Bridging Historical Subgraph Optimization and Modern Graph Neural Network Approaches in Team Recommendation*

Mahdis Saeedi
University of Windsor, ON, Canada
msaeedi@uwindsor.ca

Christine Wong
University of Windsor, ON, Canada
wong93@uwindsor.ca

Hossein Fani
University of Windsor, ON, Canada
hfani@uwindsor.ca

Abstract

Team recommendation involves selecting experts with certain skills to form a successful task-oriented team. This tutorial provides a comprehensive study of conventional graph-based and a detailed review of cutting-edge neural network-based methods through unified definitions and formulations, along with insights into future research directions and real-world applications.

CCS Concepts

• **Information systems → Recommender systems**; • **Human-centered computing → Social recommendation**; • **Computing methodologies → Neural networks**; • **Theory of computation → Graph algorithms analysis**.

Keywords

Neural Team Recommendation, Subgraph Optimization, Graph Neural Network, Social IR

ACM Reference Format:
Mahdis Saeedi, Christine Wong, and Hossein Fani. 2025. Bridging Historical Subgraph Optimization and Modern Graph Neural Network Approaches in Team Recommendation. In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining (WSDM '25), March 10–14, 2025, Hannover, Germany*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3701551.3703483

1 Tutorial Agenda (180 minutes)

We start the tutorial with a brief introduction to the team recommendation problem and offer a novel taxonomy of computational methods, as shown in Figure 1, then provide motivation by highlighting its practical significance through seemingly unrelated yet highly valuable applications of team recommendation in education, research, and healthcare.

We then proceed to explain the computational methods employed in the literature, and focus on graph construction and subgraph optimization techniques, as they comprise the pioneering body of research, and continue on learning-based methods, particularly graph neural network-based techniques, which have been building up the following and yielding to state-of-the-art performances. Then, we discuss evaluation methodologies, including the datasets, metrics, and comparative baselines that exist in the literature, followed by major lines of future research. Finally, we conclude the tutorial with hands-on experience in neural team recommendation research.

2 Introduction (10 minutes)

Increasingly, the algorithmic search for collaborative teams, also known as team recommendation, prevails in different domains such as academia, manufacturing, law, freelancing and the healthcare sector, among others, to automate the formation of teams of experts whose combined skills, applied in coordinated ways, can successfully solve difficult tasks. Recommending a successful team whose members can effectively collaborate and deliver the outcomes within the specified constraints, such as planned budget and timeline, is challenging due to the immense number of candidates with various backgrounds, personality traits, and skills, as well as *unknown* synergistic balance among them; not all teams with *best* experts are necessarily successful [12].

The formation of teams has traditionally been performed manually, relying on human intuition and experience, resulting in a process that is often tedious, error-prone, and inefficient due to *i*) hidden personal and societal biases, *ii*) a multitude of criteria to optimize, *iii*) an overwhelming number of candidates, among other reasons. Team formation can be influenced by subjective opinions that carry hidden societal biases, neglecting *diversity* in recommended experts and leading to discrimination, reduced visibility for disadvantaged *female* experts, and racial/gender disparities.

In addition, decision making is further complicated by various criteria, including project importance, budget, time constraints, and team size limitations. In addition to these, expert candidates should be examined based on individual and relative factors such as technical abilities, availability, individual cost, and personality traits, among others, which makes manual team formation on a large scale almost impossible.

Together with business sectors such as Linkedin[1], researchers have long sought to develop computational models to analyze massive collections of experts and efficiently learn relationships between experts and their skills in the context of successful and *un*successful teams and excel in recommending *almost surely* successful teams. This has resulted in a rich body of various approaches grounded in computational and social science theoretical and conceptual frameworks, wherein the problem definition of team recommendation remains the same essentially, while it has been referred to by such other names as team allocation, team composition, team selection, and team formation.

