Discovering Key Actors and Opinion Leaders on Twitter’s Start-Up and Entrepreneurship Topics Trending: A Social Network Analysis Approach

Menemukan Aktor Utama dan Opini Pemimpin pada Topik Start-Up dan Kewirausahaan yang Sedang Tren: Pendekatan Analisis Jaringan Sosial

Eristian Wibisono¹
¹Faculty of Business and Economics, University of Pecs, Hungary
¹wibisono.tian@gmail.com

Naskah diterima: 16 April 2023, direvisi: 21 Mei 2023, disetujui: 18 Juli 2023

ABSTRACT

The topic of start-ups and entrepreneurship has been widely discussed on social media to disseminate information while serving as a digital marketing platform. Unfortunately, researchers and practitioners interested in analyzing social networks to gain business advantage through digital platforms still need to be expanded. This paper analyzes networks on social media to find important actors and opinion leaders on trending start-up and entrepreneurship topics. Data analytics is performed using one of the most extensive social media, Twitter. Several visualizations of the social network of these actors are then displayed using specific network visualization applications. The results are discussed by analyzing the main attributes in the network, namely degree, centrality, modularity, and word cloud, which can lead us to discover important attributes in the network: main actors, opinion leaders, groupings of actors according to topics discussed, and specific terms that are trending topics. This paper contributes to digital marketing practitioners and social network analysis studies in related fields with a concise and duplicable methodology. Further research is recommended to work more extensively with more representative data from multiple networks to analyze these topics in more depth.

Keywords: social networks; social media; digital marketing platform; opinion leaders; start-up

ABSTRAK

Topik start-up dan kewirausahaan telah banyak dibahas di media sosial untuk menyebarkan informasi sekali-gus berfungsi sebagai platform pemasaran digital. Sayangnya, peneliti dan praktisi yang tertarik menganalisis jejaring sosial untuk mendapatkan keuntungan bisnis melalui platform digital masih perlu didorong. Makalah ini menganalisis jejaring di media sosial untuk menemukan aktor-aktor penting dan opini pemimpin (pemuka pendapat) pada topik start-up dan kewirausahaan yang sedang tren. Analisis data dilakukan dengan memanfaatkan salah satu media sosial terbesar, Twitter. Beberapa visualisasi dari jaringan sosial para aktor ini kemudian ditampilkan menggunakan aplikasi visualisasi jaringan tertentu. Hasil penelitian dibahas dengan menganalisis atribut-atribut utama dalam jaringan, yaitu degree, centrality, modularity, dan word cloud, yang dapat menunjukkan kita menemukan atribut-atribut penting dalam jaringan: aktor utama, opini pemimpin, pengelompokan aktor sesuai topik yang dibahas, dan istilah-istilah tertentu yang menjadi trending topic. Dengan metodologi yang ringkas dan dapat diduplikasi, makalah ini berkontribusi bagi praktisi pemasaran digital dan studi-studi analisis jaringan sosial pada bidang terkait. Penelitian lebih lanjut disarankan untuk bekerja lebih ekstensif dengan data

DOI: https://doi.org/10.17933/iptekkom.25.1.2023.37-56
INTRODUCTION

Nowadays, startups have to compete not only in the physical market, but also in the digital market. Digital marketing channels take many forms, one of which is social media. Social media is an online social network that disseminates information to the market (Coulter and Roggeveen 2012; Prabowo et al. 2021). One of the most popular social media platforms, such as Twitter, has been proven to significantly influence access to information for customers and potential customers (Khajeheian 2012; Rodriguez, Peterson, and Krishnan 2012). Such platforms play a role in the distribution of information in the business network, which is an essential part of entrepreneurial activity (Hajli 2014).

As stated by Stieglitz et al. (2014), social media analysis can bridge different disciplines currently developing, which use social media to study the character of their social network or prospectus, obtain specific opportunities, or increase the impact of their future works. Social media, such as Twitter and Facebook, are often used to disseminate information, bridge public discourse, and opinion, and mediate various communication interests and political dialogues. In this context, Stieglitz and Dang-Xuan (2013) emphasize the importance of monitoring or analyzing information sets from various social media, and adequate visualization is needed to increase the relevance of such information for future political purposes. In health science, millions of tweets can be found that talk about health from different perspectives. Opinion posts about health can reflect popular perceptions about topics discussed, such as medicine, lifestyle, and vaccine issues during the Covid-19 pandemic. Asghari, Sierra-Sosa, and Elmaghraby (2018) used an adaptive system specifically designed to capture the correlation of words from a network of tweets to detect trending health topics across all US states. The results obtained from social media analysis can ultimately be used in decision-making and forecasting their future work goals (Sarker 2021; Jha, Agi, and Ngai 2020).

