A Multi-seed Nodes Selection Strategy for Influence Maximization Based on Reinforcement Learning Algorithms

Gege Nie\textsuperscript{1} and Ming Tang\textsuperscript{2,*}

\textsuperscript{1}School of Communication and Electronic Engineering, East China Normal University, Shanghai 200241, China
\textsuperscript{2}School of Physics and Electronic Science, East China Normal University, Shanghai 200241, China

*Corresponding author email: mtang@ce.ecnu.edu.cn

Abstract. Identifying influential individuals in the dissemination of information is an important topic in the study of social networks. Up to now, most of the previous works of Influence Maximization on social networks has been limited to selecting seeds based on a certain structural feature of the networks. These algorithms only consider a certain structural feature and cannot effectively select suitable seeds on social networks when the network has complex and changeable structures. Most of them only get good results on a special kind of networks. In order to find the most suitable nodes as the initial seed nodes in various social networks, we designed a new seeds selection algorithm which is based on reinforcement learning (IMQ). Our approach takes advantage of the characteristics of reinforcement learning's agent that can continuously interact with the environment, this algorithm can be adapted to select the most suitable nodes as seed nodes on various social networks. It fully considers the influence of the network structure characteristics on the influence propagation process, so that this method can select the best nodes as seeds on social networks with different topologies. In order to demonstrate the superiority of the approach, we conducted comparative experiments on six real social networks. Experimental results show that IMQ can be applied to various structural social network, and has stronger universality than traditional methods.

Keywords: Influence maximization; complex network; reinforcement learning; social network.

1. Introduction

The study of influence maximization is a popular topic that has attracted researchers’ attention in many fields. Marketing and public opinion control are its practical applications. With the rapid development of the Internet in recent years, a number of large-scale online social platforms have emerged, such as Twitter, WeChat and TikTok. They are also becoming or have become a huge social platform for communication and marketing. Viral marketing is a manifestation of influence maximization in practical applications, which is a common mode of promotion or incentive users to consume. For example, a company has launched a new product and needs to do new product promotion. Nowadays, the most common promotion method is to use the characteristic of celebrities or Internet celebrities to promote their new products, or choose some individuals as free product users of their new products. The goal is to use these experience users to enhance the awareness of their new products. The problem then becomes who will be chosen as the initial seed nodes. In this paper, we propose a new algorithm to identify the groups of individuals that have the most influence in various structural social networks.
2. Related Work
At the beginning of the 21st century, Domingos and Richardson [3,6] are the first to formally propose the topic of influence maximization, and regard influence maximization as a subject of scientific research. Their work is basic and forward-looking. Kempe et al. [5] are first regarded the influence maximization problem as an algorithmic problem with an optimal solution, and they proved that this optimization is a NP-hard problem. In order to simulate the propagation process in the real world, the propagation model has been established. It can be reduced to three basic propagation models, They are cascading model [4], threshold model[12] and epidemic model [13]. In 2007, Leskovec et al. proposed a new seeds selection algorithm, which is the “Cost-Effective Lazy Forward” (CELF) scheme [7]. Wei Chen et al. [9] proposed degreeDiscountIC algorithm to select seed nodes. This method has lower complexity than the CELF algorithm, The speed of selecting seed nodes has been improved more than 800 times, and the effectiveness of the degreeDiscountIC algorithm is the same as that of CELF, which indicates that the degreeDiscountIC algorithm can be used on large networks.

3. Problem Description
In our work, we regard the social network as a directed graph network \(G=(V,E)\), where \(V\) is the set of nodes, \(E\) is the set of edges. Each directed edge \((u,v)\in E\) means that \(u\) can give an influence to \(v\), and \(v\) cannot give an influence to \(u\), which means that the influence between individuals has directionality. Our work mainly studies the problem of Influence Maximization on social networks. The purpose is to select the optimal seed node set as the initial seeds. The set of seeds is represented by \(S\), \(k\) is the number of seeds.

4. Method
In our algorithm, we quoted the idea of reinforcement learning and integrated the Markov decision process (MDP) with social networks. The Markov decision process is the theoretical basis of reinforcement learning [1]. When defining the network state, we adopted a coarse-grained strategy by defining six network features to classify the network. we use meta-learning method to defined actions. The detailed steps are described below.

4.1. Introduction to Basic Methods
Reinforcement learning is the core of our proposed algorithm. It's also a class of algorithms for machine learning, which is a kind of data driven algorithm, which uses dynamic programming to solve the optimal solution problem. MDP is an iterative process in which an Agent obtains Reward by interacting with their environment through agent’s actions [15]. The Markov decision process is also Markovian. In the next part, we will describe the fusion strategy of reinforcement learning and social network.

In this article, the method we propose uses the structural data of social networks as the driving data for reinforcement learning to find the best seed nodes. The definition of the value function is as follows:

\[
V_\pi(s) = E_\pi[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots | S_t = s] = E_\pi[R_{t+1} + \gamma V_\pi(S_{t+1}) | S_t = s] \tag{1}
\]

Where \(\gamma\) and \(R\) are the discount factor and the reward respectively, \(V_\pi(s,a)\) is the state value function given a policy \(\pi\), and \(Q_\pi(s,a)\) is the accumulate the expected value of the reward given a policy \(\pi\).

