Short Paper

A Neural Approach to Forming Coherent Teams in Collaboration Networks

Radin Hamidi Rad*
Ryerson University, Canada
radin@ryerson.ca

Shirin Seyedsalehi*
Ryerson University, Canada
shirin.seyedsalehi@ryerson.ca

Mehdi Kargar
Ryerson University, Canada
kargar@ryerson.ca

Morteza Zihayat
Ryerson University, Canada
mzihayat@ryerson.ca

Ebrahim Bagheri
Ryerson University, Canada
bagheri@ryerson.ca

ABSTRACT

We study *team formation* whose goal is to form a team of experts who collectively cover a set of desirable skills. This problem has mainly been addressed either through graph search techniques, which look for subgraphs that satisfy a set of skill requirements, or through neural architectures that learn a mapping from the skill space to the expert space. An exact graph-based solution to this problem is intractable and its heuristic variants are only able to identify sub-optimal solutions. On the other hand, neural architecture-based solutions treat experts individually without concern for team dynamics. In this paper, we address the task of *forming coherent teams* and propose a neural approach that maximizes the likelihood of successful collaboration among team members while maximizing the coverage of the required skills by the team. Our extensive experiments show that the proposed approach outperforms the state-of-the-art methods in terms of both ranking and quality metrics.

1 INTRODUCTION

With an increased demand for interdisciplinary skill sets in both academia and industry, finding a team of experts that can effectively work together plays an important role in the success of a project. Given a project and a network of experts, finding the team of experts is referred to as *team formation*. Early studies on team formation focus on graph based techniques to find a sub-graph of experts in which they collectively cover the required skills to accomplish the project. Such methods mainly optimize a monotone objective function by considering specific constraints such as past collaboration [12, 14] and personnel costs [13]. For example, Lappas et al. [10] proposed a minimum cost spanning tree objective function to minimize the communication costs while covering the required skills by the project. Kargar et al. [5, 7] solve this problem by minimizing the sum of edge weights of the extracted subgraph, where the edge weights represent the communication costs among experts. The main drawback of such techniques is that given the computationally complex nature of the problem, they only explore a portion of the input expert network, thus result in sub-optimal teams.

To address the limitations of graph-based techniques, recent works leverage neural networks to form a team of experts. In this context, the problem of team formation is formulated as building a neural model that effectively learns a mapping from skills space to the experts space. Given a set of required skills to complete a project, Spanieza et al. [14] designed an autoencoder architecture to find the best experts for each skill. However, since distribution of the skills over the experts is sparse, the proposed approach is prone to overfitting due to the non-variational nature of the proposed approach. Following on this, Rad et al. [12] proposed a variational Bayesian architecture to effectively address the sparsity problem. Techniques such as Spanieza et al. and Rad et al. mainly retrieve a ranked list of experts in which the top-k are selected as the team. This is limited due to the fact that experts are treated individually and the willingness for collaborations between experts is overlooked. Furthermore, the ranked list of experts do not necessarily present the maximum coverage of the required skills as the top ranked experts might hold overlapping expertise.

Bearing the above challenges in mind, we introduce the task of *forming coherent teams*. We propose to form a team considering two main conditions: 1) the team members collectively maximize the coverage of the input required skills, and 2) the members of the team presents a high willingness for collaboration. In summary, the contributions of the paper are listed as follows:

(1) we formulate the problem of *forming coherent teams*. The goal is to find a team whose members collectively cover the required skills and show successful collaboration in the past;
(2) We propose a novel loss function for learning to rank experts that i) minimizes the error between predicted teams and the ground truth team thus maximizes skill coverage, and ii) ensure that members of the formed team have collaborated in the past, thus the team is coherent, and;
(3) we conduct extensive experiments on a real-world dataset and show that the proposed method yields more effective results on ranking and quality metrics compared to the state-of-the-art methods.

2 COHERENT TEAM FORMATION

Given a set of skills $S = \{s_1, s_2, ..., s_m\}$, and a set of experts $E = \{e_1, e_2, ..., e_n\}$ where each expert e_i can have a limited set of skills $S_{e_i} \in S$, the goal of the team formation task is to identify a subset $E^t \subset E$ such that $\bigcup_{e^t \in E^t} S_{e^t} = S$. In other words, the aim is to form a team of experts that collectively cover a specific set of required skills S.