*https://fani-lab.github.io/OpeNTF/tutorial/wsdm25

[1] www.linkedin.com

WSDM '25, March 10–14, 2025, Hannover, Germany

ACM ISBN 979-8-4007-1329-3/25/03
https://doi.org/10.1145/3701551.3703483

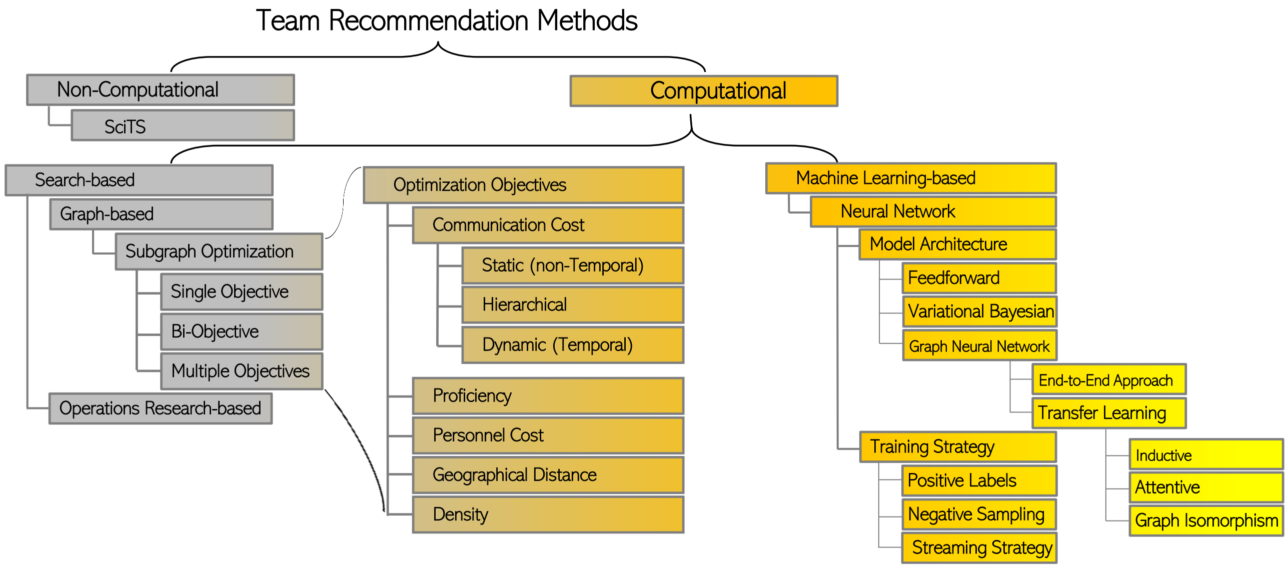

Figure 1: The hierarchy of team recommendation methods.

3 Applications (20 minutes)

In this part of our tutorial, we explain novel applications of team recommendation besides its common use cases, including:

Learning Group Recommendation: The team recommendation finds immediate application in group-based learning environments. In online classes, where physical presence and interaction are absent, team recommendation connects students to improve their social skills and combat the isolation that can sometimes accompany remote learning [4, 11]. In large classes, where individual interactions with the instructor may be limited, group work ensures that students still have ample opportunities to engage with the material. In addition, team recommendation in online and large classes promotes accountability and responsibility. These skills in collaboration and time management are invaluable in both academic pursuits and future professional endeavors, making group learning a holistic and beneficial approach for students in diverse classroom settings.

Reviewer Assignment: Another application of team recommendation is in peer-review assignments [1, 2] where a group of reviewers are paired with manuscripts within the reviewers' expertise for high-quality reviews while managing conflicts of interests. Like team recommendation, herein, research topics (skills) and reviewers (experts) are mapped into a latent space and, given a manuscript on a subset of research topics, team recommendation aims to recommend top-k optimum reviewers for the research topics.

Palliative Care: Team recommendation is also applicable in healthcare to assign a team of caregivers to patients who seek help with their daily activities due to disease or disorders [17]. The challenge lies in optimally assigning care providers in teams to address patient needs while considering factors such as communication, distance, and contract costs.

4 Proposed Methods (80 minutes)

In this tutorial section, we explore the evolution of team recommendation methods over time.

