ULTRA: Exploring Team Recommendations in Two Geographies Using Open Data in Response to Call for Proposals

Siva Likitha Valluru
svalluru@email.sc.edu
University of South Carolina
Columbia, SC, USA

Michael Widener
widenerm@bc.edu
Boston College
Chestnut Hill, MA, USA

Biplav Srivastava
biplav.s@sc.edu
University of South Carolina
Columbia, SC, USA

Sugata Gangopadhyay
sugata.gangopadhyay@cs.iitr.ac.in
Indian Institute of Technology
Roorkee, Uttarakhand, India

ABSTRACT

This paper demonstrates ULTRA (**U**niversity-**L**ead **T**eam Builder from **R**FPs and **A**nalysis), a novel AI-based recommendation system for team formation, where (1) candidate teams are formed with the goal to reach highest possible skill coverage, as demanded by an opportunity, and (2) the challenge of fair distribution of opportunities is balanced amongst all available members. Using this tool, users can explore the skills required in open data from proposal calls *(demand)* and adeptly assemble teams from candidate researcher profiles *(supply)*. The efficiency of these teams is then evaluated using an innovative goodness metric and validated through both quantitative and qualitative experiments. Beyond teaming, the tool design and evaluation of this work could interest researchers exploring the potential of *set* recommendation in other applications, rather than the well-understood traditional *single-item* recommendations. We deploy this system in two major institutions from diverse geographical regions of the world (United States and India), and in doing so, we show that ULTRA can generate good candidate teams across differing teaming contexts, and support the notion that our system is widely expandable.

1 INTRODUCTION

Forming effective and *successful* teams is a necessary business strategy (e.g., forming effective development teams in order to achieve project success in agile software development [62], establishing high-functioning surgical teams using sociometric maps [54], forming sports teams in competitive selections [4], forming a team of agents for search and rescue tasks in urban disaster scenarios [35]). In this paper, we use group recommendation strategies to focus on collaborative teaming for researchers applying to funding sources in response to calls for proposals. Funding agencies (e.g., Department of Science & Technology (DST) and Defence Research and Development Organisation (DRDO)) play a crucial role in providing support for research funding in public universities. Such opportunities also demand rapid formation of interdisciplinary teams comprising of individuals with diverse technical backgrounds. Not all teams achieve the desired level of success in fulfilling their objectives, however, primarily due to factors such as budget restrictions, time constraints, communication overhead, and disparities in skill sets among team members.

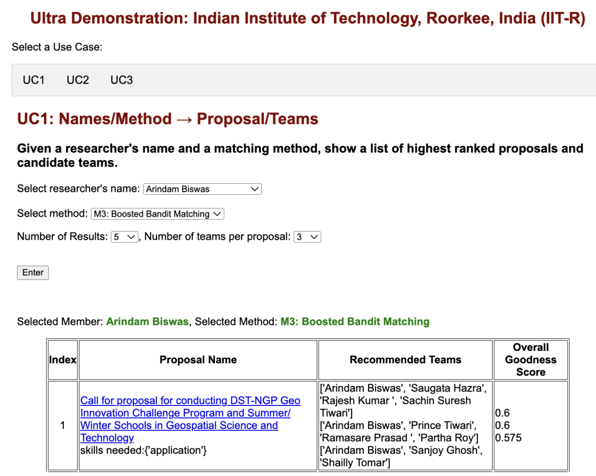

Figure 1: Demo use case UC_1: Given researcher name and teaming method, display highest ranked proposals and teams.

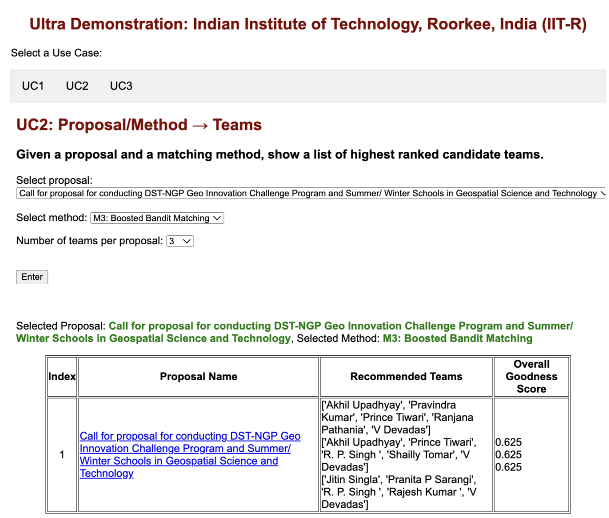

Figure 2: Demo use case UC_2: Given a call for proposal and teaming method, display highest ranked proposals and teams.

