Analyzing #FacebookDown: A Network Analysis of Information Flow and Community Structures on Twitter During Service Disruptions

Gayani Liyanage

1. Introduction

In the contemporary digital era, social media platforms serve as vital channels for global communication, and distractions to their services can have far-reaching consequences. The hashtag #FacebookDown on Twitter reflects a recent incident where users expressed their experiences during a service outage. This project aims to explore into the public sentiment and reactions encapsulated by this hashtag, using a approach of data collection, sentiment analysis, emotion classification, and network analysis. The significance of understanding user responses during such disruptions lies not only in addressing immediate user concerns but also in refining crisis management strategies in an interconnected digital environment.

The primary problem at hand is to comprehensively analyze and interpret the public sentiment and reactions associated with the #FacebookDown hashtag on Twitter. This requires gathering text-based posts from users using the hashtag and applying advanced analytics techniques, including sentiment, emotion analysis, and network analysis.

2. Methodology

2.1 Data Collection

In the pursuit of obtaining real-time data directly from Twitter, the original plan involved using the Twitter Developer API. However, the execution of this plan encountered significant challenges, primarily access restrictions and associated costs imposed by Twitter. Using the Twitter Developer API requires a paid subscription, making it less accessible for free exploration and analysis.

Twitter's restrictive policies regarding data scraping from a variety of reasons. First, Twitter places a premium on user privacy and data security. The API is designed to protect user information and prevent unauthorized access or misuse. Secondly, the sheer volume of data generated on Twitter makes it imperative to manage access efficiently to ensure fair and equitable distribution of resources. Consequently, the limitations on free access are implemented to maintain the integrity and performance of Twitter's infrastructure.

Given these constraints, an alternative approach was taken, leading to the discovery of a scrapped dataset available on GitHub. Although not as real-time as Twitter data, this dataset is still valuable for analyzing user interactions, sentiments, and network dynamics related to the #FacebookDown hashtag during service disruptions.

2.2 Preliminary Data analysis

In the data analysis phase, several key steps are executed to extract meaningful insights from the dataset. The initial step involves data cleaning, where rows with missing values in the 'tweet' column are dropped. This ensures that the subsequent analysis is conducted on a refined dataset, free from incomplete or irrelevant information.

Following data cleaning, attention turns to identifying the most liked tweets. This involves sorting the dataset based on the number of likes or reactions each tweet has received and selecting the top-performing ones. This step is essential for identifying the content that connects most with users, offering valuable insights into their preferences and engagement habits.

2.3 Description of methods

A subsequent analysis focused on visualizing the top 10 users who have tweeted the most. This plot serves as a representation of the most active contributors, giving a clear overview of user engagement and tweet volume. Understanding the top contributors helps identify influential users and provides context for the overall dynamics of the dataset.

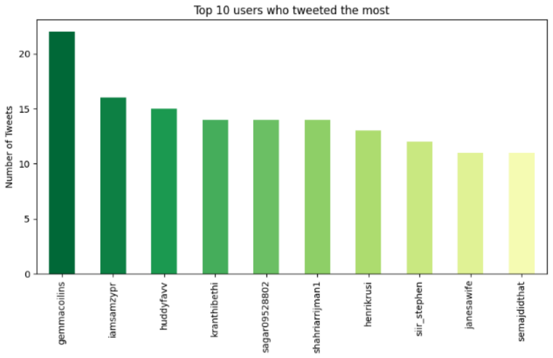

Figure 1 - Top 10 Users

Moving beyond user engagement, text data is processed further by removing stop words using Spacy, a natural language processing library. This step enhances the quality of the text data by eliminating common words that may not contribute significantly to the analysis. Following stop word removal, the analysis involves counting word frequencies to identify the most commonly used words in the dataset.

The next step is the visualization of the top 30 words. This can take the form of a bar chart and a word cloud, offering a visual representation of the most frequently occurring words in the dataset after the removal of stop words.

