Group Formation by Group Joining and Opinion Updates via Multi-Agent Online Gradient Ascent

Journal:	*IEEE Computational Intelligence Magazine*
Manuscript ID	CIM-SI-2023-0045.R1
Manuscript Type:	SI: Artificial Intelligence eXplained (AI-X)
Date Submitted by the Author:	n/a
Complete List of Authors:	Lin, Chuang-Chieh; Tamkang University, Computer Science and Information Engineering Hung, Chih-Chieh; National Chung Hsing University, Management Information Systems Lu, Chi-Jen; Academia Sinica, Institute of Information Science Chen, Po-An; National Yang Ming Chiao Tung University, Institute of Information Management
Keywords:	machine learning, game theory, online learning, gradient ascent, Nash equilibrium
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
IEEE_CIM_AIX_Group_Formation_Submission_20230604.zip	

Authors' Response

Chuang-Chieh Lin,
Tamkang University, Taiwan

Chih-Chieh Hung,*
National Chung Hsing University Taiwan

Chi-Jen Lu,
Academia Sinica, Taiwan

Po-An Chen,
National Yang Ming Chiao Tung University, Taiwan

We thank the associated editor and the three reviewers for the constructive comments and feedback. These comments, from the three reviewers, brought significant value to our manuscript. Below, we give responses to explain how we address all the comments.

Reviewer 1

1. This article makes use of several external libraries. Those need to be downloaded and embedded in the local files of the article, to ensure that this includes:

   - https://cdn.jsdelivr.net/npm/bootstrap@5.0.2/dist/css/bootstrap.min.css
   - https://cdn.jsdelivr.net/npm/bootstrap@5.0.2/dist/js/bootstrap.bundle.min.js

   Response: The mentioned external libraries are sourced from Bootstrap. In order to ensure the optimal functionality of all the features in this article, we have replaced all Bootstrap-supported classes with native JavaScript implementations. For example, the *mb-0* class in Bootstrap has been replaced with a custom-defined class, *my-auto lh-sm*. We believe that this adjustment can guarantee the smooth operation of various features.

2. <section id="introduction"> is never closed. Please add </section> before the start of the next section

   Response: Thank you for bringing this to our attention. We have addressed and resolved the issue.

3. lines 1241 and beyond: the transition property is marked with an error. I am not an expert of transitions myself, but "left" is not a valid property. You might have meant "linear" instead. Furthermore the ticks need to be removed. So consider replacing it with style="transition: linear .3s ease-in; height: 100px;"

   Response: Thanks for your comment. The `left` property is commonly used for positioning elements in CSS, determining the horizontal position relative to its containing element. As it is not directly related to the `transition` property, you cannot use it as a standalone property within `transition`. However, you can transition the `left` property indirectly by including it as one of the properties affected by the `transition` property. For example, `style="transition: left .3s ease-in; height: 100px;"`

*Corresponding author.

sets inline CSS styles for an image element. Here, `transition: left .3s ease-in` is a shorthand property for defining the transition effect. It specifies that any changes to the `left` property should have a duration of 0.3 seconds and use an ease-in timing function. The use of the `left` property within the `transition` property is not invalid in CSS. It is a common technique to achieve smooth transitions. Some examples and discussions related to this technique can be found on technical forums (e.g., Stack Overflow). You can refer to the following links for more information: 1) Using CSS transition with "left" property: `https://stackoverflow.com/questions/65100863/using-css-transition-with-left-property-doesnt-transition-smoothly` and 2) CSS transition with "top," "bottom," "left," and "right": `https://stackoverflow.com/questions/20383393/css-transition-doesnt-work- with-top-bottom-left-right`.

4. the ids "best_response", "best-response-discussion", "flexRadioDefault5", "flexRadioDefault6", "flexRadioDefault7", "flexRadioDefault8", "gaa_no_regularizers" appear multiple times.

   Response: We have reviewed the entire HTML file to ensure that each ID is unique and only appears once throughout the file. We believe the overall integrity of the HTML structure has been improved.

5. line 1410 and beyond: the svg width has a small typo. replace 250xp with 250px. The same typo appears in a later figure again.

   Response: These typos have been corrected in our revised file.

