Learning Policies for Effective Incentive Allocation in Unknown Social Networks

Extended Abstract

Shiqing Wu
University of Tasmania
Hobart, Australia
shiqing.wu@utas.edu.au

Quan Bai
University of Tasmania
Hobart, Australia
quan.bai@utas.edu.au

Weihua Li
Auckland University of Technology
Auckland, New Zealand
weihua.li@aut.ac.nz

ABSTRACT

Most existing incentive allocation approaches rely on sufficient information about users' attributes, such as their preferences, followers in the social network, and activities, to customize effective incentives. However, this may lead to failure when such knowledge is unavailable. In this light, we propose an end-to-end reinforcement learning-based framework, named Geometric Actor-Critic (GAC), to discover effective incentive allocation policies towards users in a social network. More specifically, given a limited budget, the proposed approach can extract information from a high-level network representation for learning effective incentive allocation policies. The proposed GAC only requires the topology of the social network and does not rely on any prior information about users' attributes. We use three real-world social network datasets to evaluate the performance of the proposed GAC. The experimental results demonstrate the effectiveness of the proposed approach.

KEYWORDS

Incentive Allocation; Reinforcement Learning; Unknown Social Network; Social Simulation

ACM Reference Format:

Shiqing Wu, Quan Bai, and Weihua Li. 2021. Learning Policies for Effective Incentive Allocation in Unknown Social Networks: Extended Abstract. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 3 pages.

1 INTRODUCTION

Incentivizing users to take behaviors that are profitable or beneficial to the incentive providers is a crucial problem to many fields, including promoting sales [6, 21, 22], hiring workers [2, 8, 10], encouraging beneficial behaviors [12, 15, 17]. Such processes are computationally modeled as *incentive allocation* problem, where the goal is to incentivize users with effective incentives under a budget limitation, such that the number of users who take the behavior that the incentive provider expects is maximized.

Given a finite budget, the key factor for a successful incentive allocation is whether the incentive structure and the pricing policies for users are reasonable or not, as overpricing the incentive would waste the budget, whereas underpricing could fail to incentivize the user [11]. Namely, the pricing policies should consider the utilities of both incentive providers and users. Some studies attempted to model users' attributes, such as users' preferences and skill abilities [2, 4, 8, 9, 17–19], to generate "optimal" incentives. However, these methods could be impotent when such information is lacking or unavailable.

Some studies also consider exploiting social influence to incentivize users, as information diffusion plays a crucial role in propagating persuasive information among friends in a social network [5]. However, recognizing influential users in a social network is difficult in practice. Although we can collect the topology of a social network and identify if a user is influential or not based on the number of her followers, it may contain false or weak edges that are ineffective at spreading influence [1, 7, 14]. In principle, we could conduct exhaustive surveys on all users to estimate their influential abilities. Nevertheless, such surveys are very labor-intensive and impractical [13].

Is it possible to perform effective incentive allocation for realizing user incentivization in a social network, where the information about users and influence strength among users is absent, and the budget is restricted? To tackle the incentive allocation problem on online social networks, in this paper, we propose an end-to-end reinforcement learning-based framework, named Geometric Actor-Critic (GAC), which can learn to extract information from the network for generating incentive allocation policies. GAC is able to learn both global representations for the entire network and local representations for the individual users. This allows the RL agent to estimate if the user is influential in the network, such that determine the incentive value for her.

