

Social Network Analysis

By Albert Okiri Olweny

Use Case: World Cup

Introduction

The growth of social media platforms in conveying the sentiments of the market presents an opportunity for businesses and social initiatives to make better-informed and accurate decisions. Social networks unearth opinions from relevant communities, predominant notions and popular entities within them. This document details the use of graph analysis of tweets from the Twitter API endpoint. The dataset used in the analysis is retrieved based on keywords or communities of interest.

From contemporary trends and events, *The World Cup* is chosen as an ideal use case to illustrate the effectiveness and value of this social network analysis. The data was collected on 24th October 2022 and features 12000 tweets within a span of the previous 7 days. The grouped (*world cup*) keywords search for a combination of the two words sequentially in a tweet.

Analysis

Graph Analysis

A table documents the centrality analysis for in-degree and out-degree measurements for tweet reply, quote, mention and retweet. These insights identify the most influential nodes that can be focused on depending on the interaction types. As pictured in Figure 1, a user can select the interaction type to analyze the centrality measures and network visualization. Figure 2 showcases the network visualization that permits a close examination of the network relationship between various nodes depending on interaction types.

Social Network Analysis

Interaction Analysis

Interaction types are analyzed based on tweet mentions, quotes, retweets and replies.

Select Interaction Type

Mention

In-degree measures of a node indicate the connections from other nodes towards it.

Out-degree refers to directed connections from a node.

In-Degree Measures

Name	Centrality
FIFAWorldCup	21
T20WorldCup	15
FIFAcOm	11
imVkohli	10
cryptoskullx	9
PeterTatthell	9

Out-Degree Measures

Name	Centrality
RadjaRasta_	24
DireWolf2523	17
jhoburgh	16
nasser_mo3gza	15
BilalBh10407815	14
lurv50NTwester	12



Figure 1 The user interface that features interaction type selection and tables of centrality measures.

Select the layout of the network graph

Cose

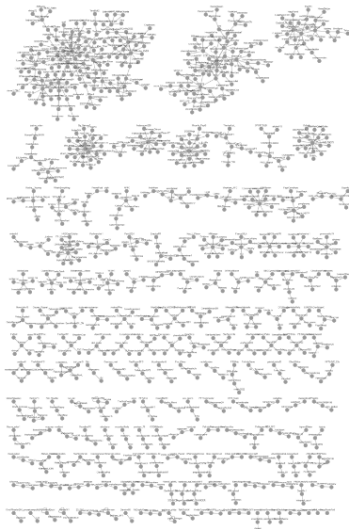


Figure 2 The user interface of interactive network visualization and selection of layout of the graph.

Topic Modelling

Topic modelling from Latent Dirichlet Allocation (LDA) algorithm reveals the salient discussions and issues that are being discussed within a given conversational network. Consistency in topics

creates an accurate perspective on the audience's sentiments. These topics are presented as a sorted list of 15 grouped topics.

Topic Modelling - LDA Analysis

Organizing tweets based on salient topics that are prevalent.

Topics

0: world; fifa; cup; documentary; hosting; <https://twitter.com/netflixuk/status/1584945385184837633/video/1>; power; politics; struggles; cup
 1: world; cup; fifa; team; iran; #worldcup2022; ban; 2022; #opiran; #opfifa
 2: world; cup; one; messi; winning; football; would; ever; greatest; last
 3: world; cup; —; «; »; think know; years; cup; jungkook
 4: cup; world; x; top; #eth; sport; new; get; called; project
 5: cup; go; form; back; world; kohli; i'm; end; t20; let's
 6: world; cup; still; saw; like; wins; de; messi; goat; tatchell
 7: cup; world; messi; games; club; leo; goals; assists; days; ⚽
 8: world; countries; african; fifa; country; cup; win; would; players; help
 9: lionel; call; people; it's; messi's; world; right; cup; something; also
 10: world; win; cup; league; goal; ballon; month; final; 3; 8th
 11: world; cup; man; starts; give; 31; spot; unbeaten; #cfc; chelsea.
 12: world; cup; qatar; 2022; fans; england; could; first; t20; cup;

* The data is retrieved from Twitter Platform

Figure 3 The list documenting the salient topics from LDA analysis of retrieved tweets.

Synthesis of Reported Data

The centrality measures depicted in the tables when combined with the network visualization reveal the most influential nodes depending on the interaction types. These can inform the marketing approach to use by identifying specific users and how they are fundamental to the networks that are of interest.

For instance, *RadjaRasta_* has the highest out-degree value for the *mention* interaction type. This indicates the user's engagement by mentioning the keywords used in retrieving the tweets. A closer examination of the network visualization in Figure 4. shows a subset of the network that is not the giant component of the community. This can be relevant in tapping into the various sub-communities that are present. Focusing on users who are prevalent on the largest network, specific sub-networks, or those with the highest centrality in all sub-networks depending on the incentives can be realized.

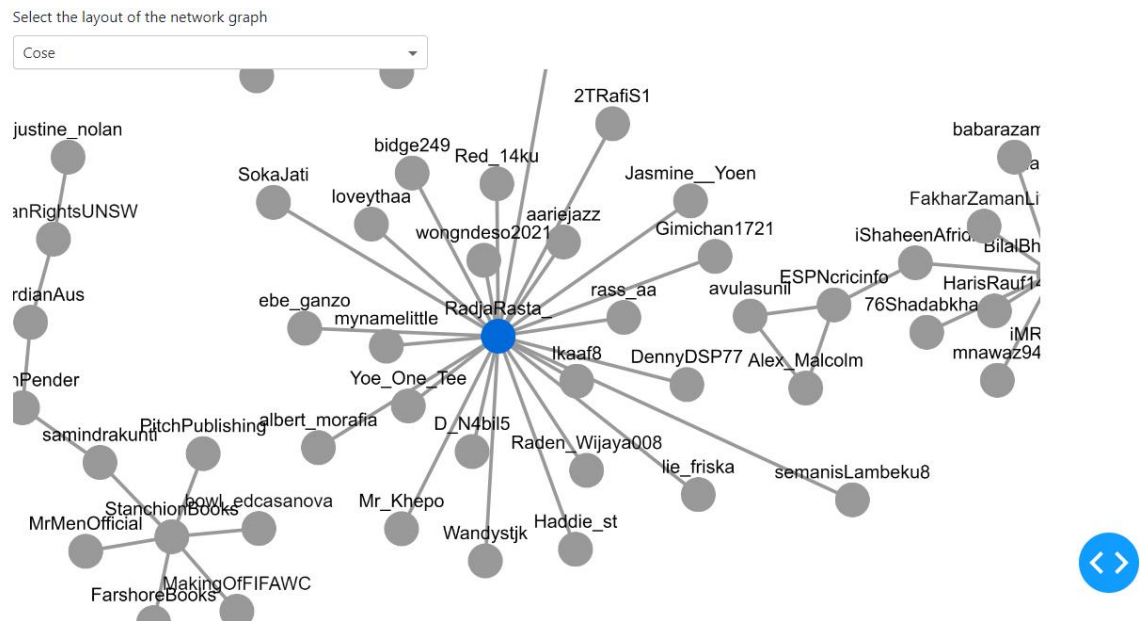


Figure 4 Interactive sub-network visualization.