

SMART 课题介绍

- 长期合作博弈中的奖励及时分配机制 2
- 社交网络上的分布式机制设计 4
- 社交网络中的公平拍卖机制（合作博弈角度） 6
- 碳市场中的机制设计 7
- 激励合作的网络成本分摊机制设计问题 9
- 社交网络中的多卖家拍卖机制 11
- 网络中防操纵机制设计——选择最具影响力的人 12
- 社交网络中的匹配问题 14
- 网络中的激励邀请分蛋糕问题 16
- 社交网络中的最优防假名拍卖 18
- 社交网络中的防假名信息传播机制设计 20

长期合作博弈中的奖励及时分配机制

葛垚鑫 张尧

课题介绍

传统的合作博弈模型中，我们研究如何将合作所得收入分配给所有人。合作博弈的研究中诞生了诸如沙普利值等经典概念，然而这些方法都有一个共同之处：在合作结束之后进行计算，再将计算出来的回报分配给每个人。有些合作博弈能够持续很久，中间不断有新的合作者加入，合作所得的价值不能等到合作结束时再分配，因此我们需要及时地根据现状进行分配。我们在最基本的合作博弈模型中添加参与者的顺序，并提出奖励机制所应该满足的性质。在此之上，我们研究了几类机制所具备的性质。

研究意义

从理论角度，本研究迭代了合作博弈的基本模型，是对合作博弈理论的发展。对于合作博弈模型的拓展有很多，包括将合作博弈与社交网络结合、增加参数表示参与者的合作程度等；对于合作博弈的分类也就也有很多，比如研究参与者之间无互补性的博弈等。本研究可以与这些理论相结合，从而使合作博弈可以用于建模更丰富的场景。

从应用角度，本研究具有广泛的应用空间。近几年来，通过深度学习、强化学习等手段训练大模型的工作层出不穷。科研工作者提出了以联邦学习为代表的方法进行分布式的训练，以在更广泛的范围内收集数据并拟合更好的模型。此类方法的前提是能够收集到大量的数据，或是能够征求到大量参与者加入训练过程，因此需要好的激励措施。除此之外，招工、融资等场景中，也可以建模成合作博弈的模型，并设计奖励分配机制用来分配利润。这些场景下，都需要研究如何在合作博弈中及时地给予参与者奖励。

拟采用的研究方法

本课题首先从理论角度入手，基于合作博弈论和机制设计的理论工具。我们将参与者加入到合作的先后顺序作为新的变量加入到合作博弈模型中，并数学化地表征了几条性质。我们在进行机制设计的过程当中，一种方法是考虑特殊的合作博弈类型，包括 submodular 博弈、XOS 博弈等，在这些特殊的博弈中研究如何设计满足所需性质的机制；另一种方法是，先考虑所需的性质，然后推出满足性质所需要的条件。理论的工作都通过数学推导作为主要工具。

本课题预计在未来采用实验手段验证所设计的机制。实验可能包含两种：1、在训练大模型的过程当中，研究数据集之间的关系和对于结果的贡献，从而研究如何给每个数据集分配价值，这一部分主要依靠深度学习和数据分析的相关技术。2、在现实场景中，依据所设计的机制，招募真实被试进行实验，从而观察所设计的机制在现实应用中的表现，这一部分除设

计实验和验证外，还有可能采用 Human-in-loop 实验方式。该方法将人类行为作为数据集训练模拟人类行为的智能体，再以智能体作为实验参与者观察其在实验中的表现，从而验证机制的效果。

困难与挑战

本研究所面临的挑战在于：1、目前国内外，尽管关于合作博弈的理论和应用工作都很丰富，但是与本课题直接相关的工作少之又少。本研究中提出的性质看似符合直觉，但实际设计机制时却有可能相互矛盾，从而设计的机制很难同时满足；若要在特定的博弈类型上进行机制设计，需要对合作博弈的分类有深入的研究，但是此方面的理论研究却主要局限于对于互补性、单调性、凸凹性等基本性质的研究。2、想要从机器学习实验中验证机制的效果，需要有大模型训练的过程，这需要算力和数据集的支持。此外，不同的场景、算法、模型对应的合作博弈不尽相同，需要分类设计机制进行验证。3、想要在现实实验中验证所设计机制的效果，需要巧妙的实验设计，从而使实验易于执行的同时，又与机制的应用场景相对应。

参考文献

- [1] Moretti S, Patrone F. Transversality of the Shapley value[J]. TOP, 2008, 16(1):1-41.
- [2] Liu Z, Chen Y, Yu H, et al. GTG-Shapley: Efficient and Accurate Participant Contribution Evaluation in Federated Learning[J]. 2021.
- [3] Faigle U, Kern W. The Shapley Value for Cooperative Games under Precedence Constraints[J]. International Journal of Game Theory, 1992, 21(3):249-266.