In the same way, the Q function given policy \(\pi\) is defined as follows:

\[
Q_\pi(s,a) = E_\pi[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots | S_t = s, A_t = a] = E_\pi[R_{t+1} + \gamma Q_\pi(S_{t+1}, A_{t+1}) | S_t = s, A_t = a] \tag{2}
\]

Proof that the state value function will converge through iteration:

\[
V_\pi(s) \leq Q_\pi(s, \pi'(s)) = E_\pi[R_{t+1} + \gamma V_\pi(S_{t+1}) | S_t = s] \leq E_\pi[R_{t+1} + \gamma^2 R_{t+2} + \ldots | S_t = s] = V_\pi'(s) \tag{3}
\]
Where $\pi'(s) = \arg\max_{a \in A} Q_\pi(s, a)$.

### 4.2. Strategies for the Integration of Reinforcement Learning and Social Networks

In order to combine reinforcement learning with graph network effectively, we designed a fusion strategy to define the environment, reward, state and action in reinforcement learning on social network. More details are as follows:

**ENVIRONMENT**: Reinforcement learning is a method of finding the optimal policy $\pi$ through the agent constant interaction with the environment. In this paper, environment is social networks.

**REWORD**: We defined the number of nodes in the activated state is reward after completion of propagation process, that is, $R = |N| - |O_{inactive}|$, $O_{inactive}$ is a set of inactive nodes after the influence spread.

**STATE**: we use a network of inactive nodes to represent the state. General speaking, the change of each node’s state will lead to the change of the network state. This method of directly representing the network state will lead to the number of network state up to $2^{10^4}$. it will result in overfitting. we use a coarse-grained strategy to represent the state. Below are the features we have designed to represent the state:

1. The number of inactive state nodes in the network.
2. The sum of the degrees of inactive nodes.
3. The average clustering coefficient of the network composed of inactive nodes.
4. The largest degree value among inactive nodes.
5. The heterogeneity of the network $H = \frac{\langle k^2 \rangle}{\langle k \rangle^2}$.
6. The density of the network.

We divide the feature value of the network uniformly into five magnitudes, which are smallest, small, medium, large, and largest. These five magnitudes are relative to the initial value of the network's features. Therefore, there are at most $6^5$ of states, and the coarse-grained strategy can limit the number of network states to an appropriate order of magnitude. The experiments show that the coarse-grained strategy satisfy this requirement.

**ACTION**: Agents interact with the environment through actions. The fusion strategy of actions in the network directly determines the way in which action select seed nodes. In this article, we use the meta-learning method, we define the action is four specific seed nodes selection method. We choose MaxDegree [1], betweenness_centrality [13], k-shell [11] and subGreedy [5] algorithms as the actions.

1. **MaxDegree**: Select the node which is the largest node degree as the seeds in the network composed of inactive nodes.
2. **betweenness_centrality**: The node with the largest betweenness centrality value is selected as the seed node of inactive nodes.
3. **k_shell**: It uses a coarse-grained node importance classification method[11], which is different from other centrality indicators to determine the importance of nodes.
4. **subGreedy**: This is a greedy method to select seed nodes. Try to use each inactive node as a seed node to propagate in the network, after two rounds of propagation. After all nodes have been tried, the node with the greatest influence is the most important node. According to the importance of this node, the seed node set can be selected.
4.3 Model Training Process

Figure 1 summarizes the process of IMQ. $\epsilon$-greedy [11] method is used to explore a new seed node selection strategy. Our Q-table finally reached a stable state. When we select the seeds according to the trained model, Q-table is no need to be modified. The update equation of Q-table is as follows [15]:

$$Q_{t+1} = (1 - \alpha)Q_t(S_t, a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1}, a_{t+1})]$$  \hspace{1cm} (4)

Among them, $\gamma$ and $\alpha$ are discount factor and learning rate respectively, $\gamma \in [0, 1]$, the term of $r_{t+1} + \gamma Q(s_{t+1}, a_{t+1})$ is fresh expected reward obtained from experience.

In the training process, the number of nodes we need to select $k$ is predefined. We trained six models on six real networks each training process is 600 episodes and choose the best one as our final model, this is a common validation method in machine learning. The other parameter is that the learning rate is 0.5, the initial value of $\epsilon$ and the discount factor $\gamma$ are 0.98. All experiments are conducted on a server with Intel(R) Xeon Gold 5118 CPU @ 2.30GHZ and 64G RAM.

5. Time Complexity of IMQ

The complexity of our proposed approach is divided into training and test process time complexity, The time complexity of each training step is $O(T(|A| + |N| + |V|))$, the complexity of the action in reinforcement learning is $T(|A|)$, The time complexity is the worst outcome. So the time complexity of the training process is $O(k|V|^2)$, and the testing process is $O(k|V|)$. The complexity of degreeDiscountIC [8] is $O(k\log|N| + |V|)$, $k$ is the size of seeds. It indicates that our algorithm can be used on large-scale networks.