4.1 Prior Techniques *(20 minutes)*

The early computational models for team recommendation were developed in operations research (OR), optimizing objectives using integer linear and/or nonlinear programming (IP). Such work, however, was premised on the mutually independent selection of experts and overlooked the organizational and/or social ties. In our tutorial, we have excluded these methods. To bridge the gap, graph-based approaches have been proposed to recommend teams via subgraph optimization of *objectives* where the different aspects of real-world teams are captured like communication cost, levels of proficiency, and geographical proximity [9]. Subgraph optimization problems are NP-hard [8]. Therefore, heuristics have been developed to solve optimization in polynomial time through greedy and/or approximation algorithms.

In our tutorial, we formalize over 13 objectives and describe the seminal heuristics in three groups: *i*) those that target minimizing communication cost only; *ii*) those that consider additional objectives such as personnel cost, expertise level, and geographical proximity jointly with communication cost; and, *iii*) those considering maximizing the teams' density only.

4.2 State-of-the-Art Methods *(60 minutes)*

Advances in machine learning, particularly graph neural networks, have led to a paradigm shift toward learning-based methods for team recommendation [13]. These methods are different in that they learn the inherent structure of the ties among experts and their skills. Learning-based methods bring efficiency while enhancing efficacy due to the inherently iterative and online learning procedure, and can address the limitations of subgraph optimization solutions with respect to scalability, as well as dynamic expert networks [15]. In our tutorial, we explain this line of research based on *i*) model architectures, *ii*) training strategies, and *iii*) evaluation methodologies.

Neural Architectures: Neural team recommendation has started with autoencoders and is being followed through other neural-based architectures like a simple feedforward network whose parameters are learned by either maximum likelihood optimization or maximizing a posterior using Bayesian neural models [14]. Naturally enough, graph neural networks have also been receiving growing attention for the team recommendation problem for their expressive performance on the vector representation of the experts and their ties, and we will lay out their details in our tutorial.

Training Strategies: We explain various strategies to train neural model parameters, including *i*) negative sampling [5] to learn from instances of teams labeled with success or failure, *ii*) streaming training [7] that encode temporal aspects in neural-based team formation methods, and *iii*) curriculum-based strategy [3] that provide an order between experts from the easy popular experts to the hard non-popular ones to overcome the neural models' popularity bias.

Evaluation Strategies: We also discuss the benchmark datasets and what has been considered as successful teams to function as the gold truth. Also, we explore quantitative and qualitative metrics utilized to measure the quality of the recommended teams.

5 Open Challenges (20 minutes)

Although remarkable progress has been made, several open issues and potential future directions are worth more research and will further unleash the great potential of recommending effective teams.

End-to-End Graph Neural Network: Neural models that accept dense vector representation of skills in the input layer outperformed the sparse occurrence vectors. However, the dense vectors are learned separately in an unsupervised manner using a graph neural network, oblivious to supervised information about successful teams [16]. The team recommendation problem can be reformulated into a link prediction in an expert graph to directly and jointly learn dense vectors of skills and experts and recommend an optimum subset of experts as a team by predicting links, avoiding the unnecessary complications of two-phase graph representation learning and neural model fine-tuning.

Diversity-Aware Team Recommendation: Existing methods primarily focus on maximizing the success rate of recommended teams, largely ignoring *diversity* in the recommended list of experts. Meanwhile, social science research provides compelling evidence about the synergistic effects of diversity on team performance; diversity breeds innovation and increases teams' success by enabling a stronger sense of community and support, reducing conflict, and stimulating more creative thinking. However, there is little to no diversity-aware algorithmic method that mitigates unfair societal biases in team recommendation algorithms. In our tutorial, we introduce notions of fairness and protected attributes and study debiasing algorithms to mitigate the potential unfairness in the models' recommendations.

Spatial Team Recommendation: In search of an optimal team, companies further look for experts in a region where the organization is *geographically* based. Existing methods use skills as a primary factor while overlooking geographical location. We conclude our tutorial by bringing forth the *spatial* team recommendation problem; that is, given a set of experts, skills and geolocations, the goal is to determine if the combination of skills and geolocations in forming teams has synergistic effects.

