

Appendix C : A Visual Analytics Tool

Imagine we are interested in comparing the popularity of a set of hashtags over time in a specific set of tweets. Fig. A shows the trends for say 200 hashtags (generated by a brownian motion). It is evidently difficult making a sense of trends when the number of time series is large. However, by using the hashtag-co-occurrence network (Fig. B) which is extracted from the tweets, one can construct an ontology (Fig.C) for the hashtags . **Here the Community labels are determined by the labels of children in that community that have the highest page rank (an example of Bottom-Up Information Diffusion).**

Fig. A

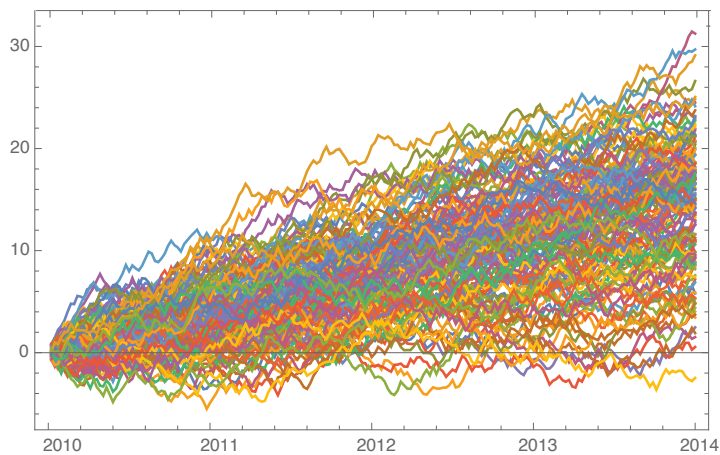
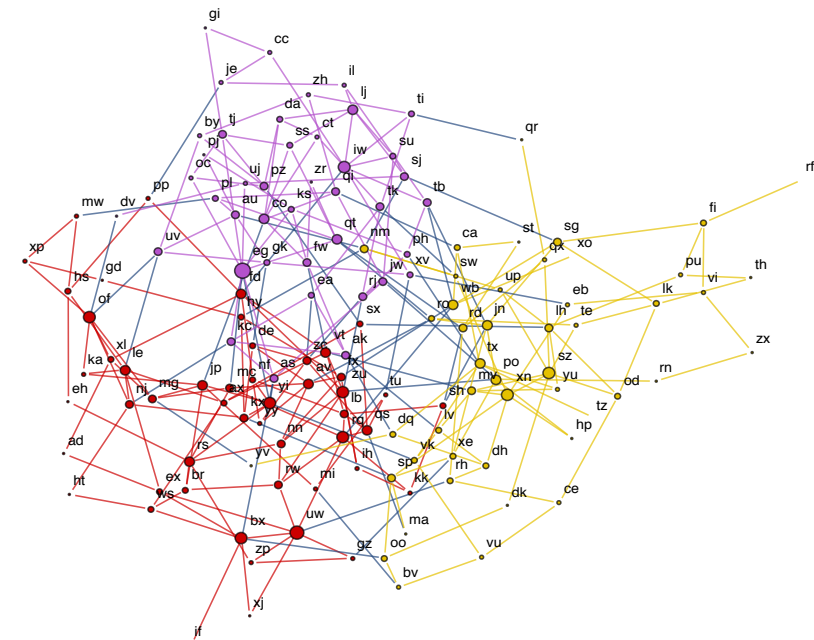
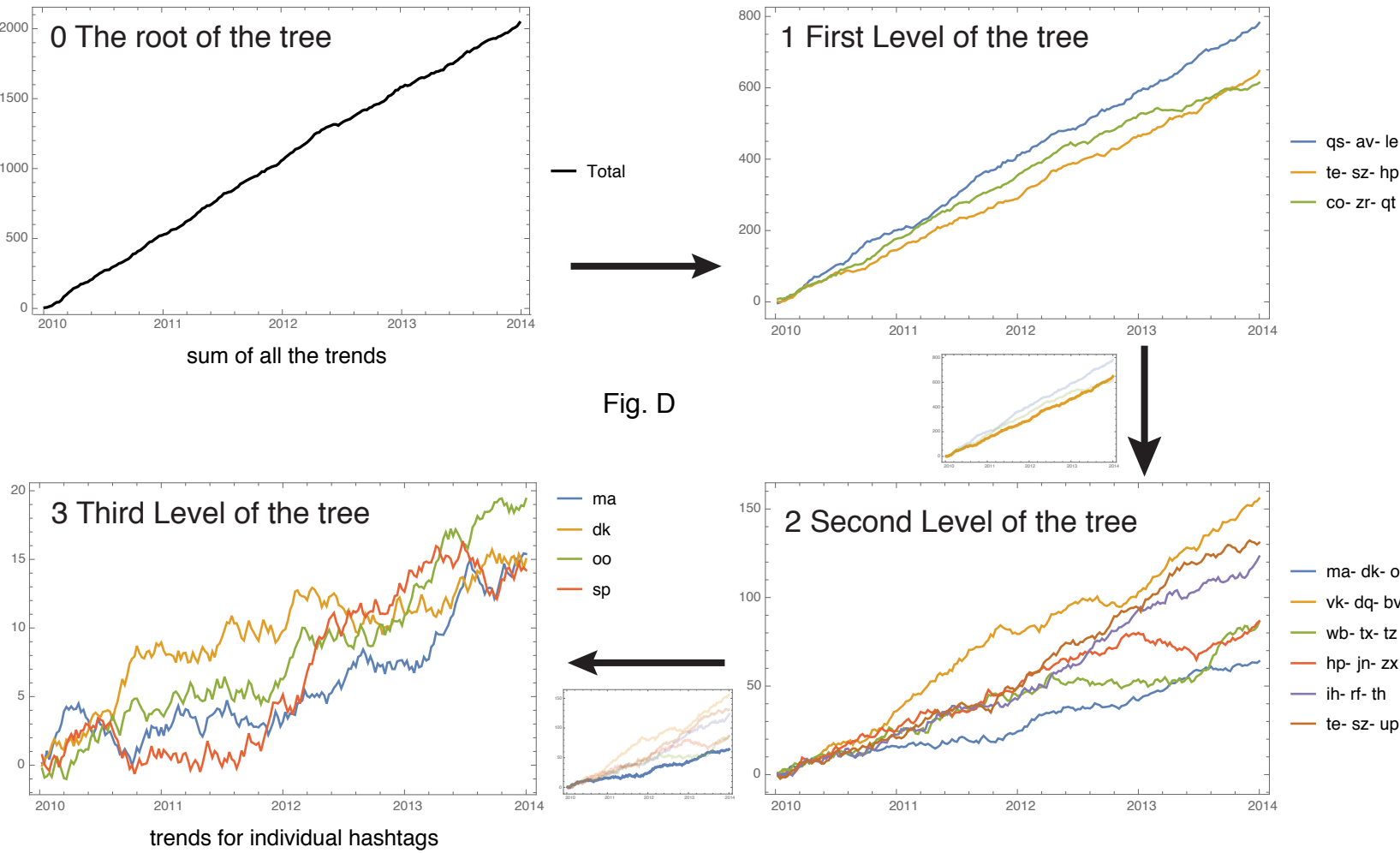
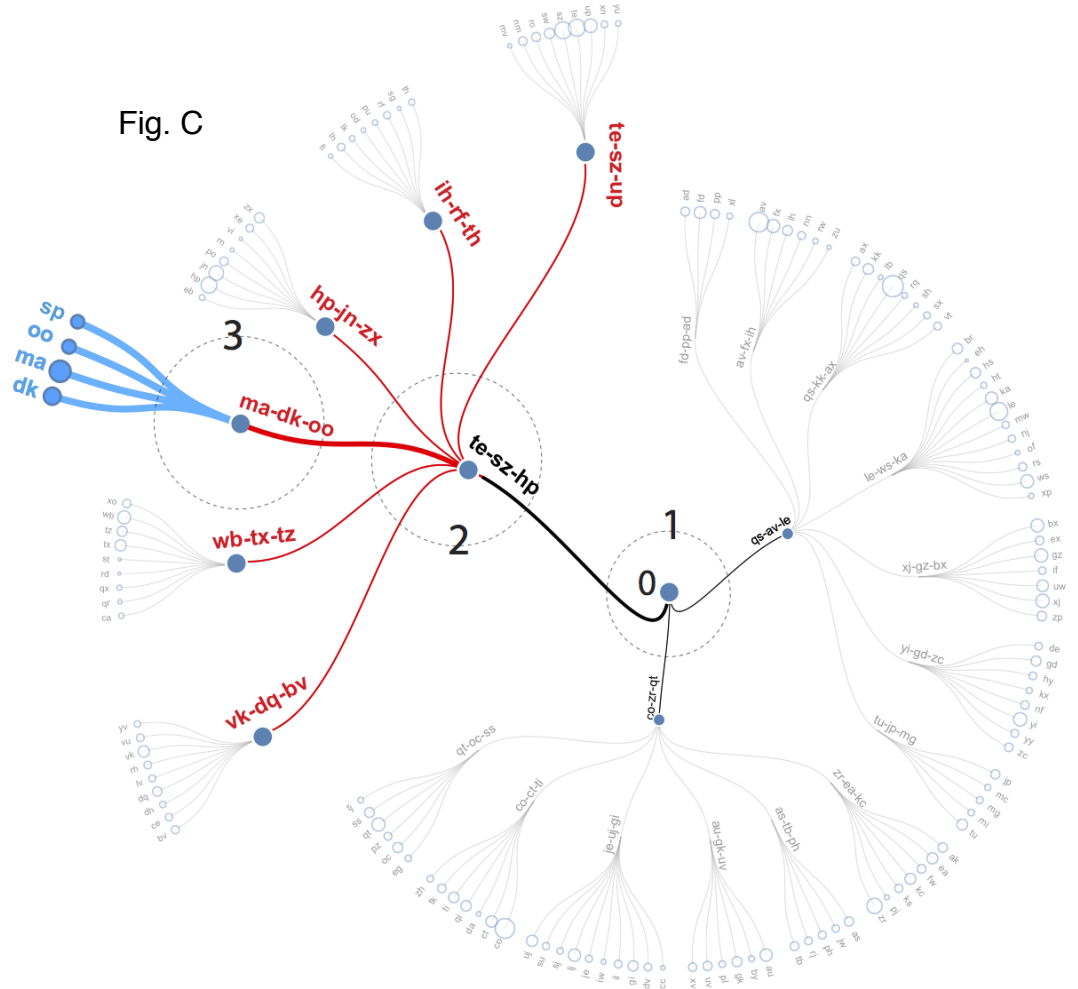


