

# ARTS1422 Data Visualization

## Lecture 10

### Network Data Visualization & Social Media Visualization

Quan Li  
Spring 2024  
2024. 03.28

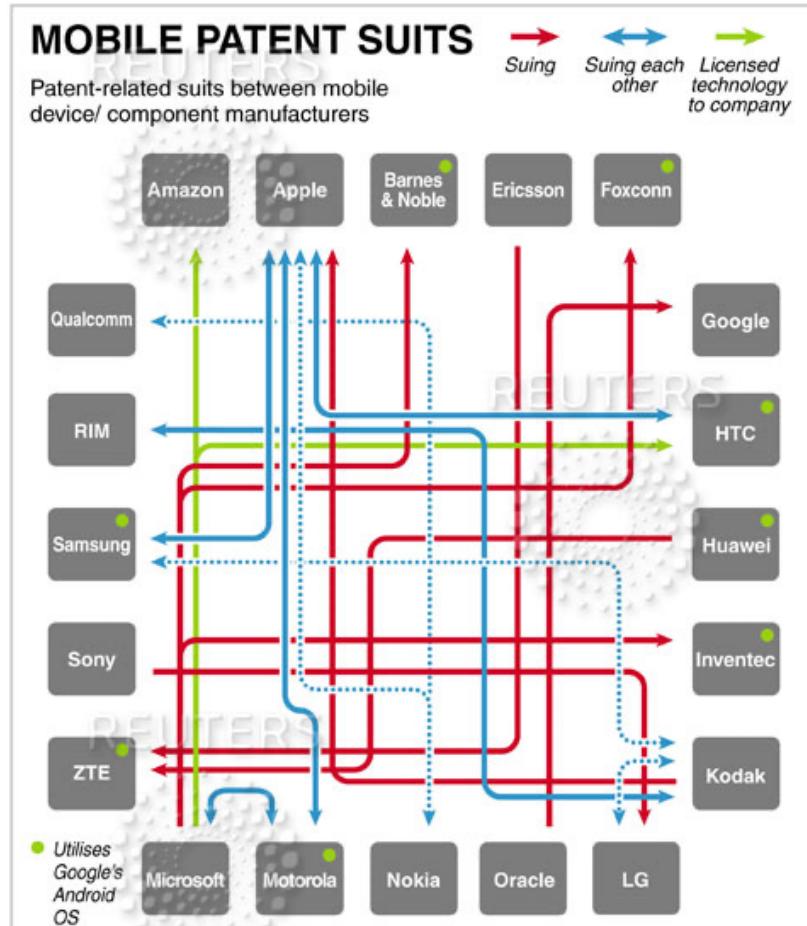


# Network Data

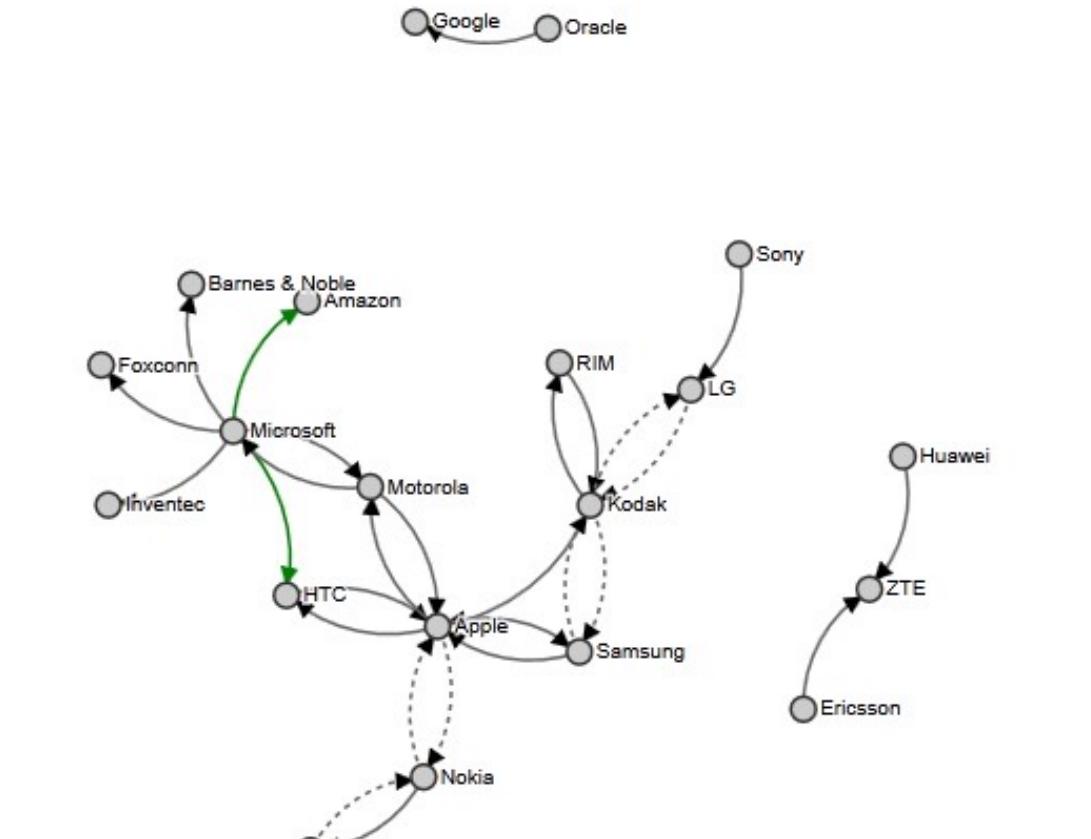
---

- Compared with hierarchies in tree visualization, network data present more **complex and flexible relations**.
  - Social network
  - Mobile network
  - Mail network
  - Collaboration network