In an ideal scenario, the members of the formed team E^t not only cover the required set of skills but also exhibit characteristics of a successful team such as having the potential for working together, which could be measured by the number of past successful collaborations. The objective of our work in this paper is not only

*These authors contributed equally.

to form a team that maximally covers a set of required skills, but also to form the team in such a way that the team members have the likelihood of effectively working with each other. We define a *coherent team* as one where the team members are likely to work well together.

This leads us to defining the task of *coherent team formation* whose objective is to retrieve a set of experts such that (1) the experts collectively cover a set of required skills and (2) the team members have successful collaboration history in the past. Let us define a Collaboration Matrix C , which is a symmetric $n \times n$ matrix, where n is the total number of experts in E. Each element c_{ij} in the matrix is set to 1 if $e_i, e_j \in E$ have previously collaborated with each other, and 0 otherwise. On the basis of the Collaboration matrix, the coherent team formation task is formally defined as identifying a set of experts $E^C = \{e^c_1, e^c{}_2, ..., e^c_k\}$ to cover a set of skills S such that (1) the team maximally covers all the skills in S defined as: $\bigcup_{e^c \in E^C} S_{e^c} = S$; and, (2) each pair of team members in E^C have collaborated with each other in the past as follows: $\forall e^c_i, e^c_j \in E^C, C_{e^c_i, e^c_j} = 1$.

3 PROPOSED APPROACH

We define the optimization problem for coherent team formation through an objective function, which optimizes two components: (1) the first component, which we refer to as *team membership component*, ensures that the retrieved set of experts have maximal overlap with the expected set of experts, and (2) the second component, referred to as *team structure component*, enforces that the retrieved set of experts have effective past collaboration with each other. As such, in our work, a coherent team is one that maximizes *team membership* for those individual experts that have the highest likelihood of possessing skills that are required within the team, and at the same time, ensures unified *team structure* by including those experts that have the highest potential to effectively work together as a part of a team. The following outlines the details of these two components .

3.1 Team Membership Component

In the Team Membership Component, the objective is to minimize the prediction error between the top-k predicted experts and the ground truth experts E^g. For every expert e^c_i in E^c, the vector is converted to a one-hot encoding representation $\hat{E} = \mathcal{G}(E^c, k)$ through a function $\mathcal{G}$ which assigns 1 for the top-k experts having the highest probabilities in E^c, and 0 otherwise. Given skill set S and the top-k predicted experts $\hat{E} = \mathcal{G}(E^c, k)$, $L_{Membership}$ is defined as the distance between $\hat{E}$ and E^g:

$$L_{Membership} = \frac{1}{T} \sum_{i=1}^{T} L(\mathcal{G}(E^c, k)), E^g_i) \qquad (1)$$

where T is the total number of teams observable in the training process. In the context of Equation 1, both the skill and expert sets are to be transformed to vector representations, and a neural network $\Phi : S \longrightarrow E$ parametrized by Θ would serve as the mapping function from the skill space to the expert space such that $E^c = \Phi(S; \Theta)$. Therefore, Equation 1 can be rewritten as follows:

$$L_{Membership} = \frac{1}{T} \sum_{i=1}^{T} L(\mathcal{G}(\Phi(S_i; \Theta), k), E^g_i) \qquad (2)$$

3.2 Team Structure Component

The Team Structure Component is responsible for making sure that the members of the team have collaborated with each other in the past. Given expert $\hat{e}_i \in \hat{E}$, we define a past collaboration score as the dot product of $\hat{E}$ to the row C_i of the Collaboration matrix which corresponds to expert i. The total past collaboration score of the retrieved set $\hat{E}$ is calculated as the normalized summation of the past collaboration scores for each expert. $L_{Structure}$ is then defined as:

$$L_{Structure} = \frac{1}{T} \sum_{i=1}^{T} [\frac{1}{K^2} \sum_{t, E[t] \neq 0} C_t . \mathcal{G}(\Phi(S_i; \Theta), k)] \qquad (3)$$

where K is the number of retrieved experts in the team, C_i is the i^{th} row of the Collaboration matrix corresponding to expert i, and T is the total number of teams observable during training.

