Modeling Team Formation in Self-assembling Software Development Teams

Mehdi Farhangian, Martin K.Purvis,
Maryam Purvis, Bastin Tony Roy Savarimuthu
Information Science Department, University of Otago, Dunedin, New Zealand
{mehdi.farhangian, martin.purvis, maryam.purvis,tony.savarimuthu}@otago.ac.nz

ABSTRACT

Many contemporary project teams are self-assembling, with potential team members operating as individual agents that self-select their own teams. Some examples include software development teams, crowdsourcing platforms, and the formation of scientific collaborative teams. In many such cases, team formation is significantly influenced by the makeup of participants' personalities and temperaments (even without considering the technical skills possessed by individuals). In this paper, we develop a model to help explain team formation processes and predict future team compositions by considering these personality aspects. We have used agent-based modeling to test a number of hypotheses on the team-formation mechanism with respect to a specific context by comparing the model's operation with data drawn from the Python Enhancement Proposal (PEP) process, an open-source software development process. In PEP operations, developers are free to select their fellow team members. We predict the future team composition of self-assembly teams by first inferring potential teammates' MBTI personalities based on analyzing their written texts expressed on social-networking sites. Once the personalities of PEP developers were identified, we simulated the team-formation process using agent-based simulation. The results were analyzed using factor analysis to examine the contribution of our hypotheses in the prediction of future team-assembly. The results indicated that a combination of four personal characteristics (knowledge of previous team performance, previous familiarity with people involved in the new team, and the degrees to which an agent is an MBTI perceiving personality and an MBTI feeling personality) improves the accuracy of the team composition prediction.

Keywords

Multi-agent Systems, Social Simulation, Personality, Team Formation, Self-assembling Teams.

1. INTRODUCTION

In recent years, the importance of focusing on teams of people rather than individuals in a team is widely being acknowledged. Today a vast proportion of teams are classified as project teams, since they are ad-hoc groups of distributed collaborators. Project teams are time-bound and are often assembled one-off and produce one-time outputs [1]. Some examples include crowdsourcing platforms, scientific collaboration teams, and open-source software development teams. In spite of their importance, there is little understanding about how these teams are assembled [2].

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9{13, 2016, Singapore.*

In this paper, we posit that personality types of individuals play an important role in the success of team work, and we consider them explicitly in the team-assembly process. Thus a model is developed on the basis of theoretical and empirical literature on personalities and team behavior. This conceptual model is then implemented as an agent-based computer simulation consisting of simple rules.

2. THE MODEL

In our model agents are of two types: requestors and contributors. Requestors send invitations to join their team to chosen contributors. Contributors choose among invitations received to join a team. The model has six factors hypothesized as important characteristics of the team assembly process: 1) the effect of familiarity of other team members on teammate selection, 2) the effect of past performance on the teammate selection, 3) the effect of the MBTI thinking-feeling personality dimension on familiarity, 4) the effect of the MBTI intuitive-sensing personality dimension on assessing past performance, 5) the effect of the MBTI extraverted-introverted personality dimension on familiarity likelihood, and 6) the effect of the MBTI perceiving-judging personality dimension on team-changing behavior.

3. DATA SET

We have chosen a specific application domain and investigated a real case study by extracting data from the Python Enhancement Proposal (PEP) process. Since the MBTI personalities of PEPs developers are not accessible, we have used techniques similar to [3] to determine the personality of people from their written texts. In this connection we developed a model to relate linguistic styles and personality.

Our method for extracting insights about the relationship between Python developers and their personality consisted of 3 main steps: 1) gathering the data; 2) finding relationships between personality and text usage; and 3) establishing the personalities of PEPs developers from the relationships found in step 2. These three steps are explained below.

Firstly we gathered texts written by users who reported their personalities in three social networking websites: Quora (393 users), Reddit (39 users), and College Confidential (185 users). In the second step, we extracted the text of the Quora users from their responses to other questions. Users' texts were analysed with the Linguistic Inquiry and Word Count (LIWC) tool [3], and the values for all the 80 LIWC dimensions were identified for each user. Secondly, after generating the value of all the variables in our Quora samples, we measured correlations between personality types and these variables. These correlations were then used to develop a computational model that can be used to determine the personalities of people from their texts. For validation, this formula was cross-tested with our Reddit and College Confidential data, and the results showed 73% and 60% accuracy, respectively. Thirdly, we applied the same formulae on the text written by Python developers involved in the creation of Python Enhancement Proposals (PEPs), which were developed in a team setting. The texts written by the team members were gathered from the PEPs' public activities on the Internet. This resulted in the identification of the personality type of the PEP developers.

