

Characteristics of online social network during a pandemic

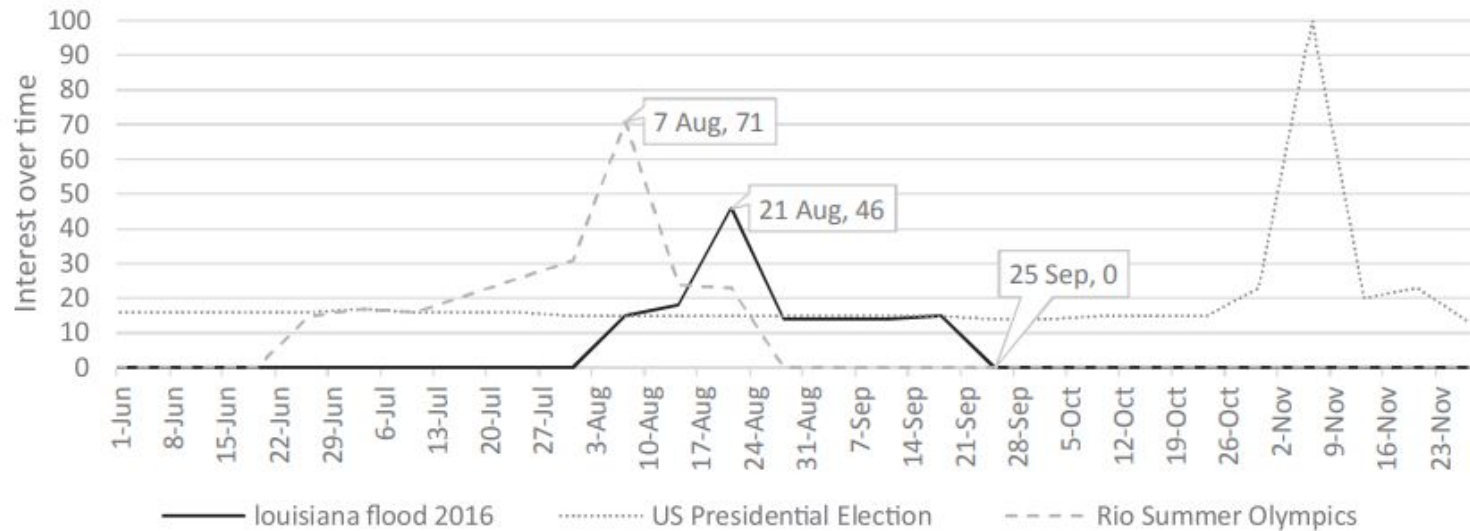
17PT17 - Muthualagesan Suryavarathan

17PT28 - Sanjay Kumar S

Roadmap

1. Problem Statement
2. Motive
3. Dataset Collection
4. Proposed Work
5. Technology Understanding
6. Existing Results
7. Conference Identification
8. Tools and Technologies

Search-term comparison during 2016 Louisiana flood in the city of Baton Rouge, Louisiana, USA



Problem Statement

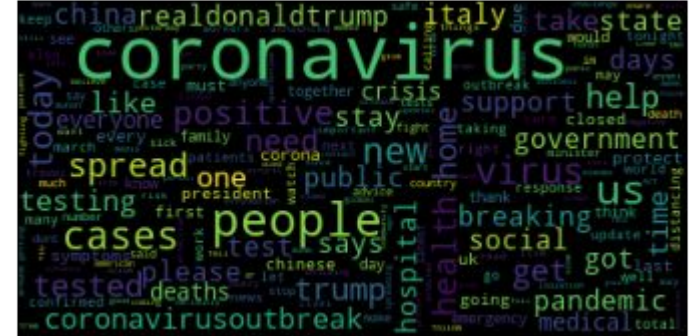
- To analyse the characteristics of popular social media such as Twitter during a pandemic.
- Apply social network analysis to convert social network data into knowledge.
- Explore patterns created by the aggregated interactions of online users on Twitter.
- Provide insights to understand the critical role of social media use for emergency information propagation.

Why Social Network Analysis?

- The core of a social network consists of numerous individuals who are actively engaged to share and update information.
- Emergency agencies and organizations are on the periphery of the social network, connecting a community with other communities.
- The results of this study will help emergency agencies develop their social media operation strategies for a pandemic response plan.

Dataset Collection

- The Twitter API has been exploited to collect tweets that mention hashtags #COVID19, #Coronavirus etc.
 - The tweets extracted are then preprocessed by removing emojis, punctuations, URLs and Reserved words such as (RT, Fav) using Regex.
 - The clean data can now be used for various analysis purposes (Sentiment, Natural Language Processing, Word Clouds)
 - Users as Social Actors (nodes in a network)
 - Comments, Retweets and Mentions as Edges.
-



Proposed work

- Our analysis should identify top topics related to COVID19, which we group into different themes.
- Find mean sentiment of these topics/terms.
- Which topic has highest and lowest mean of likes?
- Top words and pair of words used.
- Which users have the highest influence in the network?
- Analyzing community structures within the network using various SNA measures.