The above two components of our objective function allow us to ensure that we retrieve experts that (1) have similar representations to the experts of interest and hence have a higher likelihood of covering the desirable set of input skills ($L_{Membership}$) and (2) have successfully collaborated with each other in the past and hence form a coherent team ($L_{Structure}$). We linearly interpolate the two optimization components such that $L_{Membership}$ is minimized and $L_{Structure}$ is maximized during the optimization:

$$\begin{aligned} L_{Coherency} &= L_{Membership} - L_{Structure} \\ &= \frac{1}{T} \Bigg(\sum_{i=1}^{T} L(\mathcal{G}(\Phi(S_i; \Theta), k), E^g_i) \\ &\quad - \sum_{i=1}^{N} [\frac{1}{K^2} \sum_{t, E[t] \neq 0} C_t . \mathcal{G}(\Phi(S_i; \Theta), k)] \Bigg) \end{aligned} \qquad (4)$$

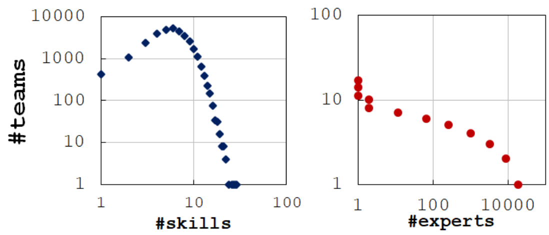

Figure 1: Distribution of articles over the skills and experts.

4 EXPERIMENTS

4.1 Neural Architecture

We employ a Variational Bayesian Neural Network [2, 4] as the architecture of the neural network Φ. The intuition behind this choice is that the distribution of the skills S over the experts E is sparse and VBNN have been shown to be able to effectively deal with sparsity [2]. Based on this architecture, we define the first component of our proposed loss function ($L_{Membership}$) as a linear interpolation of (1) the Kullback-Leiber divergence [4] between the network prediction $\hat{E}$ and the ground truth E^g as the reconstruction loss of the VBNN, and (2) the Mean Squared Error loss between the predicted experts and the ground truth E^g.

4.2 Gold Standard Dataset

Existing work on the team formation task [7, 10] have suggested that datasets such as DBLP, which consist of bibliographic information within the field of Computer Science may be a suitable gold standard dataset for team formation. These works argue that the set of authors on each paper can be seen as a team of experts

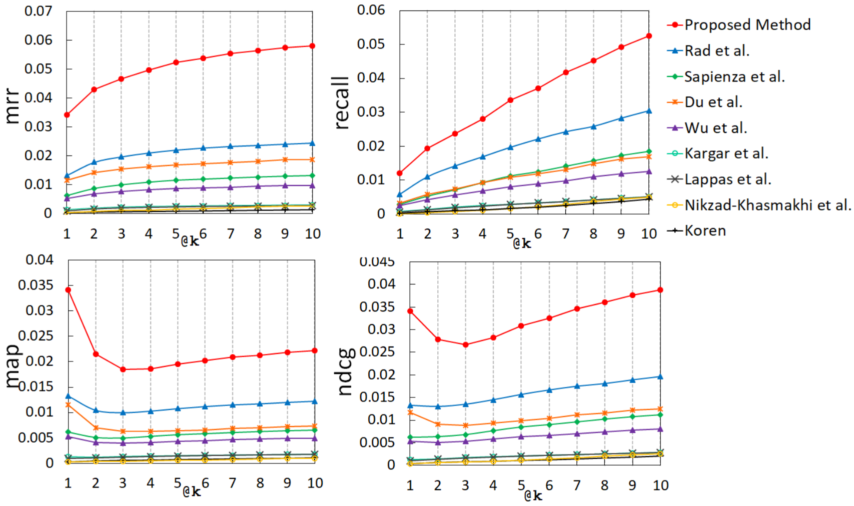

Figure 2: Comparison with baselines using ranking metrics.

who have successfully collaborated with each other to write a peer-reviewed manuscript and who collectively possessed the right set of skills to execute the research work and write the paper. We adopt the same dataset where the set of authors on each paper form a team. As suggested, we extracted the top-2000 unigrams, bigrams and trigrams with the highest tf-idf value to serve as the set of skills in our dataset, as such, each paper is also associated with a set of skills. The skill set for each author is the set of skills associated with the papers that the author has written in the past. This dataset consists of $33,002$ teams (papers) that have at least 2 authors, $1,878$ experts and $2,000$ skills. The distribution of skills over the team sizes is shown in Figure 1.