According to Men, Ji, and Chen (2017), the use of social media as a marketing strategy for startup goods and services, combined with the important role of stakeholders, is an essential strategy that should be used by startups. Analysis of content on social media such as WeChat and Weibo in China shows that information dissemination, reputation, and the involvement of opinion leaders are new ways of startup marketing. Other studies show the crucial role of opinion leaders on online social media in promoting companies’ products and services, especially for future cooperation and collaboration (Lin, Bruning, and Swarna 2018). Opinion leaders may also take time on social media to share their experiences and successes in starting and running a business. They are owners, experts, celebrities, influencers, and followers or fans who also play an important role in business communication on social media. These points make it clear that important information on social media cannot be analyzed only in terms of content. Apart from the importance of content, opinion leaders who appear with high intensity have the potential to influence the active startup environment on social media. Therefore, it is crucial for startup entrepreneurs to know and follow the trending topics in the startup business, to know what opinion content is trending and who the opinion leaders are (Kuckertz et al. 2020; Saura, Palos-Sanchez, and Grilo 2019).
Social network analysis has been widely used to analyze various economic and social phenomena (Coletto et al. 2017; Borgatti et al. 2009; Zhang 2010), including entrepreneurial activities and the recent growth of startups (Durda and Ključnikov 2019; Tajpour and Hosseini 2021; Geho and Dangelo 2012). Unfortunately, not many researchers and startup managers are interested in conducting social network analysis to manage their business, which could be because doing so sounds impractical in running an early-stage business (Siti-Nabiha, Nordin, and Poh 2021; Troise et al. 2022; Hu et al. 2019). While the concept and theory of digital entrepreneurship have developed rapidly in the last two decades (Zaheer et al. 2019; He et al. 2020), the literature addressing the importance of social network analysis in the context of early-stage entrepreneurship is still limited (Lans et al. 2015; Hayter, C. S. (2016); Lynn et al. 2020).

Based on the background description of the development of social network analysis literature in different disciplines and its urgency for the digital marketing of startups, this paper aims to position startups and entrepreneurship in the context of social media networks. A relevant research problem is how to visualize social networks when the topic of entrepreneurship and startups is trending on social media platforms. The goal is to gain a visual representation of how these topics are discussed, what the networks look like, and who the key actors and opinion leaders are in these networks. By using one of the largest social media, Twitter, this study aims to reveal the startup and entrepreneurship topics trending on this social media. To solve the research problem, social network analysis is applied by collecting data through text mining and visualizing the data set using special software to represent the research results. The applied method leads to the study results that show the specific attributes of the social network that describe the trending topics and the positions of the actors and opinion leaders involved in trending these topics. This paper provides practical benefits to prospective entrepreneurs and digital marketing practitioners seeking business and marketing opportunities on digital platforms. The replicable analysis methods presented in this paper are also expected to be used to develop or improve their businesses.

The following subsections outline the background literature related to the problem and the general objectives of this research. The second section outlines the research method to explain the methodological and analytical steps applied in this study. The visualization results and discussion are presented in the third section. A conclusion at the end concludes the paper.

Although network science is a relatively new field, it has a worldwide social impact. Successful companies such as Google, Facebook, and Twitter optimize technology and business models based on the network science. Despite Google being the most extensive network ever, social networks such as Facebook and Twitter are also ambitious to form similar social networks worldwide (Cusumano, Yoffie, and Gawer 2020; Bharadwaj et al. 2013; Dijck 2011; Kim and Kim 2018). The algorithm created by many scientists based on this media platform is about social relations and helps develop the economy and business, from small discussions for those just starting a business to using social media to boost sales through advertising (Barabási 2013).