# Mobile Patent Lawsuits

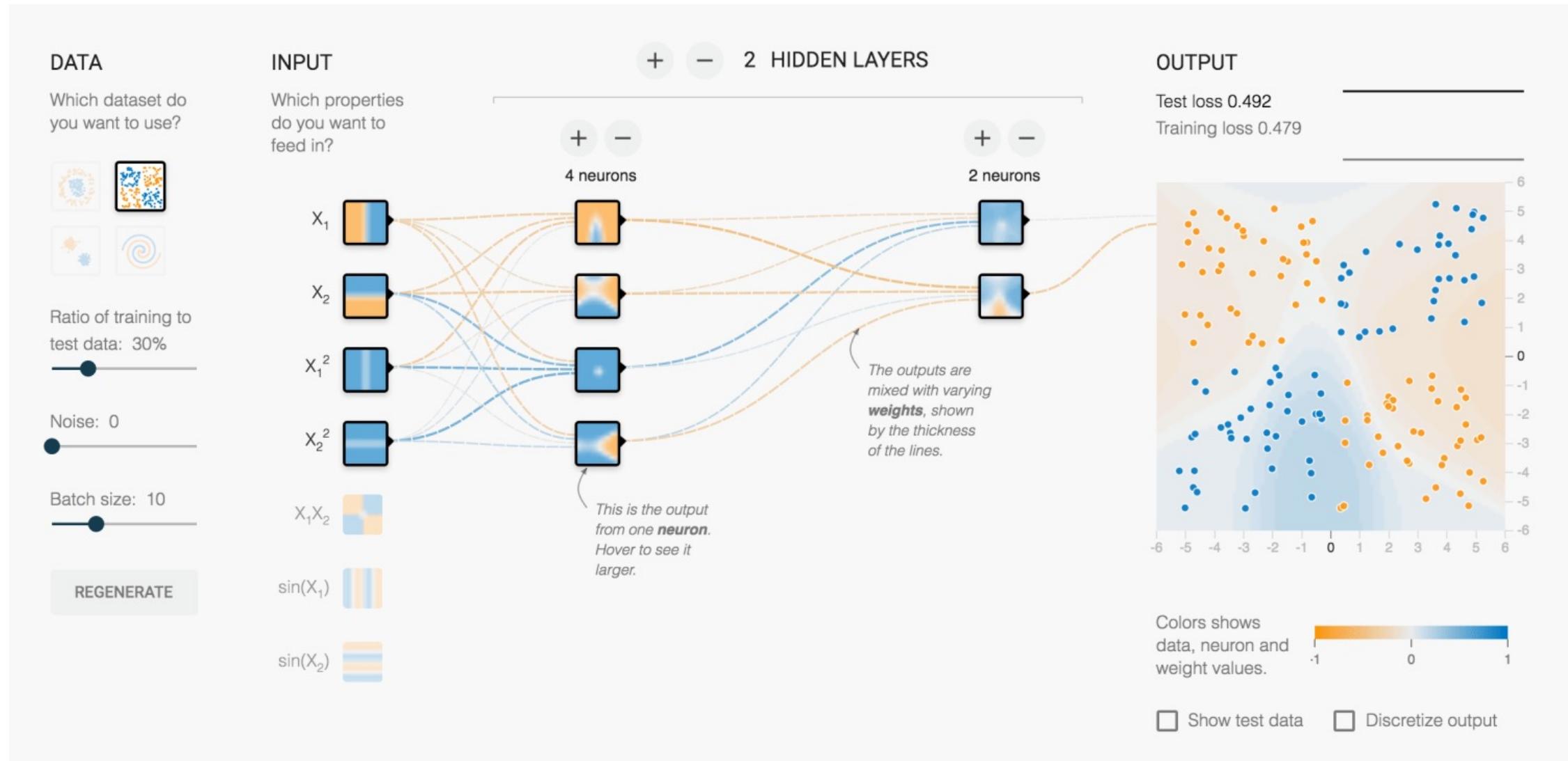


<http://blog.thomsonreuters.com/index.php/mobile-patent-suits-graphic-of-the-day/>



Based on D3.js  
<http://bl.ocks.org/1153292>

# Here's how a neural network works



# Let's play

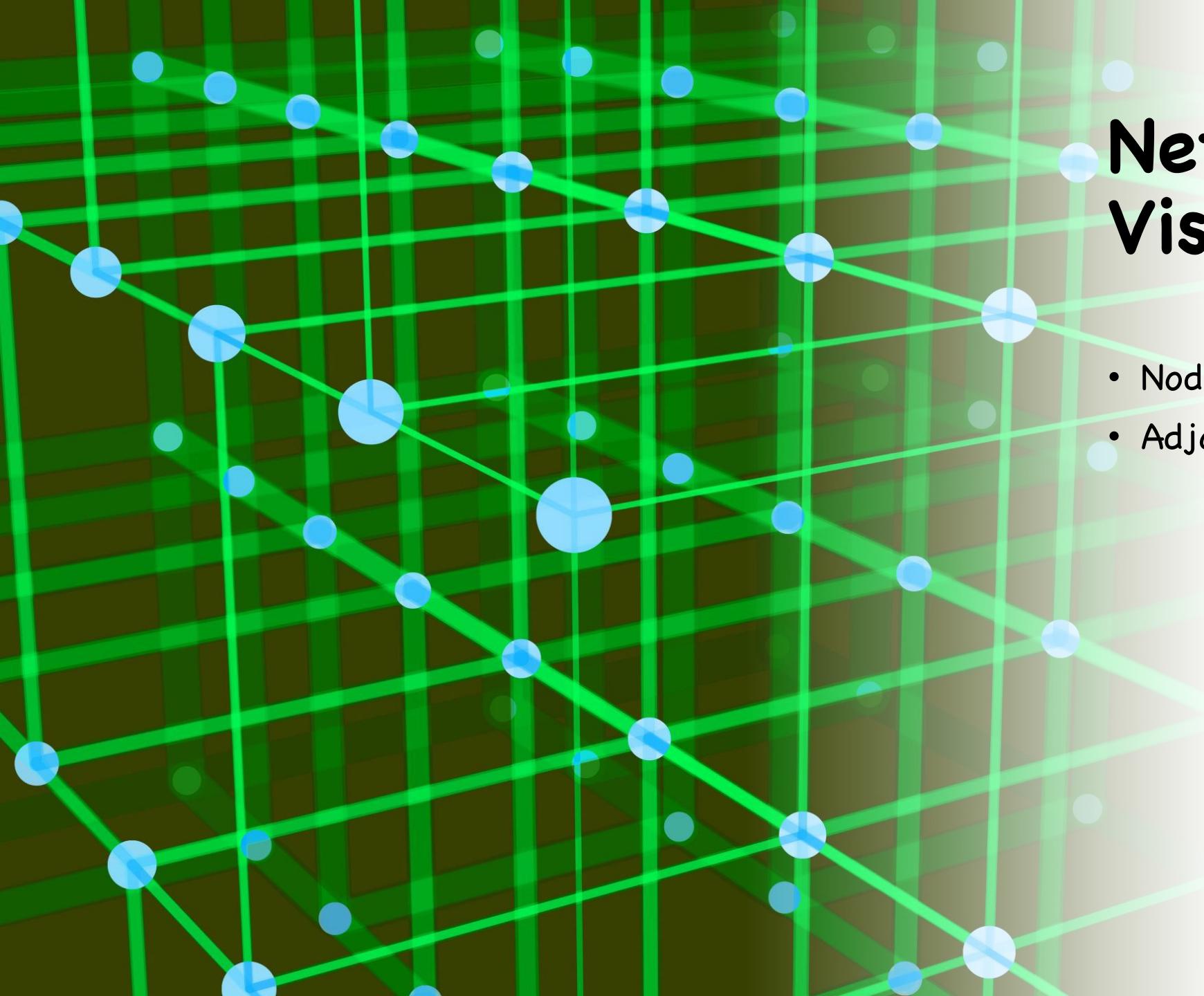
[playground.tensorflow.org](https://playground.tensorflow.org)