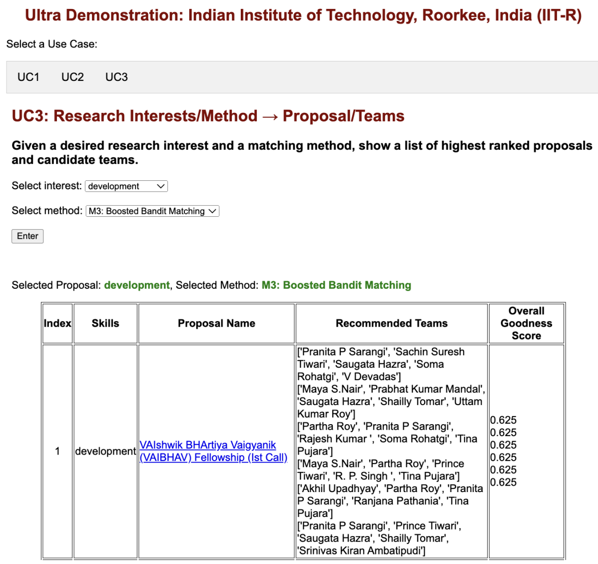

Figure 3: Demo use case UC_3: Given a research interest and teaming method, display highest ranked proposals and teams.

We expand on our prior work and show a demonstration of ULTRA[1] (**U**niversity-**L**ead **T**eam Builder from **R**FPs and **A**nalysis) [52, 58], a novel AI-driven prototype for assisting with team formation when researchers respond to requests for proposals (RFPs) from funding agencies. Our system first retrieves the technical skills required by open-data RFPs from publicly available sources (e.g., DST) and research interests provided by researchers in online profiles (e.g., personal homepages, Google Scholar publication history), as well as any other additional teaming constraints that interested users of an institution may place. Using AI and NLP techniques, ULTRA next suggests a list of candidate teams for each RFP and validates each result with a goodness score (see Section 3.2), ensuring that each team comprises of at least of two members.

ULTRA has currently been deployed at two major institutions in different geographies of the world: (1) University of South Carolina, USA (USC) [58], and (2) Indian Institute of Technology, Roorkee, India (IIT-R). The tool comprises of three practical use cases: (1) *UC1:* Given a researcher's name and matching method, display upcoming RFPs and teaming results. (2) *UC2:* For a given RFP, display the best possible teams that can be formed. (3) *UC3:* For a given research interest, display RFPs that are best aligned with that interest and the highest skilled teams that can be formed. Figures 1, 2, and 3 demonstrate all three cases at IIT-R and how the system works for an individual user who may become a potential team participant.

2 RELATED WORK

Some well-studied problems of AI in team formation come from cooperative decision-making in game theory frameworks [7, 24, 38, 49, 51, 59], where team cooperation is ensured by considering a combination of individual cost as a team cost function. Team formation has also been in considered in competitive sport leagues, where the physical and functional well-being of players is first evaluated through a multi-criteria assessment and ranking system before forming optimal teams [15, 55]. Lappas *et al.* [32] extend the problem of team formation into the social networks domain, where each node in a network is labeled with a set of skills that it possesses, and given a task that requires a certain coverage of skills to be satisfied, the goal is to next find a subgraph, where (ideally) all skills are present and the communication cost remains minimal. It is critical to note that the complexity of making multi-criteria decisions and building successful teams also increases with the number of available candidates [63], and as a result, team formation problems have become computationally expensive and NP-hard [9, 26, 29, 39, 48]. Fair allocation of tasks to candidate members should still remain unaffected, however, such that an individual is neither overassigned or neglected with tasks [5].

Existing literature systematically sheds light on the latest developments in team formation, provides an overview and classification of computational methods, identifies characteristics that are critical for effectively assigning team members to tasks, and discusses opportunities, gaps, and challenges in relevant research areas [5, 12–14, 20, 21, 30, 36, 40, 43, 53, 60]. Recruiting a group of experts to work together towards a common goal is not always easy, however, as there is no guarantee that they will *always* operate as a team [27] and may even face disadvantages that come from high task and social cohesion [22].

