Best Practices in Running IT Hackathons Based on Paragon University Dataset

Azamat Serek
Suleyman Demirel University, Kaskelen, Kazakhstan,

Meirambek Zhaparov
Paragon International University, Phnom Penh, Cambodia,

Seong-Moo Yoo (✉)
University of Alabama in Huntsville, Huntsville, USA
yoos@uah.edu

Assem Talasbek
Suleyman Demirel University Kaskelen, Kazakhstan,

Yong Kab Kim, Min-Woo Jin
Wonkwang University, Iksan, Korea

Abstract—Hackathons are special activities, which normally last about 1-3 days with teams to present their innovative solutions to the given problems in the IT domain. In this paper, we report the Paragon IT hackathon developing a web application using Facebook API and our analysis of surveys from the hackathon participants. We list and analyze the statistics of participants as their gender, age, level of satisfaction, and the willingness of attending another hackathon. Then, we mention participants' primary motivation to participate, things they learned from the hackathon, and the main challenges. Based on that, we make some recommendations how to improve this hackathon even better. The results can be extrapolated into other hackathons, especially in Asian countries.

Keywords—Hackathon, IT hackathon, data analysis, Facebook API.

1 Introduction

Hackathons are very popular globally [1-6]. They are special activities, which normally last about 1-3 days. Teams are formed to present their innovative solutions to the given problems in the IT domain.

Paragon Hackathon was carried out on January 10-11, 2020 at Paragon International University in Cambodia. ParagonIU Hackathon was a two-day coding sprint that participants challenged to develop a web application using Facebook API. We asked questions on the hackathon to get ideas on what organizational adjustments

should be made to develop the hackathon even better. The obtained results have the potential to be extrapolated into other hackathons, particularly in Asian countries, due to the higher likelihood of the similarity of participants.

Paragon International University (once in the past known as Zaman University) was built up in 2010. Zaman University was rebranded and was formally perceived as Paragon International University by a sub-declaration of the Royal Government of Cambodia on 28 January 2019. With a promise to greatness, Paragon International University endeavors to be the main college in Cambodia as far as showing quality, research and network administrations.

In this paper, we present a strong analysis of the hackathon together with some recommendations to improve it even better. The specific contributions of the paper include:

  i.   Extrapolatory data analysis of the hackathon survey for the evaluation purpose
 ii.   Recommendations for improvement of the hackathon

This paper is organized as follows: Section 2 introduces previous work. In section 3 we mention method and materials used in this hackathon. In section 4 we explain exploratory data analysis. In section 5 we make some recommendations. Finally, section 6 makes conclusion.

2   Previous Work

Decker et al. [1] reported the result from their Think Global Hack Local (TGHL) hackathon, that connected non-profit organizations with student developers spending a weekend to address the issue that females are participating hackathons with less frequency as males. They made the hackathon non-competitive and community-based realm and observed the divergences in the hackathon. Tandem et al. [2] conducted a university hackathon to combat malaria and zika with land-based and flying drones. It was a one-day event at California State University, where a lot of participants worked on the projects. The authors found that students had an overall increase in interest in science and engineering after participating in the hackathon.

Saravi et al. [3] described a hackathon-style system engineering process as an approach to the rapid generation and development of early design concepts of complex engineered products. They showed that the hackathon method offers significant benefits to stakeholders. A significant benefit of this process was to achieve useful results in a very short timeframe compared to using regular internal methods. Gama et al. [4] proposed a hackathon methodology in an educational setting. The authors showed that learning is one of the main motivators of hackathon participants, streamlining the ideation process with regular deliverables in short time frames and leading to objective discussion and quick decision making.

Chandrasekaran et al. [5] organized Oak Ridge Leadership Compute Facilities GPU Hackathon. The authors reported the training format adopted for the hackathon and discussed the reasons for the successes and failures of teams based on case studies from more than 15 hackathons. They summarized the outcomes and takeaways for

participants, and then looked at the hackathon format from the perspective of educators. Raatikainen et al. [6] organized a three-days hackathon to access a device-centric cloud ecosystem in industrial setting. They showed that a hackathon is a promising approach in software engineering due to an increase in the relevance of speed in software engineering. They also showed encouraging experience about the hackathon among the participants in terms of the social benefits, such as collaboration, inspiration, and work motivation.

