How the Minimal Degree of a Social Graph Affects the Efficiency of an Organization

Ilya Samonenko\textsuperscript{1(✉)}, Tamara Voznesenskaya\textsuperscript{1}, and Rostislav Yavorskiy\textsuperscript{2}

\textsuperscript{1} Higher School of Economics, Moscow 101000, Russia
\{isamonenko, tvoznesenskaya\}@hse.ru
\textsuperscript{2} Tomsk Polytechnic University, Tomsk 634050, Russia
ryavorsky@tpu.ru

Abstract. This paper continues our research on how the communication structure of an organization affects its efficacy. We assume that the organization consists of several types of professionals (researchers, engineers, testers etc.) and each project team must include at least one professional of each type. We study a dynamic simulation model, which defines a team creation process, based on the social graph of the organization. The greatest lower bound of the average utilization rate is found under (is based on) the assumption that each node in the social graph has a degree of at least $\delta$. This theoretical result is supported by a series of computer simulations.

Keywords: Professional network · Self-organizing teams · Team model · Social graph · Utilization rate · Simulation modelling

1 Introduction

The topic of team modelling and performance analysis generates regular interest among researchers, see [5]. Analysis and improvement of employee utilization are widely covered in business administration literature, but papers, which use mathematical modelling for that are quite rare. In our previous paper [8] a literature overview of team modelling approaches and the utilization rate analysis was presented. An important insight of the literature study is that a meaningful modelling approach should take into account the skill profiles of the employees, their roles and communications.

A simple example which we have in mind is a group of researchers and engineers working at a rather big academic research center, see [7]. Some of the group members know each other, and these connections are used to form project teams when information about a new grant, a call for papers or other opportunity occurs. We use a simulation modeling approach to study the process of information distribution and team forming in response to that new opportunity. We are especially interested in the efficiency of such organization as the average number of well-formed project teams per opportunity. The same model type
could also be applied to other cases, where the project teams are self-organized by the employees on the basis of existing social connections.

An average utilization rate of a group is the proportion of specialists involved in projects to the total number of specialists. We are interested in studying this important aspect of the organization performance. In particular, we want to analyze how the minimal degree of the social graph affects the utilization rate of an organization.

The horizontal connection between members of an organization is an important factor. Many researchers consider different characteristics of these connections in team modeling. For example, team sports performance analysis (see e.g. [4]) pay attention not only to social graph topology but also to the complexity of interpersonal interactions of players. In [2], a model of multi-agent team performance for the RoboCup Rescue domain is considered. The authors simulated the behavior of small teams of 2–3 robots, working on civilian rescue tasks to study the synergy among them. The different characteristics of intra-team interaction are considered in [3]. Computer simulation to support the systematic design of organization was also studied in [1].

The paper is organized as follows. In Sect. 2, we provide formal definitions, then different restrictions and conditions on the model are considered. The main part of the paper consists of theorems in Sect. 3, which formalize some properties of the model with respect to these conditions. Section 4 provides the simulation analysis and Sect. 4 concludes.

2 Definitions

We consider here a simplified version of a more general simulation model developed for analysis of team performance in a research centre [7,8]. Input parameters of the model are the following:

- \( P = A \sqcup B \) is a finite set of group members (employees, professionals) of two categories: type \( A \) (researchers) and type \( B \) (engineers), where \( |A| = n_1 \) and \( |B| = n_2 \);
- \( G \subseteq P \times P \) is an undirected social graph of an organization, so \( G(p_1, p_2) \) indicates that employees \( p_1, p_2 \in P \) have established relations enough to invite each other to a new project.

In general, we assume that for a successful project one needs a team, which includes \( k_1 \) researchers and \( k_2 \) engineers. The focus of this paper is where \( k_1 = k_2 = 1 \).

The simulation process is executed in the following way (it is described in detail in [6]):

- **Team initialization.** When an employee decides to initiate a new project, a new team is created, which consists of only the initiator.
- **Invitation process.** Each team member sends an invitation to join his project to each of his contacts in the social graph of the organization. An employee always accepts the invitation if their competencies are needed for this team and they have time available. Each employee can only join one project.

- **Team finalization:**
  - **Success.** A project team is formed where $k_1$ researchers and $k_2$ engineers have joined it.
  - **Failure.** It may be that the combined skills of members are not sufficient and there is no one else available with the required skill set. Then the team formation process is stopped and members who already joined become available again.

We assume that the receiver of an invitation makes the decision and responds immediately (from a practical point of view this means that all the communications are fast and decisions are instant in comparison to the duration of projects).

The simulation process sequentially runs the rules above for all the agents (employees) until the formation of new teams is impossible.

For a given graph $G$, let $w$ denote the number of employees who are finally involved in a project team. Then the utilization rate is defined as a fraction:

$$\rho(G) = \frac{w}{n_1 + n_2}.$$ 

Note that the team building is a non-deterministic process, so value $\rho(G)$ will differ for different simulations, see detailed [6].

### 3 Upper and Lower Bounds for the Utilization Rate

We assume that each project needs exactly one researcher and one engineer.