Identifying and evaluating opportunities and synthesizing various resources to start a business (start-up) is essential for success. This entrepreneurial process occupies a strategic position because it involves all the functions, activities, and actions related to starting a business (Cornwall and Naughton 2003). After a new business is formed, the entrepreneurial process becomes an effective cycle to continuously increase the added value of the business (Kodithuwakku and Rosa 2002). A recent study conducted by Sukumar, Jafari-Sadeghi, and Xu (2021) stated that entrepreneurial activity generally carried out on social media is the collection
and analysis of information aimed at decision making or to boost promotion and sales. However, there are challenges of information overload and false information.

Marketing strategies on social media are applied to attract consumers’ attention by throwing various issues that might be hotly discussed in the market, talking about them as a problem, and then throwing out several solutions. Responses on social media that appear from discussions, especially if a solution seems, usually lead to customer interest in the product or service being discussed. The response will be more fantastic, along with the number of discussions related to the issues discussed (Evans 2010; Gillin 2007; Toder-Alon, Brunel, and Fournier 2014a). In the definition of social analysis, digital interactions, relations of actors, and some ideas can be analyzed, measured, and interpreted through Social Network Analysis (SNA). Social Network Analysis is also part of text mining, where text data is processed or processed to produce information that can be processed further (Sarkar 2016a).

How a business operates is also inseparable from the influence of community activities. In this case, the role of social media is likened to an electronic word-of-mouth system (Bruns and Burgess 2012; S. B. Park, Jang, and Ok 2016). There is a significant increase in the use of social media today so that it is used as an opportunity for observers of social behavior and communication to see various phenomena that occur in society in the form of specific communication patterns or trends (Hidayat, Rafiki, and al Khalifa 2018; Garg, Swami, and Malhotra 2018; Cao et al. 2018; Dlamini and Johnston 2018; Kaur, Shri, and Mital 2018; Stieglitz et al. 2018; Gandomi and Haider 2015).

The use of social media has expanded worldwide and accelerated the diffusion process of various types of information (Xiang et al. 2015; Almotairy, Abdullah, and Abbasi 2020). As one of the world’s largest social media, Twitter has successfully strengthened relationships with customers and reinforced public perception of product brands (Sindhani et al. 2019). Compared to several other types of information media, social media is currently the choice of more people (Alalwan 2018; Simon, Goldberg, and Adini 2015; Nisar, Prabhakar, and Patil 2018). Studies related to the use of social media in the startup business environment prove that this media can improve the relationship of various parties in the business environment and even improve business performance (Almotairy, Abdullah, and Abbasi 2020).

Twitter is seen as one of the most extensive data and information sources on social media platforms, which is also seen as an effective marketing medium that improves business performance at a low cost (Malhotra, Malhotra, and See 2012; Garg, Swami, and Malhotra 2018). Twitter connects people’s behavior and sentiment and helps connect them with the government (Khan et al. 2014; Lakiwal and Kar 2016). Many recent studies have applied community behavior analysis to Twitter social media, including on new product development (Rathore and Ilavarasan 2020), climate change (Dahal, Kumar, and Li 2019), community communication and networking (Goodier 2018), and startup company strategy (Singh and Dhir 2019). Early research has discussed the role of microblogs in monitoring social phenomena through geo-tagging techniques, and then the spatial correlation of networks within them can be identified (Takhteyev, Gruzd, and Wellman 2012). The study of Joseph et al. (2017) examined the Internet of Things searched by this keyword search method on Twitter. The social media analysis conducted by Hussain et al. (2018) on immigration in Canada utilizes social media Twitter to analyze the attribute values of the network using the R studio and Gephi analysis tools. This research proves that using these two devices is very good for mining analysis and exploring big data.
RESEARCH METHOD

This study utilizes the social media Twitter to obtain all the data (text mining). The data was collected using an application based on the Netlytic website (www.netlytic.org). A Netlytic account allows the connection to a personal account on Twitter. First, Twitter data was crawled on Netlytic. To generate the data, the keywords "(startup) AND (entrepreneurship)" were used and only English tweets were filtered. This means that only accounts that talk about startups and entrepreneurship were pulled together.