In our experiments, we adopted a 10-fold cross validation strategy. For the sake of testing, the set of skills of each paper were given as input, and a ranked list of authors of the paper were the generated as output. We note that all of our code, models, and results are publicly available online on Github website[1]. The quality of the ranked list of authors were evaluated using the following metrics.

4.3 Metrics

In order to evaluate the quality of the suggested teams by our method as well as the baselines, we adopt two sets of metrics as suggested in [6, 12]. These metrics evaluate team quality from two complimentary perspectives, namely *ranking* and *quality* metrics.

Ranking metrics are the standard metrics that evaluate the quality of a ranked list of items. In the context of team formation, earlier authors [1] have used ranking metrics to evaluate the performance of team formation where the top-k members of the ordered list of experts at a certain cut off retrieved by each method is viewed as the team. We report four ranking metrics, namely `MAP`, `NDCG`, `Recall` and `MRR`.

From the quality perspective, we measure two metrics as suggested by [15, 17], namely *team formation success* and *communication cost*. The team formation success metric measures the percentage of the desired skills that have been covered by the proposed team. In other words, team formation success shows the extent to which the formed teams were able to cover the required set of skills. Furthermore, communication cost computes the shortest path distance between pairs of experts in each formed team. An ideal team is one where all authors have already collaborated with each other and therefore, the author nodes are neighbors and hence pairwise distance between neighbors is zero. A lower communication cost is desirable that would show the members of the team had closer working relationship with each other in the past. Higher metric values for ranking and team formation success metrics and lower communication costs are desirable.

4.4 Baselines

We adopt three classes of baseline methods for team formation, namely *graph-based*, *neural-based*, and *collaborative filtering-based* methods. In the class of *graph-based* techniques, we consider the work by `Kargar et al.` [6] and `Lappas et al.` [10], which propose heuristics to traverse the expert collaboration graph in order to identify relevant subgraphs and sub-tress within the graph, respectively. Within the context of *neural-based* methods, we include the work by `Rad et al.` [12], which trains a variational Bayesian neural architecture to map skill and expert spaces for team formation, the method by `Sapienza et al.` [14] that uses an autoencoder architecture to learn associations between experts within a collaboration network, and the work by `Nikzad-Khasmakhi et al.` [11] that learns neural expert representations that allow for computing similarities between experts in order to form teams. Finally, we formulate the problem of team formation as one of a recommendation problem and adopt three strong recommendation methods, namely the work by `Wu et al.` [16] that offers a recurrent neural recommender model based on an LSTM autoregressive method, the method by `Du et al.` [3] that offers a Bayesian group ranking model, and the widely known svd++ method by `Koren` [9].

4.5 Findings

We report our findings based on the two classes of evaluation metrics, namely *ranking* and *quality* metrics.

4.5.1 Ranking Metrics. In the context of the ranking metrics, we make several observations based on the results in Figure 2:

[1]https://github.com/radinhamidi/Forming-Coherent-Teams-in-Collaboration-Networks

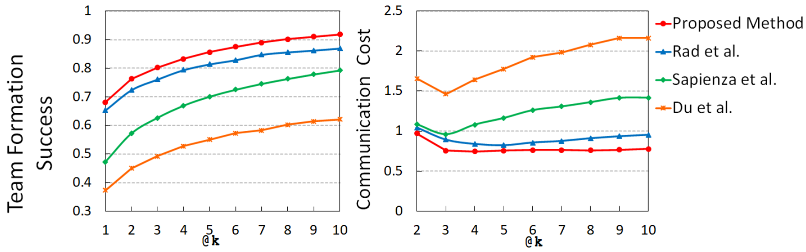

Figure 3: Comparison with strong baselines using quality metrics.