Social network analysis and tools

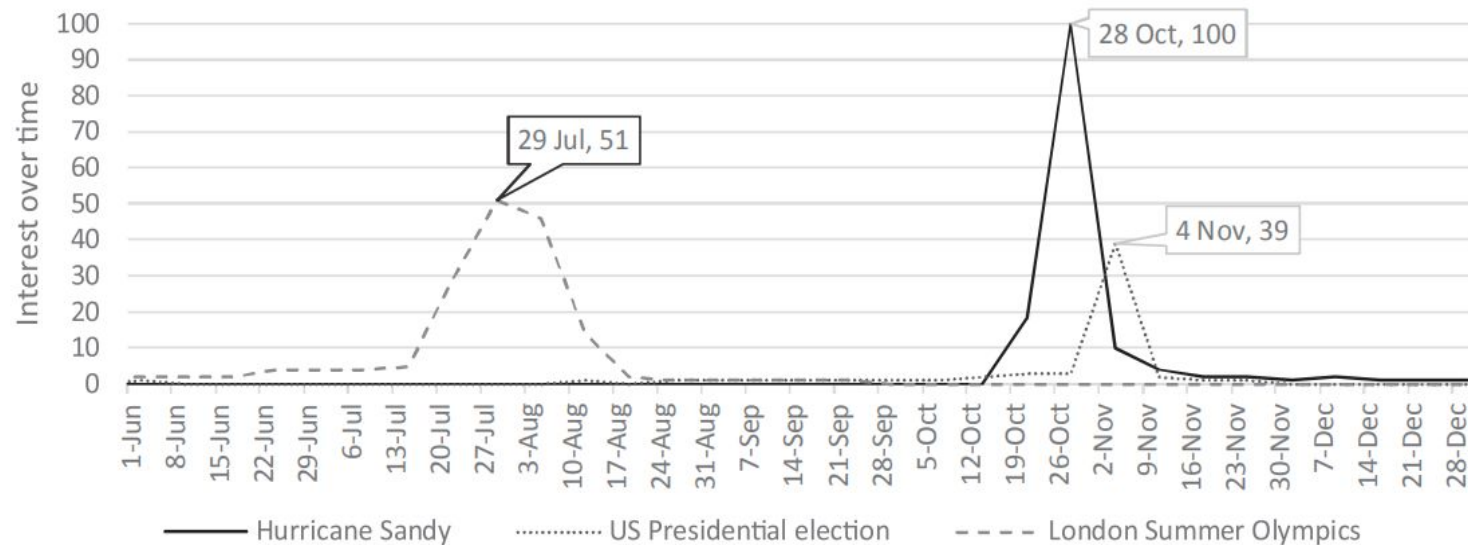
- Software and tools have been developed to fulfill the increasing need for social network data mining and visualization technology.
- GUESS - Gython (an extension of Python, or more specifically Jython)
- Prefuse - Java. Data modeling, visualization, and interaction.
- JUNG - Java Universal Network/Graph Framework.
- NODEXL - Advanced tool. Not open source.
- Gephi - Easy creation of social data connectors to map community organizations

Google Trends

- Trends data is an unbiased sample of our Google search data. It's anonymized (no one is personally identified), categorized (determining the topic for a search query) and aggregated (grouped together).
- This allows us to measure interest in a particular topic across search, from around the globe, right down to city-level geography.

<https://trends.google.com/trends/explore?date=2012-01-01%202012-12-31&q=London%20summer%20olympics>

Search-term comparison during 2012 Hurricane Sandy in the U.S.



Comparison of Facebook and Twitter

- As of Feb 2021, Facebook has 2.60 billion monthly users, while Twitter has 353 million monthly active users.
- 66% of Facebook users are on the platform on a daily basis, compared with 53% of Twitter users.
- More users implies more information which results in easy spread of false information.
- Twitter is relatively a better space for information sharing.
- Hence Twitter is chosen for analysis.

Degree centrality

- **Degree centrality** refers to the number of edges a vertex has to other vertices.
- **In-degree** is the number of incoming edges incident to the vertex .
- **Out-degree** is the number of outgoing edges incident to the vertex.
- **Betweenness** quantifies the number of times a vertex acts as a bridge along the shortest path between two other vertices.