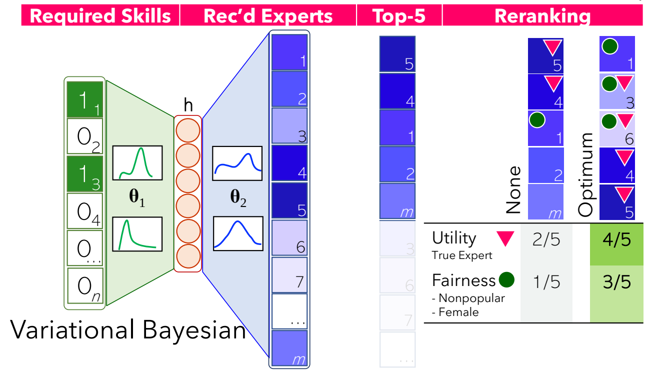

Figure 2: Post-processing fairness-aware reranking.

6 Hands-On (50 minutes)

We introduce publicly available libraries for team recommendation. Notably, we provide hands-on experience with `OpeNTF`[2] [6], an open-source benchmark library for neural models that: *i*) can efficiently preprocess large-scale datasets, *ii*) can be easily extended or customized to new neural methods, and *iii*) is extensible to experiments on new datasets from other domains. We also introduce `Adila`[3] [10], that enables further in-processing female-advocate loss regularization and/or post-processing reranking (Figure **??**) to the list of recommended experts to ensure the desired fair outcome. `Adila` is equipped with fairness metrics, which, in tandem with *utility* metrics, allows exploring the synergistic trade-offs between notions of fairness and success rate.

7 Relevance to the Community

The team recommendation problem is a domain within social information retrieval (**Social IR**), where the right group of experts are searched to solve the tasks either independently or with the help of social resources. In this tutorial, *i*) we target *beginner* or *intermediate* researchers, industry technologists and practitioners with a broad interest in developing recommender systems, willing to have a whole picture of team recommendation techniques. *ii*) Furthermore, this tutorial targets audiences from the graph neural network community for a comprehensive review of subgraph optimization objectives and calls them for further development of effective yet efficient graph neural networks with a special focus on team recommendation. Last, having regard to the unified comparative analysis, this tutorial enables *iii*) organizations and practitioners to compare different models and readily pick the most suitable one for their application to form *almost surely* guaranteed successful teams.

Background Knowledge: The target audience for this tutorial will be those *familiar* with graph theory and neural network architectures. Where appropriate, the tutorial will make no assumption about the audience's knowledge of more advanced techniques, such as subgraph bi-objective optimization methods or graph convolutional networks. As such, sufficient details will be provided as

[2] `https://github.com/fani-lab/OpeNTF`

[3] `https://github.com/fani-lab/Adila`

appropriate so that the content will be accessible and understandable to those with a fundamental understanding of such principles.

8 Prior Tutorials

Despite the substantial number of computational models for team recommendation, there is, however, yet to be a comprehensive tutorial with comparative analysis and critical reviews on approaches' applicability in real-world scenarios. To fill the gap, we provided a pioneering tutorial at `UMAP24`[4] centered on a narrowed scope of subgraph optimization with experts being *online* skilled users. Building on this framework, we extended our scope at `SIGIR-AP24`[5] by exploring emerging methods based on graph neural network to leverage the expert collaborative network structure for generating dense vector representation of skills. To foster further research, we propose this tutorial to systematically review seminal solutions through a novel taxonomy from a computational perspective, with a particular emphasis on state-of-the-art representation learning approaches and training strategies. In our tutorial, we bring forth a unifying and vetted methodology to the various definitions in this realm, criticize assumptions and comparative benchmarks, and point out shortfalls to smooth the path for future directions.