Fig. B



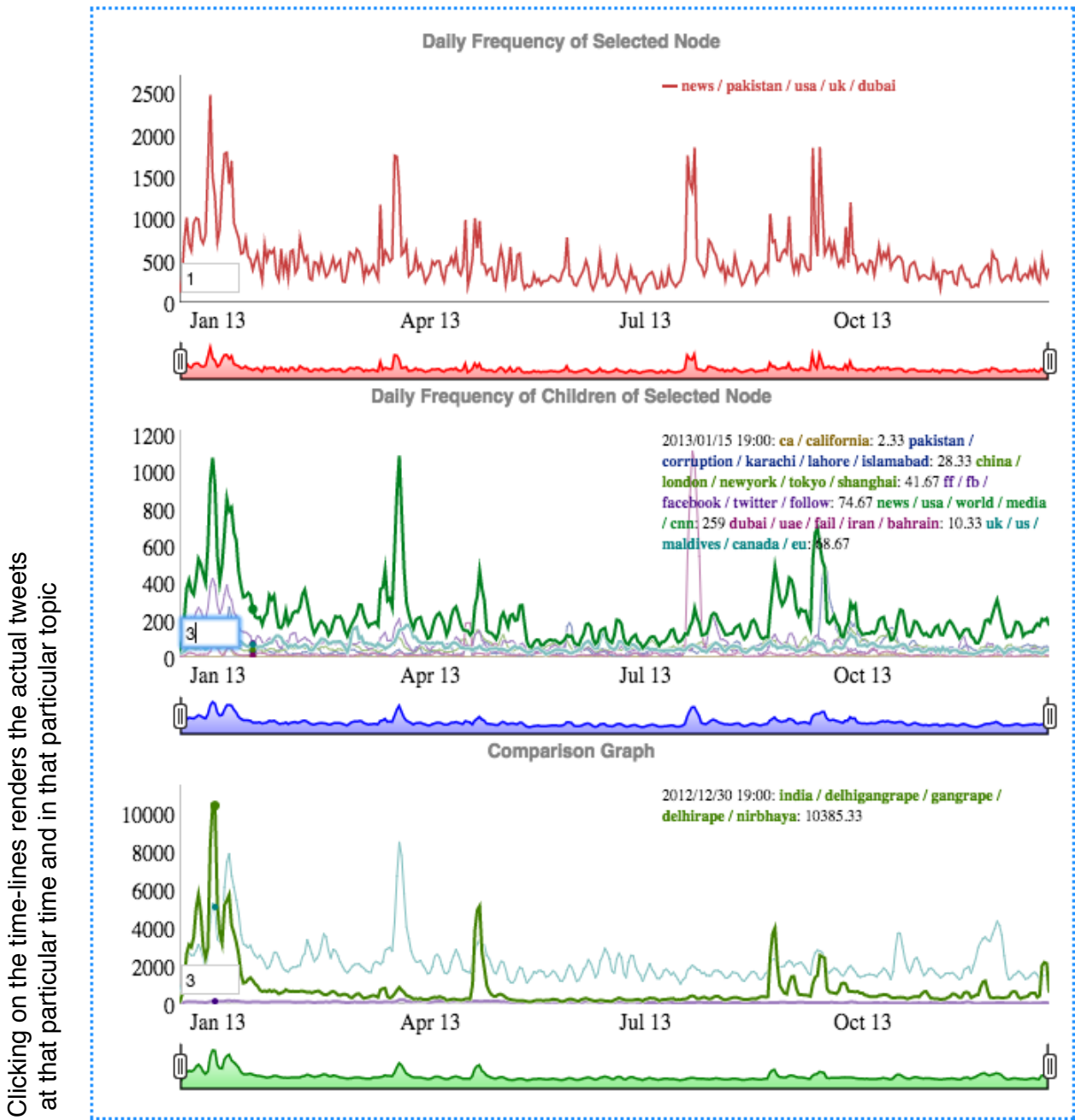
Here the size of the nodes shows their pagerank. Label of each community is composed of the labels of the nodes with top 3 pagerank.

Fig. C



Now by simply taking the trends of the hashtags as an attribute for each tag and applying BUID by recursively aggregating them over the communities we will enable the user to obtain an overall image of the trends by navigating through the tree(Fig.C). In particular, when the user expands a node on the tree , the aggregated trends for the children of that node will be shown. Fig.D demonstrates the resulting trends from the four stages of navigation marked on Fig.C.

Based on this simple idea, I built a tweet explorer app that could aggregate the trends of tweets based on their assigned topic (obtained by community detection). The Screen-Shot below shows the tree obtained by analysis of 50 Billion tweets from a year-long activity related to the indian rape case of 2012.



Clicking on the time-lines renders the actual tweets at that particular time and in that particular topic