(1) within the baselines, we find that the two strongest baselines are the work by `Rad et al.` [12] and `Sapienza et al.`, which are both based on neural architectures. The additional strong baseline is the work by `Du et al.`, which considers a Bayesian group recommendation model, which considers group dynamics when forming teams. Based on this observation, it seems that team formation models that adopt a neural architecture are more effective, which could be due to their ability to effectively work with sparse network structures such as that of a collaboration network [12].
(2) In contrast, graph-based methods, such as the state of the art work by `Kargar et al.`, are not able to show competitive performance compared to neural method primarily since they operate based on local heuristics that explore graph subsets that may not lead to global optimal teams. These methods are designed in such a way since graph traversal is expensive and identifying specific subgraphs has shown to be an NP-hard problem [8].
(3) We also find that our proposed approach shows a substantially higher performance compared to all other baselines on the ranking metrics. The main distinguishing aspect of our work is its focus on ensuring that experts within the same team have had past collaboration history. This not only will make sure that the experts in one team are more coherent and have a higher likelihood of forming a successful team, but it will also increase the probability of the retrieved experts being relevant for the given input skill sets. This is because experts who collaborated with each other in the past are likely to focus on synergistic topical areas, thus maximizing expert past collaboration history within each team can indirectly also ensure higher relevance of the experts for the input skills.

4.5.2 Qualitative Metrics. Now in order to report the qualitative metrics and to ensure the results are observable clearly in the figure, we select the top-3 best performing baselines based on the ranking metrics and compare them with our proposed approach in Figure 3. We interpret the findings of the quality metrics by analyzing the behavior of the models from the perspectives of team formation success and communication cost in tandem. Ideally, when the expected team is formed, team formation success is maximized and communication cost is minimized. Figure 3 shows that our proposed approach is increasing team formation success but at the same time reducing communication cost. This means that even in cases when our proposed approach is not able to identify the expected team, it is not selecting experts who would only cover the expected skills, but rather it retrieves experts with relevant expertise and with past effective collaboration history. The results show that our proposed approach is able to outperform the other strong baselines on both team formation success and communication cost metrics.

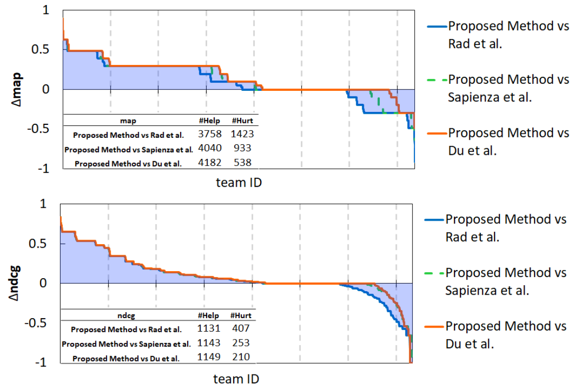

Figure 4: Help-Hurt diagram comparing proposed method versus top-three competitors.

4.5.3 Finer-grained Analysis. We present a more in-depth analysis of the model performance using help-hurt diagrams in Figure 4 based on `MAP` and `NDCG` metrics (the other two metrics are comparable and omitted due to space limitation). In this figure, each point on the x-axis represents an instance of a team and the y-axis shows the percentage of the difference between the `MAP` or `NDCG` of our proposed approach compared that of a baseline. A larger area under the curve on the top-left side of each diagram compared to the area under the curve on the bottom-right shows that our approach has been able to improve (1) a larger number of teams, and (2) to a greater extent compared to the baseline methods. The figure shows that the improvements reported for our method are observable consistently on a large number of teams and over the strongest baselines.

5 CONCLUDING REMARKS

This paper offers a novel neural approach that maximizes the likelihood of successful collaboration between team members as well as the coverage of required skills by them. The proposed approach addresses the optimization problem with two components: 1) minimizing the prediction error between the top-k predicted experts and the ground truth experts in order to maximize skill coverage, and 2) taking past collaborations between team members into account so to maximize team coherency. Through a range of experiments, we show that our model achieves superior results in terms of both ranking and quality metrics. The improvements are as high as 3 folds in terms of `MRR` and 2 folds based on `NDCG` compared to our best team formation baseline at a team size of 10. The results also indicate that the teams formed by our method are more collaborative and have higher skill coverage.

REFERENCES

[1] Krisztian Balog, Yi Fang, Maarten De Rijke, Pavel Serdyukov, and Luo Si. 2012. Expertise retrieval. *Foundations and Trends in Information Retrieval* 6, 2–3 (2012), 127–256.