In all the above-mentioned hackathons, the responses were mostly positive on the level of satisfaction of the participants and on the increase of knowledge from the hackathon. It is also shown that the hackathons are useful in many domains of life.

3   Methods and Materials

3.1   Dataset description

We analyze the dataset about the survey results of 31 hackathon participants. We made two surveys collected. The first survey was made from one part of participants, whereas the second survey was made from the other part. The questions were about gender, age, the level of satisfaction, the willingness of attending another hackathon, the biggest challenge for them, whether they obtained new methods and knowledge during the hackathon, what they learned newly about the Facebook developer, and any comments or suggestions.

3.2   Used technology

To make dataset analysis, we used Python programming language, Pandas library.

- Python is an interpreted, high level, universally useful programming language. Its object-arranged methodology intends to assist software engineers with composing clear, intelligent code for both little and huge scope ventures.
- In PC programming, Pandas is a product library composed of the Python programming language for information control and examination. Specifically, it offers information structures and activities for controlling numerical tables and time arrangement. It is free programming discharged under the three-proviso BSD license.

3.3   Justification of chosen methodology

Python is considered one of the best programming languages for scientific purposes. It is easy to write a script on it and execute it. Python can be successfully launched on different platforms as Google Colab, Jupyter Notebook, Deepnote, etc. Pandas is a Python library for data wrangling and analysis. Lots of data analytics functions are packed into separate modules on it, which makes the process of writing data analysis code smooth.

3.4 Limitations

Our collected data size is not large due to a limited number of participants. It can lead to biased results.

4 Exploratory Data Analysis

4.1 Categorical columns exploration

In the first survey, there were 16 rows and 43.75% of missing values. In the second survey, there were 15 rows and 6.66% of missing values. Table 1 lists the number of participants in two surveys. We had 16 participants in the first survey and 15 participants in the second survey, total 31 participants. Among them, 25 were male students, five were female students, and another student did not want to mention own gender. The number of female students was very small; thus, we did not analyze the data by gender in this paper. This small number of female students matches to the literature [1] which mentioned that females participate in hackathons with much less frequency as males.

Table 1. Number of participants

	Male student	Female student	No response	sum
1st survey	13	2	1	16
2nd survey	12	3	0	15
sum	25	5	1	31

Table 2. Age of participants

Age	17	18	19	20	21	No response	sum	Average age
1st survey	1	2	3	8	1	1	16	19.4
2nd survey	1	1	4	6	2	1	15	19.5

Table 3. Two preliminary questions

	yes	no	sum
Did the participant hear about Facebook Developer API before?	8	7	15
Is this the first hackathon for the participant?	12	3	15

Table 2 shows the age range of participants. We observe that most participants are 19 or 20 years old. Table 3 lists answers to two preliminary questions from the participants in the second survey. These questions were not asked in the first survey; therefore, they were not added to the analysis. For most of the participants, this was the first hackathon to them. Half of the participants heard about Facebook API developer, and the other half did not.

4.2 Numerical columns exploration

Table 4 lists participants' answers to two postmortem questions, the level of satisfaction in the first survey and the willingness of attending another hackathon in both surveys. The data on the level of satisfaction in the second survey was not collected. Here, 1, 2, 3, 4, 5 represents "strongly disagree", "disagree", "neutral", "agree", "strongly agree", respectively.

The average level of satisfaction is 3.88 out of 5, and the average willingness of attending another hackathon is 3.88 (the first survey) and 3.80 (the second survey) out of 5. From these averages, we can say that the participants showed a great interest in the hackathon. Here, the level of satisfaction is almost normally distributed with a little bit of sparsity. The willingness of attending another hackathon is almost exponentially distributed in the first survey, but almost normally distributed in the second survey.