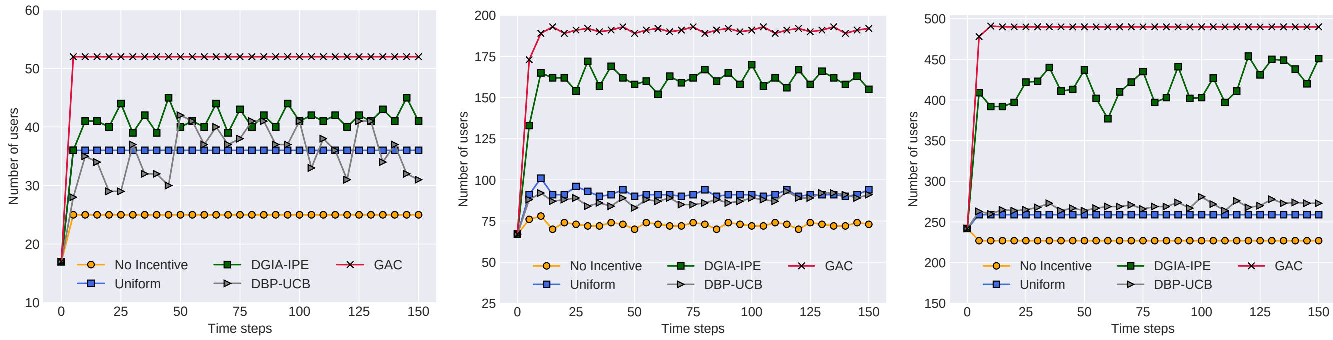
2 PROBLEM FORMULATION

Let $G = (V, E)$ represents a social network, where G denotes a directed graph, $V = \{v_1, ..., v_i\}$ denotes a set of user agents, and E denotes a set of edges in the network. Each edge implies that v_i can influence v_j's behaviors, and the strength of influence is from 0 to 1. We also define v_i's one-hop neighbors who can influence v_i directly as in-neighbors, and those who are directly influenced by v_i as out-neighbors. Each user v_i would take a behavior based on own preferences at every time step. Let the range of an incentive be from 0 to 1. Given a finite budget B, the objective is to maximize the number of users who take z^* as behaviors by providing users incentives, where z^* represents the behavior expected by the incentive provider.

3 GEOMETRIC ACTOR-CRITIC

To tackle the incentive allocation problem in a social network, it is essential to take both the structure of the network and users'

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.



(a) Dolphins, B=3 (b) Twitter, B=20 (c) Wiki-Vote, B=40

Figure 1: Performance comparison of GAC and baseline approaches

features into consideration when generating incentive policies. Due to the complexity of a social network, it is necessary to encode the information about a network in a low-level graph representation.

The input of GAC includes the user feature matrix and two different adjacency matrices, i.e., the in-adjacency matrix and the out-adjacency matrix. The reason why we input two adjacency matrices is that a social network is typically a directed graph structure, and its adjacency matrix is not symmetric. To keep all structural information about the graph, we deploy two independent components consisting of Graph Neural Networks (GNNs) to encode both the out-adjacency and the in-adjacency matrices, respectively. In each network encoding component, a Mean "variant" of GraphSage [3] is deployed first to learn refined user features by aggregating features from their neighbors. Through the first GraphSage layer, we can obtain local representation for all users, i.e., node embedding. After learning the node embeddings for all users, we use two Diffpoll layers [20] to learn the global representation of the entire social network, which can aggregate user features in a hierarchical manner.

Once the local and global representations are obtained, i.e., node embeddings for all users and graph embedding, they can be combined as a vector containing information from both the local and global representations via a matrix-vector product. These two newly generated vectors would be concatenated. The concatenated vector would be normalized by using L2-Norm and subsequently fed through three fully connected layers with nonlinear activation functions. The first two layers use *ReLU* as activation functions, while the last layer uses *tanh* as the activation function. Due to the range of each entry in the output is from -1 to 1, we need to re-scale it to the range from 0 to 1 before allocating incentives for users.

4 EXPERIMENT

To evaluate the performance of the proposed GAC, we conduct the simulation by deploying three real-world social networks, i.e., dolphins (62 nodes and 159 edges)[1], Twitter(236 nodes and 2478 edges)[2], and Wiki-Vote(889 nodes and 2914 edges)[3]. Meanwhile, we utilize the Agent-based Decision Making (ADM) model [16] to simulate users' behaviors, in which the values of users' preferences are randomly derived from the uniform distribution $U(0, 1)$. In addition, we also assign a random value from 0 to 1 for the influence strength of each edge, and ensure that the sum of influence strength from the user's in-neighbors would not exceed 1.