9 Presenters

Mahdis Saeedi (She/Her) is a Postdoctoral Fellow at the School of Computer Science and a Lecturer at the Department of Mathematics, University of Windsor. She has taught graph theory, linear algebra, and graph representation learning at undergraduate and graduate levels. During her PhD, she focused on theoretical aspects of graph mining, including edge ideals in bipartite and glued graphs. She has published in the field's premier venues, such as ECIR, WISE, UMAP, Transactions of Combinatorics, and the Journal of Algebra.
Christine Wong (She/Her) is an *undergraduate* student in the School of Computer Science, University of Windsor, with outstanding research contributions resulting in publications at SIGIR, CIKM, and UMAP conferences. She has received the Graduate Assistance Excellence Award, the Presidents's Renewable Entrance Scholarship, and the Silver Medal at the UWill Discover 2023 competition.
Hossein Fani (He/Him) is an Assistant Professor at the School of Computer Science, University of Windsor. His research is at the intersection of Social Network Analysis, User Modeling, and Information Retrieval, with a diverse team of 15+ HQP, funded by NSERC-DG, NSERC-RTI, and CFI-JELF. Fani's research appears in Elsevier's IP&M, ACM's TOIS, Wiley's JASIST, and SIGIR, CIKM, WSDM and ECIR. He translates his research into techniques for the industry while leading R&D funded by NSERC Alliance and Mitacs Accelerate. He has been granted patents by the USPTO for his research, including `US10,885,131`, `US11,768,522`, and `US12,067,625`. Fani's teaching experience spans over 15 years in countries with different cultures and educational systems. He has taught core courses, including natural language communication and graph representation learning, in multi-section classes ranging from 30 (graduate) to 200+ (undergraduate) with in-person, hyflex, and online modalities. He has developed two graduate courses, including big data analytics and the reciprocal role of AI, science, and society.

[4] `https://www.um.org/umap2024/tutorials/`
[5] `https://www.sigir-ap.org/sigir-ap-2024/tutorial/`