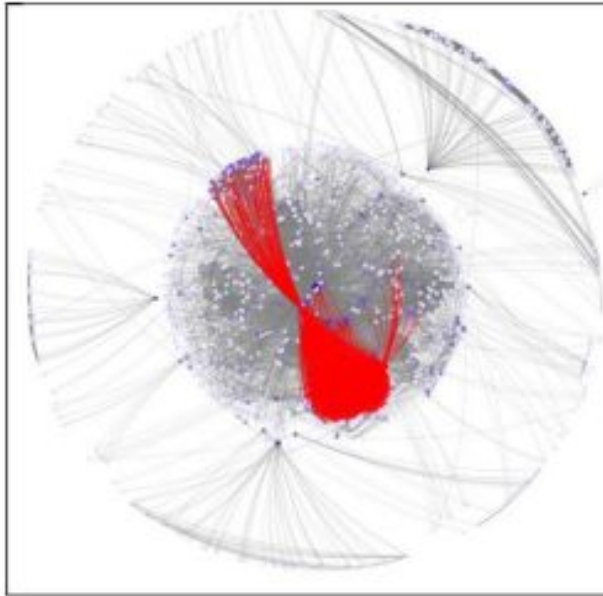
Girvan–Newman algorithm

Provide an informal description of the 10 largest groups, which account for about 38% of the entire network.

Top 10 largest communities in the social network.

Rank	Size	Description
G1	144	Flood inundation map, information of debris separation, shelter locations
G2	63	Commenters on the flood inundation map – e.g., map update requests and sharing map information
G3	43	Donations and supports
G4	40	Road conditions (road closed/open)
G5	34	Locations of debris removal, debris collection status map
G6	33	Ordinances to help Baton Rouge residents; housing, noise ordinance waivers, waiving permit fees for structures damaged, policy changes
G7	30	Debris separation, Louisiana Department of Environmental Quality
G8	28	Reactors to hiring workers to help with debris removal efforts
G9	16	Commenters on the debris removal hiring event
G10	15	City events after final debris collection

Top 10 out-degree centrality and degree distribution



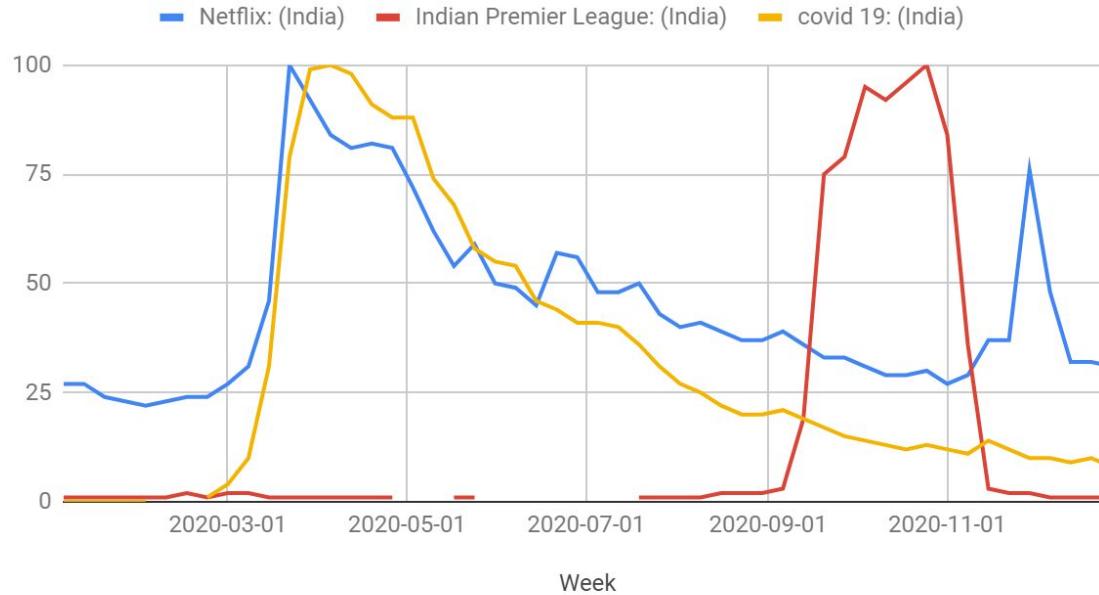
Users	Out-degree
CBR	228
User#1	153
User#2	150
User#3	149
User#4	147
User#5	147
User#6	146
User#7	145
User#8	144
User#9	143

Top words and word pairs

- Text analysis identified 77% of the posts during emergency responses as having positive words.
- The top five words during disaster responses are map, water, thanks, GIS and flooded.
- Most word pairs are related to flood, disaster recovery team and disaster debris removal in the city.

Google Trend for COVID-19

Google Trends



Conferences Identified (So Far)

- ASONAM 2021 - The 2021 IEEE/ACM International Conference on *Advances in Social Networks Analysis and Mining* 08-11 November 2021, The Hague, Netherlands

Tools and Technologies

- GUESS - <http://graphexploration.cond.org/>
- NODEXL - <https://nodexl.com/>
- Prefuse - <http://kneecap.ist.berkeley.edu/>
- JUNG - <http://jung.sourceforge.net/>
- Gephi - <https://gephi.org/>

Bibliography

1. Jooho Kim, Makarand Hastak, **Social network analysis: Characteristics of online social networks after a disaster**, *International Journal of Information Management*, Volume 38, Issue 1, 2018, Pages 86-96, ISSN 0268-4012, <https://doi.org/10.1016/j.ijinfomgt.2017.08.003>.