Touched by the Hackathon: A Study on the Connection between Hackathon Participants and Start-Up Founders

Alexander Nolte
alexander.nolte@ut.ee
University of Tartu
Tartu, Estonia
Carnegie Mellon University
Pittsburgh, PA, USA

ABSTRACT

Time-bounded events such as hackathons, code fests and others have become a global phenomenon. Entrepreneurial hackathons in particular have gained wide spread popularity because they come with the prospect to being the grounds where the next billion dollar enterprise is born. There is however limited insight into whether and how hackathons participants and start-up founders are connected beyond studies on singular events focusing on hackathons as a starting point for start-ups. To address this gap we conducted a study on a dataset covering 44 hackathons over three years and 489 start-ups in the North-Eastern European country of Estonia. Our findings indicate that hackathons are not always the start of an entrepreneurial endeavor but can also be useful through later stages as a means to develop future products, find future employees and others. The results presented in this paper are based on an initial analysis of this rich dataset and thus present the starting point of a larger study on the connection between the hackathon and start-up communities which is currently in planning.

1 INTRODUCTION

Starting as coding competitions in the early 2000s time-bounded events such as hackathons\(^1\), code fests, hackdays and others have since proliferated into various other domains including companies of different sizes \([17, 26, 31]\), (higher) education institutions \([14, 29]\), civic engagement groups \([15, 16, 20]\), (online) communities \([8]\) and others. During such events participants typically form teams and engage in intensive collaboration to complete a project of interest \([28]\). Hackathons are organized by individuals with different aims \([22]\) such as fostering the development of new and innovative products and services \([4, 32]\), tackling civic and environmental issues \([1, 10, 30]\), spreading knowledge \([13, 18, 25]\) and expanding communities \([24, 34]\). Entrepreneurial hackathons in particular have gained popularity because they are perceived by participants and organizers as the grounds where next big idea that can be turned into a billion dollar enterprise is born \([6, 7, 9, 17]\). Estonia is at the forefront of this development due to its vivid (IT) start-up scene that has spawned many successful companies which have raised millions of Euros in investments and created over 2300 jobs\(^2\) with no sign of slowing down\(^3\). The proliferation of the IT sector in Estonia has been attributed to “Tiger Leap” program which was created in the late 1990s after Estonia declared its independence from the Soviet Union with the aim to equip all schools with computers and internet access \([23]\). Since then the Estonian start-up scene has matured and established itself as a major asset to the overall economy of the country \([11]\).

Despite interest on hackathons continuously growing in recent years \([27]\) there is still a profound lack of understanding of whether and how hackathons contribute to fostering entrepreneurship and how they can be purposefully integrated into existing entrepreneurial practice. Most studies on hackathons focus on the event itself covering aspects such as how teams organize during a hackathon \([35]\), how to deal with diverse audiences \([12]\) and how to encourage participation \([34]\). There is work on potential outcomes of hackathons e.g. in the context of start-ups \([6, 7]\) and larger corporations \([26]\) but these studies focus on singular events only. A larger scale overview of the connection between hackathons and start-ups is missing at this point. Moreover most existing studies perceive hackathons as the starting point of an entrepreneurial endeavour \([6, 9]\) without ever questioning this assumption. We aim to avoid this assumption by conducting an exploratory study focusing on

\(^1\)We will use the term hackathon as a substitute for other similar events throughout the remainder of this article.

\(^2\)https://www.slideshare.net/mvavulski/estonian-startup-scene

\(^3\)http://summit.startupnations.co/estonian-startup-statistics-first-half-2017
the time-line between when a founder of a start-up participated in
a hackathon and when s/he founded her/his company thus asking
the following first research question:

RQ1: What is the temporal connection between founding a start-up
and participating in one or multiple hackathons for founders of
Estonian start-ups?

Despite there being no agreed upon process of how start-ups
mature and become established companies there is a general un-
derstanding that start-ups go through multiple stages [3, 19, 21].
Not limiting our work to perceiving hackathons as the starting
point for entrepreneurs as discussed before thus naturally leads
to the question if there are other phases during the maturation
of start-ups in which hackathons might be perceived to be useful by
their founders. We thus also ask the following connected research
question:

RQ2: What are potential differences between start-ups whose founders
participate in hackathons at different times before or after they
founded their company?