Challenges in Single-Item Recommendation. A common goal of many commercial recommender systems is to provide top-k highest relevant suggestions of items that are most likely of interest to a particular user [11, 31]. However, many single-item recommender systems also suffer from known issues such as popularity bias [1], cold start problem [45], data sparsity [31], user privacy [18], and scalability [61].

Motivation. We therefore expand on our prior work, ULTRA (**U**niversity-**L**ead **T**eam Builder from **R**FPs and **A**nalysis) [52, 58], and consider a group recommendation setting that promotes research collaboration using novel AI methods to recommend optimal teaming suggestions. Our work promotes multi-functional and interdisciplinary teams to form, bringing together individuals with diverse work responsibilities from various disciplines.

3 PROBLEM SETTING

In this section, we briefly explain the problem and the metrics and methods used to display teaming suggestions.

3.1 Problem Definition

Funding agencies regularly release Requests for Proposals (RFPs) centered around specific themes, seeking innovative ideas for potential funding. Researchers, in response, submit proposals outlining their ideas and presenting a comprehensive actionable plan to reach said goals within certain constraints, including budget limitations and time frames. While doing so, researchers frequently seek collaboration with colleagues to jointly respond to these calls. Consequently, we examine a collaborative environment where the

[1] A full demo interaction with ULTRA can be found at https://www.youtube.com/watch?v=CUk_Rgvd5kg. Additional details about usecases, experiments, and survey resources are at [57].

availability of potential participants may fluctuate, often unpredictably.

We identify at least two distinct user groups for our system: (1) administrators within researchers' organizations (e.g., universities) aiming to promote more collaborations, proposals, and diversity at their institutions, and (2) prospective team members responding to a specific RFP and seeking collaborative opportunities. Various environments call for different teams to be formed and matched with relevant opportunities. Furthermore, candidate team members may also evolve over time, along with their skills and research directions. Each user, i.e., university administrator or potential team participant, will interact with the system when an RFP is announced. The system will first leverage the requirements set forth by the call, match them with the skill sets and interests extracted from candidate researcher profiles, and algorithmically propose potential team formations, which users may either accept or reject. We then evaluate the effectiveness of our algorithms and validate the teaming results through a goodness score (see Section 3.2).

3.2 Metrics

For each candidate team formed, we measure its effectiveness towards a call for proposal using a goodness score G. The score indicates the likelihood of a team succeeding in meeting the requirements of a given RFP, and is determined by calculating the weighted mean of the following four configurable metrics: *redundancy (m_r), set size (m_s), coverage (m_c),* and *k-robustness (m_k).* Redundancy is defined as the percentage of overlapping or duplicate skills within a team of multiple researchers. Set size is defined as the number of members present within each candidate team. Coverage refers to the extent to which a candidate team possesses a diverse range of skills that collectively cover the requirements of a given RFP. Lastly, we borrow the k-robustness metric from [44], where a team can still satisfy all teaming constraints even in the absence of at most k members. Table 1 gives a brief summary of the metrics and goodness score.

To compute the overall goodness score, we first normalize the aforementioned metrics to make their values query-independent, and assign each of the metrics a predefined weight by default. Given our use cases (Section 4.1), these weights are initialized with the intuition to optimize both maximum profit (i.e., project completion) and credibility (i.e., project quality) a team can achieve [25, 47]. As efficient team formation also often seeks to minimize skill redundancy while ensuring a well-balanced set of complementary skills within the team [42, 64], high coverage and robustness are, therefore, more desired for overall project success, whereas high redundancy and set size are less prioritized. For each candidate team, we consequently penalize the latter two metrics, and reward the former. The penalized metrics are set to a negative weight of -1, whereas the rewarded ones are set to a positive weight of +1. The team's overall goodness score is then subsequently determined by computing the weighted mean across all four metrics. For additional reference, we make our metrics tool publicly available on GitHub[2].