References

[1] Meltem Aksoy, Seda Yanik, and Mehmet Fatih Amasyali. 2023. Reviewer Assignment Problem: A Systematic Review of the Literature. *J. Artif. Intell. Res.* 76 (2023), 761–827. https://doi.org/10.1613/JAIR.1.14318
[2] Negar Arabzadeh, Sajad Ebrahimi, Sara Salamat, Mahdi Bashari, and Ebrahim Bagheri. 2024. Reviewerly: Modeling the Reviewer Assignment Task as an Information Retrieval Problem. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, CIKM 2024, Boise, ID, USA, October 21-25, 2024*, Edoardo Serra and Francesca Spezzano (Eds.). ACM, 5554–5555. https://doi.org/10.1145/3627673.3679081
[3] Reza Barzegar, Marco Nicola Kurepa, and Hossein Fani. 2025. Adaptive Loss-based Curricula for Neural Team Recommendation. In *WSDM*.
[4] Gonzalo Candel, Víctor Sánchez-Anguix, Juan M. Alberola, Vicente Julián, and Vicent J. Botti. 2023. An Integer Linear Programming Model for Team Formation in the Classroom with Constraints. In *Hybrid Artificial Intelligent Systems - 18th International Conference, HAIS 2023, Salamanca, Spain, September 5-7, 2023, Proceedings (Lecture Notes in Computer Science, Vol. 14001)*, Pablo García Bringas, Hilde Pérez García, Francisco Javier Martínez de Pisón, Francisco Martínez-Álvarez, Alicia Troncoso Lora, Álvaro Herrero, José Luís Calvo-Rolle, Héctor Quintián, and Emilio Corchado (Eds.). Springer, 397–408. https://doi.org/10.1007/978-3-031-40725-3_34
[5] Arman Dashti, Saeed Samet, and Hossein Fani. 2022. Effective Neural Team Formation via Negative Samples. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17-21, 2022*, Mohammad Al Hasan and Li Xiong (Eds.). ACM, 3908–3912. https://doi.org/10.1145/3511808.3557590
[6] Arman Dashti, Karan Saxena, Dhwani Patel, and Hossein Fani. 2022. OpeNTF: A Benchmark Library for Neural Team Formation. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17-21, 2022*, Mohammad Al Hasan and Li Xiong (Eds.). ACM, 3913–3917. https://doi.org/10.1145/3511808.3557526
[7] Hossein Fani, Reza Barzegar, Arman Dashti, and Mahdis Saeedi. 2024. A Training Strategy for Future Collaborative Team Prediction. In *ECIR (Lecture Notes in Computer Science)*. Springer.
[8] Richard M Karp. 1972. Reducibility among combinatorial problems. In *Complexity of Computer Computations: Proceedings of a symposium on the Complexity of Computer Computations*. Springer, 85–103.
[9] Theodoros Lappas, Kun Liu, and Evimaria Terzi. 2009. Finding a team of experts in social networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 28 - July 1, 2009*, John F. Elder IV, Françoise Fogelman-Soulié, Peter A. Flach, and Mohammed Javeed Zaki (Eds.). ACM, 467–476. https://doi.org/10.1145/1557019.1557074
[10] Hamed Loghmani and Hossein Fani. 2023. Bootless Application of Greedy Re-ranking Algorithms in Fair Neural Team Formation. In *Advances in Bias and Fairness in Information Retrieval - 4th International Workshop, BIAS 2023, Dublin, Ireland, April 2, 2023, Revised Selected Papers (Communications in Computer and Information Science, Vol. 1840)*, Ludovico Boratto, Stefano Faralli, Mirko Marras, and Giovanni Stilo (Eds.). Springer, 108–118. https://doi.org/10.1007/978-3-031-37249-0_9
[11] Larry K Michaelsen, Arletta Bauman Knight, and L Dee Fink. 2023. *Team-based learning: A transformative use of small groups in college teaching*. Taylor & Francis.
[12] Satyam Mukherjee, Yun Huang, Julia Neidhardt, Brian Uzzi, and Noshir Contractor. 2019. Prior shared success predicts victory in team competitions. *Nature human behaviour* 3, 1 (2019), 74–81.
[13] Hoang Nguyen, Radin Hamidi Rad, Feras N. Al-Obeidat, Ebrahim Bagheri, Mehdi Kargar, Divesh Srivastava, Jaroslaw Szlichta, and Fattane Zarrinkalam. 2023. Learning heterogeneous subgraph representations for team discovery. *Inf. Retr. J.* 26, 1 (2023), 8. https://doi.org/10.1007/S10791-023-09421-6
[14] Radin Hamidi Rad, Hossein Fani, Ebrahim Bagheri, Mehdi Kargar, Divesh Srivastava, and Jaroslaw Szlichta. 2024. A Variational Neural Architecture for Skill-based Team Formation. *ACM Trans. Inf. Syst.* 42, 1 (2024), 7:1–7:28. https://doi.org/10.1145/3589762
[15] Radin Hamidi Rad, Hossein Fani, Mehdi Kargar, Jaroslaw Szlichta, and Ebrahim Bagheri. 2020. Learning to Form Skill-based Teams of Experts. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19-23, 2020*, Mathieu d'Aquin, Stefan Dietze, Claudia Hauff, Edward Curry, and Philippe Cudré-Mauroux (Eds.). ACM, 2049–2052. https://doi.org/10.1145/3340531.3412140
[16] Anna Sapienza, Palash Goyal, and Emilio Ferrara. 2019. Deep Neural Networks for Optimal Team Composition. *Frontiers Big Data* 2 (2019), 14. https://doi.org/10.3389/FDATA.2019.00014
[17] Kalyani Selvarajah, Pooya Moradian Zadeh, Ziad Kobti, Mehdi Kargar, Mohd Tazim Ishraque, and Kathryn Pfaff. 2018. Team Formation in Community-Based Palliative Care. In *2018 Innovations in Intelligent Systems and Applications, INISTA 2018, Thessaloniki, Greece, July 3-5, 2018*. IEEE, 1–7. https://doi.org/10.1109/INISTA.2018.8466272