Netlytic only pulls data from the last seven days (January 8-16, 2022). Meanwhile, the maximum amount of data obtained is 10000 data. This data was taken on January 16, 2022 and got the last data of 7873. Then the data was downloaded in ‘.gexf’ format for analysis in Gephi. In Netlytic, the initial description of the data can be browsed, such as statistical records, trending posts over time, top users by centrality, and network visualization.

The next step is to visualize the network in Gephi. From the Twitter data retrieved through Netlytic, 4383 nodes (actors) and 6458 edges (relationships) were formed. The graph type "directed" was chosen to see the direction of the connections. All network overview options were run to visualize the network based on specific statistical groupings. In addition, nodes were clustered and separated based on color, size, and label. Given the large amount of data, the Yifan Hu Proportional algorithm/layout option was used to display the network visualization and filter out nodes/actors that have no relationship with other actors (Weighted Degree min = 1). The last step is to describe the network attributes according to the existing theory.

RESULTS AND DISCUSSION

Before importing the data into Gephi, some descriptions from the Nelytic web-based application were noticed. Nelytic provided 7873 data points, 4383 nodes (actors/accounts), and 6458 edges (relations/links). The data from Nelytic was collected from January 8 to January 16, 2022, covering the last seven days. The report published by Netlytic shows the trend of tweets about startups and entrepreneurship being discussed at the time. The trends show a significant increase in startup and entrepreneurship topics on Twitter during the analysis period. In the first six days, there were about 500-1500 tweets discussing and creating startup and entrepreneurship hashtags, although there was a decrease on the last day, as shown in Figure 1.

![Fig 1. Posts/tweets about Start-Up and Entrepreneurship over time on Twitter](Source: output of Netlytic, own work)
The terms start-up and entrepreneurship are the dominant terms in this tweet. But it also coexists with several other words such as entrepreneur, business, innovation, marketing, small business, ecosystem, etc. It indicates that the topic of start-up and entrepreneurship is related to each other, and in general, it is also related to several different terms. These are the business terms that are most often discussed on social media (Fig 2).

![Fig 2. The most frequent words appear on Twitter](image)

Source: output of Netlytic, own work

Furthermore, Netlytic also displays users or actors based on in-degree (top figure) and out-degree (bottom figure) centrality who talk the most about start-up and entrepreneurship topics, as shown in Fig 3 while the Fig 4 shows the initial SNA network that was formed in Netlytic.
To visualize the network in detail, the data obtained from Netlytic was imported into the Gephi application. Next, before creating an appropriate visualization, I ran through all the options in the Statistics section, including Network Summary, Nodes Summary, and Edges Summary. Yifan Hu's proportional layout was chosen to visualize the network layout. The resulting algorithm
shows the attraction between nodes. Nodes with strong attraction will appear closer and move away when disconnected. Now, based on the overview that has been performed, we can view what these network attributes look like.

![Fig 5. Visualization of start-up and entrepreneurship networks on Twitter using Gephi](source)

The node degree in a network shows the number of edges connected to that node. Meanwhile, the degree distribution shows the degree distribution in the entire network. The degree distribution shows the number of connections contained in this graph. Value shows the value of the nodes, and count indicates the number of nodes (Newman 2008).

Based on Fig 6, the results obtained are Average Degree: 1,130. The degree report graph shows that at least 3000 nodes have a value less than 5, and around 300 and 420 nodes have a value less than 5. The in-degree distribution shows the number of connections that enter or lead to nodes, while the out-degree distribution shows how nodes relate to other nodes (Broder et al. 2011). The in-degree graph shows at least 200 and 400 nodes with a degree value of less than 5 and at least one node with values around 400, 240, 120, 75, and 90. The out-degree graph shows there are at least 200, 1200 and 2800 nodes with a value less than 5 and there is at least one node with a value of around 80, 120, 180, 240 and 360.
Fig 6. Degree distribution of the nodes
Source: output of Gephi, own work
The degree rankings are shown in Table 1:

<table>
<thead>
<tr>
<th>Id</th>
<th>Label</th>
<th>Indegree</th>
<th>Outdegree</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>n3216</td>
<td>princearihan</td>
<td>550</td>
<td>0</td>
<td>550</td>
</tr>
<tr>
<td>n3244</td>
<td>mygovindia</td>
<td>408</td>
<td>2</td>
<td>410</td>
</tr>
<tr>
<td>n3923</td>
<td>superteamup</td>
<td>0</td>
<td>358</td>
<td>358</td>
</tr>
<tr>
<td>n2214</td>
<td>bjp4india</td>
<td>243</td>
<td>0</td>
<td>243</td>
</tr>
<tr>
<td>n3869</td>
<td>productier1</td>
<td>0</td>
<td>241</td>
<td>241</td>
</tr>
</tbody>
</table>

Closeness and betweenness centrality and eccentricity are distinguished by Freeman (1995). Betweenness centrality shows the frequency with which a node appears on the closest path between other nodes in the network. In contrast, closeness centrality shows the average distance from the starting node to other nodes in the network. Eccentricity is the farthest distance of a node from its starting point. Nodes with a high betweenness centrality value can be considered opinion leaders.

From the results of the centrality analysis in Fig 6, the diameter value is 4, and the Average Path length value is 1.47. It means that a starting node has the furthest distance after passing through 4 other nodes. The betweenness centrality graph shows that more than 4000 nodes do not have a relationship or do not become a bridge that connects one node to another node. In addition, one node has a betweenness centrality value of more than 350 and 900, 2 nodes that have a betweenness centrality value between 200-250, 2 nodes that have a betweenness centrality value between 100-150. Closeness centrality that is close to the value indicates a node that has a stronger relationship with other nodes. In the centrality graph, there are about 1300 nodes that are not intermediaries between one node and another node (having a value of 0), but more than 1800 nodes have a value of 1. The eccentricity graph shows a value of diameter four which to reach the farthest point, a node will pass four other nodes.
Fig 7. Centrality distribution of the nodes
Source: output of Gephi, own work
The centrality rating is shown in Table 2:

<table>
<thead>
<tr>
<th>Id</th>
<th>Label</th>
<th>Eccentricity</th>
<th>Closeness</th>
<th>Betweenness</th>
</tr>
</thead>
<tbody>
<tr>
<td>n3244</td>
<td>mygovindia</td>
<td>1</td>
<td>1</td>
<td>931.83</td>
</tr>
<tr>
<td>n3065</td>
<td>agazdecki</td>
<td>3</td>
<td>0.5</td>
<td>366</td>
</tr>
<tr>
<td>n2971</td>
<td>microacquire</td>
<td>2</td>
<td>0.75</td>
<td>249</td>
</tr>
<tr>
<td>n3898</td>
<td>smmaholic</td>
<td>1</td>
<td>1</td>
<td>204.5</td>
</tr>
<tr>
<td>n2012</td>
<td>startupindia</td>
<td>2</td>
<td>0.67</td>
<td>134</td>
</tr>
</tbody>
</table>

In this case, the actor/user that has the highest betweenness centrality is the node "mygovindia," so the mygovindia account is the opinion leader on the topic.

Modularity in network analysis explains the clustering of nodes and what topics are being discussed (topical clusters) (Cheng et al. 2020; Sadri et al. 2022; Chouchani and Abed 2020; Azaouzi, Rhouma, and ben Romdhane 2019; Wallace, Gingras, and Duhon 2009; A. Park, Conway, and Chen 2018). From the analysis results, the data shows the value of Modularity: 0.783, Modularity with resolution: 0.783, and Number of Communities: 648, as shown in Fig 8. It means that there are 648 groups discussing topics about start-ups and entrepreneurship on Twitter in this network.

Using the visualization and network analysis results in Gephi, several essential network attributes are described, such as degree, centrality, modularity, and word cloud. From the degree report, the results show that at least five critical actors in the network have a degree value above 200. The three actors in the top 5 networks (princearihan, mygovindia, and bjp4india) have high values on the in-degree distribution, which means that these three actors have the most response
from the outside. In comparison, the other two actors (superteamup and productier1) have this value in their out-degree distribution, indicating the magnitude of the relationship between this actor and other actors in the network.

The centrality analysis shows that only one actor has a betweenness centrality value significantly different from the other four top actors, namely the "mygovindia" account. Although the degree distribution of this account is in second place after the "princearihan" account and only has a 20%-degree difference, the "mygovindia" account is the account that appears most often and is close to other accounts as intermediaries. It can be said that in this analysis period, the "mygovindia" account is the opinion leader. In addition, this account also has a closeness value = 1, which indicates its crucial position in this network. The diameter value (=1) indicates that the information submitted by this account will quickly get a response from other accounts/actors in the network.