[2]Our metrics tool, *Ultra-Metric*: https://github.com/ai4society/Ultra-Metric

Metric	Weight w_i	Definition
m_r	w_r	M out of N total skills that are satisfied by more than one candidate member.
m_s	w_s	Team size M out of a max allowed N members.
m_c	w_c	M out of N total skills that are satisfied by a team, in response to an RFP.
m_k	w_k	Represented using a binary indicator of $\{0, 1\}$, where 0 indicates a 0-robust team and 1 indicates the team is *at least* 1-robust.
G	–	A weighted mean of the aforementioned four metrics and their respective weights: $G = \sum_{i=1}^{4} m_i w_i$

Table 1: Evaluating a team's overall goodness G using four metrics, where each is assigned a configurable weight w_i and normalized afterwards, such that $\sum_{i=1}^{4} w_i = 1$.

3.3 Methods

We used four different methods to suggest teaming choices: (1) *M0: Random Team Formation,* (2) *M1: Team Formation Using String Matching,* (3) *M2: Team Formation Using Taxonomical Matching,* and (4) *M3: Team Formation Using Boosted Bandit.*

We consider *M0* as our baseline, where candidate teams are formed in a randomized manner, without any adherence to the skills in demand, and matched to an arbitrary proposal, regardless of relevance. In *M1,* given a call for proposal and list of N available researchers, we extract the skills needed from each call and match with those provided by researchers. Using a string matching threshold, we verify if there is any overlap between the two sets and form candidate teams accordingly. To further improve the accuracy and precision of *M0* and *M1*, we next consider *M2,* a query-based *semantic matching* method, combined with the use of a *taxonomy*. We use a poly-hierarchical, subject-based ontology, provided by the *ACM Computing Classification System (CCS)* [3], determine if two research skills may be matched semantically rather than only string-wise, and form candidate teams using researchers with the highest taxonomical matches. Finally, *M3* extracts rules automatically from data, unlike the previous three methods which applied manually-crafted rules to match researchers to teams. With more facts and data provided, this method uses contextual bandits [28] to learn more complex rules automatically and display teaming recommendations.

4 DEMONSTRATION OF ULTRA

4.1 Use Cases

We identify at least three practical use cases (UCs) that demonstrate the utility of our system. Each use case has various input prompts to select from and includes different algorithms that can be used to recommend or suggest proposals and teams to interested users.

For the first use case, UC_1, given a researcher's name and a teaming method, we display a list of k highest ranked proposals and possible teams). Similarly, for UC_2, given a proposal call from a list of recently announced proposals (ideally refreshed in real-time or regularly), we display the best possible teams available for the call. And the final use case, UC_3, takes input in the form of a research interest and teaming method, and displays respective matching

Method	Average Quality	Average Volume
M0	0.0883 ± 0.0106	10
M1	0.1063 ± 0.1001	10
M2	0.2734 ± 0.1017	8
M3	0.5229 ± 0.0035	3

Table 2: Average quality (G) and volume of teams (#T) shown *per* researcher (r_j) at IIT-R. This was done for each method M_i, across 227 RFPs and 141 researchers. For average quality, we report the mean and standard deviation, denoted as `mean`±`STD`.

proposals and teams for those parameters. We empirically evaluate the three use cases and the functionality of the four methods used in Section 4.2.

4.2 Computational Evaluation of Output

4.2.1 Quality vs. Volume of Teams. For each method, we assess the quality (goodness) and volume (size) of each teaming suggestion that every researcher receives per every call for proposal. Our experiments iterate across an initial dataset of 227 RFPs and 141 researchers. The RFP data has been gathered from the Department of Science & Technology (DST), a division of Research and Development (R&D) programmes belonging to and funded by the Government of India's Ministry of Science and Technology [16]. We collect candidate researcher information from IIT-R's faculty member directly and extract their respective research profiles from Google Scholar.

For each call, each researcher has a maximum cutoff of 10 teams. For each method, we then find the average goodness of teams a researcher has been recommended. Table 2 shows the computational evaluation of ULTRA on IIT-R's data. We observe another unique trade-off as a result, where teams formed algorithmically led to an increased precision and quality, and a notable decrease in quantity of teams. *M0*, random team formation, showed poorest quality in results with an average goodness of only 0.0883, despite the number of teams being abundantly available. While *M3*, on the other hand, shows a decrease in the number of teaming choices available for a researcher, there is nonetheless a visible increase in quality, compared to the average goodness received for *M0*.