4. SIMULATION AND RESULTS

By using an agent-based simulation, we investigated the hypotheses of our proposed model. The main research question is to what extent those hypotheses explain the behavioral outcome in the PEPs data. Experiments have been conducted using NetLogo [4].

In order to evaluate our model, we used a cross-validation procedure. A similar approach for validation of agent-based model can be seen in [5] and [6]. Our simulations start with 10 percent of teams; so among 78 PEP teams, the compositions of the 8 first teams are known and the model predicts the composition of the remaining 70 teams. The performance of the model is evaluated by counting the number of correctly predicted teams. The next iteration of the process occurs by selecting the first 9 PEPs teams to predict the remaining 69 teams. We continue with this process until the final step, when only the last two teams are unknown. The accuracy of the model is then the average of all of the experimental iterations.

The performance of the model is evaluated by counting the number of correctly predicted teams. We compared our simulation results with the real data from PEPs. Our model in average predicted 8.3% of the teams, correctly. When the proposed model is not employed and we ask the requestors and the contributors to randomly select each other, correctly chosen teams only occur 2.9% of the time. That shows the potential of the model to be used for prediction of compositions of self-assembled teams.

Furthermore, we test each hypothesis separately, to investigate the influence of each factor on the final results. Each of these factors either does or does not have an effect on the team formation, resulting in $2^6 (i.e. 64)$ separate conditions. Each condition was modelled as a separate experiment and was repeated 100 times. A factor analysis on the average performance of experiments reveals the contribution of each factor on the results. The results show that four factors appear to significantly affect the team formation mechanisms. These are 1) previous performance 2) teammate familiarity 3) MBTI feeling personality and 4) MBTI perceiving personality. The other factors did not have significant effects on the team formation in this particular context. In other words, including factors such as the effect of extraversion on the familiarity connections and the effect of sensing on the previous performance, though plausible from a psychological context, did not improve the accuracy of the model in connection with PEP team formation.

5. CONCLUSIONS

This paper offers a contribution to the relatively unexplored area of using simulation to study the impact of personality while forming self-assembly teams. We have developed an agent-based model to explore how different variables (e.g. personality types, familiarity of team members, and past team performance experience), impact the formation of self-assembly teams, using PEPs as a concrete example. To achieve that, we created a computational model to determine the relationship between MBTI personality and the LIWC dimensions.

Using simulation results, we found certain characteristics (i.e. previous performance of teams, familiarity of members in a team, and the impact of feeling personality dimension on the familiarity and the impact of perceiving dimension personality on the likelihood of changing teams) helped to predict future team compositions in PEPs, better than teams that are formed randomly. This highlights the potential of simulation for understanding team formation.

Note that our results are representative of PEPs, and we urge caution about over-generalizing our model for other data. With our application, we were working with a limited number of developers' teams in a particular domain. In addition our approach for predicting personality from text use also has some limitations. The LIWC tool has a bias against individuals whose first language is not English, and we did not separate non-English users and developers. Social roles, gender, age, and other demographic factors which are not covered in this study might be involved in the linguistic styles. Further experiments and validations should be performed before our correlations and results can be generalized.

6. REFERENCES

[1] S. G. Cohen, "What Makes Teams Work: Group Effectiveness Research from the Shop Floor to the Executive Suite," J. Manage., vol. 23, no. 3, pp. 239–290, Jun. 1997.

[2] N. Contractor, "Some assembly required: leveraging Web science to understand and enable team assembly.," Philos. Trans. A. Math. Phys. Eng. Sci., vol. 371, no. 1987, p. 20120385, Mar. 2013.

[3] J. Pennebaker and L. King, "Linguistic styles: language use as an individual difference.," Journal of personality and social psychology, 1999.

[4] S. Tisue and U. Wilensky, "NetLogo : Design and Implementation of a Multi-Agent Modeling Environment," in Proceedings of Agent, 2004.

[5] P. G. and R. Crutzen, "An Agent-Based Social Network Model of Binge Drinking Among Dutch Adults." JASSS, 31-Mar-2013.

[6] P. J. Giabbanelli, A. Alimadad, V. Dabbaghian, and D. T. Finegood, "Modeling the influence of social networks and environment on energy balance and obesity," J. Comput. Sci.,vol.3,no.1–2,pp.17–27,Jan.2012.