Reviewer 2

The authors focus their study on the best response dynamics and multi agent online learning by introducing a group formation approach. Specifically, the authors study how the online learning algorithms can help stabilize the entire system composed by strategic agents that perform their best response. The manuscript is overall well written and easy to follow and the authors have well thought out their main contributions. The provided theoretical analysis concrete, complete, and correct and the authors have provided all the intermediate steps in order to enable the average reader to easily follow it. Furthermore, the limited provided numerical results show the success of the proposed online gradient ascent approach with regularization. The authors are highly encouraged to consider the following suggestions provided by the reviewer in order to improve the scientific depth of their manuscript, as well as they need to address the following minor comments in order to improve the quality of presentation of the manuscript. Initially, the concept of satisfaction games has been recently proposed in the literature, A Paradigm Shift Toward Satisfaction, Realism and Efficiency in Wireless Networks Resource Sharing, doi: 10.1109/MNET.011.2000368, and provides a new perspective on the multi agent systems in terms of decision making. In Section 1, the authors need to discuss this aspect and clarify why the authors focus their research on maximizing the payoff rather than satisfying minimum constraints which is the current approach in the state-of-the-art. In Section 2, the physical meaning of the public preference vector and the private preference vector needs to be provided and tied with some indicative applications to enable the average reader to follow the rest of the analysis. In Section 4, the authors need to definitely provide the computational complexity of the multi agent online gradient descent algorithm and also provide some indicative numerical results in order to quantify its complexity towards realizing the pure strategy Nash equilibrium. Finally, the overall manuscript needs to be checked for typos, syntax, and grammar errors in order to improve the quality of its presentation.

1. Initially, the concept of satisfaction games has been recently proposed in the literature, *A Paradigm Shift Toward Satisfaction, Realism and Efficiency in Wireless Networks Re-*

source Sharing, doi: 10.1109/MNET.011.2000368, and provides a new perspective on the multi agent systems in terms of decision making. In Section 1, the authors need to discuss this aspect and clarify why the authors focus their research on maximizing the payoff rather than satisfying minimum constraints which is the current approach in the state-of-the-art.

Response: We thank the reviewer for the comment. We find that Papavassiliou et al.'s work "A Paradigm Shift Toward Satisfaction, Realism and Efficiency in Wireless Networks Resource Sharing, doi: 10.1109/MNET.011.2000368" is interesting and relevant to our work. Hence, we add a paragraph in subsection "RELATED WORK" of Section 1 and discuss the "satisfaction equilibrium" where each player targets satisfaction of quality of service requirements rather than individual utility maximization. We point out why we focus on utility maximization and why the satisfaction equilibrium is not applicable in our case.

2. In Section 2, the physical meaning of the public preference vector and the private preference vector needs to be provided and tied with some indicative applications to enable the average reader to follow the rest of the analysis.

   Response: We thank the reviewer for this helpful comment. In the first paragraph of Section 2, we add physical meaning of the preference vectors and provide examples and arguments to motivate the our discussion and facilitate the readers to follow the rest of the analysis as well.

3. In Section 4, the authors need to definitely provide the computational complexity of the multi agent online gradient descent algorithm and also provide some indicative numerical results in order to quantify its complexity towards realizing the pure strategy Nash equilibrium

   Response: We thank the reviewer for the comment. The analysis of computational complexity of our algorithm is added in Section 4. Furthermore, we've also added indicative numerical results to exemplify our proposed multi-agent online learning algorithm just before the analysis of the computational complexity.

4. Finally, the overall manuscript needs to be checked for typos, syntax, and grammar errors in order to improve the quality of its presentation.

   Response: We thank the reviewer for the suggestion. We have sent the manuscript to a professional academic English editing service (the receipt is attached as a certificate) and improved its presentation quality. In addition, The REFERENCES are listed in alphabetical order now.

Reviewer 3

This paper provides an interesting interactive application to facilitate audiences to grasp the idea of pure-strategy Nash equilibrium and how the system converges to a stable state by means of decentralized online gradient ascent. However, even though this work has some contributions, both theoretically and practically, the main contribution is not clearly described. Moreover, the details of this research cannot be found in this paper. Here are some of the issues that the authors need to address to improve the quality of the paper.

Overall, some contents of this paper are not mature enough to merit for publication in this journal. I suggest that this paper be revised to improve its quality.

1. The related work section should be improved. More related works need to be added in this paper to give a complete discussion of such system.

Response: We thank the reviewer for the comment. We have added subsection "RELATED WORK" in Section 1. Four categories of related literature, such as "On Group Formation and Collective Behavior", "On Opinion Formation", "On Satisfaction Game", and "On Multiagent Online Learning", are discussed there.