The upper bound for $\rho(G)$ is rather straightforward. Since any project needs exactly one researcher and one engineer, if $n_1 \neq n_2$ then at least $|n_1 - n_2|$ employees will always be idle. So,

$$\rho(G) \leq \frac{2 \min(n_1, n_2)}{n_1 + n_2}.$$ 

This upper bound is the lowest, see [8] for the detailed proof.

Next we prove the lower bound for the utilization rate for the model defined above under the assumption that each node in the social graph has a degree greater than $\delta$. This generalizes our previous results, see Fig. 1 and [8], which correspond to case $\delta = 1$.

For $k_1 = k_2 = 1$ connections between employees of the same category do not affect the team formation. So, we assume that the social graph is bipartite.
Theorem 1 (lower bound for the utilization rate). Let $G = (A \sqcup B, E)$ denote a bipartite social graph of an organization, $|A| = n_1, |B| = n_2$. Assume that each node in $G$ has a degree equal to or greater than $\delta$, then:

$$\rho(G) \geq \frac{2\delta}{n_1 + n_2}$$

and this lower bound is the greatest.

Proof. Under the assumption of the two skills model, every project team consists of exactly two members, a researcher (from $A$) and an engineer (from $B$).

First, we explicitly construct a social graph and a simulation run for which

$$\rho(G) \geq \frac{2\delta}{n_1 + n_2}.$$ 

Then, it will be shown that for any graph and simulation, $\rho(G)$ cannot be lower.

Consider Fig. 2. It is a bipartite graph with $n_1$ nodes of type $A$ and $n_2$ nodes of type $B$. The nodes are divided into four groups:

- group I consists of $(n_2 - \delta)$ nodes of type $B$;
- group II consists of $\delta$ nodes of type $A$;
- group III consists of $\delta$ nodes of type $B$;
- group IV consists of $(n_1 - \delta)$ nodes of type $A$.

Each node from group I is linked to each node from group II. Similarly, each node from group II is linked to each node from group III, each node from group III is linked to each node from group IV. It is clear that the degree of nodes in groups I and IV are equal to $\delta$, and the degree of nodes in groups II and III are equal $n_2$ and $n_1$ correspondingly, which is greater that $\delta$.

For this structure of the social graph, it may happen that nodes from groups II and III form $\delta$ pairs and all the remaining nodes will be idle. For this run the utilization rate equals

$$\rho = \frac{2\delta}{n_1 + n_2}.$$ 

Let’s now show that this is the strict lower bound and any simulation makes at least $\delta$ teams. This could be shown by induction on $\delta$. Case $\delta = 1$ is trivial,
How the Minimal Degree of a Social Graph Affects the Efficiency

Fig. 2. The worst case social graph of a two-skilled organization: group I consists of \((n_2 - \delta)\) nodes, groups II and III contain \(\delta\) nodes, group IV consists of \((n_1 - \delta)\) nodes since each node has at least one connection, at least one pair will be formed. Consider any bipartite graph such that each degree of every node is equal or greater than \(\delta\). Take an arbitrary node of any type, and any one from its connections to form a team. In the subgraph, all the nodes have a degree \(\delta - 1\) or greater, so at least \(\delta - 1\) teams will be formed according to the induction hypothesis.

From a practical point of view this result looks quite frustrating. For example, assume \(n_1 = n_2 = 15\) and \(\delta = 3\), this has a balanced group of 30 workers in which everyone is linked to at least three potential collaborators. According to the reason above it may happen that only 6 teams are formed and the remaining 18 members are forced to be idle. The group utilization rate for this example is \(\rho = 40\%\), which is far from intuitively expected \(\rho = 100\%\), which is potentially reachable when the social graph is more dense.

One of the useful implications is the fact that measures aimed at establishing more connections in a team do not always guarantee a substantial growth of the average utilization rate.

4 Simulation Analysis

In this section we describe in detail the simulation process because its parameters strongly influence the observed results.

We implement function GetRandomGraph\((n, \delta)\) that creates a random bipartite graph in the following way. First, for given \(n\) and \(\delta\) graph \(G = (A \sqcup B, \emptyset)\) with \(n\) vertices and no edges is created, where \(A\) stands for a set of vertices corresponding to researchers, \(B\) is the set of vertices corresponding to engineers, and \(|A| = |B| = n/2\). Then, each vertex \(a \in A\) is connected to \(\delta\) different vertices from \(B\) with equal probability, and vice versa. Each vertex \(b \in B\) is connected to \(\delta\) different vertices from \(A\) with equal probability. It may happen that some
edges are created twice, so we remove the duplicates. Such a procedure guarantees that every researcher is connected with no less than $\delta$ engineers, and every engineer is connected with no less than $\delta$ researchers. As a result we have a random bipartite graph

$$G = (A \sqcup B, E) = \text{GetRandomGraph}(n, \delta),$$

with the degree of each vertex not less than $\delta$.

The first series of experiments was aimed at finding dependencies between guaranteed minimal node degree $\delta$ and utilization rate $\rho(\text{GetRandomGraph}(n, \delta))$ for different values of $n = 20, 30, 40$ and 50, see Fig. 3.

