

**Hauptseminar**

**Mining Volatile Data**

**Wintersemester 2012/13**

**Dr. Eirini Ntoutsis**

**November 07, 2012 Presentations Part 1**

[http://www.dbs.ifi.lmu.de/cms/Hauptseminar "Mining Volatile Data" WS1213](http://www.dbs.ifi.lmu.de/cms/Hauptseminar%20Mining%20Volatile%20Data%20WS1213)

- **Topic detection & Topic-Sentiment detection**

1. “Discovering evolutionary theme patterns from text: an exploration of temporal text mining”, KDD 2005
2. “Topic sentiment mixture: modeling facets and opinions in weblogs”, WWW 2007

Michael Stockerl

Liangchen Fan

- **Community detection**

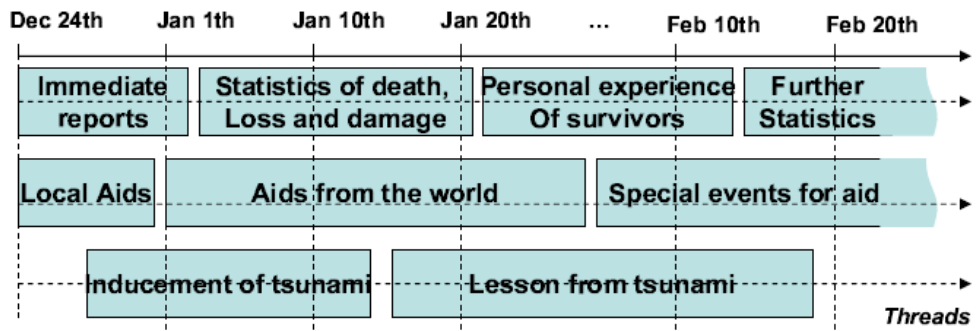
3. “Finding and evaluating community structure in networks”, Psys. Rev. 2003
4. “SCAN: A Structural Clustering Algorithm for Networks”, KDD 2007

Sigrid Zaenkert

Katharina Rausch

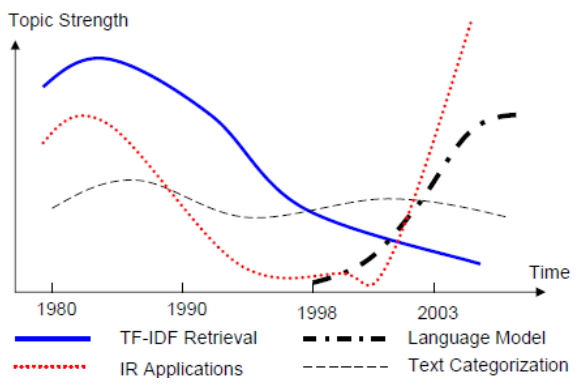
# Discovering evolutionary theme patterns from text: an exploration of temporal text mining

- Works upon a temporal annotated text stream
- Goal 1: find the theme structure



Example of theme thread structure from the Asian tsounami, 2004



- Goal 2: find the strength of the theme over time (theme life cycle)



Example of theme strength in IR

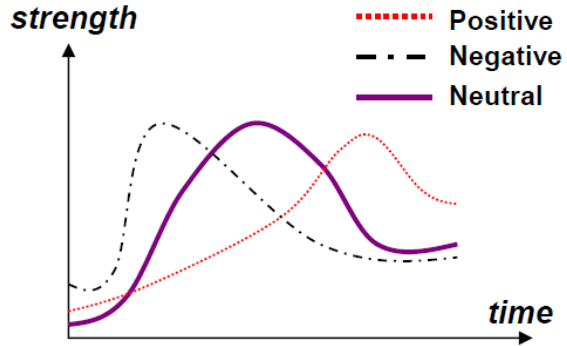
# Topic sentiment mixture: modeling facets and opinions in weblogs

- Works upon a temporal annotated opinionated text stream
- Goal 1: find opinionated topics

Query: <i>Dell Laptop</i>			
	positive	negative	neutral
<b>Topic 1</b> <i>(Price)</i> 	<ul style="list-style-type: none"> <li>• it is the best site and they show Dell coupon code as early as possible</li> </ul>	<ul style="list-style-type: none"> <li>• Even though Dell's price is cheaper, we still don't want it.</li> <li>• .....</li> </ul>	<ul style="list-style-type: none"> <li>• mac pro vs. dell precision: a price comparis..</li> <li>• DELL is trading at \$24.66</li> </ul>
<b>Topic 2</b> <i>(Battery)</i> 	<ul style="list-style-type: none"> <li>• One thing I really like about this Dell battery is the Express Charge feature.</li> </ul>	<ul style="list-style-type: none"> <li>• my Dell battery sucks</li> <li>• Stupid Dell laptop battery</li> <li>• .....</li> </ul>	<ul style="list-style-type: none"> <li>• i still want a free battery from dell..</li> <li>• .....</li> </ul>