社交网络上的分布式机制设计

刘浩新

课题介绍

该课题旨在设计出社交网络上不依赖于任何中心节点并且满足所需性质的分布式机制。不同于传统的中心化机制需要由一个中心节点收集所有参与者的信息并且计算与执行分配结果，分布式机制旨在将机制的所有计算与执行分散给所有参与者，每位参与者各自作为一个独立的个体参与到信息的传递和机制的执行，可以按照自己的利益任意地执行动作，如何让这些参与者执行我们预期的动作以实现全局的期望性质是本课题的关键所在。

研究意义

现存的机制基本都是中心化机制，这些机制都要求一个可信任的中心节点（通常为卖家）来执行整个机制并承担计算任务，如此一来，该中心节点会知道所有参与者的私人社交关系，造成潜在的隐私泄露问题。并且，由于社交网络天然是一个分布式的结构，所以社交网络是分布式机制的天然环境，社交网络上的每一位用户都可以成为机制发起人，但不可能要求所有用户都是可信任的，社交网络也不希望依赖于第三方平台，因此，不依赖于任何中心节点的分布式机制在现实世界中具有更高价值。

困难与挑战

不同于传统的中心化机制，分布式机制将机制的执行分散到所有参与者。因此，参与者需要执行的任务不仅仅是中心化机制里向中心节点汇报他们的 type 这一种动作，这也意味着参与者拥有更大的动作空间，想要阻止他们恶意操控运行过程也会变得更加困难。设计分布式机制的难点通常包括两个部分，一是要在分布式问题设定下，即没有节点能够获取全局的结构信息，设计出正确的机制运行流程。二是要保证在无人监管且没有中心节点处理信息的前提下，所有人能够按照事先的预期行动，没有人能够通过恶意操控来获取更高收益，通常需要设计一个均衡点来解决该问题。

研究方向

方向一 分布式拍卖机制

社交网络上的拍卖机制以 IDM 机制为代表，而在分布式拍卖机制中，我们希望社交网络上的用户不需要将自己的 type 汇报给卖家，因此卖家无法获知图的全部结构信息来确定关键节点和后续分配，需要设计出一种新的分布式机制。

已发表的相关成果： Distributed Mechanism Design in Social Networks (AAMAS 2023)

后续可以继续探索如何对社交网络上的分布式拍卖机制进行性质刻画, 以及将目前的分布式机制扩展到普适性更强的一类机制。

方向二 分布式匹配机制

目前课题组所设计出的 Leave and Share、CTC 等机制均在社交网络上中心化的场景取得了不错的结果, 由于匹配市场中每个人都带着自己的物品参与机制, 每个人在市场中承担的角色都是等价的, 不存在卖家和买家的身份差别, 因此更加适合进行分布式实现与拓展。分布式匹配机制这一方向正是去设计出上述机制的分布式实现方法并拓展归纳为一类 (class) 机制。

其他方向:

和组内的其他方向结合, 探索已有机制的分布式机制或者按照分布式的理念设计出新的机制。

参考文献

- [1] Joan Feigenbaum, Michael Schapira, and Scott Shenker. 2007. Distributed algorithmic mechanism design. In *Algorithmic Game Theory*. Vol. 14. Cambridge University Press Cambridge, 363–384.
- [2] Jeffrey Shneidman and David C Parkes. 2004. Specification faithfulness in networks with rational nodes. In *Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing*. 88–97.
- [3] Bin Li, Dong Hao, Dengji Zhao, and Tao Zhou. 2017. Mechanism design in social networks. In *Thirty-First AAAI Conference on Artificial Intelligence*. AAAI Press, 586–592.
- [4] Haoxin Liu, Yao Zhang, and Dengji Zhao. "Distributed Mechanism Design in Social Networks." *arXiv preprint arXiv:2303.03077* (2023).