Although the total diameter of this network is 4, the average path length is 1.47, which indicates that information that occurs in this network can be conveyed quickly over a short distance. Modularity analysis shows the clustering of actors in the network and the topic being discussed. The study results show that there are 648 communities in the network grouped by topic.

At this stage, this study has shown how entrepreneurial and startup activities can be analyzed by collecting or mining large amounts of data from social media. This dataset is then analyzed to obtain specific information according to its use, for example, to support decision-making and improve promotion. On the other hand, it is necessary to be careful in filtering and understanding the information obtained from social media (Sukumar, Jafari-Sadeghi, and Xu 2021b). The issues thrown into social media are very diverse, but the goal remains the same: to attract the public's or consumers' attention. Sometimes problems are raised by someone hoping certain social media actors will provide solutions. This type of discussion continues to evolve through various social media responses, such as tweets, tweet replies, likes, or hashtags on specific topics (Toder-Alon, Brunel, and Fournier 2014). According to Sarkar (2016), these digital interactions, actor relationships, and synthesis of ideas can be further interpreted for a purpose.

Finding key topics and the actors who talk about them the most can help provide new ideas for startup entrepreneurs. Startup entrepreneurs should consider the issues that develop on social media if they want to exist on social media. Taking advantage of developing positive points and counteracting the risk of negative business issues is like a double-edged sword for entrepreneurs. On the one hand, they want to benefit from social media, but on the other hand, adverse effects must always be accounted for. The faster process of information diffusion through social media is similar to electronic word-of-mouth marketing (S. B. Park, Jang, and Ok 2016). Specific communication patterns or trends can be identified if a businessperson has particular skills in performing data or text mining, visualizing the dataset, and using the analysis results (Amado et al. 2018; Shao et al. 2022). Unfortunately, these skills have not been an important concern for early-stage businesses. Therefore, this paper can provide a practical perspective for early-stage business entrepreneurs by offering less complex methodological skills and analytical tools that can be easily found. Many other social media such as Instagram, Youtube, and TikTok can be analyzed, which have more interesting algorithms that combine audio-visual formats. However, other methodological challenges are beyond the scope of this paper.
CONCLUSION

This study aims to analyze social networks on Twitter that discuss trending topics related to startups and entrepreneurship. In particular, this social network analysis attempts to identify critical actors and opinion leaders in these trending topics. The network visualization first shows several terms closely related to startups and entrepreneurship, such as entrepreneur, business, innovation, marketing, small business, and ecosystem, where these terms are theoretically and conceptually related to the terms startup and entrepreneurship. Next, social network analysis identifies vital network attributes such as centrality, modularity, and degree, which indicate which key actors are trending on a topic, how many network groups are discussing a trending topic, and the extent of the relationship between these actors and other actors in the network. Because of their central and vital position in the network, key actors who are trending topics and opinion leaders can be found.

In conducting this study, the author is aware of several limitations. Firstly, data collection from Netlytic, which has several limitations, ultimately cannot represent the actual state of the social network of actors who play an essential role in the trending topic of startups and entrepreneurship on broadly social media. Therefore, using complex data sources with a massive amount of data over a more extended period is highly recommended to obtain more robust research results. Secondly, this research is limited to the network attributes mentioned above. With richer data sources, future research can expand the identification of other network attributes that can help optimize decision-making for startup entrepreneurs who want to use social media and social network analysis to improve their digital marketing strategies.

ACKNOWLEDGEMENT AND DISCLAIMER

The author gratefully acknowledges the support provided by POLISS (https://poliss.eu), the funded project under European Union’s H2020 Research and Innovation Programme, Grant Agreement No. 860887. Views and opinions are only those of the author and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

The author would also like to thank Professor Tamás Sebestyén, editorial team, and anonymous reviewers for their constructive criticism and suggestions.

The Author declares that there is no conflict of interest.
REFERENCES


Eristian Wibisono

Discovering Key Actors and Opinion Leaders on Twitter’s Start-Up and Entrepreneurship Topics Trending: A Social Network Analysis Approach