2. I also suggest that authors need to give more details about the proposed algorithm (or system) to attract the attention of researchers.

   Response: We thank the reviewer for the comment. We exemplify our algorithm by providing an illustrating example of how the algorithm computes the utilities, winning probabilities, reward, gradients, updates, etc., of agents. We've also improved the statements in Section 4 (the paragraph after Figure 7 and the second paragraph after "Regularization") to better explain and motivate our multiagent online learning algorithm.

3. More comparisons between the proposed algorithm (or system) and other recent state-of-the-art algorithms should be given to show the importance of this research.

   Response: We thank the reviewer for the comment. In the new added subsection "RELATED WORK" in Section 1, We discuss other recent state-of-the-art algorithms or approaches and point out the connection and difference between our work and the related work. We also point out the importance of our work there.

Scribendi Inc.
304-405 Riverview Drive
Chatham, Ontario N7M 0N3 Canada

Receipt

Customer

Po-An Chen
1001 University Rd
East District
Hsinchu

Date: 2023-05-30
Task No.: 951422
Returned

Words	Description	TOTAL
6,192	Academic Editing in 3 days Submitted files: Group Formation by Group Joining and Opini....pdf	USD 228.77

Sub Total		USD 228.77
Shipping		USD 0.00
Tax Rate(s)	None	USD 0.00
TOTAL		USD 228.77

Payments

Methods: Stripe
Amount: US$ 228.77
Date: Tue, 30 May 2023 16:58:09 GMT

Office use only

PAID IN FULL

The client is responsible for remitting all local taxes and fees to the relevant authorities.

Thank you for your business!

All orders subject to our terms of service: https://www.scribendi.com/terms

GST/HST #: 89219 5702 RT0001 QST #: 1225989199 TQ0001

Group Formation by Group Joining and Opinion Updates via Multi-Agent Online Gradient Ascent

Chuang-Chieh Lin, *Tamkang University, Taiwan*,
Chih-Chieh Hung, *National Chung Hsing University, Taiwan**
Chi-Jen Lu, *Academia Sinica, Taiwan*
Po-An Chen, *National Yang Ming Chiao Tung University, Taiwan*

***Abstract*—This article aims to exemplify best-response dynamics and multi-agent online learning by group formation. This extended abstract provides a summary of the full paper in IEEE Computational Intelligence Magazine on the special issue *AI-eXplained* (AI-X). The full paper includes interactive components to facilitate interested readers to grasp the idea of pure-strategy Nash equilibria and how the system of strategic agents converges to a stable state by the decentralized online gradient ascent with and without regularization.**

I. Introduction

GAME theory has been applied in a variety of situations due to its predictability of outcomes in real world. It can also be used in solving problems, such as saddle-point optimization which has been used extensively in generative adversarial network models [1]. A *game* consists of strategic agents, each of which acts rationally to maximize its own reward (or utility) or minimize its cost. A *Nash equilibrium* is a stable state composed of the strategies of all agents such that none of the agents wants to change its own strategy unilaterally. Therefore, such a stable state is possibly achievable or even predictable. However, how to achieve a Nash equilibrium in a game may not be quite straightforward, especially when agents behave in a "decentralized" way. Indeed, when an agent's reward function depends on the strategies of the other agents, the maximizer of one agent's reward function is not necessarily a maximizer for any other agent, and it may change whenever any other agent changes its strategy.

In this article, we consider the group formation of strategic agents as an example to demonstrate the strategic behaviors of the agents. A strategic agent can either join a group or change its opinion to maximize its reward. The eventual equilibrium of the game hopefully suggests predictable outcomes for the whole society. For the case in which agents apply group-joining strategies, we consider *pure-strategy Nash equilibria* (PNE) as the solution concept, where a pure strategy means a strategy played with a probability of 1. For the case in which agents change their opinions, we assume that each agent plays an *online gradient ascent algorithm*, which guarantees the time-average convergence to a hindsight optimum for a single agent (see [2] for the cost-minimization case), in a decentralized way, and then we examine the possibly convergent state of the system.

*Corresponding Author: Chih-Chieh Hung (smalloshin@nchu.edu.tw).