社交网络中的公平拍卖机制（合作博弈角度）

顾子欣 葛垚鑫

课题介绍

社交网络中的拍卖机制设计通过潜在收益激励买家互相邀请，买家可能因为邀请从拍卖中获益。社交网络中“邀请”的行为促成了本为竞争关系的拍卖参与者之间的合作。本课题旨在从合作博弈的角度出发，研究如何更加公平地为“亦敌亦友”的拍卖参与者分配收益，设计一个满足激励相容、个体理性等性质，且更公平的社交网络拍卖机制。

研究意义

以 IDM 为代表的社交网络拍卖机制，虽然将钱分给了一些非赢家参与者，但能获得收益的人都集中于赢家的关键传播序列上。但事实上，其他买家也为拍卖社会福利的提升做出了一定贡献。我们希望设计一种新的社交网络拍卖机制，让所有买家都有机会因为自己对拍卖做出的贡献获得收益。

拟采用的研究方法

以 IDM 和 Shapley Value 这两个在社交网络拍卖和合作博弈中非常经典的方法为基础，尝试提取他们的核心思想并加以结合，从而设计出更加公平的社交网络拍卖机制，并使其满足激励相容、预算平衡等性质。

困难与挑战

传统的合作博弈中并没有“报价”这一步骤，因而智能体们也无法通过“误报”的操作来影响机制结果，提升自己的收益。基于这一特点，经典的合作博弈方法（如 Shapley Value）并不考虑真实性这一性质，计算方法本身也强依赖于特征函数。然而，在社交网络拍卖机制设计中，需要非常严格精密的设计才能使智能体们无法通过“误报”的操作来操控机制结果。从合作博弈的角度入手设计社交网络拍卖机制时，如何在融入合作博弈的一些优秀特性的同时保证激励相容的性质是为一大难点。

参考文献

- [1] Bin Li, Dong Hao, Dengji Zhao, and Tao Zhou. 2017. Mechanism design in social networks. In Thirty-First AAAI Conference on Artificial Intelligence. AAAI Press, 586–592.
- [2] Zhang Yao, Dengji Zhao. Incentives to invite others to form larger coalitions. Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems, 2022: 1509–1517.

碳市场中的机制设计

顾子欣 徐大来

课题介绍

力争 2030 年前实现碳达峰、2060 年前实现碳中和，是以习近平同志为核心的党中央经过深思熟虑作出的重大战略决策。全国碳排放权交易市场是落实国家碳达峰目标、碳中和愿景的重要政策工具。如何更高效、更合理地分配碳排放配额，如何利用市场机制控制和减少温室气体的排放，推进绿色低碳发展，是当下非常值得研究探讨的一大问题。

方向一 碳排放配额分配机制

我国的碳排放配额是由国家分配到地方再分配到企业的。如何合理分配是一个非常有意义的问题。我们将碳排放配额的分配抽象为一个资源分配问题，希望建立一个更优的分配机制，从而更高效地推进绿色低碳发展。

方向二 碳排放配额交易市场机制

近年，我国在逐步建立碳排放权交易市场。2021 年 7 月 16 日，全国碳排放权交易市场上线交易启动仪式顺利举行。为实现国家碳中和的重大战略目标，如何更好地利用市场机制达到减排的目的，是我们的一大目标。

方向三 碳市场仿真模型搭建

碳市场中有很多不同的智能体参与，如公司、政府、第三方检测机构等等。将碳市场抽象化为一个仿真模型，一方面可以模拟未来的发展，预测碳中和目标的；另一方面也可以通过仿真测试我们设计的机制，并基于此进一步优化。我们打算以 Agent Based Modeling 为基础，尝试搭建碳市场模型。

困难与挑战

- 1) 市场中的机制设计不同于传统的机制设计，其视角更为宏观，需要考虑更多因素。
- 2) 以市场为背景进行机制设计，需要深入了解市场本身，才能贴近市场，使研究不悬浮。
- 3) 碳市场不同于一般的股票市场，政府在其中起了非常重要的作用，且有着需要实现的大目标——“2030 碳达峰，2060 碳中和”。政府的策略也是该课题的一大关键。

参考文献

[1] 人民银行研究局和清华大学能源环境经济研究所联合课题组：浅谈基于主体的模型在碳中和路径和金融支持政策研究中的应用。中国人民银行工作论文 No.2022/2, 2022 年 3 月

17 日。

[2] 清华大学能源经济环境调研所：全国碳排放交易体系实务手册。

[3] Chen B, Zhang H, Li W, Du H, Huang H, Wu Y, Liu S. Research on provincial carbon quota allocation under the background of carbon neutralization. Energy Reports. 2022 Jul 1;8:903-15.