5 DISCUSSION

A challenge of interdisciplinary team formation problems arises when candidate researchers are restricted to work for only a limited number of RFPs. Another challenge arises when a team of researchers who possess all skill requirements are recommended for an RFP, which may limit the quality of available choices of researchers for other teams. Fairness is therefore an aspect of team formation that should be also taken into account, in such a way that good members could be assigned to all needed teams [37, 46].

In current fairness-AI literature, there exist two frameworks to determine algorithmic bias: individual fairness [17] and group fairness [33]. Individual fairness measures ask that algorithms make teaming decisions about individuals in such a way that similar individuals are treated similarly [50]. Group fairness requires an equitable treatment of all users in protected demographics [50]. In

Gender	Pop. Bias	*M0*	*M1*	*M2*	*M3*
Male	0.199	0.198	0.060	0.190	0.075
Female	0.050	0.049	0.035	0.047	0.031
Unknown	0.752	0.751	0.905	0.764	0.894

Table 3: For IIT-R data, we group it into three possible gender categories: *male*, *female*, and *unknown*, and display the population and model bias. (The gender distribution is calculated for all methods, across all teaming results.)

the work at hand, we study both population and model bias in the context of unfairness with regards to a single protected attribute: *gender* of a researcher. For input data, we gather all candidate researchers and the teams recommended to them. We infer each researcher's gender using an open-source Python library called `gender_guesser`. Out of 141 researchers, 28 are classified as *male*, 7 as *female*, and 106 as *unknown*. Table 3 shows the bias present in ULTRA (IIT-R), across all four methods.

Additionally, fairness metrics, tools, and principles have also been discussed in literature [2, 6, 8, 10, 17, 19, 23, 34, 41, 56], and they can be optimized by adding fairness terms to our teaming objectives. We leave such a study for future work.

6 CONCLUSION

To conclude, we presented the problem of building teams for funding that allows for collaboration opportunities, and demonstrated an AI-assisted tool for team formation. We then created and implemented AI methods using string, semantic, and relational boosted bandit methods in ULTRA, and quantitatively demonstrated them to be useful for potential team participants, where informed methods show an increase in recommendation quality and precision, and a notable decrease in the average number of available teams. Since our methods are reliant on open data, however, this dependency can both be a source of strength and weakness. While data has the potential to encourage teaming without human bias, it can also lead to suboptimal recommendations if the data is obsolete or insufficient (e.g., a research interest no longer accurately reflects a researcher's current and future research interests). Similarly, any feedback on the success of a teaming suggestion is only possible when an RFP has been awarded, and this information is usually not available until months or years after the recommendation.

Our work also inspires several interesting future extensions: considering a variety of domain knowledge including but not limited to fairness constraints, teaming constraints, domain-specific constraints in various collaborative settings (e.g., healthcare, finance), etc. and developing a knowledge-driven learning system that can both exploit both the data and such knowledge. Another potential direction involves developing methods for interactive teaming, allowing the system to not only present recommendations but also provide explanations and further refine and fine-tune the learned models with human inputs.

REFERENCES

[1] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. 2019. Managing popularity bias in recommender systems with personalized re-ranking. *arXiv preprint arXiv:1901.07555* (2019).

[2] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. 2020. The connection between popularity bias, calibration, and fairness in

recommendation. In *Proceedings of the 14th ACM Conference on Recommender Systems*. 726–731.

[3] ACM. 2012. ACM Classification Scheme. In *https://www.acm.org/publications/computing-classification-system/how-to-use*.

[4] Faez Ahmed, Kalyanmoy Deb, and Abhilash Jindal. 2013. Multi-objective optimization and decision making approaches to cricket team selection. *Applied Soft Computing* 13, 1 (2013), 402–414.

[5] Aris Anagnostopoulos, Luca Becchetti, Carlos Castillo, Aristides Gionis, and Stefano Leonardi. 2012. Online team formation in social networks. In *Proc. 21st international conference on WWW*.

[6] Ashwathy Ashokan and Christian Haas. 2021. Fairness metrics and bias mitigation strategies for rating predictions. *Information Processing & Management* 58, 5 (2021), 102646.

[7] Coralio Ballester. 2004. NP-completeness in hedonic games. *Games and Economic Behavior* 49, 1 (2004), 1–30.