To answer these two main research question we conducted an
exploratory case study comparing a dataset of 44 entrepreneurial
hackathons that took place over a span of more than three years
with information about 489 currently active start-ups in Estonia.
Comparing the time-lines of the start-ups that were connected to
hackathons we found that founders who participated in hackathons
do not always found their companies afterwards. Most founders
that participated in hackathon(s) in fact had founded a start-up
before they went to their first hackathon. Moreover our findings
point towards potential use cases for hackathons beyond founding a
start-up such as finding employees and developing future products.
During the course of this paper we will mainly focus on the tempo-
ral aspects of the connection between hackathon participants and
start-up founders. Other aspects such as the motivation of founders
to participate in one or multiple hackathons and the perception
of hackathon participants on entrepreneurial hackathons will be
subject of an (extended) future study.

2 EMPIRICAL METHOD

In order to answer the research questions stated in the introduction
we conducted an exploratory case study ?? by connecting two
datasets containing information about hackathons that took place
in Estonian and start-up communities that were founded there
respectively (c.f. Figure 1 for an overview). We will elaborate on the
specifics of the datasets (section 2.1) and on the analysis procedure
(section 2.2) in the following.

2.1 Data

For this work we collaborated with the two Estonian based orga-
nizations Garage48\(^4\) and Startup Estonia\(^5\). Garage48 is a for-profit
company that organizes entrepreneurial hackathons across Europe
with a strong focus on the Baltic states and Estonia in particular.
They are the premier organizer of hackathons in the region and
they are closely connected to the Estonian start-up scene as evi-
dent by the frequent participation of well known individuals from

that scene as speakers, mentors and judges in their hackathons.
Garage48 started in 2010 and has since then organized more than
130 hackathons across 44 cities in 22 countries\(^6\). Startup Estonia
is a government initiative that – among other functions – curates
a public database of active start-ups in Estonia. At the point of
data collection in March 2019 this database contained information
about 489 active start-ups which were founded between 2001 and
2018. Comparing this dataset with information about 44 hackathons
with a total of 3535 unique participants – which was provided by
Garage48 and is part of their registration procedure for hackathons
– allowed us to draw connections between hackathon participants
and start-up founders.

\[^4\]http://garage48.org/

\[^5\]https://startupestonia.ee

\[^6\]http://garage48.org/blog/garage48-is-thrilled-to-show-its-new-fresh-face

\[^7\]https://www.e-krediidiinfo.ee

\[^8\]https://github.com/seatgeek/fuzzywuzzy

The two datasets we analyzed contained information about
where and when hackathons took place and who the participants
were including information about their current profession. They
also contained information about when start-ups were founded,
who founded them and in which sector they operate. In addition we
obtained information about the number of employees and the tax-
able income of start-ups from Creditinfo Eesti\(^7\) which hold publicly
available information about companies registered in Estonia.

2.2 Procedure

To study the connection between hackathon participants and start-
up founders we compared the names of founder of active start-ups
with the names of hackathons participants. For the comparison
we used the Python library fuzzywuzzy\(^8\) which supports Fuzzy
string matching based on the Levenshtein distance metric. This
approach appeared reasonable since participants enter their names
manually into a hackathon registration form which is error prone.
After the comparison we conducted a manual cleaning procedure
confirming the identified names by comparing them with the lists
of registered hackathon participants and start-up founders. We
removed potential duplicates and false-positives such as obvious
nicknames and individuals that registered using their first name
or initials only and that were identified as the same person by the
matching algorithm. We also removed start-ups for which we could
not acquire monetary and employee related information because
they were not reported to the Creditinfo Estonia database.
After establishing the connection between start-ups founders and hackathon participants, we analyzed the time between the founding of a start-up and the different hackathons each founder participated in (section 3.1). Based on these results, we created three clusters for (1) start-ups that were founded before founders participated in a hackathon, (2) start-ups that were founded after founders participated in a hackathon and (3) start-ups where founders participated before and after they founded a start-up. We then compared the three clusters with respect to their relationship to hackathons, taxable income and their number of employees (section 3.2).

3 PRELIMINARY FINDINGS

In this section, we will first provide an overview of the characteristics of the dataset we analyzed (cf. Table 1). Afterwards, we will discuss temporal connections between hackathon participation and the founding of start-ups (section 3.1) thus answering RQ1 before clustering and comparing them (3.2) to answer RQ2.

