Reddit Hyperlink Network Analysis
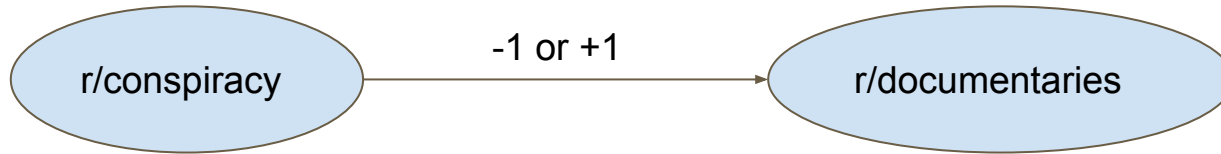
Kritika Gupta (201601129)

reddit.com

- **Reddit** : social news aggregation, web content rating, and discussion website

- Divided into communities ("subreddits") formed on the basis of areas of interest

    - Eg. r/politics, r/worldnews, r/music, r/networkscience

- Posts in a subreddit can contain hyperlinks to other subreddits.

    - *"Come look at all the brainwashed idiots in r/Documentaries. Seriously, none of those people are willing to even CONSIDER that our own country orchestrated the 9/11 attacks. They are all 100% certain the "turrists" were behind it all, and all of the smart people who argue it are getting downvoted to the depths of hell. Damn shame. Wish people would do their research. Here's the link."*
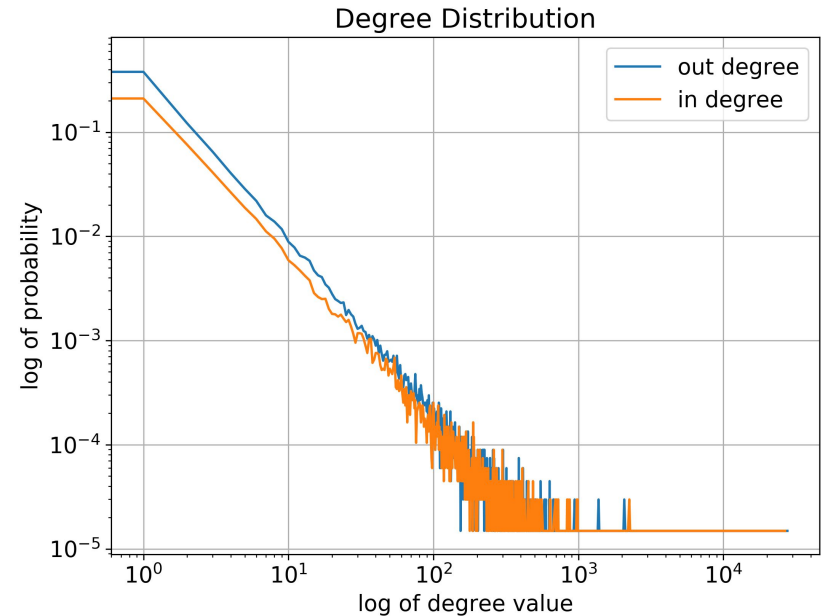
Reddit Hyperlink Network (from SNAP)

- Nodes : Subreddits

- Edges : Posts that create a hyperlink to another subreddit

    - directed links from source subreddit to target subreddit

- Edge weight : Sentiment of the post. It can be negative (hateful attack) or neutral/positive (appreciation)

