Uncovering Teamwork in Networks – Prediction, Optimization and Explanation

Liangyue Li and Hanghang Tong

Arizona State University

1 Introduction

A. Motivations. Teams are defined as people who are readily identifiable to each other by role and position. Individuals work together in teams to develop synergy. The ultimate outcome is expected to be greater than the sum of each individual. The performance of a team depends upon many factors, e.g., individual's ability, levels of commitment and their ability to work with others, etc. It is critical to understand the mechanism that drives the performance and provide measures to help struggling teams or enhance an existing team's performance.

B. Problem Description. In this work, we consider teams as subgraphs embedded in a larger social network and tackle three complementary research tasks, namely, prediction, optimization and explanation. See Fig. 1 for an overview.

Task 1: Team Performance Prediction. In this task, we aim to predict a given team's long-term performance (e.g., in 10 years) or performance trajectory (e.g., performance in the consecutive 10 years). Such prediction can help track a team's performance and invoke an early intervention if necessary.

Task 2: Team Performance Optimization. In this task, we focus on optimizing the performance of an existing team. For example, if a team member becomes unavailable for the given task, who is the best alternate to replace that person's role in the team (i.e., *team replacement*); if the team leader perceives the need to enhance certain expertise of the entire team, who shall we bring into the team (i.e., *team expansion*); if we need to reduce the size of an existing team, who shall leave the team with least disruption (i.e., *team shrinkage*).

Task 3: Team Performance Explanation. This task is complementary to the above two tasks by providing explanation to the prediction results as well as the optimization tasks. Such explanation can help users assess trust, which is fundamental if one plans to take high-stake actions based on the model.

C. Background and Related Work. Classic team performance and team formation problems have been extensively studied in psychology and operation research communities [9]. In [3], Lappas et al. first study the team formation problem in the context of social networks to form a team with the desired skills and a strong team cohesion. As follow-up work, Anagnostopoulos et al [1] studies forming teams to accommodate a sequence of tasks arriving in an online fashion.

D. Novelty and Significance Relative to the State of the Art. In the context of social networks, our work will establish effective algorithms and tools for the performance prediction and optimization of teams along with explanations.

Fig. 1: Research Tasks Overview.

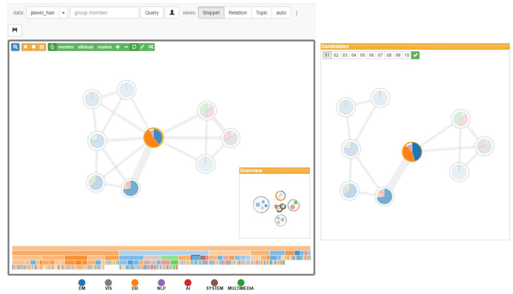


Fig. 2: Our visualization system TeamOpt.

We have also developed a visual system for team optimizations [6,2] (see Fig. 2). To our best knowledge, this is the first comprehensive effort that integrates interactive visualization mechanisms, machine learning models and advanced network analysis algorithms for optimizing teams. The preliminary results, including publications and prototype systems, are available at `team-net-work.org`.

2 Our Work

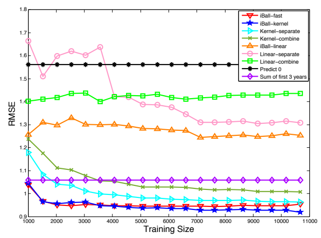
A. Proposed Solutions. We introduce our proposed solutions for performance prediction and optimization. The explanation model is described in our work in progress. For long-term team performance prediction, we propose a joint predictive model, iBall [4], which is essentially a network of prediction/regression models. The key of the model is to leverage the correlations among teams in different domains or departments. The formulation is as follows:

$$\min_{\mathbf{w}^{(\mathbf{i})}, i=1,\ldots,n_d} \sum_{i=1}^{n_d} \mathcal{L}[f(\mathbf{X}^{(\mathbf{i})}, \mathbf{w}^{(\mathbf{i})}), \mathbf{Y}^{(\mathbf{i})}] + \theta \sum_{i=1}^{n_d} \sum_{j=1}^{n_d} \mathbf{A}_{\mathbf{ij}} g(\mathbf{w}^{(\mathbf{i})}, \mathbf{w}^{(\mathbf{j})}) + \lambda \sum_{i=1}^{n_d} \Omega(\mathbf{w}^{(\mathbf{i})})$$