Table 1: Descriptive statistics for hackathons (n = 44) and start-ups (n = 241) in our dataset.

<table>
<thead>
<tr>
<th></th>
<th>max</th>
<th>mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td># participants per hackathon</td>
<td>193</td>
<td>95.50</td>
<td>38.41</td>
</tr>
<tr>
<td>taxable income (EUR) per start-up</td>
<td>8665 550.49</td>
<td>287 740.71</td>
<td>971 026.01</td>
</tr>
<tr>
<td># employees per start-up</td>
<td>333</td>
<td>12.01</td>
<td>28.13</td>
</tr>
</tbody>
</table>

The dataset we analyzed contains information about 44 hackathons which were frequented by 3335 unique participants between February 2016 and March 2019. The largest hackathon had 103 participants with an average of 96. Most participants only went to one hackathon. 13.10% of the study population participated in more than one event with one individual participating in 10 hackathons. The participants came from diverse backgrounds covering different professions within and beyond the IT domain such as front-end and back-end development, marketing, design engineering and others (cf. Table 2 for an overview).

Table 2: Top reported backgrounds of hackathon participants with at least 100 mentions.

<table>
<thead>
<tr>
<th>Participant background</th>
<th># of participants with this background</th>
</tr>
</thead>
<tbody>
<tr>
<td>marketing</td>
<td>264</td>
</tr>
<tr>
<td>back-end developer</td>
<td>243</td>
</tr>
<tr>
<td>front-end developer</td>
<td>165</td>
</tr>
<tr>
<td>project manager</td>
<td>148</td>
</tr>
<tr>
<td>designer</td>
<td>124</td>
</tr>
<tr>
<td>engineer</td>
<td>115</td>
</tr>
</tbody>
</table>

In addition, our dataset contains information about 241 start-ups with an average of 12 employees (SD = 28.13) that were founded between February 2001 and December 2018. The largest start-up has 333 employees. Most of the start-ups in our dataset are related to difference aspects of IT providing technology and services for education, transport, health and others. A large portion of start-ups in our dataset focus on software development but some also develop hardware or provide software and sales related services. They generate an average taxable income of more than 250,000 EUR per year but the large standard deviation (SD = 971 026.01) points towards a considerable disparity between the different start-ups in terms of their taxable income.

3.1 Temporal Connection Between Start-up Founding Date and Hackathon Participation

In order to analyze the temporal connection between start-ups and hackathons (RQ1), we started by identifying start-ups whose founders participated in one or multiple hackathons. Out of the 241 start-ups that were contained in the cleaned dataset 28 (11.62%) were connected to hackathons i.e. were founded by one or multiple individuals who participated in one or multiple hackathons. Most of these start-ups were established before the individuals who founded them participated in a hackathon (85.71%). Only 10.71% were established after their founders had participated in a hackathon and we also identified one case in which a founder participated in a hackathon before and after s/he founded her/his start-up (cf. row “all” in Table 3). Most founders only participated in one hackathon with an average participation rate of 1.75 hackathons per founder (SD = 1.73).

Based on this finding we divided our dataset into three clusters for the following analysis (section 3.2). These clusters consist of (1) start-ups which were founded before any of their founders participated in a hackathon (C1), (2) start-ups which were founded after any of their founders participated in their first hackathon and (C2) (3) start-ups where their founders participated in a hackathon before and after they founded it (C3).

Before focusing on the comparison, we calculated the distance between hackathons and the founding date of connected start-ups (cf. Figure 2). We found that the longest distance between the founding of a company and the last hackathon one or multiple founders participated in is more than 11 years (135 months). The longest distance between a founder participating in a hackathon before founding her/his start-up in comparison is only 6 months. The shortest time between a hackathon and the foundation of a start-up is three months. There is an average of more than 3 and a half years between the different events (m = 42.96 months) with a large disparity between the different start-ups (SD = 36.82).

Related to answering RQ1, our analysis thus revealed that the temporal connection between founding a start-up and founders participating in one or multiple hackathons was not limited to start-ups being founded after a hackathon in our case. We found start-ups that were created before, after and between one or multiple founders participating in a hackathons. We also found that – in contrast to the generally accepted assumption that start-ups are founded after a hackathon – most founders in our dataset created their start-up before participating in a hackathon. Finally, our analysis also
showed a large variety related to the time between founding a start-up and participating in a hackathon as well between participating in different hackathons in our dataset.

