

Top Trending of Twitter Friends

● Basic Functionality of Twitter

- Write Status as a Tweet.
- See and Reply to Friends' status.
- Discuss over the same topic through Hashtag.
- Top Trending of a particular location.

● Motivation

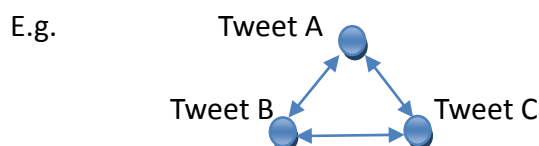
- Deficiency: Worldwide or location based Top Trending may not be interesting to Users.
- Fact: User cares more about the current trending of their friends.
- Proposition: Allow user to see the top trending of friends.
- Define "Hot" Topic among friends:
 1. High Term Frequency: Appear in friends' tweets more times.
 2. High Tweet Rank: Share Hashtag(s) with many other tweets; Replied by many users.

● Top Trending based on Tweet Rank

- For all the tweets, we construct a Tweet Graph to represent their relationship. Construct Tweet Graph as follows:
 1. A tweet is a node, the reply and Hashtag sharing will be represented as edges.
 2. If TweetA reply to TweetB, there will be an edge pointing from TweetA to TweetB.

E.g. Tweet A   Tweet B

3. For all the tweets sharing the same Hashtag, there will be a two direction edge pairwise assign to these tweets. Tweet A, Tweet B and Tweet C share the same Hashtag. Their graph will be build as follows:



- Rank Tweet using Tweet Graph

1. Algorithm is similar to PageRank Algorithm.
2. Random Walk
 - 1) Start at a random tweet.
 - 2) Follow an out edge with equal probability.
3. For each tweet, assign a value between 0 and 1 as its Rank Value.
4. Outcome: Rank Vector of all tweets.
5. Since tweets are within 24 Hours, the calculated rank has a real time popularity rate.
6. Teleport for Isolated Tweet:
 - 1) Reason: Even though these tweets are not replied, not reply to others or not containing Hashtag, it's still possible these tweet are related to some hot topic.
 - 2) Operation: Assign a random probability for an isolated tweet to all other tweets.

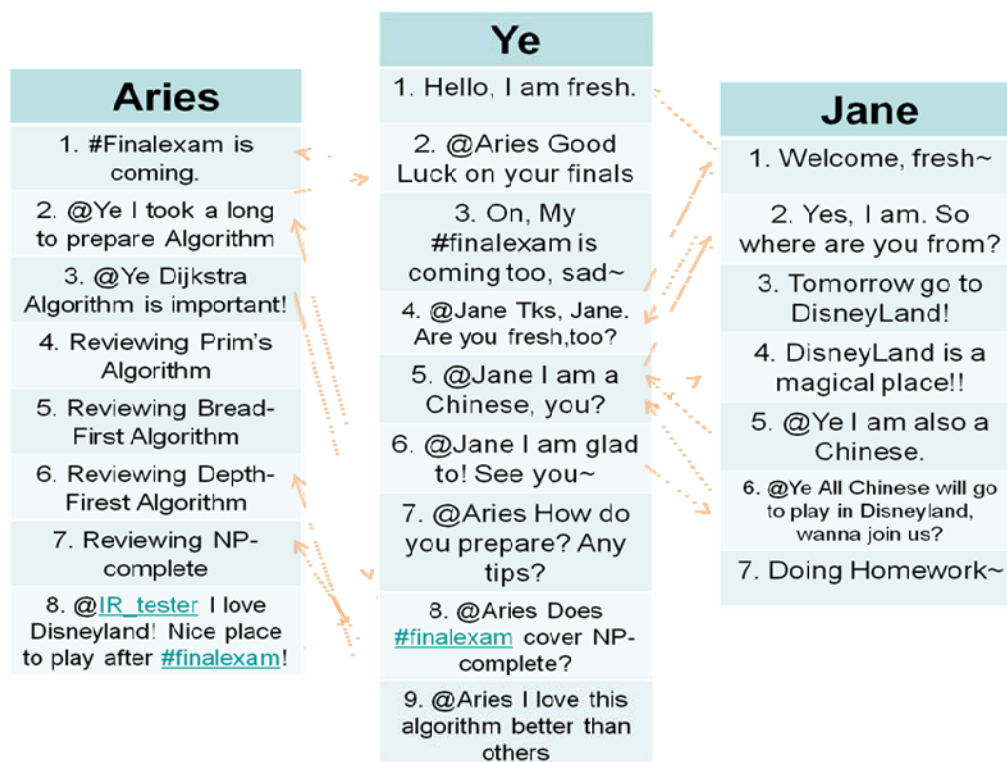
- Term-Frequency in Tweets
 1. Intuition: Terms appears more frequently in tweets tend to be more important too.
 2. Term_Term-Frequency Adjacency List: Parse all terms in all tweets.
Data Structure: [term, a list of <term_id_i, term-frequency>]
- Top Trending based on the score of terms
 1. Intuition: Terms with top ranking are Top Trending (i.e. hot topic)
 - 1) More frequently appear in User and user's friends tweets
 - 2) Hot Hashtag may be cited by many tweets.
 - 3) Replied by many people.
 2. Scheme: For each term, the score is the sum of the multiplication of term frequency and TweetRank for each tweet on the tweet list.