[2] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra. 2015. Weight Uncertainty in Neural Networks. (2015). http://arxiv.org/abs/1505.05424

[3] Yulu Du, Xiangwu Meng, Yujie Zhang, and Pengtao Lv. 2020. GERF: A Group Event Recommendation Framework Based on Learning-to-Rank. *IEEE Trans. Knowl. Data Eng.* 32, 4 (2020), 674–687. https://doi.org/10.1109/TKDE.2019.2893361

[4] Alex Graves. 2011. Practical Variational Inference for Neural Networks. In *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011. Proceedings of a meeting held 12-14 December 2011, Granada, Spain*, John Shawe-Taylor, Richard S. Zemel, Peter L. Bartlett, Fernando C. N. Pereira, and Kilian Q. Weinberger (Eds.). 2348–2356. https://proceedings.neurips.cc/paper/2011/hash/7eb3c8be3d411e8ebfab08eba5f49632-Abstract.html

[5] Mehdi Kargar and Aijun An. 2011. Discovering top-k teams of experts with/without a leader in social networks. In *Proceedings of the 20th ACM Conference on Information and Knowledge Management, CIKM 2011, Glasgow, United Kingdom, October 24-28, 2011*, Craig Macdonald, Iadh Ounis, and Ian Ruthven (Eds.). ACM, 985–994. https://doi.org/10.1145/2063576.2063718

[6] Mehdi Kargar, Lukasz Golab, Divesh Srivastava, Jaroslaw Szlichta, and Morteza Zihayat. 2020. Effective Keyword Search over Weighted Graphs. *IEEE Transactions on Knowledge and Data Engineering* (2020). https://doi.org/10.1109/TKDE.2020.2985376

[7] Mehdi Kargar, Morteza Zihayat, and Aijun An. 2013. *Finding Affordable and Collaborative Teams from a Network of Experts*. 587–595. https://doi.org/10.1137/1.9781611972832.65

[8] Richard M. Karp. 1972. Reducibility Among Combinatorial Problems. In *Proceedings of a symposium on the Complexity of Computer Computations*. Plenum Press, New York, 85–103.

[9] Yehuda Koren. 2009. Collaborative filtering with temporal dynamics. 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, 447–456. https://doi.org/10.1145/1557019.1557072

[10] 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

[11] Narjes Nikzad-Khasmakhi, Mohammad Ali Balafar, Mohammad-Reza Feizi-Derakhshi, and Cina Motamed. 2021. ExEm: Expert embedding using dominating set theory with deep learning approaches. *Expert Syst. Appl.* 177 (2021), 114913. https://doi.org/10.1016/j.eswa.2021.114913

[12] Radin Hamidi Rad, Ebrahim Bagheri, Mehdi Kargar, Divesh Srivastava, and Jaroslaw Szlichta. 2021. Retrieving Skill-Based Teams from Collaboration Networks. In *SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021*, Fernando Diaz, Chirag Shah, Torsten Suel, Pablo Castells, Rosie Jones, and Tetsuya Sakai (Eds.). ACM, 2015–2019. https://doi.org/10.1145/3404835.3463105

[13] H. Rahman, S. Roy, S. Thirumuruganathan, S. Amer-Yahia, and Das G. 2019. Optimized Group Formation for Solving Collaborative Tasks. *VLDBJ* 28, 1 (2019), 1–23.

[14] 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

[15] Kalyani Selvarajah, Pooya Moradian Zadeh, Mehdi Kargar, and Ziad Kobti. 2019. Identifying a team of experts in social networks using a cultural algorithm. *Procedia Computer Science* 151 (2019), 477–484.

[16] Chao-Yuan Wu, Amr Ahmed, Alex Beutel, Alexander J. Smola, and How Jing. 2017. Recurrent Recommender Networks. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, WSDM 2017, Cambridge, United Kingdom, February 6-10, 2017*, Maarten de Rijke, Milad Shokouhi, Andrew Tomkins, and Min Zhang (Eds.). ACM, 495–503. https://doi.org/10.1145/3018661.3018689

[17] Morteza Zihayat, Aijun An, Lukasz Golab, Mehdi Kargar, and Jaroslaw Szlichta. 2018. Effective Team Formation in Expert Networks.. In *Alberto Mendelzon International Workshop on Foundations of Data Management (AMW)*.