[8] Alex Beutel, Jilin Chen, Tulsee Doshi, Hai Qian, Allison Woodruff, Christine Luu, Pierre Kreitmann, Jonathan Bischof, and Ed H Chi. 2019. Putting fairness principles into practice: Challenges, metrics, and improvements. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 453–459.

[9] Avradeep Bhowmik, Vivek Borkar, Dinesh Garg, and Madhavan Pallan. 2014. Submodularity in team formation problem. In *Proceedings of the 2014 SIAM international conference on data mining*. SIAM, 893–901.

[10] Sarah Bird, Miro Dudík, Richard Edgar, Brandon Horn, Roman Lutz, Vanessa Milan, Mehrnoosh Sameki, Hanna Wallach, and Kathleen Walker. 2020. Fairlearn: A toolkit for assessing and improving fairness in AI. *Microsoft, Tech. Rep. MSR-TR-2020-32* (2020).

[11] Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. 2013. Recommender systems survey. *Knowledge-based systems* 46 (2013), 109–132.

[12] Bart H Boon and Gerard Sierksma. 2003. Team formation: Matching quality supply and quality demand. *European Journal of Operational Research* 148, 2 (2003), 277–292.

[13] Chad Crawford, Zenefa Rahaman, and Sandip Sen. 2016. Evaluating the efficiency of robust team formation algorithms. In *International Conference on Autonomous Agents and Multiagent Systems*. Springer, 14–29.

[14] Wilmax Marreiro Cruz and Seiji Isotani. 2014. Group formation algorithms in collaborative learning contexts: A systematic mapping of the literature. In *Collaboration and Technology: 20th International Conference, CRIWG 2014, Santiago, Chile, September 7-10, 2014. Proceedings 20*. Springer, 199–214.

[15] Stanislav Dadelo, Zenonas Turskis, Edmundas Kazimieras Zavadskas, and Ruta Dadeliene. 2014. Multi-criteria assessment and ranking system of sport team formation based on objective-measured values of criteria set. *Expert Systems with Applns.* 41, 14 (2014), 6106–6113.

[16] DST. 2023. Department of Science & Technology. https://dst.gov.in/. *Government of India | Ministry of Science and Technology* (2023).

[17] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*. 214–226.

[18] Arik Friedman, Bart P Knijnenburg, Kris Vanhecke, Luc Martens, and Shlomo Berkovsky. 2015. Privacy aspects of recommender systems. *Recommender systems handbook* (2015), 649–688.

[19] Pratyush Garg, John Villasenor, and Virginia Foggo. 2020. Fairness metrics: A comparative analysis. In *2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 3662–3666.

[20] Matthew E. Gaston and Marie desJardins. 2005. Agent-Organized Networks for Dynamic Team Formation. In *Proc. AAMAS* (The Netherlands). ACM, 230–237. https://doi.org/10.1145/1082473.1082508

[21] Matthew E Gaston, John Simmons, and Marie desJardins. 2004. Adapting Network Structure for Efficient Team Formation.. In *AAAI Technical Report (2)*. 1–8.

[22] James Hardy, Mark A Eys, and Albert V Carron. 2005. Exploring the potential disadvantages of high cohesion in sports teams. *Small group research* 36, 2 (2005), 166–187.

[23] J Henry Hinnefeld, Peter Cooman, Nat Mammo, and Rupert Deese. 2018. Evaluating fairness metrics in the presence of dataset bias. *arXiv preprint arXiv:1809.09245* (2018).

[24] Martin Hoefer, Daniel Vaz, and Lisa Wagner. 2015. Hedonic coalition formation in networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 29.

[25] Jiuchuan Jiang, Bo An, Yichuan Jiang, Chenyan Zhang, Zhan Bu, and Jie Cao. 2019. Group-oriented task allocation for crowdsourcing in social networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 51, 7 (2019), 4417–4432.

[26] Julio Juárez and Carlos A Brizuela. 2018. A multi-objective formulation of the team formation problem in social networks: preliminary results. In *Proceedings of the genetic and evolutionary computation conference*. 261–268.