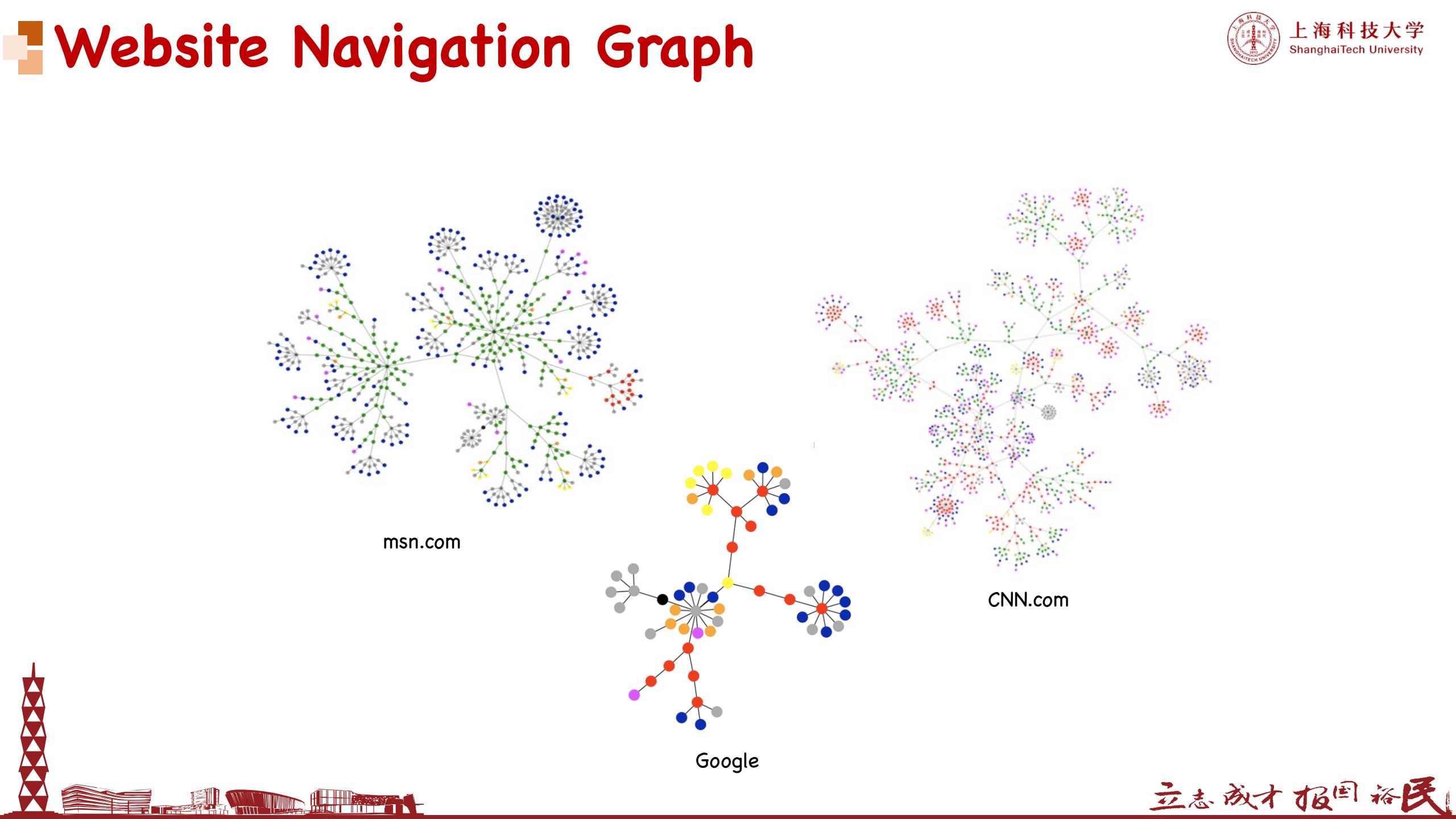
# Properties of Network

- Complex relations
  - Directions, weights
- Network centrality
  - Degree
  - Closeness
  - Betweenness
  - Eigenvector

# Network Data Visualization

- Node-link diagram
- Adjacency matrix



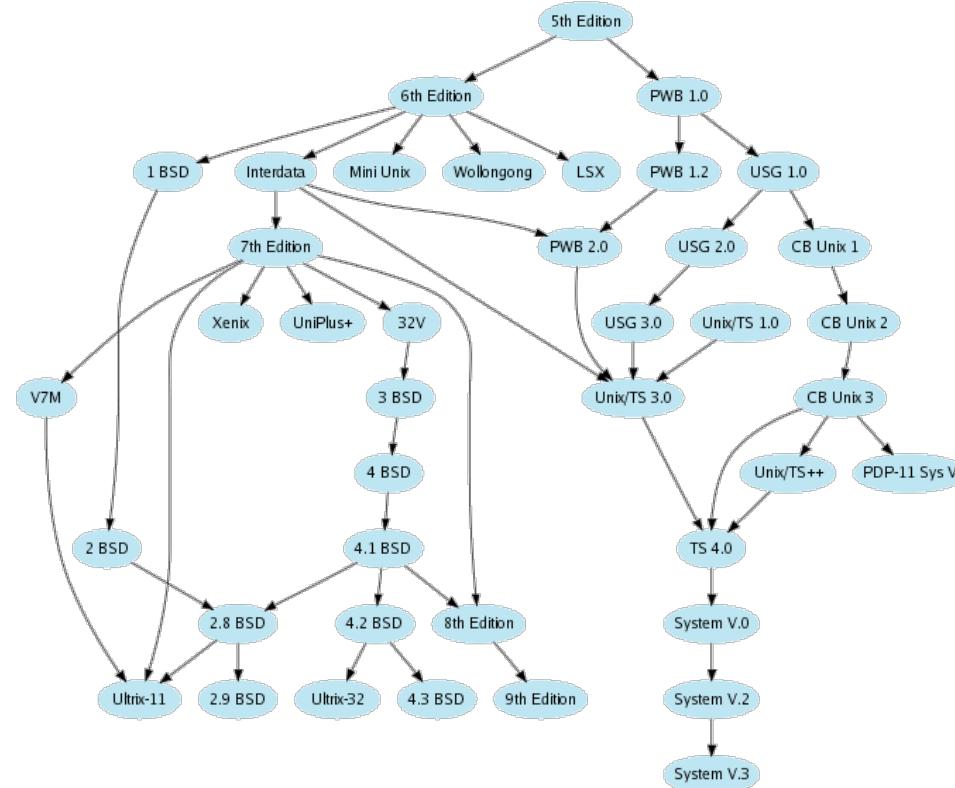


# Graph Representation

- Node-link diagram
  - Sugiyama
  - Force-directed layout
  - MDS layout
- Adjacency matrix
- Attribute-based representation