Table 4. Two postmortem questions

	1	2	3	4	5	sum	average
Are you satisfactory at this hackathon? (1st survey)	0	1	4	7	4	16	3.88
Will you attend another hackathon? (1st survey)	1	1	3	5	6	16	3.88
Will you attend another hackathon? (2nd survey)	0	1	4	7	3	15	3.80

4.3 Looking at varied data

Figure 1 shows the correlation matrix for the first survey, and Figure 2 shows the matrix for the second survey. The correlational matrix is an n × n grid, where n is the number of quantitative features in the input dataset. Each cell of a matrix represents a color, where color represents how strong the correlation is. The darker is the color, the lower is the correlation, and the lighter is the color, the higher is the correlation. At the right side of Figure 1 and Figure 2, the values of colors are explained in a separate bin. There are 3 × 3 blocks in Figure 1 (from the first survey) and 2 × 2 blocks in Figure 2 (from the second survey). Figure 2 does not contain a feature "the level of satisfaction"; thus, there are only two quantitative variables, whereas there are three quantitative variables in Figure 1.

In Figure 1, we observe a strong correlation between the willingness of attending another hackathon and the level of satisfaction (around 0.7), almost medium correlation between age and level of satisfaction (around 0.4) and 0 correlation between age and the willingness of attending another hackathon. In Figure 2, because there were only two quantitative features in the second survey, we observe almost no correlation between age and the willingness of attending another hackathon.
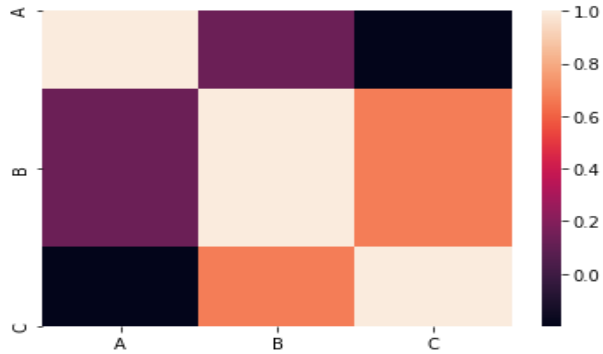
Fig. 1. Correlation matrix for the first survey (A: Age, B: level of satisfaction, C: Willingness of attending another hackathon)
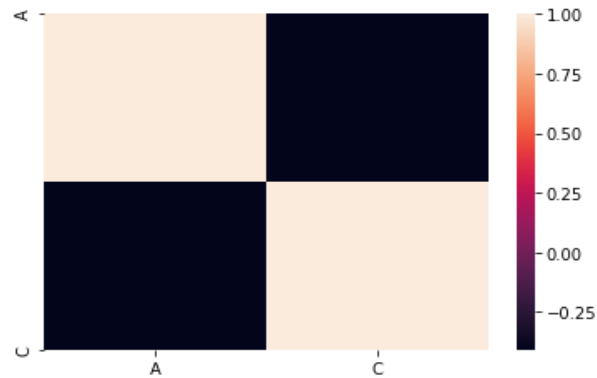

Fig. 2. Correlation matrix for the second survey (A: Age, C: Willingness of attending another hackathon)

5 Recommendations

The usefulness of the hackathon from the perspective of participants is confirmed in [1-6]; however, we obtain new insights based on our hackathon. The primary motivations to participate in this hackathon are new experience (most of the participants wrote this was their primary motivation), friends encouraged to attend, to be a part of a social group, and work on real-world problems. The things that the participants learned according to the surveys are construction of Chatbot, Spark AR, messenger AI bot, Facebook API usage and mechanism, and the way to build and upload game using SDK, implementation, and training of chatbot.

The main challenges that participants encountered are understanding the business concept to use in a project, many things to learn all at once, whether to focus on technology or business in a project, time limit, lack of coding skills, and distractions.

Generally, we can conclude that the participants were satisfied with the organization of the hackathon.

We make some recommendations to improve the hackathon considering the comments and suggestions written by the participants as follows:

- The hackathon should be more focused on the information provided to students. They did not know exactly what to do and what the hackathon expects from them.
- A team should not have a lot of participants. Two- or three-persons team seems to be enough.
- Number of participants should be increased.
- Number of categories should be increased

6   Conclusion

Organizing hackathons takes considerable time and effort. In this paper, we report Paragon IT hackathon. Most of the survey results were positive. However, some organizational changes should happen to make it even better to increase participants and categories, to decrease the member size per team, to provide detailed information about hackathon, etc.