$$score = \sum_{i=n} (TermFrequency) * (TweetRank), i \in TweetList$$

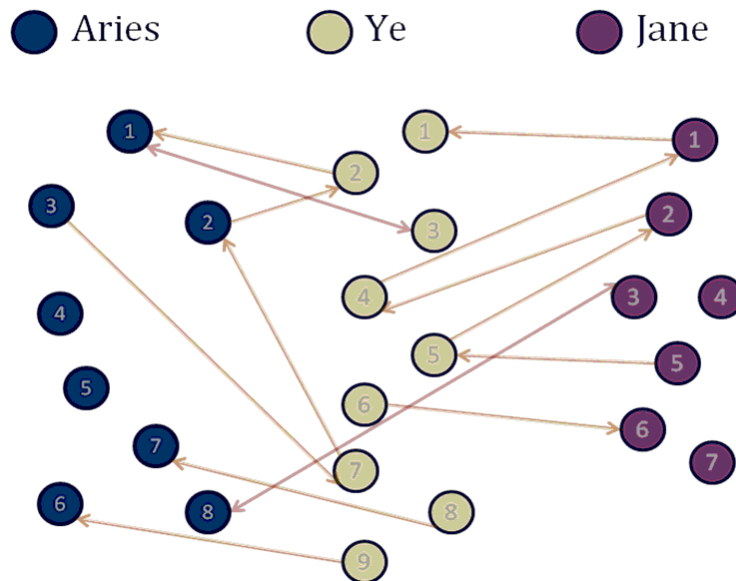
- Implementation with Twitter API
 1. Jtwitter: A small library providing easy access to the Twitter API
<http://www.winterwell.com/software/jtwitter.php>
 2. Twitter
Method: getFriendsTimeline()
 3. Twitter.Status
 - 1) Field: id
 - 2) Field: inReplyToStatusId

● Demo

- Scenario: Three Real User in Twitter, and they are following each other. Their tweets' relationship created as follows. Their central topics are #finalexam, fresh, Disneyland , Algorithm, Chinese(nationality).



- The Tweet Graph based on this scenario is:



- Running Results: The ranked Top Trending is
 1. #finalexam
 2. Fresh
 3. Disneyland
 4. Algorithm
 5. NP-complete
- Observation: The result is very similar with the actual topics among user's status and conversations (we were talking about the final exam and what we were going to do after the finals). Our tool is able to find the Top trending of our twitter friends.

● Limitations

- Twitter Limit Rate
Twitter API only allows clients to make a limited number of calls in a given hour. API Requests: 150 per hour.
- Common Words Elimination
Small corpus for Common Words, current capacity is 60 common words.
- Top Treading is based on single words.
It's better to use phrase.

● Conclusion

In this final project, we proposed a concept of Tweet Rank to rank tweet using Tweets Graph which is build on the connections among users and their friends' tweets. Considering the frequent term is also important, we rank terms with Tweet Rank combined with term frequency in tweets. In this way, we calculate Top Trending of friends of a particular user. This is a feature that current Twitter doesn't provide, but may be useful and meaningful to users.