COSC 2671 Social Media and Network Analytics

Tute Week 6

Event Detection

Learning outcomes:

- Reinforce and revise Event Detection concepts and ideas.

Tutorial Questions

1. Consider the following expected and observed frequencies for the following set of words:

Terms	Expected (frequency)	Observed (frequency)
abc	20	30
hoop	1	10
laptop	3	10
tiger	60	120
phone	60	50

   Compute the frequency ratio for each. Then compute the burstiness score for each. What are the top 2 bursty terms using the frequency ratio. What about the burstiness scores?

2. For the previous question, apply the Yates adjustment to each of the computed burstiness scores. Does it make a difference?

3. Why are we interested in finding clusters of keywords or hashtags that have bursty behaviour and close spatial similarity when detecting unknown, local events?

4. Contrast the differences between a method that ranks the clusters and takes the top ones as significant events versus using a classifier to determine whether a cluster is an event or not. When would one use one over another?

5. Consider the following scenario. We are asked to analyse a stream of tweets and try to detect unknown events about Metro (the company running Melbourne trains). What are some factors that we might want to consider? How would we approach this task?