# Sugiyama Representation

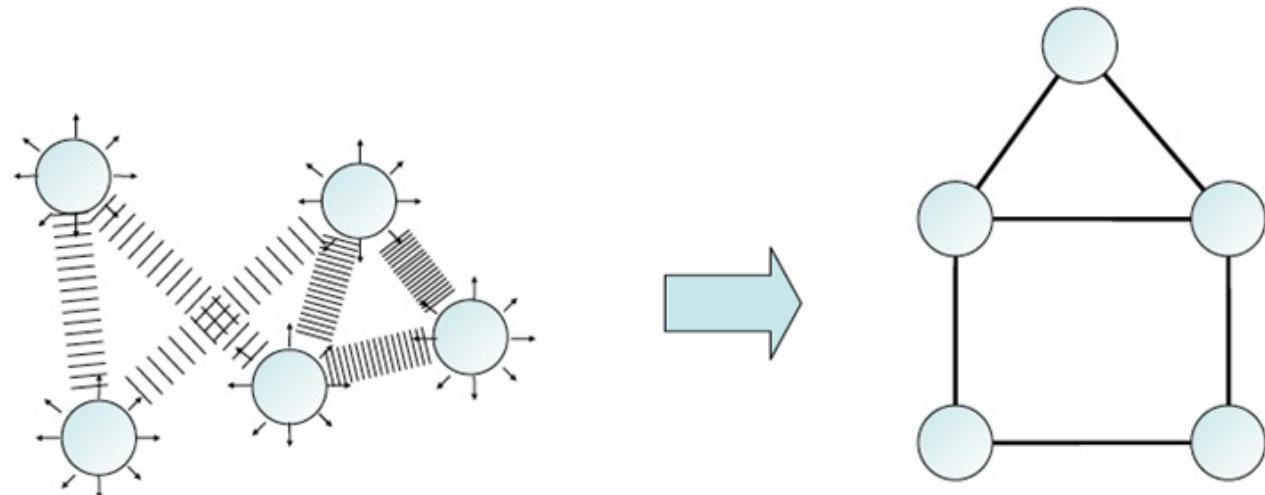
- Good for **intrinsically ordered** graphs
- Depth of graph can be mapped to axes
- High readability
- Fast (depends on the heuristic method)
- Not so good for graph without **intrinsic order**
- Easy to implement (free library graphviz, <http://www.graphviz.org>)



UNIX Family

# Force-based algorithm

- What about graphs without intrinsic order?
- Physics model:
  - edge → spring
  - node → mass point



# Force-Directed Layout

- Peter Eades introduced the layout in 1984 in the paper A Heuristic for Graph Drawing.
- With the effects of **spring force and gravitation**, nodes far away will be dragged near and vice versa.
- The layout reaches a **balance** and becomes stable after **iterations**.
- Spring Model:
- Energy Model:

$$E_s = \sum_{i=1}^n \sum_{j=1}^n \frac{1}{2} k(d(i,j) - s(i,j))^2$$

$$E = E_s + \sum_{i=1}^n \sum_{j=1}^n \frac{r w_i w_j}{d(i,j)^2}$$





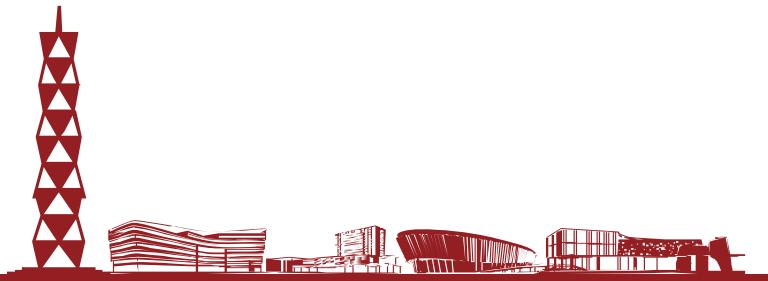
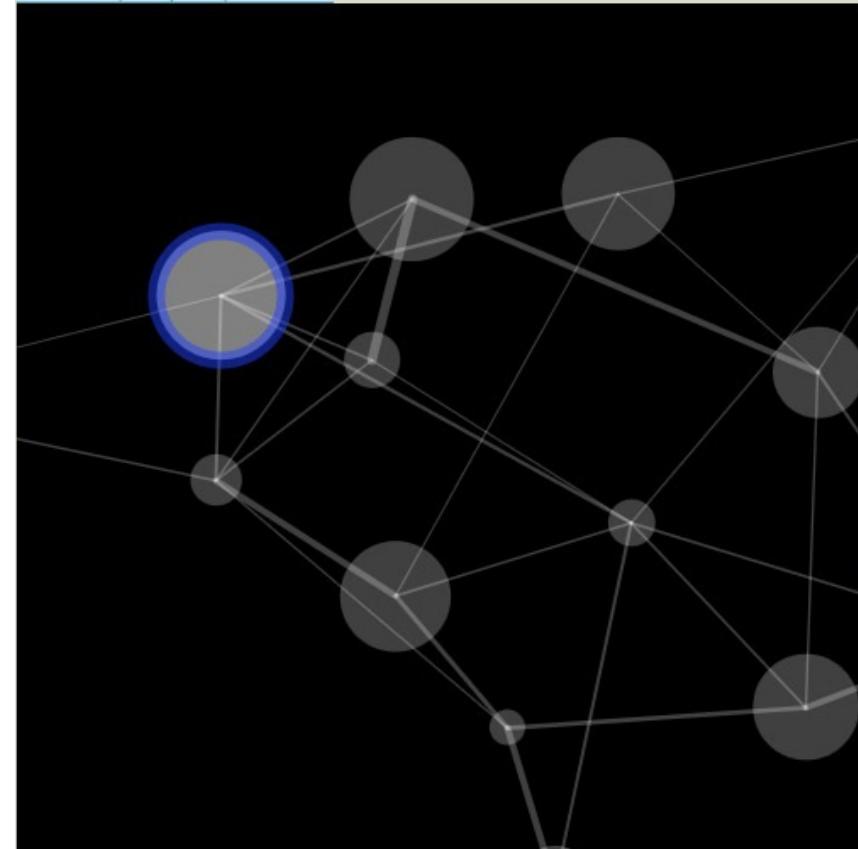
# Nodes Position Computation\*

- From random or initial **configuration**
- Loop:
  - Compute **the repulsion and attraction** force for every pair of nodes.
  - Accumulate the **force** (vector) for every node.
  - Update nodes position **step by step** according to their forces.
- Loop stops when the layout is “good enough”