激励合作的网络成本分摊机制设计问题

张天一 张俊宇

假设有一个源提供某种服务，分散在不同位置的智能体都想要通过与源连通来享受该服务。但是任何两个智能体之间的连通都需要一个成本。连通所有智能体的总成本需要智能体分摊。此外，每个智能体有两种策略行为：汇报其对该服务的估值及其与其他智能体之间的连通关系。我们的目标是设计一个能激励智能体真实汇报其估值和连通关系的成本分摊机制。具体来说，我们把以上场景建模成一张带权重的无向图：其中智能体用节点表示，源用一个特殊的节点表示，两个智能体或者智能体与源之间的连通关系用边表示，连通成本用边上的权重表示。目标是设计一个成本分摊机制使得每个智能体愿意真实地汇报其估值并且提供其所有的连通关系。

方向一 防合谋的网络成本分摊机制设计

智能体除了具有汇报自己的估值和连通关系的行为外，还存在更高级的行为，比如多个智能体之间可以结盟来形成小团伙，通过协商来集体修改自己的行为，以此通过操控图的结构来获得更少的成本分摊份额，从而让小团伙中的成员获利。因此，如何设计出可以防止智能体的高级策略行为的成本分摊机制是需要进一步考虑的问题。该问题的挑战在于智能体的高级行为和汇报连通关系的行为可以同时改变图的结构，在二者的共同作用下，智能体具有更大的操控空间，因此设计激励相容的成本分摊机制就变得很有挑战。

方向二 多源网络成本分摊机制设计

在原来的问题中，只有一个源节点提供服务。然而，在现实场景中，可能有多个源同时提供不同的服务。此时，不同的节点（智能体）可能会有不同的需求，例如：有些节点仅需要其中某些源提供的服务，而另一些节点则需要其他源提供的服务。因此，我们需要设计在该场景下的成本分摊机制来保证节点能诚实地汇报其行为。为了简化场景，我们可以假设每个节点都需要所有源提供的服务，即每个节点都需要与所有源连通。同时，每个节点依旧具有汇报其估值和连通关系的行为。我们的目标是设计出可以激励节点诚实地汇报其行为的成本分摊机制。此外，所设计的机制还需要具有其他良好的性质，比如：预算可行性（每个节点的成本分摊份额不能超过其汇报的估值）和预算平衡性（所有节点的成本分摊份额之和等于连通所有节点的总成本）等。问题的挑战是：由于每个节点都需要与所有源连通，因此节点的策略行为对图结构的影响更大，导致设计出满足激励相容性的成本分摊机制变得更困难。

参考文献

[1] Zhang, Tianyi, Junyu Zhang, Sizhe Gu, and Dengji Zhao. "Cost Sharing under Private Valuation and Connection Control." arXiv preprint arXiv:2303.03083 (2023).

[2] Bergantiños, Gustavo, and Adriana Navarro-Ramos. "A characterization of the folk rule for multi-source minimal cost spanning tree problems." *Operations Research Letters* 47, no. 5 (2019): 366-370.

社交网络中的多卖家拍卖机制

张尧 郑珊珊

课题介绍

在 IDM 的设计中，我们看到，可以通过社交网络传播拍卖信息获得更多买家，以提高社会福利和卖家收益。而在这个过程中，也许同时会有这样的事件发生：1) 通过社交网络找到了其他同质的卖家；2) 同时有另一个同质的卖家在通过社交网络传播拍卖等。此时，社交网络中将有不止一个卖家存在，所对应的机制设计问题便是“社交网中的多卖家拍卖机制”。这将比 IDM 的设计更具有挑战性。

在这个问题中，多个卖家可能出现社交网络中的任意位置。我们将先考虑这家卖家都是同质（即他们所卖的物品是一样的，并且每个卖家只卖一件），以及每个买家都至多买一件物品的情况。在此基础上，我们重点研究两个问题。