- Graph is a directed, weighted, multigraph.
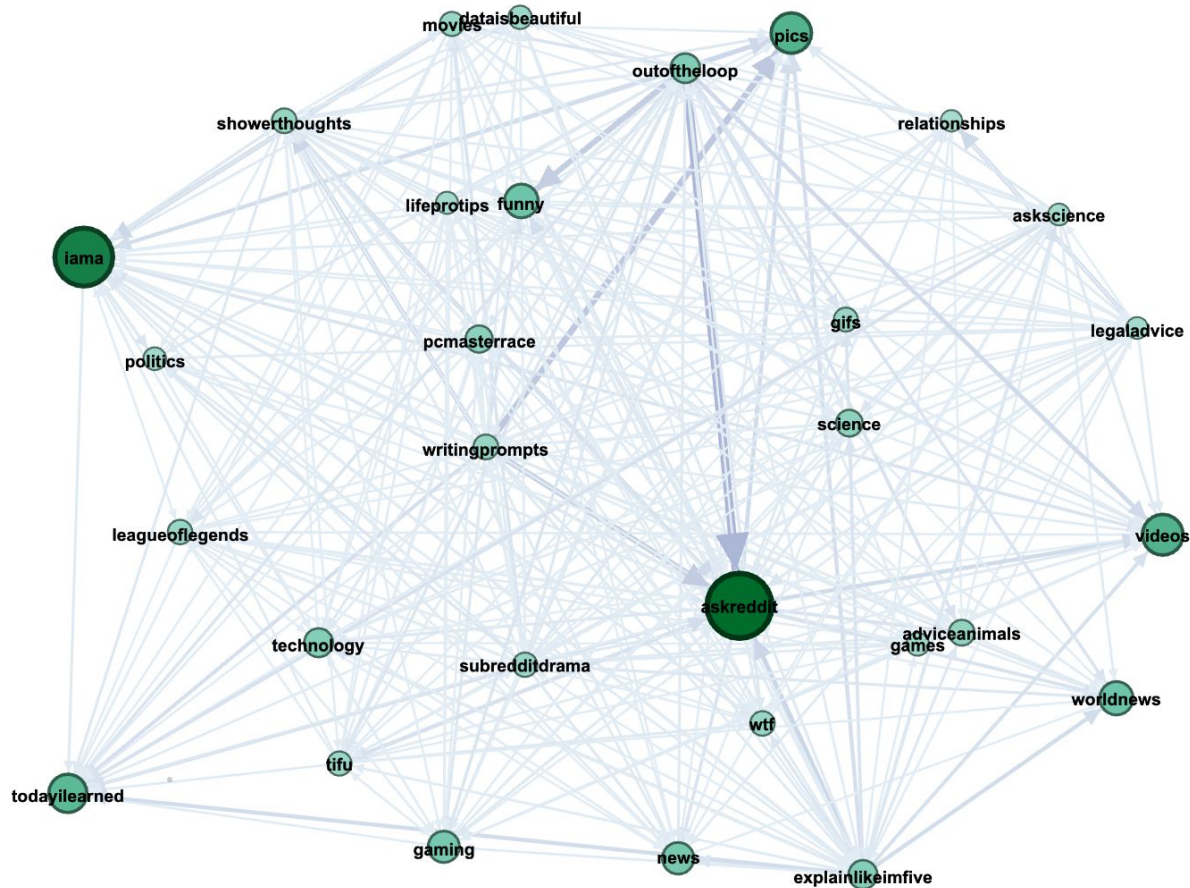
Network Properties

- Nodes = 67,180

- Edges = 8,58,488

- # of connected nodes = 3,39,643

- Network density = 0.00019023 →
  Sparse network

- Weakly connected components = 712

  - 97% of nodes in giant weakly connected
    component

Degree Distribution

PageRank

Subreddit	PageRank Centrality
askreddit	0.021
iama	0.018
pics	0.011
funny	0.009
videos	0.009
todayilearned	0.006
gaming	0.006
worldnews	0.005
science	0.004
news	0.004

TABLE I: PageRank Centrality

Who spreads and receives the most hate?

- Only 10% of the subreddits cause 80% of the negative sentiment posts.

Subreddit	% of attacks caused
subredditdrama	10.932
bestof	4.214
drama	3.272
circlebroke2	2.152
shitpost	1.591

TABLE IV: Subreddits that cause the most negative links.

Subreddit	% of attacks received
askreddit	4.094
worldnews	2.073
news	1.975
pics	1.794
todayilearned	1.714

TABLE V: Subreddits that are hit by the most negative links.

- Above lists are determined by absolute number of negative outgoing and incoming edges.

- Normalizing by total degree, subreddits such as r/*mongolianhatewatch, r/redditliars* post only negative sentiments, but number of posts is negligible.

- Similarly, subreddits r/*databreaches, r/zionismexposed* have only negative incoming edges, but the number of posts is negligible.

Reciprocity

- Negative reciprocity : does a subreddit hate those which hate it? Is it hated by those which it hates?

Negative Reciprocity	Number of nodes
0	91%
< 0.1	75%
1	0.6%

Positive Reciprocity	Number of nodes
0	78%
< 0.3	75%
1	2%

All negative edges are birdirectional

All positive edges are birdirectional

- Negativity is not mutual in most cases .

Negative Reputation of Subreddits

- Negative sentiment posts can target subreddits and be means to discredit or defame them in the eyes of users.

- Model user behavior as a **random walk** across Reddit by following hyperlinks in posts.

- Rules

  - A browsing session consists of following hyperlinks found with positive/neutral sentiments (don't follow negativity).

  - Encountering a negative sentiment towards a subreddit forms a negative impression of the target subreddit.

- Which subreddit will gain the worst reputation?

- How many steps (~time spent in browsing subreddit) are required?

Results

Subreddit	Reputation	Steps per user
askreddit	-282	19.0
pics	-129	18.0
worldnews	-120	18.0
funny	-117	17.0
videos	-113	19.0
todayilearned	-102	17.0
adviceanimals	-84	19.0
wtf	-82	18.0
news	-72	22.0
iama	-71	17.0

TABLE VI: Random Walk Simulation Results