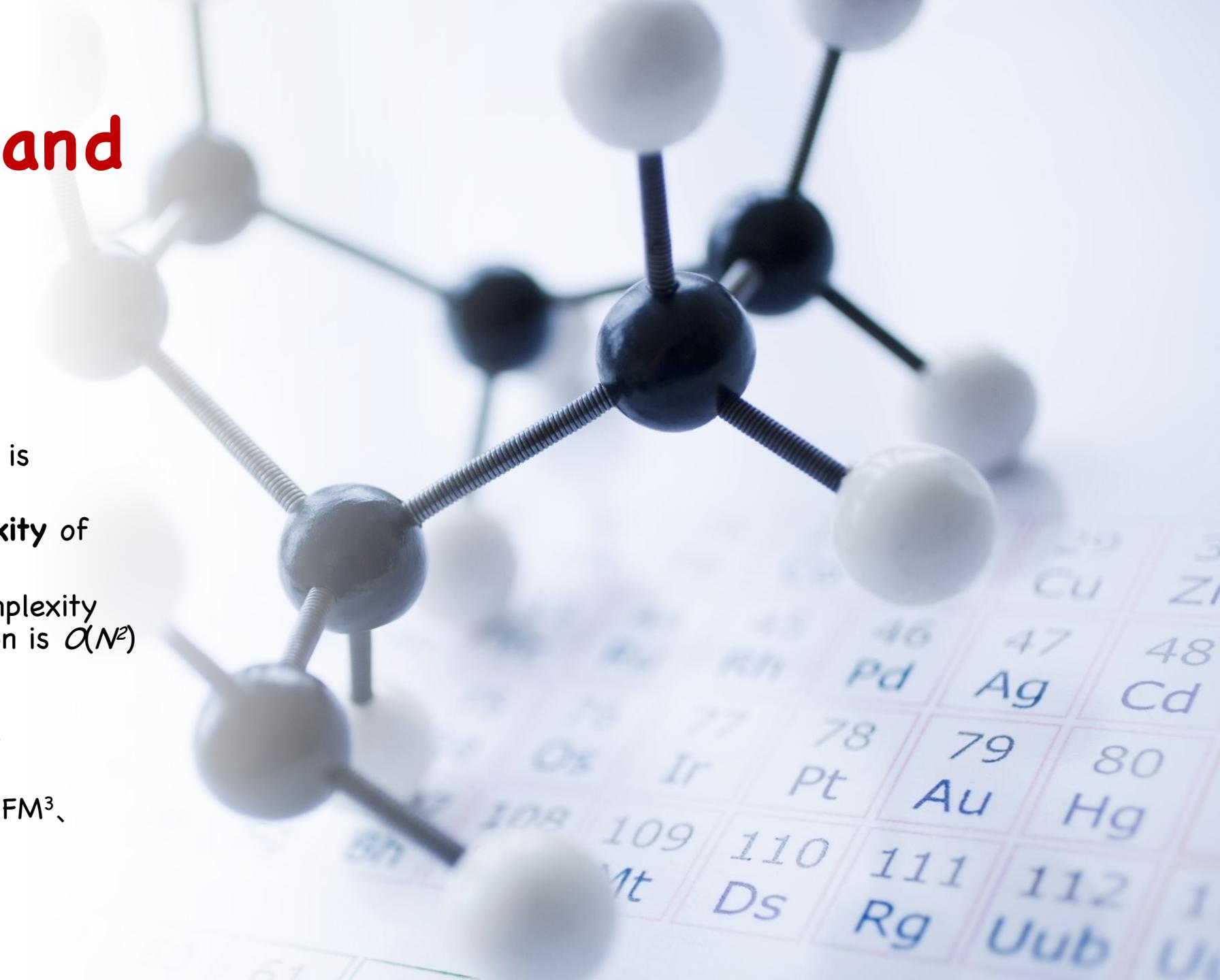
# Advantages

- Very flexible for any type of graphs
- Forces can be customized
- Easy to implement



# Limitations and Extensions

- Limitations
  - Local optimal
  - Initial configuration is important
  - Computation complexity of iterative algorithm
    - Computation complexity for each iteration is  $\mathcal{O}(N^2)$
- Extensions
  - Barnes-Hut quadtree decomposition
  - FADE, GRIP, FMS, FM<sup>3</sup>, GVA





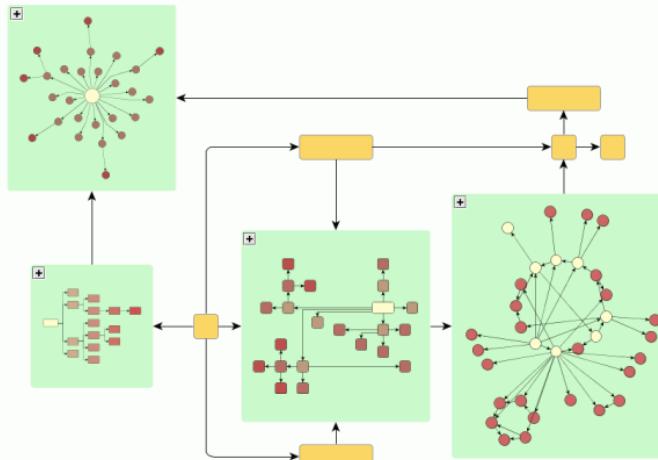
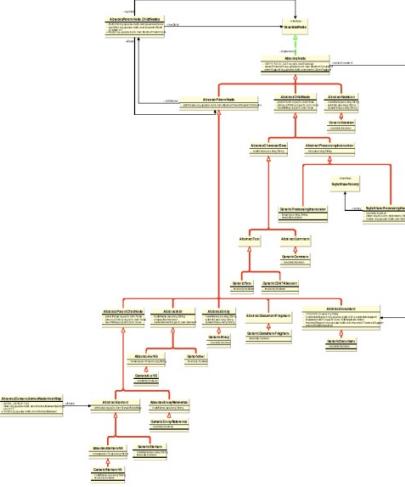
# Results of Force-directed Layout