The performance of the proposed GAC is compared with four baseline approaches: • **No Incentive.** Here all users would receive no incentives and make decisions only based on their preferences and social influence from neighbors. • **Uniform Allocation.** This is a Naïve approach, which allocates the same incentives to users. The incentive values are determined by the number of users and the budget amount. • **DGIA-IPE.** DGIA-IPE is an adaptive incentive allocation approach which estimates influential relationship among users and then determines the incentive values based on the observation of users' behaviors [16]. • **DBP-UCB.** This is a bandit-based dynamic pricing algorithm, which models each price option as a discrete option [12]. In the experiments, the default metric used to evaluate the performance of approaches is the number of users who take z^*.

In this experiment, as we can observe from Figure 1(a), the proposed GAC is able to incentivize more users than the other baseline approaches in all three networks. Meanwhile, the performance of DGIA-IPE is slightly better than Uniform allocation and DBP-UCB in Dolphins network, while outperforms these two approaches significantly in both Twitter and Wiki-Vote networks. The reason is that 1) the IPE algorithm would estimate the influential relationships among users while Uniform allocation and DBP-UCB cannot, and 2) the size and density of Dolphins network is much smaller than the other two networks. We also notice that the performance of DGIA-IPE appears not very stable in Wiki-Vote network. A possible reason is that the average degree of Wiki-Vote is much lower than that of the Twitter network, and the IPE algorithm fails to estimate influence strength well in such a sparse social network. On the other hand, DBP-UCB slightly outperforms the Uniform allocation in Wiki-Vote network, but performs worse in Twitter network. This might also be caused by the different sparsity of two networks, as DBP-UCB does not take social influence into consideration when generating incentives.

[1] http://networkrepository.com/soc-dolphins.php
[2] https://snap.stanford.edu/data/ego-Twitter.html
[3] http://networkrepository.com/soc-wiki-Vote.php

REFERENCES

[1] Soheil Eshghi, Setareh Maghsudi, Valerio Restocchi, Sebastian Stein, and Leandros Tassiulas. 2019. Efficient Influence Maximization under Network Uncertainty. In *INFOCOM 2019 - IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS 2019*. IEEE, New York City, NY, USA, 365–371.

[2] Xiaoying Gan, Xiong Wang, Wenhao Niu, Gai Hang, Xiaohua Tian, Xinbing Wang, and Jun Xu. 2017. Incentivize Multi-Class Crowd Labeling Under Budget Constraint. *IEEE Journal on Selected Areas in Communications* 35, 4 (2017), 893–905.

[3] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive Representation Learning on Large Graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Curran Associates, Inc., New York, NY, USA, 1024—-1034.

[4] Juan Li, Yanmin Zhu, and Jiadi Yu. 2019. Redundancy-Aware and Budget-Feasible Incentive Mechanism in Crowd Sensing. *Comput. J.* 63, 1 (01 2019), 66–79.

[5] Yuchen Li, Ju Fan, Yanhao Wang, and Kian-Lee Tan. 2018. Influence Maximization on Social Graphs: A Survey. *IEEE Transactions on Knowledge and Data Engineering* 30, 10 (2018), 1852–1872.

[6] Romain Lopez, Chenchen Li, Xiang Yan, Junwu Xiong, Michael I. Jordan, Yuan Qi, and Le Song. 2020. Cost-Effective Incentive Allocation via Structured Counterfactual Inference. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*. AAAI Press, Palo Alto, CA, USA, 4997–5004.

[7] Shodai Mihara, Sho Tsugawa, and Hiroyuki Ohsaki. 2015. Influence maximization problem for unknown social networks. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. Association for Computing Machinery, New York, NY, USA, 1539–1546.