II. Group and Opinion Formation

Given a set V of n agents $v_1, v_2, \ldots, v_n$. Each agent v_i is represented as a *public preference vector* z_i and a *private preference vector* s_i, such that the former (we call it an *opinion*) corresponds to the preference revealed to all the agents while the latter corresponds to its *belief*, which is unchangeable. We consider $s_i, z_i \in \mathcal{K}$ such that $\mathcal{K} := \{x \in [-1,1]^k : \|x\|_2 \leq 1\} \subset \mathbb{R}^k$ is the feasible set. One can realize that each dimension of the domain stands for a certain social issue, such that -1 maps to the far-left politics, while 1 maps to far-right politics. The bounded 2-norm constraint is in line with the bounded rationality of a person, or the bounded budget for a group. We use $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ and $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ to denote two profiles that include each agent's opinion and belief, respectively. Each agent is initially regarded as a group. The *opinion of a group* is the average of the opinions of its members. Similar to the *monotone* setting in [3], a group wins with higher odds if its opinion brings more utility to all the agents. The *reward* (i.e., payoff) of an agent is the expected utility that it can get from all the groups. Specifically, assume that we currently have $m \leq n$ groups $G_1, G_2, \ldots, G_m$, and denote by $|G_i| = n_i$ the number of members in group G_i. Let $\mathcal{G} = (G_1, G_2, \ldots, G_m)$ denote the profile of the groups. To ease the notation, we denote by $\tau = (\mathbf{z}, \mathbf{s}, \mathcal{G})$ the *state* of the game. The reward function of agent i is $r_i(\tau) = \sum_{j=1}^{m} p_j(\tau)\langle s_i, \bar{g}_j \rangle$, where $\bar{g}_j = \sum_{v \in G_j} z_i/|n_j|$ represents the average opinion of group G_j and the winning probability $p_j(\tau)$ of group G_j is

$$p_j(\tau) = \frac{e^{n_j \langle \bar{g}_j, \sum_{v_r \in V} s_r \rangle}}{\sum_{i \in [m]; n_i > 0} e^{n_i \langle \bar{g}_i, \sum_{v_r \in V} s_r \rangle}},$$

where $[m]$ denotes $\{1, 2, \ldots, m\}$. We consider the following strategic behaviors of an agent in such a game:

- Group Joining:
  - Seeking a specific group that hopefully maximizes the agent's reward and joins the group.
- Opinion Updating without Regularization:
  - Each agent in a certain group tries to maximize its reward by changing its own opinion.
- Opinion Updating with Regularization:
  - Each agent in a certain group tries to maximize its reward by changing its own opinion, while the reward includes the regularization $-\|s_i - z_i\|_2^2$, which

hopefully limits how strategic an agent can be by preventing it from moving too far from its own belief.

III. Group Joining and Pure-Strategy Nash Equilibria

A. 1D Representation

When the opinions and beliefs are assumed to be in $[-1, 1] \subset \mathbb{R}$, we can illustrate these vectors as well as the dynamics of changes on a real line. For example, in Fig. 1 we have five agents v_1, v_2, v_3, v_4, v_5. By assuming v_1, v_2, v_3, and v_4 to have their opinions z_1, z_2, z_3, z_4 fixed, we can observe the changes in the winning probabilities and the rewards of all the agents by moving z_5 from -1 to 1.

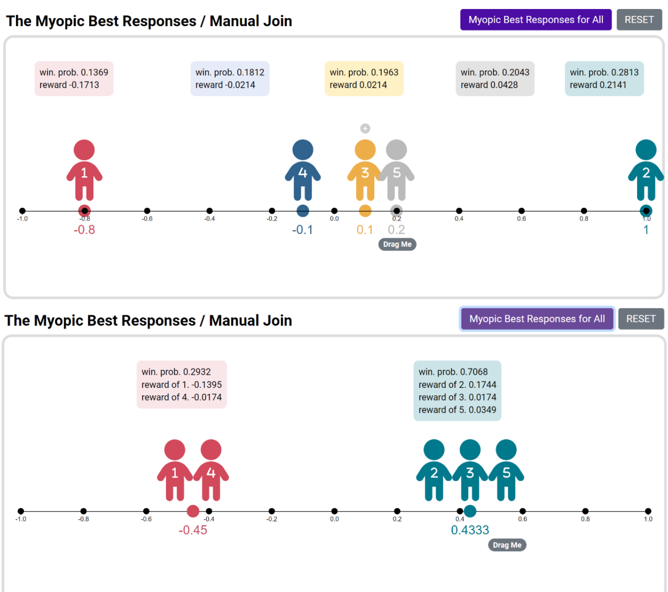

Fig. 1. 1D Representation: myopic best-responses and a PNE.