[27] Julio Juárez, Cipriano (Pano) Santos, and Carlos A. Brizuela. 2021. A Comprehensive Review and a Taxonomy Proposal of Team Formation Problems. *ACM Computing Survey* 54, 7, Article 153 (jul 2021), 33 pages. https://doi.org/10.1145/3465399

[28] Ashutosh Kakadiya, Sriraam Natarajan, and Balaraman Ravindran. 2021. Relational boosted bandits. In *Proceedings of the AAAI Conference on AI*, Vol. 35. 12123–12130.

[29] Mehdi Kargar, Aijun An, and Morteza Zihayat. 2012. Efficient bi-objective team formation in social networks. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2012, Bristol, UK, September 24-28, 2012. Proceedings, Part II 23*. Springer, 483–498.

[30] Hyeyoung Ko, Suyeon Lee, Yoonseo Park, and Anna Choi. 2022. A survey of recommendation systems: recommendation models, techniques, and application fields. *Electronics* 11, 1 (2022), 141.

[31] Madhusree Kuanr and Puspanjali Mohapatra. 2021. Recent challenges in recommender systems: a survey. In *Progress in Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2019, Volume 2*. Springer, 353–365.

[32] Theodoros Lappas, Kun Liu, and Evimaria Terzi. 2009. Finding a team of experts in social networks. In *Proc. 15th ACM SIGKDD*. 467–476.

[33] Yunqi Li, Hanxiong Chen, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang. 2021. User-oriented fairness in recommendation. In *Proceedings of the Web Conference 2021*. 624–632.

[34] Yunqi Li, Hanxiong Chen, Shuyuan Xu, Yingqiang Ge, Juntao Tan, Shuchang Liu, and Yongfeng Zhang. 2023. Fairness in Recommendation: Foundations, Methods and Applications. *ACM Transactions on Intelligent Systems and Technology* (2023).

[35] Somchaya Liemhetcharat and Manuela Veloso. 2012. Modeling and learning synergy for team formation with heterogeneous agents. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. 365–374.

[36] Somchaya Liemhetcharat and Manuela Veloso. 2014. Weighted synergy graphs for effective team formation with heterogeneous ad hoc agents. *Artificial Intelligence* 208 (2014), 41–65.

[37] Lucas Machado and Kostas Stefanidis. 2019. Fair team recommendations for multidisciplinary projects. In *IEEE/WIC/ACM International Conference on Web Intelligence*. 293–297.

[38] Hannan Amoozad Mahdiraji, Elham Razghandi, and Adel Hatami-Marbini. 2021. Overlapping coalition formation in game theory: A state-of-the-art review. *Expert Systems with Applications* 174 (2021), 114752.

[39] Anirban Majumder, Samik Datta, and KVM Naidu. 2012. Capacitated team formation problem on social networks. In *Proceedings of the 18th ACM SIGKDD international conference on knowledge discovery and data mining*. 1005–1013.

[40] Naseebah Maqtary, Abdulqader Mohsen, and Kamal Bechkoum. 2019. Group formation techniques in computer-supported collaborative learning: A systematic literature review. *Technology, Knowledge and Learning* 24 (2019), 169–190.

[41] Staša Milojević. 2014. Principles of scientific research team formation and evolution. *Proceedings of the National Academy of Sciences* 111, 11 (2014), 3984–3989.

[42] E Molleman and J Slomp. 1999. Functional flexibility and team performance. *International Journal of Production Research* 37, 8 (1999), 1837–1858.

[43] Chinasa Odo, Judith Masthoff, and Nigel Beacham. 2019. Group formation for collaborative learning: A systematic literature review. In *Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part II 20*. Springer, 206–212.

[44] Tenda Okimoto, Nicolas Schwind, Maxime Clement, Tony Ribeiro, Katsumi Inoue, and Pierre Marquis. 2015. How to Form a Task-Oriented Robust Team.. In *AAMAS*. 395–403.

[45] Deepak Kumar Panda and Sanjog Ray. 2022. Approaches and algorithms to mitigate cold start problems in recommender systems: a systematic literature review. *JIIS* 59, 2 (2022), 341–366.

[46] Evaggelia Pitoura, Kostas Stefanidis, and Georgia Koutrika. 2022. Fairness in rankings and recommendations: an overview. *The VLDB Journal* (2022), 1–28.

[47] Radin Hamidi Rad, Shirin Seyedsalehi, Mehdi Kargar, Morteza Zihayat, and Ebrahim Bagheri. 2022. A Neural Approach to Forming Coherent Teams in Collaboration Networks.. In *EDBT*. 2–440.