## Force-directed graph with elliptic forces

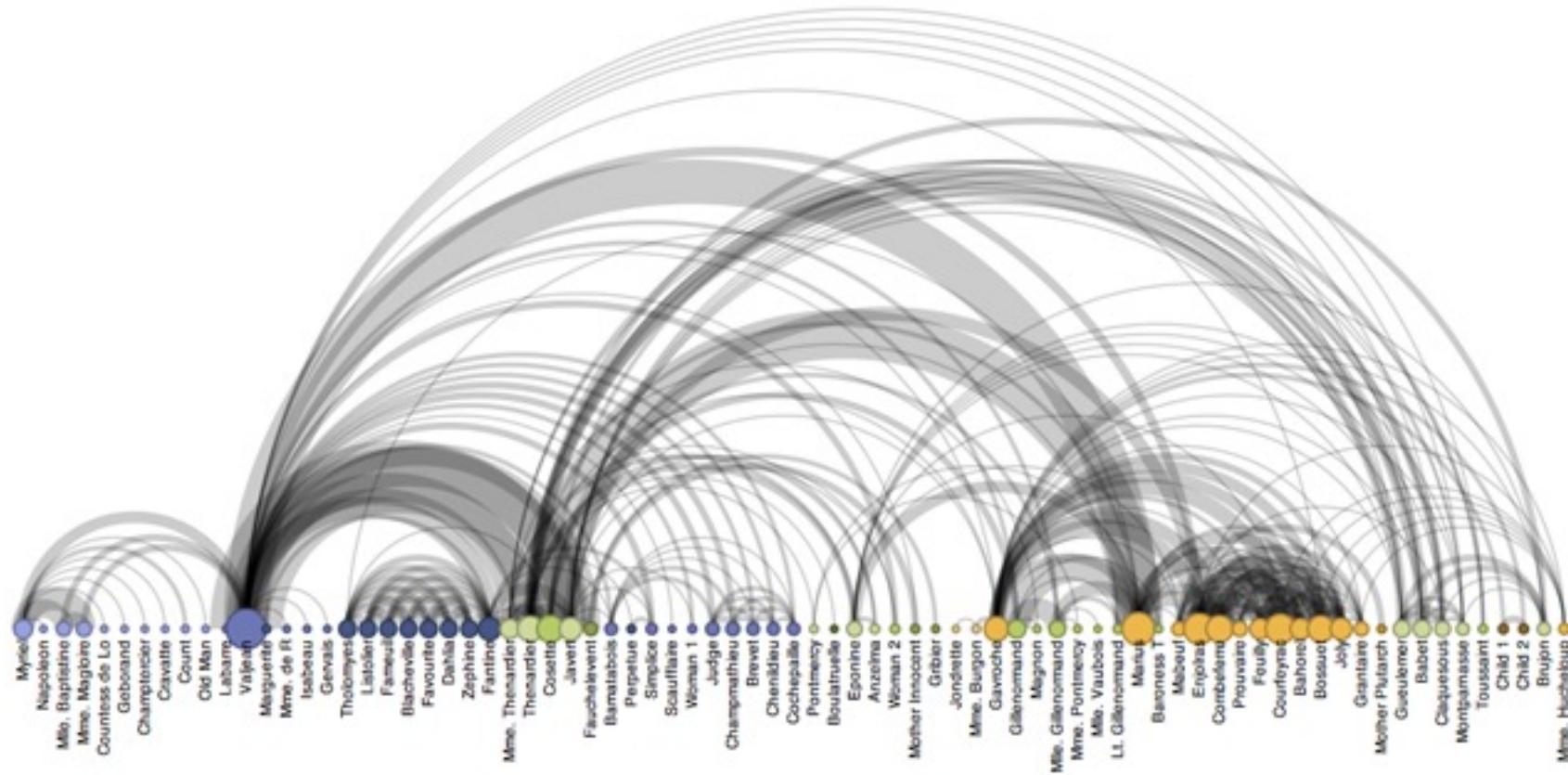


# Other presentations of node-link diagram

- Orthogonal Diagram
  - UML diagram
- Ring ordered
  - For ring topology
  - Widely used in social network
- Nested ordered
  - Recursively applying nested layout
  - For intrinsic ordered topology
- Arc Diagram



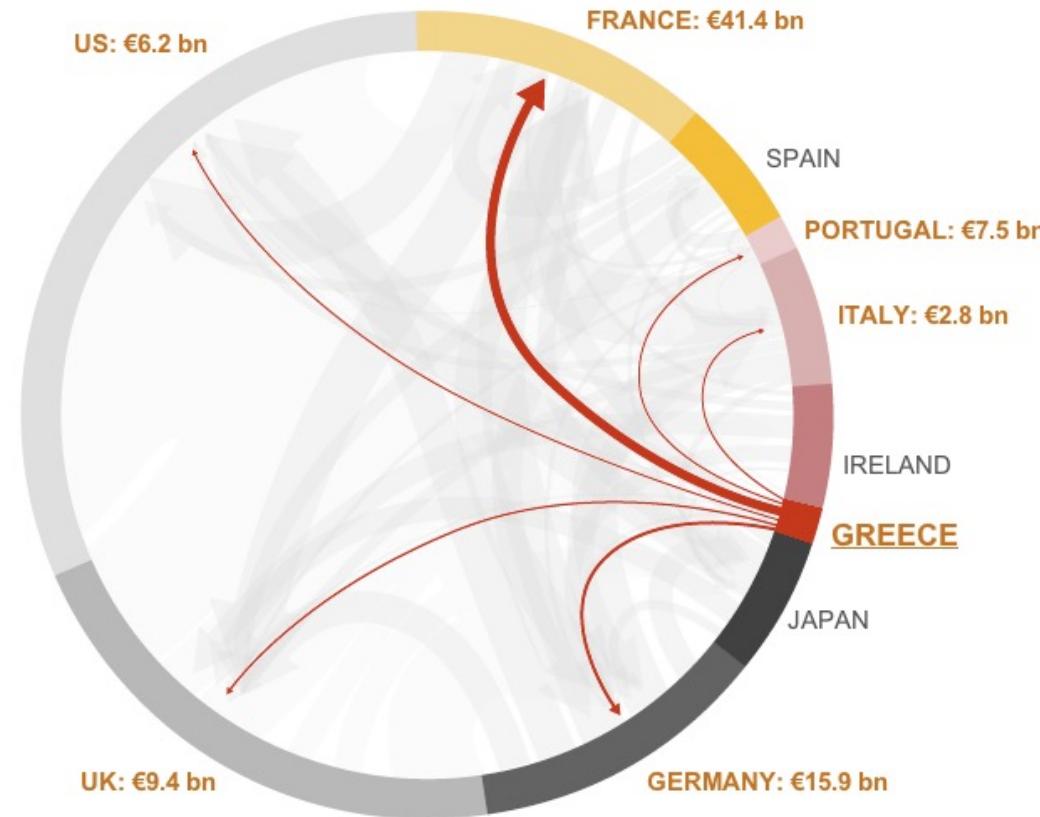
# Arc Diagram



Character relations in *Les Misérables*

<http://hci.stanford.edu/jheer/files/zoo/>

# Arc Diagram



## GREECE

GDP: €0.2 tn

Foreign debt: €0.4 tn

38,073

Foreign debt per person

252%

166%

Foreign debt to GDP Govt debt to GDP

Risk Status: HIGH

Greece is heavily indebted to eurozone countries and is one of three eurozone countries to have received a bailout. Although the Greek economy is small and direct damage of it defaulting on its debts might be absorbed by the eurozone, the big fear is "contagion" - or that a Greek default could trigger a financial catastrophe for other, much bigger economies, such as Italy.