The second series of experiments was aimed at finding dependencies between the size of graph $n$ and the utilization rate $\rho(\text{GetRandomGraph}(n, \delta))$ for different values of $\delta = 1, 2, \ldots, 6$, see Fig. 4.

For each fixed value of $n$ and $\delta$, $T = 500$, graphs $G_i(n, \delta)$ where $i = 1, \ldots, T$, were generated randomly. Then, for each graph $G_i(n, \delta)$ we have run $N = 200$ simulations $\rho_{sj}(G_i(n, \delta))$, $j = 1, \ldots, N$, and compute the following characteristics.

**Fig. 3.** Dependencies between $\delta$ and $\rho$ for $n = 20, 30, 40$ and 50.
How the Minimal Degree of a Social Graph Affects the Efficiency

Fig. 4. Dependencies between \( n \) and \( \rho \) for \( \delta = 1, 2, \ldots, 6 \)

\[
\text{Average of } \rho(n, \delta) = \frac{1}{T} \sum_{i=1}^{T} \frac{1}{N} \sum_{j=1}^{N} \rho_{sj}(G_i(n, \delta)),
\]

\[
\text{Average of } \rho_{\text{max}}(n, \delta) = \frac{1}{T} \sum_{i=1}^{T} \max_{j=1..N} \rho_{sj}(G_i(n, \delta)),
\]

\[
\text{Average of } \rho_{\text{min}}(n, \delta) = \frac{1}{T} \sum_{i=1}^{T} \min_{j=1..N} \rho_{sj}(G_i(n, \delta)),
\]

\[
\text{Maximum of } \rho_{\text{max}}(n, \delta) = \max_{i=1..T} \max_{j=1..N} \rho_{sj}(G_i(n, \delta)),
\]

\[
\text{Minimum of } \rho_{\text{min}}(n, \delta) = \min_{i=1..T} \min_{j=1..N} \rho_{sj}(G_i(n, \delta)).
\]
As a result, we found the following.

1. Typically, random simulation results are far better than the worst possible bound established, according to our theorem. For $\delta = 1$, the average value of utilization rate $\rho(n, \delta)$ is around 80%, while the theoretical minimum is below 10% for $n > 20$, see blue and dashed graphs on Fig. 3 and Fig. 4.

2. Average $\rho(n, \delta)$ increases with $\delta$, which complies with the intuition that adding more connections increases utilization, see Fig. 3. The graphs show that for $\delta = 3$ (each employee knows at least three colleagues from the other category), the average utilization is almost certainly greater than 80%.

3. For a fixed $\delta \in \{3, 4, 5, 6\}$ average the utilization $\rho(n, \delta)$ is rather stable and does not depend on $n$, see Fig. 4.

4. At $\delta \geq 3$, the average value is about 90%, and with the increase of $\delta$, does not grow significantly. Therefore, for a high probability of getting on a team, it is enough for an engineer or researcher to have three contacts. This means a simple practical recommendation: build a team so that each specialist has at least three contacts with specialists with other skills.

5 Conclusion

In this paper, an analysis of the utilization rate for an organization with a guaranteed minimum degree of a bipartite social graph has been provided. We generalized our previous results [8] for the greatest lower bound of the utilization rate. Simulation analysis gave us some ideas and possible directions for future research. In particular, we plan to consider different distributions of node degrees for this graph generation algorithm. Besides, it is interesting to study the generalization of this model with a larger number of skills and the evolving topology of social graphs.

References

1. Jin, Y., Levitt, R.E., Kunz, J.C., Christiansen, T.R.: The virtual design team: a computer simulation framework for studying organizational aspects of concurrent design. Simulation 64(3), 160–174 (1995)
2. Liemheetcharat, S., Veloso, M.: Weighted synergy graphs for effective team formation with heterogeneous ad hoc agents. Artif. Intell. 208, 41–65 (2014)
3. Peltomäki, M., Alava, M.: Correlations in bipartite collaboration networks. J. Stat. Mech. Theory Experiment 2006 (2005). https://doi.org/10.1088/1742-5468/2006/01/P01010
4. Ribeiro, J., Silva, P., Duarte, R., Davids, K., Garganta, J.: Team sports performance analysed through the lens of social network theory: implications for research and practice. Sports Med. 47, 1–8 (2017)
5. Salas, E., Goodwin, G.F., Burke, C.S.: Team Effectiveness in Complex Organizations: Cross-disciplinary Perspectives and Approaches. Routledge, Milton Park (2008)
6. Samonenko, I., Voznesenskaya, T.: The influence of self-organizing teams on the structure of the social graph. In: TMPA: International Conference on Software Testing, Machine Learning and Complex Process Analysis. Springer (2019, to appear)

7. Voznesenskaya, T.V., Krasnov, F.V., Yavorsky, R.E., Chesnokova, P.V.: Modeling self-organizing teams in a research environment. Bus. Inform. 13, 7–17 (2019). https://doi.org/10.17323/1998-0663.2019.2.7.17

8. Yavorskiy, R., Voznesenskaya, T., Samonenko, I.: Effect of social graph structure on the utilization rate in a flat organization. In: AIST 2019. CCIS, vol. 1086, pp. 214–224. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-39575-9_22