ren H. Efficient greedy algorithms for influence maximization in social networks. J Inform Process Syst. 2014;10(3):471-82.
[68] Chen S, Fan J, Li G, Feng J, Tan Kl, Tang J. Online topic-aware influence maximization. PVLDB. 2015;8(6):666-77.
[69] Derakhshan B. Influence maximization in large scale graphs. Helsinki: University of Helsinki; 2015.
[70] Jia S, Chen L. Sampling based influence maximization on linear threshold model. J Phys: Conf Ser. 2018;989:012013.
[71] Tang J, Zhang R, Yao Y, Zhao Z, Chai B, Li H. An adaptive discrete particle swarm optimization for influence

maximization based on network community structure. Int J Mod Phys C. 2019;30(6):1950050.
[72] Hua Y, Chen B, Yuan Y, Zhu G, Ma J. An influence maximization algorithm based on the mixed importance of nodes. Computers, Materials & Continua. 2019;59(2):517-31.
[73] Arora N, Banati H. IM-GSO: A community directed group search optimization approach for influence maximization. Cybern Syst. 2018;49(7-8):497-520.
[74] Panagopoulos G, Vazirgiannis M, Malliaros FD. Influence maximization via representation learning. arXiv:1904.08804. 2018:1-13.
[75] Yang W, Brenner L, Giua A. Influence maximization in independent cascade networks based on activation probability computation. IEEE Access. 2019;7:13745-57.
[76] He X, Kempe D. Robust influence maximization. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 Aug 13-17; San Francisco, USA. New York: ACM; 2016. p. 885-94.
[77] Mohammadi A, Saraee M. Finding influential users for different time bounds in social networks using multi-objective optimization. Swarm Evol Comput. 2018;40:158-65.
[78] Cheng S, Shen H, Huang J, Chen W, Cheng X. IMRank: influence maximization via finding self-consistent ranking. Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval; 2014 Jul 6-11; Gold Coast, Australia. New York: ACM; 2014. p. 475-84.
[79] Gunnec D, Raghavan S, Zhang R. Tailored incentives and least cost influence maximization on social networks. USA: University of Maryland; 2017.
[80] Jing D, Liu T. Context-based influence maximization with privacy protection in social networks. J Wireless Com Network. 2019;2019:142.
[81] Shan X, Chen W, Li Q, Sun X, Zhang J. Cumulative activation in social networks. Sci China Inf Sci. 2019;62:52103.
[82] Kuhnle A, Alim MA, Zhang H, Li X, Thai MT. Multiplex influence maximization in online social networks with heterogeneous diffusion models. IEEE Trans Comput Soc Syst. 2018;5(2):418-29.
[83] Rao A, Spasojevic N, Li Z, DSouza T. Klout Score: measuring influence across multiple social network. BIG DATA '15: Proceedings of the 2015 IEEE International Conference on Big Data (Big Data); 2015 Oct 29 - Nov 1; Santa Clara, USA. USA: IEEE; 2015. p. 2282-9.
[84] Li Y, Fan J, Ovchinnikov GV, Karras P. Maximizing multifaceted network influence. 35th IEEE International Conference on Data Engineering (ICDE 2019); 2019 Apr 8-11; Macau, China. USA: IEEE; 2019. p. 446-57.
[85] Wang W, Street WN. Modeling and maximizing influence diffusion in social networks for viral marketing. Appl Netw Sci. 2018;3(6):1-26.
[86] Aslay C, Lakshmanan LV, Lu W, X Xiao. Influence maximization in online social networks. Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining; 2018 Feb 5-9; Los Angeles, USA. New York: ACM; 2018. p. 775-6.
[87] Sun J, Tang J. Models and algorithms for social influence analysis. Proceedings of the sixth ACM international conference on Web search and data mining; 2013 Feb 4-8; Rome, Italy. New York: ACM; 2013. p. 775-6.
[88] Christakis NA, Fowler JH. Connected: the surprising power of our social networks and how they shape our lives. New York: Little, Brown Spark; 2011.
[89] Segev N, Avigdor N, Avigdor E. Measuring influence on instagram: a network-oblivious approach. arXiv: 1806.00881. 2018:1-5.
[90] Erlandsson F, Bródka P, Borg A, Johnson H. Finding influential users in social media using association rule learning. Entropy. 2016;18(164):1-15.
[91] Weskida M, Michalski R. Finding influentials in social networks using evolutionary algorithm. J Comput Sci. 2019;31:77-85.
[92] Lynn CW, Lee DD. Statistical mechanics of influence maximization with thermal noise. Europhys Lett. 2017;117(6):66001.
[93] Erkol S, Faqeeh A, Radicchi F. Influence maximization in noisy networks. Europhys Lett. 2018;123(5):58007.
[94] Gonga M, Yan J, Shen B, Ma L, Cai Q. Influence maximization in social networks based on discrete particle swarm optimization. Inform Sci. 2016;367:600-14.