研究方向一

当多个卖家同处一个市场管理之下，即有一个统一的机制处理多卖家如何分配物品（注意：此时需要考虑卖家可能在传播过程中进行操纵）。从多个卖家出发运行 IDM 并非一个行之有效的办法。特别地，当多个卖家和网络中其他节点的连接方式完全相同时，该问题相当于多物品传播拍卖机制设计（每个买家需求量为1）。所以，该问题或许可以通过以此为突破点寻求可行的机制。

研究方向二

当多个卖家互相独立地运行自己选择的拍卖机制时，我们考虑多个卖家的互相博弈。比如，如果有两个卖家，当其中一方使用 IDM 进行拍卖的时候，另一方的最佳选择是什么？双方同时使用 IDM 是否是一个纳什均衡？这些问题将会更贴近现实场景中的复杂市场，也可以通过这个问题预测卖家的行为以及理解不同传播拍卖机制的关系。这个问题的难点在于涉及多个卖家的博弈以及可能的传播拍卖机制的空间较大。该问题可以通过先研究两个卖家，以及限制传播拍卖机制类型（比如必须激励相容以简化买家的行为，但要注意买家可能会对双方有选择性；再比如进一步限制机制是属于特殊的一类机制集合）出发，进行研究和拓展。

参考文献

[1] Paes Leme R, Sivan B, Teng Y. Why do competitive markets converge to first-price auctions?[C]//Proceedings of The Web Conference 2020. 2020: 596-605.

网络中防操纵机制设计——选择最具影响力的人

张尧 赵雨昕

课题介绍

本课题以设计出满足激励相容的同行互选机制为目标，以后代数作为衡量影响力的标准，希望能在防止参与者操纵的条件下选择最具有影响力的人或群体。在参与者构成的网络中，参与者集合是固定的，但网络结构只能通过参与者的汇报信息来得知。参与者可以通过谎报信息操纵最终的选择结果，使得自己被选中的概率增加，因此我们需要防操纵的机制。

研究意义

小至论文投稿的同行评议，大至政治党派的选举，选择问题在现实场景十分常见。随着网络技术的兴起，此类问题更是比比皆是，如在引文网络中选出对某领域贡献最大的一篇文章、在社交媒体中确定出热度最高的用户等。我们统一将这些被选择者的特征归纳为影响力，并希望最终的被选择者具有最大的影响力，这对于相关领域的价值重大，例如选出的文章是未来学术发展的一个风向标，选出的用户可与社交平台合作从而产生巨大收益等。

拟采用的研究方法

我们目前延续了之前工作的研究方法。具体而言，先将能通过断边操作让自己有机会被选中的参与者构成一个集合（只有在集合里的人才有被选择的可能性），然后为集合中的每个参与者分配一定的选择概率。在满足激励相容的前提下，我们用近似比来衡量此类机制的好坏程度（机制输出的期望值与最优值之比）。

困难与挑战

方向一、扩展到选择多个人的情况

当以后代数作为影响力衡量标准时（在满足激励相容情况下，希望被选中的人后代数越多越好），我们已经找到了选择一个最有影响力的人的最优机制。并且扩展到了最多选择两个最有影响力的人的情况。随着选择人数的增多，网络结构会愈加复杂。我们现在正在考虑如何设计选择三人及以上的激励相容机制。

方向二、其他衡量影响力的方式

除了后代数，还有其他衡量影响力的方式，如直接入度（希望被选中的人直接入度越多越好）。可以考虑一些新的 setting，如在直接入度的衡量标准下，选择两个人使得他们的并集越大越好。

参考文献

- [1] Babichenko, Yakov, Oren Dean, and Moshe Tennenholtz. "Incentive-compatible diffusion." *Proceedings of the 2018 World Wide Web Conference*. 2018.
- [2] Babichenko, Yakov, Oren Dean, and Moshe Tennenholtz. "Incentive-compatible selection mechanisms for forests." *Proceedings of the 21st ACM Conference on Economics and Computation*. 2020.
- [3] Zhang, Xiuzhen, Yao Zhang, and Dengji Zhao. "Incentive compatible mechanism for influential agent selection." *Algorithmic Game Theory: 14th International Symposium, SAGT 2021, Aarhus, Denmark, September 21–24, 2021, Proceedings 14*. Springer International Publishing, 2021.