Example of topic-sentiment summary

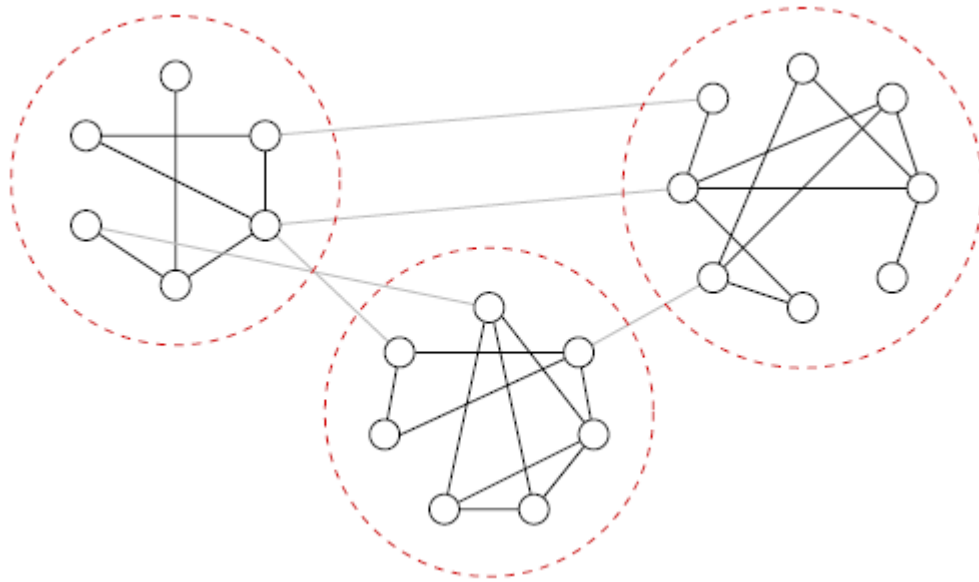
- Goal 2: find the topic-sentiment dynamics



Example for the theme price

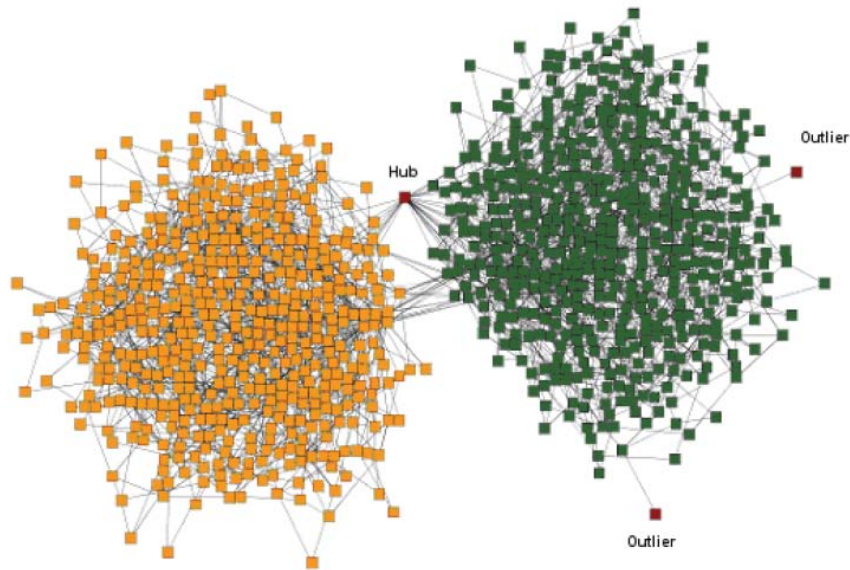
# Finding and evaluating community structure in networks

- Works upon a (static) network/ graph
- Goal : community detection (hierarchical clustering)



# SCAN: A Structural Clustering Algorithm for Networks

- Works upon a (static) network/ graph
- Goal : community detection (density-based clustering)



2<sup>nd</sup> part of the presentations,  
next Wednesday, 14/11/2012.