### 3.2 Differences Between Connected Start-ups

In order to answer our second research question \( \text{(RQ2)} \) we compared the start-ups in our dataset along the three aforementioned clusters. We focused on the time between different events, the taxable income of start-ups and their number of employees (c.f. Table 3 for an overview). We focused on a qualitative comparison based on mean values and standard deviations rather than conducting a more sophisticated statistical analysis since the large difference in cluster size and the small size of two of the three cluster can be expected to disproportionally affect outcomes.

First it should be noted that there is a disparity between start-ups that were founded after a hackathon and all other clusters in terms of distance between hackathons as pointed to in the previous section. The temporal distance between hackathons and start-ups that were founded after their founders participated in a hackathon is only four months on average with a relatively low standard deviation \( (SD = 1.73) \). The temporal distance between events for start-ups that were founded before their founders participated in a hackathon is almost 50 months on average \( (m = 49.04, SD = 36.30) \) with a large standard deviation. This finding indicates that start-up founders might come with intentions beyond the creation of a start-up. It could however also be partially attributed to our dataset containing data of start-ups from the last seventeen years while it only covers data from hackathons over the past three.

Our analysis also pointed towards a considerably higher average taxable income and considerably more employees on average for start-ups that were founded after a hackathon compared to those that were founded before. The difference is almost 200,000 EUR and 5 employees on average (c.f. Table 3). Moreover we found that both the taxable income and the number of employees is vastly different among the start-ups that were founded after a hackathon pointing towards one company being largely more successful than the others. This result can however also be skewed because of the large difference in cluster size.

The only start-up where founders participated in a hackathon before and after founding had a small taxable income compared to the overall average (c.f. Table 3) but it had considerable more employees (27 compared to an average of 9.82). The average time between hackathons was closer to the average time of start-ups that were founded after a hackathon than to those that were founded before (14 months compared to an average of 42.96).

Our analysis thus revealed a variety of differences between start-ups whose founders participated in hackathons at different times before or after they founded their company in the cases we studied \( \text{(RQ2)} \). Our findings point towards start-up founders coming to hackathons with intentions beyond the creation of a start-up. Moreover they indicate differences between the different clusters related to average taxable income and employees with start-ups that were created after their founders participated in their first hackathon showing considerably larger growth. It should however be noted that these findings are based on clusters that greatly vary in size.

---

### Table 3: Descriptive statistics for start-ups that are connected to hackathons.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>N hackathons</th>
<th>Hackathon distance (months)</th>
<th>Taxable income (EUR)</th>
<th>Employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>24</td>
<td>( m = 1.83, SD = 1.86 )</td>
<td>( m = 136007.86, SD = 233096.49 )</td>
<td>( m = 4.33, SD = 4.04 )</td>
</tr>
<tr>
<td>C2</td>
<td>3</td>
<td>( m = 1.00, SD = 0.00 )</td>
<td>( m = 310829.18, SD = 1072816.10 )</td>
<td>( m = 9.79, SD = 17.18 )</td>
</tr>
<tr>
<td>C3</td>
<td>1 and 2</td>
<td>14</td>
<td>99045.95</td>
<td>27</td>
</tr>
<tr>
<td>All</td>
<td>28</td>
<td>( m = 1.75, SD = 1.73 )</td>
<td>( m = 284534.64, SD = 994379.38 )</td>
<td>( m = 9.82, SD = 16.34 )</td>
</tr>
</tbody>
</table>
4 DISCUSSION

Our analysis pointed towards hackathons potentially serving as a starting point for founders to create a start-up but they also appear to be perceived as useful by founders after they had already created their start-up (RQ1). In fact most of the start-ups that were connected to hackathons in our case had been founded before their founders participated in a hackathon. It thus appears necessary to extend our current understanding about the role hackathons in the process of the inception and growth of a start-up beyond perceiving them as a starting point [6, 7]. The work presented in this paper takes an initial step into this direction.

It also appears reasonable for founders to participate in hackathons after they already founded their start-up because it is unrealistic to expect that the short timespan of a hackathon is sufficient for participants to come up with an idea and build a presentable product. It rather appears reasonable that an individual forms an idea for a start-up and subsequently participates in a hackathon to take – maybe even initial – steps towards an initial / presentable prototype or product. This assumption is supported by the findings of Nolte el al. [26] who found that participating in a hackathon with a well formed project idea can positively influence project continuation after a hackathon in a corporate context. It thus appears reasonable to expect a similar effect in this case since the focus of both entrepreneurial and corporate hackathons is typically on taking steps towards the creation of a sustainable product.