社交网络中的匹配问题

杨天易 宋欣薇 李淼 翟宇翔

课题介绍

人们或许很难给出一件物品的价值,但是总是能够比较出两件物品的优劣。在算法博弈论中,人们对于物品的偏好就来自于这个比较。匹配问题研究的是,如何根据一群人的偏好,将某一个集合的人或物分配给这群人。很自然的,我们希望这个最终分配足够“好”,那么如何定义什么样的分配是好的呢?稳定性、最优性等定义便被提出了。

择校问题就是一个常见的匹配问题,学生对学校有不同的偏好,而学校通过他们的整体表现来对学生进行排序。除此之外,婚姻也是一个匹配问题,聪明地选择伴侣或许是“稳定婚姻”的关键。在一些更复杂的情况下,参与者的偏好可能随着时间而演变,我们可以应用 Learning 方法解决动态匹配问题。

带初始物品的单边匹配: 每个参与者带一个初始物品参与匹配,参与者对所有物品的严格偏好排序、以及参与者之间的社交网络关系是自己的私有信息。参与者之间彼此交换,以期拿到更喜欢的物品。

双边匹配: 参与者分为两边,每一边都只对另一边的参与者和自己有偏好。参与者的偏好排序、以及参与者之间的社交网络关系是自己的私有信息。参与者之间可以两两配对,或者不配对。

(传统) 稳定性: 若当前匹配结果满足稳定性,则不存在任意一组参与者,他们能够自行配对使得他们中有些人匹配结果变好,其他人不变。

(传统) 最优性: 若当前匹配结果满足最优性,则不存在另一个匹配方式,使得一些参与者得到的匹配变好,其他人的匹配结果不变。

研究方向一：设计最优机制

对于单边匹配和双边匹配,如何定义最优性、设计机制产生最优分配是一个重要的问题。目前,我们有了一些基于参与者间结构、偏好的最优性定义方式,但是并不能证明它们就是能达到的理论边界。

怎么做: 基于现有的理论框架探寻边界,以及特定社交网络和偏好下的可行分配。

参考文献

[1] Yang T, Zhai Y, Zhao D, et al. One-sided matching with permission[J]. arXiv preprint arXiv:2201.05787, 2022.

- [2] Kawasaki T, Wada R, Todo T, et al. Mechanism design for housing markets over social networks[C]//Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems. 2021: 692-700.
- [3] Cho S H, Todo T, Yokoo M. Two-Sided Matching over Social Networks[C]//31st International Joint Conference on Artificial Intelligence, IJCAI 2022. International Joint Conferences on Artificial Intelligence, 2022: 186-193.

研究方向二：构建匹配问题实验平台

以网站的方式动态呈现目前社交网络中匹配问题的已有机制，让更多人能够亲自体验不同机制的运行过程。除展示外，平台提供实验入口，便于后续开展真实实验。观察分析参与者的决策，分析并解释理论与实验的差距。

怎么做：构建展示平台、思考实验设计问题。

参考文献

- [1] Ferris J, Hosseini H. MatchU: An Interactive Matching Platform[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(09): 13606-13607.
- [2] Koster, R., Balaguer, J., Tacchetti, A. et al. Human-centred mechanism design with Democratic AI. Nat Hum Behav 6, 1398–1407 (2022). <https://doi.org/10.1038/s41562-022-01383-x>

研究方向三：设计整合市场的匹配规则

整个市场可以视作是由多个子市场组成，有研究表明对于单边匹配，在执行最优机制（TTC）时，市场整合可能带来极大的社会福利损失。我们试图设计新的单边匹配机制，解决市场整合带来的社会福利损失。

怎么做：分析执行不同机制的最差情况、平均情况；重新定义合理的稳定性，设计机制考察其在新的稳定性下的表现。

参考文献

- [1] Kumar R, Manocha K, Ortega J. On the integration of Shapley–Scarf markets[J]. Journal of Mathematical Economics, 2022, 100: 102637.