6. Experiment and Result

6.1. Experimental Data

Table 1 is the experimental data, which are six real social networks. Which can be downloaded from Stanford University Data Collection website [9], the structural properties value of six real-world networks are list in Table 1. $|V|$ and $|N|$ are the number of edges and nodes respectively, $|D|$ is the diameter of the network, $<c>$ is the average clustering coefficient of the network.

| name          | $|N|$  | $|V|$  | $|D|$ | $<c>$  |
|---------------|-------|-------|-------|-------|
| email-Eu-Core | 1005  | 25571 | 7     | 0.3994 |
| Facebook      | 4039  | 88234 | 8     | 0.6055 |
| Wiki-vote     | 7115  | 103689| 7     | 0.1409 |
| CollegeMSeg   | 1899  | 20296 | 10    | 0.1093 |
| GR-QC         | 5242  | 14496 | 17    | 0.5296 |
| p2p Network   | 10876 | 39994 | 9     | 0.0062 |
6.2. Experimental Setting
To show the superiority of our algorithm for IMQ, we have designed six comparative experiments to test the validity of our method. The SIR model can be used to compare the superiority of different methods. In our experiment, the IMQ algorithm is compared with the four algorithms mentioned above. We used the degree-discountIC method to replace the subGreedy in the comparison experiment [8]. At the beginning of the experiment, the nodes of seeds will be set to state I, and the other nodes will be set to state S. The process of spreading influence takes place in the SIR model [14]. The recovery rate of nodes in state I is 1. In order to highlight the difference between the influence of nodes selected by different methods, we set different infection rates $\beta$ on different networks. The results in Figure 2 demonstrate that the seeds selected by the IMQ algorithm have greater influence.

![Figure 2. Influence spread of different algorithms on six real networks under different k values.](image)

We also designed experiments to investigate the effect of varying infection rates $\beta$ with different methods. In this experiment, we select 7 nodes in each network as seed nodes. The experimental results are shown in Figure 3. It shows that our proposed algorithm can select better seed nodes in real networks with different structural characteristics, and our algorithm has strong stability and can select the best set of seed nodes on different social networks.

![Figure 3. Influence spread of different algorithms on six real networks under different $\beta$ values.](image)
7. Discussion and Conclusion

The result of Experiments show that our new algorithm is suitable for social networks with various topologies. For a trained model, we only need to input the network state at a certain moment to quickly get the best seed nodes selection strategy. The process is to select the best seed node selection strategy through Q-table. Our method can also be extended to select multiple nodes. In the next work, we believe that we can do some follow-up work to make our work better. First, what we are doing now is each round of single node selection. We select a seed node after each round of propagation. In the next follow-up work, we can try to limited or fixed k, so when we actually select the seed nodes, we must consider the number of seeds selected in each training step. The algorithm in this paper has not considered the relationship between the magnitude of influence and time. In order to better simulate real social networks, we believe that the method improved can be used to study the problem of maximizing influence on time-varying networks.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 11975099), the Natural Science Foundation of Shanghai (Grant No. 18ZR1412200).

References

[1] Barabási, A.-L. & Albert, R. Emergence of Scaling in Random Networks. Science, 1999, pp 509–512
[2] Lin S C, Lin S D, Chen M S. A Learning-based Framework to Handle Multi-round Multi-party Influence Maximization on Social Networks[J]. 2015, pp 695-704
[3] P. Domingos and M. Richardson. Mining the network value of customers. In Proceedings of the 7th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2001, pp 57–66
[4] Goldenberg J, Libai B. Using Complex Systems Analysis to Advance Marketing Theory Development: Modeling Heterogeneity Effects on New Product Growth through Stochastic Cellular Automata[J]. Academy of Marketing inference Review, 2001, p 9
[5] D. Kempe, J. M. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In Proceedings of the 9th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2003, pp 137–146
[6] M. Richardson and P. Domingos. Mining knowledge-sharing sites for viral marketing. In Proceedings of the 8th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2002, pp 61–70
[7] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. S. Glance. Cost-effective outbreak detection in networks. In Proceedings of the 13th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2007, pp 420–429
[8] W. Chen, Y. Wang, and S. Yang, “Efficient influence maximization in social networks,” in ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2009, pp 105-108
[9] Stanford Large Network Dataset Collection [Online]. Available: http://snap.stanford.edu/data/.
[10] Granovetter, Mark. Threshold Models of Collective Behavior[J]. American Journal of Sociology, 1978, pp1420-1443
[11] Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E. & Makse, H. A. Identification of influential spreaders in complex networks. Nat. Phys.6, 2010, pp 888–893
[12] Zareie A, Sheikhhahmad I, Khamforoosh K. Influence Maximization in Social Networks Based on TOPSIS[J]. Expert Systems with Applications, 2018, p 108
[13] Buscarino, A., Fortuna, L., Frasca, M., & Latora, V. Disease spreading in populations of moving agents. EPL, 2008, p 82
[14] C. Watkins, “Learning from delayed rewards,” Ph.D.dissertation, Cambridge University, 1989s
[15] Mitchell T M. Machine Learning[M]. McGraw-Hill, 2003