Moreover founders might perceive participation in a hackathon as an opportunity to receive feedback about an idea or prototype and expose it to a larger audience. Networking has been found to be a motivation for individuals to participate in entrepreneurial hackathons [6] and hackathons of scientific communities [5]. Participating in a hackathon to expose an idea or product and expand the founder’s network appears particularly relevant in our case since the hackathons we studied were organized by an organization that is well connected within the Estonian start-up ecosystem. It thus appears reasonable to expect a similar effect in this case since the focus of both entrepreneurial and corporate hackathons is typically on taking steps towards the creation of a sustainable product.

Our analysis however also provided hints towards the potential that start-ups that are founded after a hackathon were more successful in the case we studied. One possible explanation for this finding – despite it potentially being based on the small size of this particular cluster – could be that participation in a hackathon can allow future founders to expose their idea to an interested audience, reflect it critically and subsequently found a start-up based on a mature idea that already had exposure to parts of the local start-up ecosystem. This can be perceived as a contradiction the previous finding that having an idea prior to a hackathon could be beneficial for its sustainability of a in the long run. Discussing an idea first might however also be beneficial for its continuation since most successful start-ups are not built on an initial idea. They rather go through multiple ideas and concepts before a suitable product-market-fit can be achieved [2, 3]. This potential contradiction however needs to be clarified by further investigation.

4.1 Limitations

The work presented in this paper was based on an analysis of two connected dataset which poses certain limitations. Our datasets only contained information about a subset of hackathons that took place during the analyzed time span and they only contained information about start-ups that were founded in Estonia. These limitations are however only marginally significant for the presented study since our aim was to explore the temporal connection between founders participating in a hackathon and founding a start-up (RQ1) as well as the potential differences between founders that participate in hackathons at different points in time (RQ2) rather than providing a comprehensive overview of all possible scenarios. Moreover, our datasets do not contain information about the motivation of start-up founders to participate in a hackathon. Their participation of founders might thus have been motivated by means beyond the operation of their current or future start-up which might affect the conclusions we draw in the discussion related to both research questions. It should however be noted that most hackathons organized by the organization that provided the dataset aim to foster entrepreneurship which makes it probable that participants indeed came with the motivation to create or foster their respective start-up.

Comparing the names of founders and start-up participants to establish connections also poses a threat to the reliability of our findings. In order to mitigate this threat we abstained from drawing any causal conclusions about the identified connections but rather discuss the diversity of connections between founders who participated in hackathons at different points in time thus going beyond the current perception of this connection in academic work which is limited to perceiving hackathons as a starting point for start-ups. This finding holds true even if some of the connections we identified might turn out to be false-positives.

5 CONCLUSION AND FUTURE WORK

During the course of this paper we presented preliminary results from an exploratory case study on the connection between hackathon participants and start-up founders in Estonia. Mainly focusing on temporal aspects of this connection we found that hackathons do not always mark the start of an entrepreneurial path but might...
also be useful for start-ups at later stages of their existence. Participation in these cases might have been motivated by aspects such as finding new employees, acquiring feedback for future products and others. Moreover start-ups that were founded after founders participated in a hackathon were more successful in the case we studied pointing towards the potential usefulness for hackathons to develop a mature idea, gather feedback and create exposure before developing the final product. The dataset we analyzed for this paper however does not allow us to draw causal conclusions because reasonable interpretations as presented in the discussion (section 4).

Building on the previously presented work we are currently planning a mixed-method case study that combines a quantitative survey instrument with qualitative interviews to complement the datasets that formed the basis for this paper. We will target both founders of start-ups and hackathon participants including founders that did not participate in a hackathon and hackathon participants that did not found a start-up despite potentially participating in a hackathon with this goal in mind. Focusing on the perception of founders and hackathon participants we aim acquire a rich set of relevant data that allows us to gain insights into the potential role of hackathons in entrepreneurial practice beyond serving as a starting point for start-ups.

ACKNOWLEDGMENTS

We would like to thank Garage48 and Startup Estonia for collaborating with us on this work.

REFERENCES