TextBlob's sentiment analysis, part of the project's implementation, involves using the TextBlob library to assign sentiment polarity scores to #FacebookDown tweets. It quantifies sentiments as positive and negative providing a concise overview of the emotional tone in the dataset.

Applied Natural Language Processing (NLP) techniques to predict emotions in the 'tweet' column. This involved analysis, capturing emotional expressions within the text data, contributing to a richer understanding of user sentiments during the #FacebookDown period.

After filtering the dataset to include the top 100 tweets based on 'likes_count,' a network graph was constructed. This involved mapping relationships and interactions among users who generated highly-liked tweets. The graph provides a visual representation, highlighting influential nodes, communication hubs, and broader network dynamics, contributing to a comprehensive understanding of user engagement during the #FacebookDown event.

3 Results

The presented bar chart illustrates the visualization of the top 30 words used in tweets during the breakdown period. The chart clearly shows that the words "Facebook," "Instagram," and "WhatsApp" are used most in the dataset. These three terms emerge as the most frequently utilized words in the tweets, suggesting a enhanced focus on these social media platforms during the specified period of service disruption.

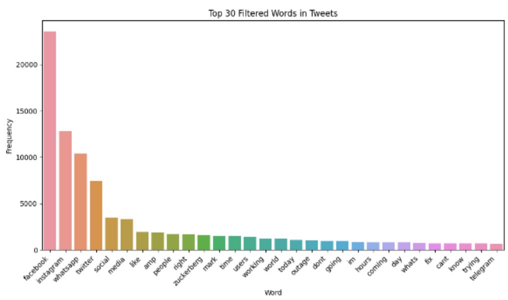

Figure 2 - Top 30 Filtered Words in Tweets

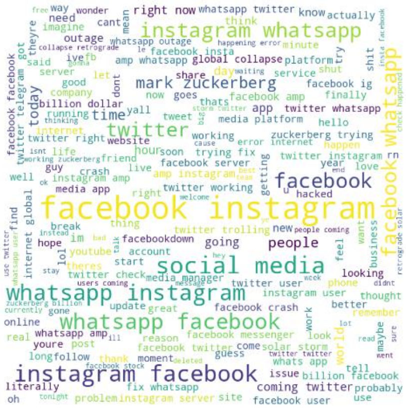

Figure 3 - Word Cloud

3.1 Sentiment Analysis

The sentiment analysis shows that people using the hashtag express a mix of emotions. Most of the discussions, around 68%, seem to convey dissatisfaction, frustration, or other negative reactions. On the positive side, about 32% of the content reflects positive sentiments, which could include expressions of resilience, optimism, or other positive reactions.

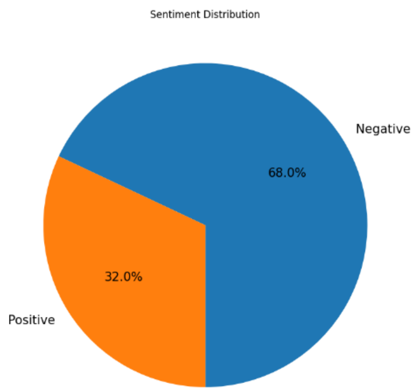

Figure 4 - Sentiment Distribution

The sentiment analysis looks at specific words and phrases that indicate different emotions. This helps us understand the discussions more thoroughly, pinpointing key themes, and finding the reasons behind the expressed sentiments. These insights give us a clear picture of how users feel during the period related to the hashtag Facebookdown.