网络中的激励邀请分蛋糕问题

葛煦

课题介绍

该课题研究了在社交网络下的分配蛋糕问题。一个发起者拥有一份资源，并想利用这份资源作为奖励将该某个特定的信息（如产品广告、合作任务等）传播给社交网络中尽可能多的参与者。其主要涉及社交网络中预算分配和分蛋糕两个问题。

社交网络中的预算分配问题，课题组已发表多项研究成果，其重要性质是邀请激励，即设计合理的分配奖金规则，使得参与者愿意邀请尽可能多的邻居。分蛋糕问题是算法博弈论的研究领域之一，其主要关注点是公平性，它研究如何将一个异质物品公平地分配给若干个参与者，比例性和无嫉妒性是两个重要的公平性定义。

研究意义

发起者想将自己的信息进行最大化传播，一种方式就是利用一定预算做广告；如果发起者想要利用现金以外的奖励（如产品、优惠券等）激励邀请，那么就需要考虑这些奖品的异质性。目前，预算分配问题在社交网络中已有相关成果。预算分配场景是分蛋糕场景的一种特例，即所有人对于蛋糕的估值均相同的情况，该课题将奖励从同质物品（钱）扩展到异质物品（蛋糕）。机制不仅要给参与者足够的邀请激励，还需要使得蛋糕的分配尽量公平。

研究方向

目前已经发现无法同时满足激励邀请和比例性，而现有的机制以保证邀请激励优先，将公平性进行了弱化。目前，该课题有如下两个思路进行优化：

一、提升机制的公平性

社交网络分蛋糕问题目前已有相关成果，但该机制的公平性比较弱，因此可以从预算分配问题出发，重新寻找公平性的提升方法。预算分配问题已经有多项工作，它们均能保证邀请的激励相容，这些工作使用的思路有所不同，均可以作为分蛋糕机制设计的出发点。对于每种方法，可以探索将它们扩展至异质物品时，能满足何种公平性，进而找到更合理的公平性定义，提升机制的公平性。

二、在已有的公平性定义下，寻找激励相容的机制

目前已有其它工作，发表了在社交网络中的分蛋糕公平性定义，（例如无嫉妒性要求各参与者不嫉妒自己的邻居），这些定义弱于传统分蛋糕问题中对应的公平性，可以作为该工作中的公平性。在这个公平性基础上，目前还没有找到能够满足激励相容的分配机制，因此它的基础上寻找激励相容的机制，是该项工作的一个可行方向。

参考文献

- [1] Bei X, Qiao Y, Zhang S. Networked fairness in cake cutting[C]. Proceedings of the 26th International Joint Conference on Artificial Intelligence.
- [2] Shi, H., Zhang, Y., Si, Z., Wang, L., Zhao, D.: Maximal information propagation with budgets. In: ECAI 2020 - 24th European Conference on Artificial Intelligence.

社交网络中的最优防假名拍卖

江天

课题介绍

互联网和社交媒体的发展，使得交易发起者（卖家）能够以较低的成本将交易信息通过社交联系网的方式传播。然而，由于互联网的匿名性，买家可以通过在社交网络中创造并扮演虚假节点的方式，提高自己获得的收益，这种操纵行为被称作假名攻击（Sybil attack）。

我们首次设计了一系列在社交网络中的防假名（Sybil-proofness）拍卖机制，并希望找到在所有防假名机制中，使拍卖者获得理论上最高收益的机制。

研究意义

假名攻击和最优性问题，是目前传播拍卖领域中最困难和亟待解决的两个问题。其一是机制有效性的重要保证，是传播拍卖机制在现实场景中实际应用的必然要求；其二是机制设计的核心之一，是机制优劣比较中最重要的标准。

如能将以上两个问题结合，对于具体的社交网络情景，激励人们真实汇报信息，杜绝假名攻击行为的出现，并在此基础上实现卖家收益最大化，也即解决社交网络中的最优防假名机制问题，则大体上相当于完全解决了传播拍卖领域。并且该机制也可能在现实中取得直接、有效的应用。

拟采用的研究方法

以往防假名攻击研究主要有两个思路。一是通过智能体的行为空间进行约束，但在社交网络传播拍卖的场景下，该方法与希望智能体尽可能多地传播的激励直接矛盾；另一种常见方法则是，利用先验或其他参与者汇报的信息，对智能体的汇报进行事中或事后的校验，对假名攻击的行为或其可能性施加惩罚，但是这种方法只有在拍卖发起者能采取的惩罚空间足够大的特定场景下才能起效，否则很容易使机制失去事中个体理性或传播激励相容的性质。