[48] Syama Sundar Rangapuram, Thomas Bühler, and Matthias Hein. 2013. Towards realistic team formation in social networks based on densest subgraphs. In *Proceedings of the 22nd international conference on World Wide Web*. 1077–1088.

[49] Elham Semsar-Kazerooni and Khashayar Khorasani. 2009. Multi-agent team cooperation: A game theory approach. *Automatica* 45, 10 (2009), 2205–2213.

[50] Nasim Sonboli, Masoud Mansoury, Ziyue Guo, Shreyas Kadekodi, Weiwen Liu, Zijun Liu, Andrew Schwartz, and Robin Burke. 2021. librec-auto: A Tool for Recommender Systems Experimentation. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 4584–4593.

[51] Matthew Spradling, Judy Goldsmith, Xudong Liu, Chandrima Dadi, and Zhiyu Li. 2013. Roles and teams hedonic game. In *Algorithmic Decision Theory: Third International Conference, ADT 2013, Bruxelles, Belgium, November 12-14, 2013, Proceedings 3*. Springer, 351–362.

[52] Biplav Srivastava, Tarmo Koppel, Sai Teja Paladi, Siva Likitha Valluru, Rohit Sharma, and Owen Bond. 2022. ULTRA: A Data-driven Approach for Recommending Team Formation in Response to Proposal Calls. In *IEEE ICDM Workshops 2022*. https://doi.org/10.1109/ICDMW58026.2022.00130

[53] Georgios Stavrou, Panagiotis Adamidis, Jason Papathanasiou, and Konstantinos Tarabanis. 2023. Team Formation: A Systematic Literature Review. (2023).

[54] Mark Sykes, Brigid M Gillespie, Wendy Chaboyer, and Evelyn Kang. 2015. Surgical team mapping: implications for staff allocation and coordination. *AORN journal* 101, 2 (2015), 238–248.

[55] Madjid Tavana, Farshad Azizi, Farzad Azizi, and Majid Behzadian. 2013. A fuzzy inference system with application to player selection and team formation in multi-player sports. *Sport Management Review* 16, 1 (2013), 97–110.

[56] Tzu-Liang Bill Tseng, Chun-Che Huang, How-Wei Chu, and Roger R Gung. 2004. Novel approach to multi-functional project team formation. *International Journal of Project Management* 22, 2 (2004), 147–159.

[57] Siva Likitha Valluru, Biplav Srivastava, and Sai Teja Paladi. 2023. ULTRA Resources Github. In *https://github.com/ai4society/ULTRA-Team- Recommendation-Resources*.

[58] Siva Likitha Valluru, Biplav Srivastava, Sai Teja Paladi, Siwen Yan, and Sriraam Natarajan. 2022. ULTRA: A Data-driven Approach for Recommending Team Formation in Response to Proposal Calls. *The Thirty-Sixth Annual Conference on Innovative Applications of Artificial Intelligence (IAAI/AAAI)*.

[59] Michael Van De Vijsel and John Anderson. 2004. Coalition formation in multi-agent systems under real-world conditions. *Proceedings of association for the advancement of artificial intelligence* (2004).

[60] Hyeongon Wi, Seungjin Oh, Jungtae Mun, and Mooyoung Jung. 2009. A team formation model based on knowledge and collaboration. *Expert Systems with Applications* 36, 5 (2009), 9121–9134.

[61] Yu Xin et al. 2015. *Challenges in recommender systems: scalability, privacy, and structured recommendations*. Ph. D. Dissertation. Massachusetts Institute of Technology.

[62] P Zainal, D Razali, and Zulkefli Mansor. 2020. Team formation for agile software development: a review. *Int. J. Adv. Sci. Eng. Inf. Technol* 10, 2 (2020), 555–561.

[63] Edmundas Kazimieras Zavadskas, Algimantas Zakarevicius, and Jurgita Antucheviciene. 2006. Evaluation of ranking accuracy in multi-criteria decisions. *Informatica* 17, 4 (2006), 601–618.

[64] Xi Zhu and Douglas R Wholey. 2018. Expertise redundancy, transactive memory, and team performance in interdisciplinary care teams. *Health services research* 53, 6 (2018), 4921–4942.