3.2 Emotion Classification

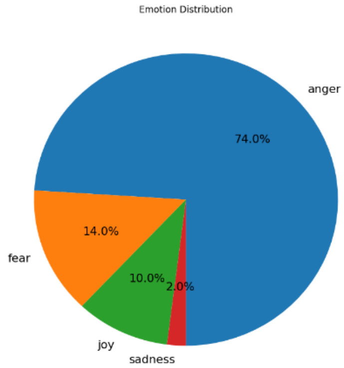

Figure 5- Emotion Distribution

The emotions in the hashtag discussions mostly revolve around anger, making up about 74% of what users express. This suggests a lot of frustration and dissatisfaction. Fear is in about 14% of the emotions, indicating some concerns or uncertainties. On the positive side, joy makes up 10%, showing that despite the overall negativity, there are moments of happiness or satisfaction in users' posts.

Moreover, there's no explicit mention of sadness (0%), meaning that during this period, users didn't explicitly express sadness in their hashtag-related discussions. So, in a nutshell, these emotion insights give us a quick look into the varied emotions users share, with anger being the dominant one, some fear and joy.

3.3 Network Analysis

After selecting the top 100 tweets based on the 'likes_count,' a network graph was drawn, revealing a network with 42 nodes and 147 edges. This network provides a visual representation of the interactions among users who generated the most-liked tweets. Each node represents a user, and edges between nodes indicate interactions or connections, potentially in the form of replies, retweets, or mentions.

Within this network, several hashtags prominently stand out, reflecting the recurring themes in the discussions. The most frequently used hashtags include "whatsappdown," "facebookdown," "serverdown," and "instagramdown." These hashtags serve as key markers during the specified period, pointing towards service disruptions or outages on popular social media platforms.

The dominance of hashtags like "whatsappdown" and "facebookdown" suggests a focus on user reactions and discussions related to temporary breakdowns or disruptions on these platforms. The inclusion of "serverdown" emphasizes the broader technical aspect, indicating discussions around server-related issues.

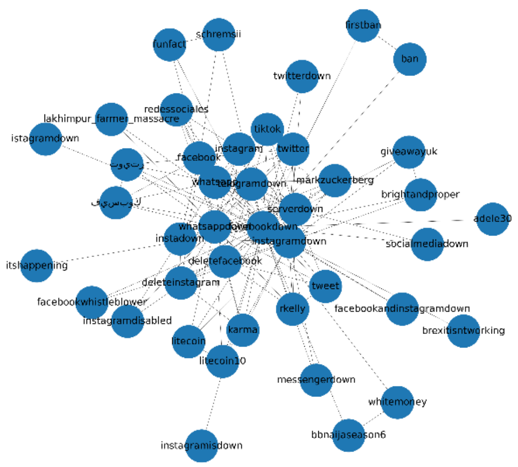

Figure 6 - Network Analysis

4 Conclusion

In conclusion, this project expolred into the public sentiment and reactions encapsulated by the #FacebookDown hashtag on Twitter, on user experiences during a service outage. Despite the initial plan to obtain real-time data directly from Twitter using the Twitter Developer API, limitations and access restrictions led to an alternative approach utilizing a scrapped dataset available on GitHub. The subsequent data analysis, spanning from the identification of the most liked tweets to sentiment and emotion analysis, provided a best understanding of user interactions, preferences, and emotional expressions during the specified period.

The sentiment analysis revealed a mixed emotional landscape, with approximately 68% of discussions expressing negative sentiments and 32% conveying positivity. Further emotion classification highlighted that anger dominated, comprising about 74% of the expressed emotions, followed by fear (14%) and joy (10%). Lastly, the network analysis, based on the top 100 liked tweets, showed a dynamic interaction graph with 42 nodes and 147 edges. Moreover, hashtags like "whatsappdown," "facebookdown," and "serverdown" emerged as key markers, reflecting user discussions around service disruptions on various platforms.

In essence, this comprehensive analysis provides valuable insights into user behaviors and sentiments during service outages, contributing to a deeper understanding of crisis management strategies and the broader digital landscape. The visual representations, including network graphs and emotion distributions, offer a clear snapshot of the user interactions, enhancing comprehension of the collective user experience during instances of social media disruptions.