Using Social Annotations for Trend Discovery in Scientific Publications

Meiqun Hu, Ee-Peng Lim and Jing Jiang
School of Information Systems, Singapore Management University

Introduction

- Given a publication, one may ask:
  - ***How much*** interest does the IR (for example) community have on this work?
  - ***When*** did such interest emerge?
  - ***How fast*** was the emergence?
- Research questions:
  - How to find emerging trends for publications?
  - How to use emerging trends to help information seekers explore the social annotation space?

Definitions

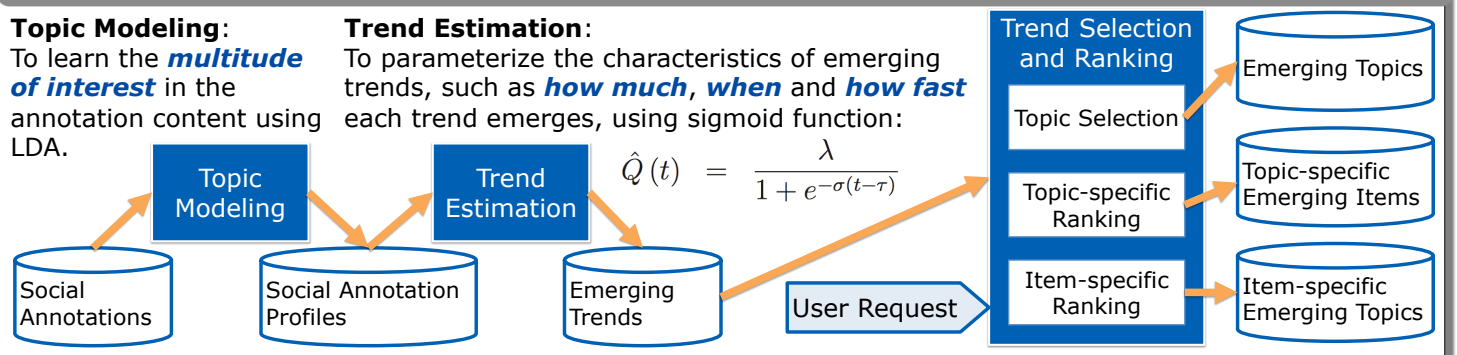
- **Social Annotation Documents**
  - Social tagging community → bookmarks
  - Research community → citing documents
- **Social Annotation Profiles**
  - A stream of social annotation documents
- **Social Annotation Time Series**
  - A series of consecutive disjoint time windows and each time window counts the number of social annotation documents within the time window

Proposed Trend Discovery Process

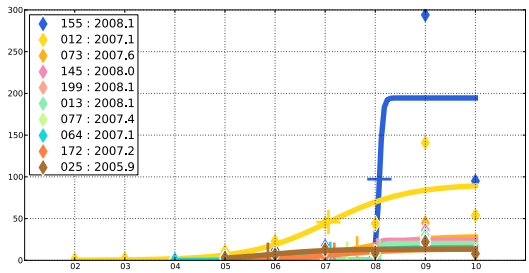
Topic Modeling:
To learn the ***multitude of interest*** in the annotation content using LDA.

Trend Estimation:
To parameterize the characteristics of emerging trends, such as ***how much***, ***when*** and ***how fast*** each trend emerges, using sigmoid function:

$$\hat{Q}(t) = \frac{\lambda}{1 + e^{-\sigma(t-\tau)}}$$



Publication-specific Emerging Topics



For Book: Convex Optimization

ID	Topic Keywords	How much	When	How fast
155	channel capacity	194.5	2008	1378.9
012	optimization problem	91.4	2007	28.7
073	wireless networks	28.6	2007	11.6
145	sensor networks	24.0	2008	190.9
199	noise signal filters	20.5	2008	110.0

Topic-specific Emerging Publications

For Topic 155: channel capacity

Title	How much	When	How fast	Citations
Elements of Information Theory	200.0	2008	1018.6	2410
Convex Optimization	194.5	2008	1378.9	1239
On limits of wireless communications in a fading environment when using multiple antennas	146.5	2008	782.5	487
NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey	93.0	2008	562.5	242
Matrix Computations (3rd ed.)	43.5	2008	224.2	1121

- Information seekers may explore emerging trends in the social annotation space from ***topic to publications***, and from ***publication to topics***, **iteratively**.
- Emerging trends characterized by the estimated parameters enable trend selection and ranking.