[8] Chenxi Qiu, Anna Squicciarini, and Benjamin Hanrahan. 2019. Incentivizing Distributive Fairness for Crowdsourcing Workers. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 404–412.

[9] Sotsay Sengvong and Quan Bai. 2017. Persuasive public-friendly route recommendation with flexible rewards. In *2017 IEEE International Conference on Agents (ICA)*. IEEE, New York City, NY, USA, 109–114.

[10] Yaron Singer and Manas Mittal. 2013. Pricing Mechanisms for Crowdsourcing Markets. In *Proceedings of the 22nd international conference on World Wide Web*. Association for Computing Machinery, New York, NY, USA, 1157–1166.

[11] Adish Singla and Andreas Krause. 2013. Truthful Incentives in Crowdsourcing Tasks Using Regret Minimization Mechanisms. In *Proceedings of the 22nd international conference on World Wide Web*. Association for Computing Machinery, New York, NY, USA, 1167–1178.

[12] Adish Singla, Marco Santoni, Gábor Bartók, Pratik Mukerji, Moritz Meenen, and Andreas Krause. 2015. Incentivizing Users for Balancing Bike Sharing Systems. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*. AAAI Press, Palo Alto, CA, USA, 723–729.

[13] Thomas W Valente and Patchareeya Pumpuang. 2007. Identifying opinion leaders to promote behavior change. *Health Education & Behavior* 34, 6 (2007), 881–896.

[14] Bryan Wilder, Nicole Immorlica, Eric Rice, and Milind Tambe. 2018. Maximizing Influence in an Unknown Social Network. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*. AAAI Press, Palo Alto, CA, USA, 4743–4750.

[15] Shiqing Wu and Quan Bai. 2019. Incentivizing Long-Term Engagement Under Limited Budget. In *PRICAI 2019: Trends in Artificial Intelligence*. Springer International Publishing, Cham, Switzerland, 662–674.

[16] Shiqing Wu, Quan Bai, and Byeong Ho Kang. 2019. Adaptive Incentive Allocation for Influence-Aware Proactive Recommendation. In *PRICAI 2019: Trends in Artificial Intelligence*. Springer International Publishing, Cham, Switzerland, 649–661.

[17] Shiqing Wu, Quan Bai, and Sotsay Sengvong. 2018. GreenCommute: An Influence-Aware Persuasive Recommendation Approach for Public-Friendly Commute Options. *Journal of Systems Science and Systems Engineering* 27, 2 (2018), 250–264.

[18] Yue Wu, Fan Li, Liran Ma, Yadong Xie, Ting Li, and Yu Wang. 2019. A Context-Aware Multiarmed Bandit Incentive Mechanism for Mobile Crowd Sensing Systems. *IEEE Internet of Things Journal* 6, 5 (2019), 7648–7658.

[19] Shuai Xiao, Lei Lv, Le Guo, Yuanbo Chen, Shuang Yang, Zaifan Jiang, and Jun Zhu. 2019. Model-based constrained MDP for budget allocation in sequential incentive marketing. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. Association for Computing Machinery, New York, NY, USA, 971–980.

[20] Rex Ying, Jiaxuan You, Christopher Morris, Xiang Ren, William L Hamilton, and Jure Leskovec. 2018. Hierarchical Graph Representation Learning with Differentiable Pooling. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Curran Associates, Inc., New York, NY, USA, 4805–4815.

[21] Wen Zhang, Dengji Zhao, and Yao Zhang. 2020. Incentivize Diffusion with Fair Rewards. In *Proceedings of the 24th European Conference on Artificial Intelligence*. IOS Press, Washington, DC, 251–258.

[22] Dengji Zhao, Bin Li, Junping Xu, Dong Hao, and Nicholas R Jennings. 2018. Selling Multiple Items via Social Networks. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 68–76.