在传播拍卖的具体场景下，转卖者及其儿子节点之间的假名攻击几乎无法被机制发现或阻止，同时成本也极低。因此，我们在设计机制时需要拓展智能体在博弈时的行为空间，将这类假名攻击合理化为转卖者的一种权利。再根据显示原理，解出在智能体行为空间拓展后的新博弈下达成的贝叶斯纳什均衡，并在与之相对应的机制中将该均衡实现。

困难与挑战

（1）为了实现防假名，同时使最优机制存在，我们引入了对个体估计的先验信息，将机制拓展到贝叶斯场景下，这在之前的研究中是鲜有被考量的。

（2）目前的最优机制在可行性上（时空间复杂度、鲁棒性等）的实现存在一定困难，期待找到具有现实可操作性且卖家收入次优的机制。

参考文献

- [1] Myerson R B. Optimal auction design[J]. Mathematics of operations research, 1981, 6(1): 58-73.
- [2] Li B, Zhao D, et al. Diffusion auction design[J]. Artificial Intelligence, 103631. 2021.

社交网络中的防假名信息传播机制设计

郑俊杰 葛煦

课题介绍

假名攻击又被称为女巫攻击 (Sybil attack), 主要是指社交网络中的少部分恶意节点通过伪造多个虚假身份的方式, 干扰和影响其它真实节点的正常活动的行为, 会造成整个系统的破坏性影响。在社交网络的信息传播场景中, 例如在线求助某些问题的答案、分包大量的人工标注任务、利用社交网络贩卖商品等, 社交网络中可能存在一些恶意的参与者, 他们通过伪造假名身份参与到网络中以谋取更高的收益, 这要求机制能够防止假名攻击带来的危害。同时, 机制设计要激励参与者自发的传播信息并做出贡献, 这被称为激励相容性。激励相容和防止假名攻击这两个性质存在一些冲突, 如何设计同时满足这两个性质的机制是我们的目标。

研究意义

在传统的信息传播的场景中, 组织者往往只是借助自身的人际关系或者通过广告等渠道方式来进行宣传。这些手段面临着参与者覆盖面不够广, 虚假参与者盛行, 额外支出成本过高等问题。传统的机制设计方案通常假设参与者是事先已知的, 对这类问题束手无策, 因此本课题则希望发挥所有参与者的社交关系的作用, 利用参与者的人际关系做到自发性的信息传播, 同时能够解决参与者伪造身份的问题。

拟采用的研究方法

一、利用其他的额外的信息, 来设计相关的机制。例如, 利用参与者做出贡献的信息、参与到任务中的先后顺序的不同 (即时间序列信息) 等。

二、 社交网络图中的图结构 (不同人之间的社交连接关系) 包含了丰富的信息, 如何利用这些信息也是一个重要的话题。一个自然的想法是, 如果一个参与者的邀请者数量非常多, 可以通过很多路径从组织者找到他, 那么他的真实程度就越高。

三、 恶意参与者除了单独创建假身份实施假名攻击之外, 还可以串通多个其他的参与者进行合作伪造, 这被称为合谋攻击, 一个简单的例子就是多个参与者合成为一个参与者比他们

分开能获得更多的奖励。考虑抵御假名攻击、抵御合谋攻击和实现激励相容性这三个方面的要求，给我们留下了更为有限的设计空间。

困难与挑战

一个重要的结论是，如果不引入参与者的额外信息，激励相容性和防假名性之间存在一个不可能性。为了满足激励相容性，我们应该给那些邀请更多人加入的参与者额外的奖励。另一方面，我们应该减少伪造身份的参与者的奖励以满足防假名性。

参考文献

- [1] ZHANG Y, ZHANG X, ZHAO D. Sybil-proof answer querying mechanism[C]//Proceedings of the Twenty-Ninth International Conference on Artificial Intelligence.2021:422-428.
- [2] BABAI OFF M, DOBZINSKI S, OREN S, et al. On bitcoin and red balloons[C]//Proceedings of the 13th ACM conference on electronic commerce.2012:56-73.
- [3] EMEK Y, KARIDI R, TENNENHOLTZ M, et al. Mechanisms for multi-level marketing[C]//Proceedings of the 12th ACM conference on Electronic commerce.2011:209-218.