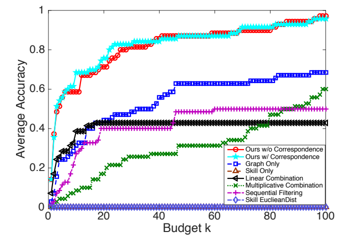
Omitting the notation definitions, we want to mention that the key is the second term where $\mathbf{A}_{ij}$ models the relationship between the i-th and j-th domains. The predictive models for the two domains are encouraged to be similar if $\mathbf{A}_{ij}$ is of large value. For the performance trajectory prediction, we propose iPath [8] to leverage the smoothness between adjacent time steps and omit the details here.

For team performance optimization, we start with *team replacement* [5] and identify two design objectives for a good team replacement strategy, i.e., *skill matching* and *structural matching*. The new member should bring a similar skill sets and have a similar network structure as the person leaving. Such design calls for a similarity measure between two individuals in the context of the team itself. Mathematically, we adopt the random walk based graph kernel to simultaneously capture both the skill and structural matching. To speed up the computation, we propose scalable algorithms by carefully designing the pruning strategies and metric analyses for exploring the smoothness and correspondences between the existing and the new teams. The algorithms designed for team replacement can be adapted to other team enhancement operations, e.g., *team expansion*. The details are described in our journal paper [7].

B. Datasets. We use a variety of real-world team datasets for evaluations. (1) Entertainment teams: we use a movie dataset and treat each film crew as a team. The team network is constructed based on co-starring relationship and the



(a) Performance prediction error.



(b) Average accuracy of team replacement.

Fig. 3: Some results of team performance prediction and optimization.

performance is measured based on box office and IMDb ratings. (2) Research teams: we use DBLP data where we have meta-data including title, authors, venues, year of the research papers. (3) Sports teams: we use NBA data where we have the information of the players and teams season by season.

C. Evaluations and Results. We use cross-validation techniques to evaluate RMSE (root mean squared error), precision/recall for performance prediction and optimization. The efficiency of the proposed algorithms is evaluated by their running time and scalability. We also perform user studies for both the quantitative and qualitative evaluations of team optimization.

D. Work in Progress. Our ongoing work consists of three directions. (1) Predictive Optimization: if a team is predicted to be under-performing in the future, we recommend a team optimization action that can prevent the performance degradation. (2) Explanation for Prediction: provide an explanation for the reason why a team is predicted to perform better or worse in the future. (3) Explanation for Optimization: provide an explanation for why we recommend a certain candidate for replacement, expansion or shrinkage.

References

1. Aris Anagnostopoulos, Luca Becchetti, Carlos Castillo, Aristides Gionis, and Stefano Leonardi. Online team formation in social networks. In *WWW*, 2012.
2. Nan Cao, Yu-Ru Lin, Liangyue Li, and HangHang Tong. g-miner: Interactive visual group mining on multivariate graphs. In *CHI*, 2015.
3. Theodoros Lappas, Kun Liu, and Evimaria Terzi. Finding a team of experts in social networks. In *KDD*, pages 467–476, 2009.
4. Liangyue Li and Hanghang Tong. The child is father of the man: Foresee the success at the early stage. In *KDD*, pages 655–664. ACM, 2015.
5. Liangyue Li, Hanghang Tong, Nan Cao, Kate Ehrlich, Yu-Ru Lin, and Norbou Buchler. Replacing the irreplaceable: Fast algorithms for team member recommendation. In *WWW*, pages 636–646. ACM, 2015.
6. Liangyue Li, Hanghang Tong, Nan Cao, Kate Ehrlich, Yu-Ru Lin, and Norbou Buchler. Teamopt: Interactive team optimization in big networks. In *CIKM*, 2016.
7. Liangyue Li, Hanghang Tong, Nan Cao, Kate Ehrlich, Yu-Ru Lin, and Norbou Buchler. Enhancing team composition in professional networks: Problem definitions and fast solutions. *TKDE*, 29(3):613–626, 2017.
8. Liangyue Li, Hanghang Tong, Jie Tang, and Wei Fan. *iPath: Forecasting the pathway to impact*, pages 468–476. 2016.
9. Staša Milojević. Principles of scientific research team formation and evolution. *Proceedings of the National Academy of Sciences*, 111(11):3984–3989, 2014.