7   References

 [1] A. Decker, K. Eiselt, and K. Voll, "Understanding and improving the culture of hackathons: Think global hack local," IEEE Frontiers in Education Conference, El Paso, TX, USA, Oct. 2015. https://doi.org/10.1109/fie.2015.7344211
 [2] J. Tandon, R. Akhavian, M. Gumina, and N. Pakpour, "CSU East Bay hack day: A university hackathon to combat malaria and zika with drones," IEEE Global Engineering Education Conference, Athens, Greece, April 2017. https://doi.org/10.1109/educon.2017.7 942968
 [3] S. Saravi, D. Joannou, R. S. Kalawsky, M. R. N. King, I. Marr, M. Hall, P. Wright, R. Ravindranath, and A. Hill, "A Systems Engineering Hackathon – a methodology involving multiple stakeholders to progress conceptual design of a complex engineered product," IEEE Access, 2018. https://doi.org/10.1109/access.2018.2851384
 [4] K. Gama, B. Alencar, F. Calegario, A. Neves, and P. Alessio, "A hackathon methodology for undergraduate course projects," IEEE Frontiers in Education Conference, San Jose, CA, USA, Oct. 2018. https://doi.org/10.1109/fie.2018.8659264
 [5] S. Chandrasekaran et al., "Best practices in running collaborative GPU hackathons: Advancing scientific applications with a sustained impact," Computing in Science & Engineering, July/August 2018, pp. 95-106. https://doi.org/10.1109/mcse.2018.042781332
 [6] M. Raatikainen, M. Komssi, V. d. Bianco, K. Kindstom, and J. Jarvinen, "Industrial experiences of organizing a hackathon to assess a device-centric cloud ecosystem," IEEE 37th Annual Computer Software and Applications Conference, Kyoto, Japan, July 2013. https://doi.org/10.1109/compsac.2013.130

8   Authors

Azamat Serek is a member of the computer science department in engineering and natural sciences faculty in Suleyman Demirel University. His research interests include but not limited to parallel computing, distributed programming, and big data analytics.

Meirambek Zhaparov is a Vice Rector of student Affairs, dean of Information Communication Technologies faculty at Paragon International University, Phnom Penh, Cambodia. He was an Assistant Professor of Computer Science and Dean of Engineering and Natural Sciences Faculty at Suleyman Demirel University, Kaskelen, Kazakhstan in 2018. He earned a BS degree in Computer Science at Suleyman Demirel University in 2003, MS degree in Computer Science at D. Serikbayev East Kazakhstan State University, Ust-Kamenogorsk, Kazakhstan in 2006, and PhD degree in Computer Science at Suleyman Demirel University in 2013. His research interests include microlearning and IT education.

Seong Moo-Yoo is an Associate Professor of Electrical and Computer Engineering at the University of Alabama in Huntsville (UAH), Huntsville, Alabama, USA. Before joining UAH, he was an Assistant Professor at Columbus State University, Columbus, Georgia, USA. He earned a BA degree in Economics at Seoul National University, Seoul, Korea, and MS and PhD degrees in Computer Science at the University of Texas at Arlington, USA. His research interests include computer network security, malware detection, and wireless networks. He has authored/co-authored over 100 scientific articles in refereed journals and international conferences. He is a life member of IEEE and a senior member of ACM.

Assem Talasbek is a Senior Lecturer of Computer Science department at Suleyman Demirel University (SDU), Kaskelen, Kazakhstan and a PhD student at SDU. She earned a BS degree in Computer Science at Suleyman Demirel University, Kaskelen, Kazakhstan, in 2015, MSc degree in Computer Science at M.Auezov South Kazakhstan State University, Shymkent, Kazakhstan in 2017. She is an Oracle Academy instructor at Suleyman Demirel University. Her research interests include IT education, Machine Learning.

Yong-Kab Kim received his Ph.D degree in Electrical & Computer Engineering in North Carolina State University. He is currently in Professor at School of Electrical Information Communication Engineering, Wonkwang University, Korea. His research interests are Remote sensing for Visible Communication, Optical Fiber Sensing, Intelligent Control, and Advanced Computational Intelligence.

Min-Woo Jin is master`s student in Information & Communication Engineering in Wonkwang University. His research interests are Electronic Circuit, Smart Sensor, and Block Chain.