B. Group Joining: Myopic Best Responses

We assume that agent i decides to join G_j, for which $j = \arg\max_\ell p_\ell(\tau) \cdot \langle \bar{g}_\ell, s_i \rangle$. We call this strategy a *myopic best response*. An agent joins a group by considering not only its winning probability but also the utility that the agent can get from the group before joining. The state in the bottom of Fig. 1 is an example of a PNE.

IV. Opinion Updates by Online Learning

A. 2D Representation

We illustrate the opinions and beliefs as well as the dynamics of opinion changes in $\mathcal{K} := \{x \in [-1, 1]^2 : \|x\|_2 \le 1\} \subset \mathbb{R}^2$. The 2-norm constraint that $\|z_i\|_2, \|s_i\|_2 \le 1$ correlates the dimensions. A projection of the opinion is required if the constraint is not satisfied.

B. Online Gradient Ascent

We consider the setting that each agent tries to maximize its own reward by "changing its opinion" without deviating from the group to which it belongs. Each agent runs the online gradient ascent algorithm to iteratively update its opinion so as to maximize its reward. The update is done by adding a certain quantity (tuned by the learning rate η) toward the direction of the gradient. A "projection" $\Pi_{\mathcal{K}}(x)$ which projects x onto the feasible set $\mathcal{K}$ by dividing its 2-norm is performed if necessary.

Algorithm: Multi-Agent Online Gradient Ascent

Input: feasible set $\mathcal{K}$, T, learning rate η.
for $t \leftarrow 1$ to T **do**
  for each agent i **do**
    observe reward $r_i(\tau)$, where state $\tau = (\mathbf{z}, \mathbf{s}, \mathcal{G})$
    $z_{i,t+1} \leftarrow \Pi_{\mathcal{K}}(z_{i,t} + \eta \nabla_{z_i} r_i(\tau))$
  end for
end for

C. Online Gradient Ascent with Regularization

The reward function for agent i including the regularizer, is defined as $r_i(\tau) = \sum_{j=1}^{m} p_j(\tau)\langle s_i, \bar{g}_j \rangle - \|z_i - s_i\|_2^2$. Since $-\|z_i - s_i\|_2^2$ is always non-positive, an agent will be constrained to consider "not being too far from its belief." Our experimental illustrations show that such a regularization helps the game converge to a state where agents' opinions will not be too far from their beliefs (see Fig. 2).

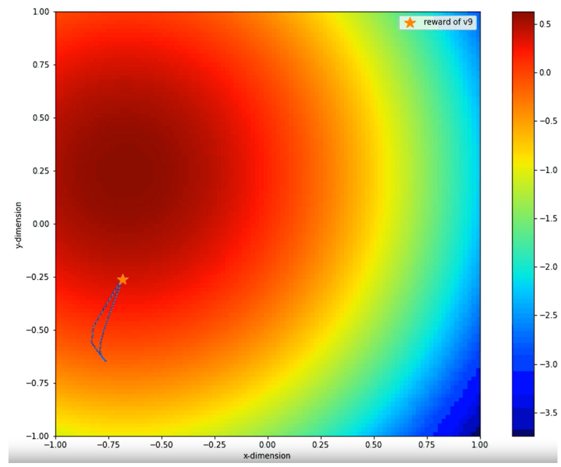

Fig. 2. Opinion updates via the online gradient ascent with regularization.

V. Conclusion

This article presents a preliminary study on the dynamics of group formation. From the illustrations, readers can realize what a pure-strategy Nash equilibrium in a system of multi-agents is and also learn how an online gradient ascent algorithm as one of the dynamics can reach a stable state.

Acknowledgments

This work is supported by the National Science and Technology Council under grant nos. NSTC 110-2222-E-032-002-MY2, NSTC 111-2221-E-005-047- and NSTC 111-2410-H-A49-022-MY2. We also thank Victorien Yen for helping implement the JavaScript codes in the immersive article.

References

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, pp. 139–144, 2020.
[2] F. Orabona, "A modern introduction to online learning," *arXiv:1912.13213*, 2022.
[3] C.-C. Lin, C.-J. Lu, and P.-A. Chen, "On the efficiency of an election game of two or more parties: How bad can it be?" *arXiv:2303.14405*, 2023.