[Back to introduction](#)

Source: Bank for International Settlements, IMF, World Bank, UN Population Division

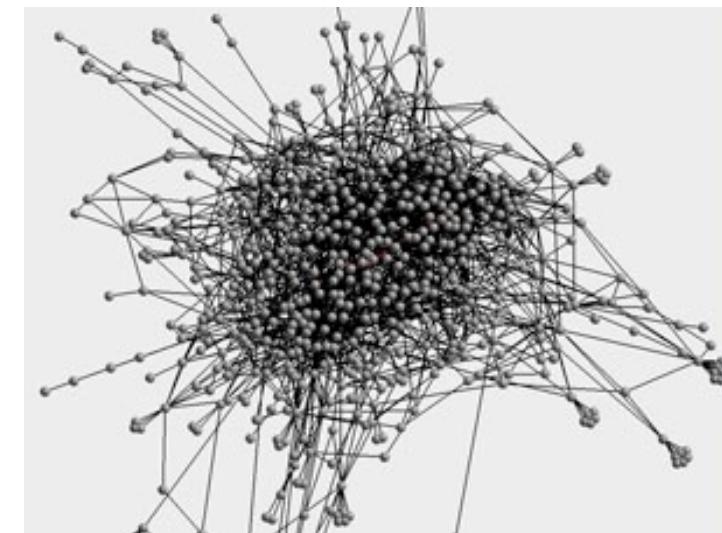
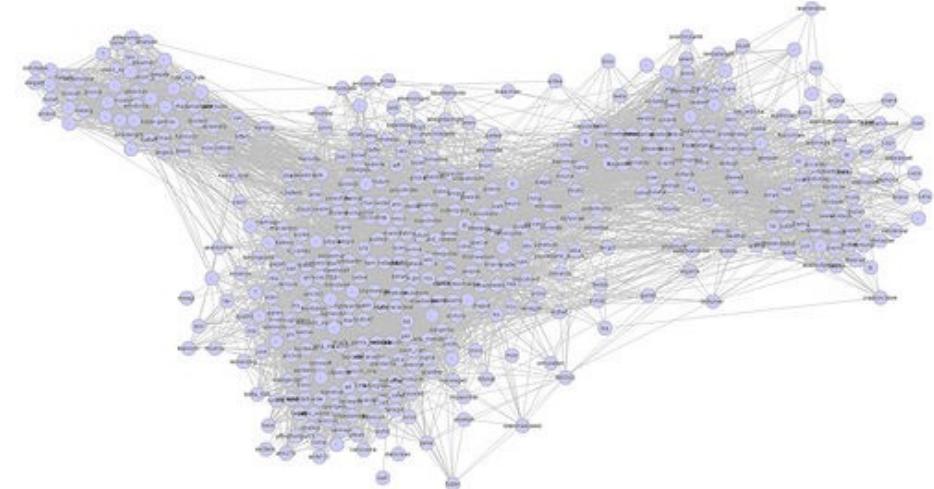
Eurozone debt web: Who owes what to whom?

<http://www.bbc.co.uk/news/business-15748696>

# Node-Link Diagram Summary

---

- Intuitive visual interpretation
- Good representation of topology, clusters and paths
- Flexible, many variants
- Almost for all algorithms, time complexity  $> \mathcal{O}(N^2)$
- Not so good for cluttered graphs (especially edge cluttered graphs)



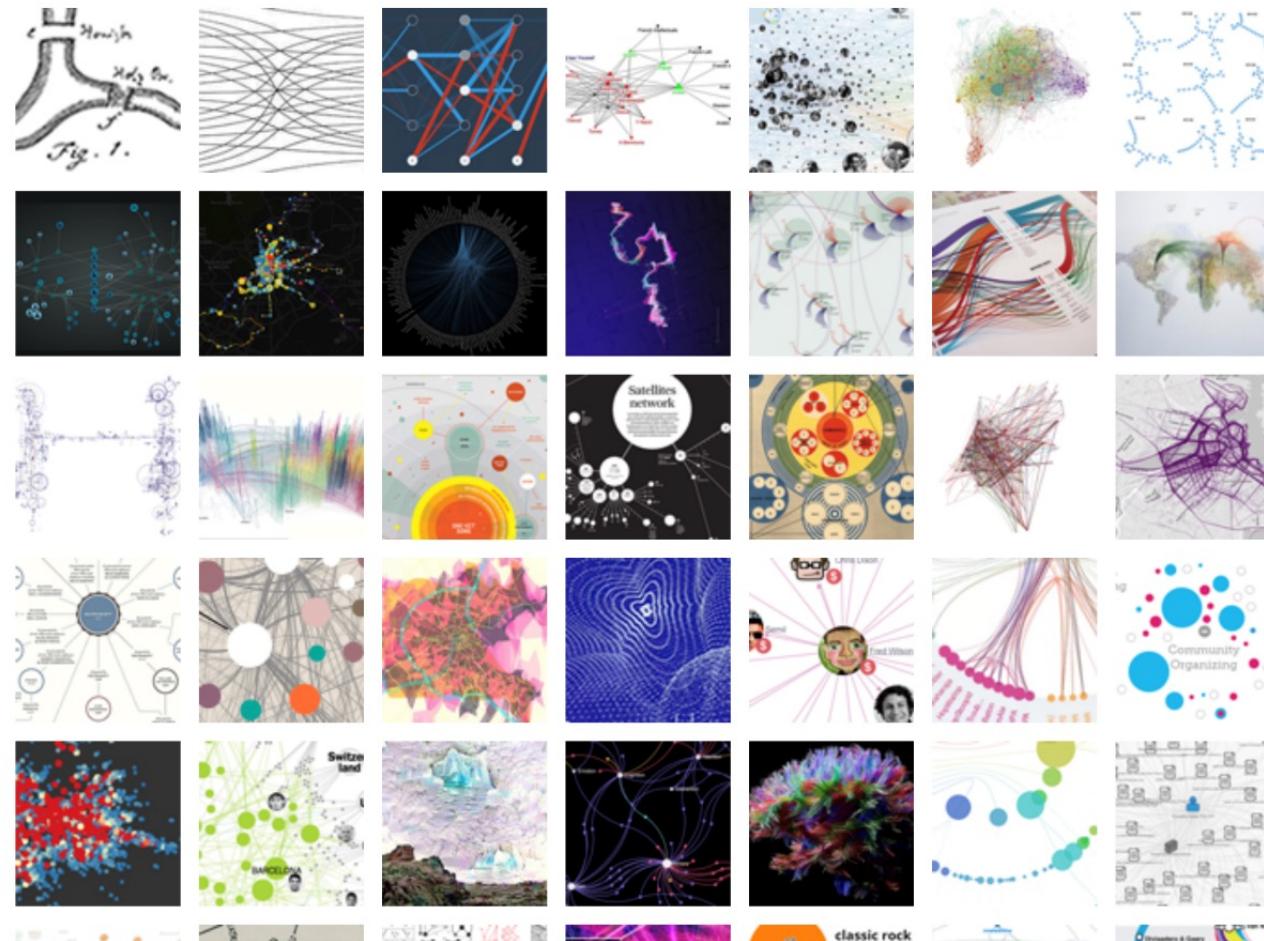


# visual complexity

Search the VC database:

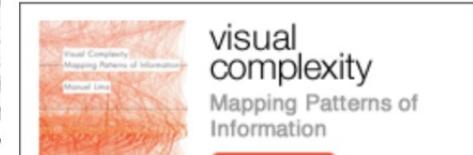
## Latest Projects:

Indexing **1000** projects

## Filter by:

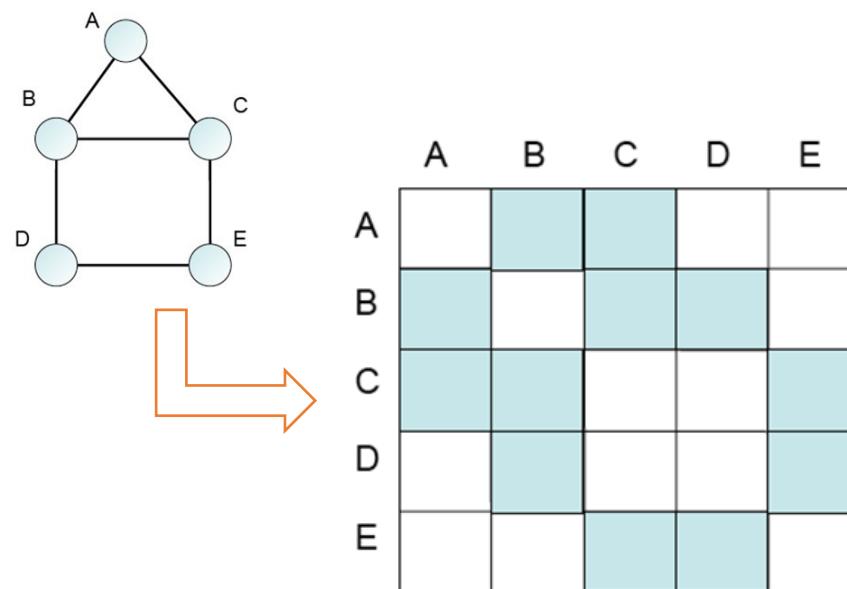
SUBJECT ▾

- Art (74)
- Biology (60)
- Business Networks (50)
- Computer Systems (39)
- Food Webs (16)
- Internet (35)
- Knowledge Networks (141)
- Multi-Domain Representation (70)
- Music (47)
- Others (77)
- Pattern Recognition (53)
- Political Networks (34)
- Semantic Networks (44)
- Social Networks (135)
- Transportation Networks (70)
- World Wide Web (55)

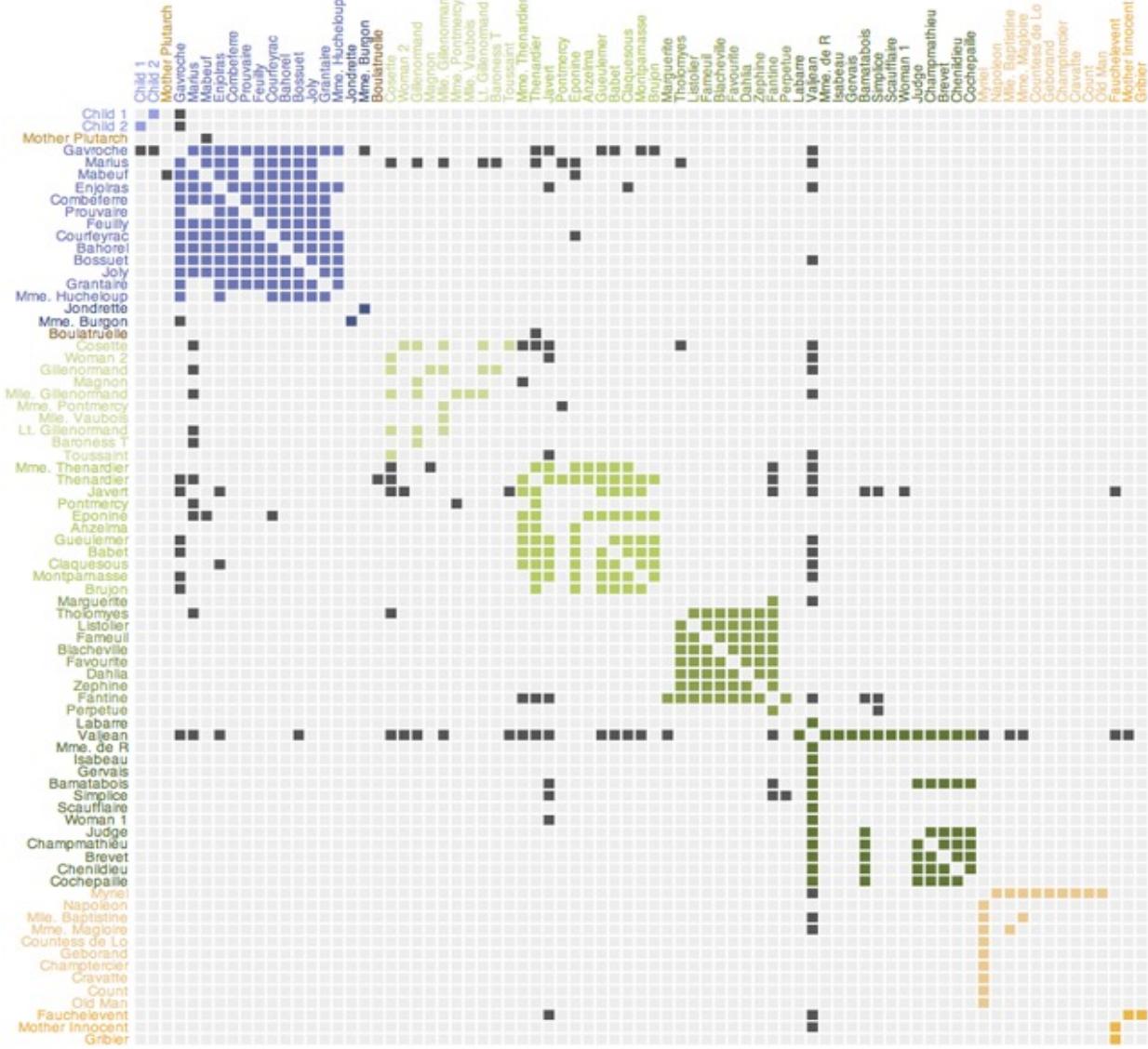
[See All \(1000\)](#)

# Adjacency Matrix

- $N \times N$  matrix, representing relations among  $N$  objects.
- Position  $(i, j)$  represents the relation between the  $i$ th and the  $j$ th object,
  - Weight
  - Direction
  - Self-reflexivity
- Related issues
  - Ordering
  - Path finding



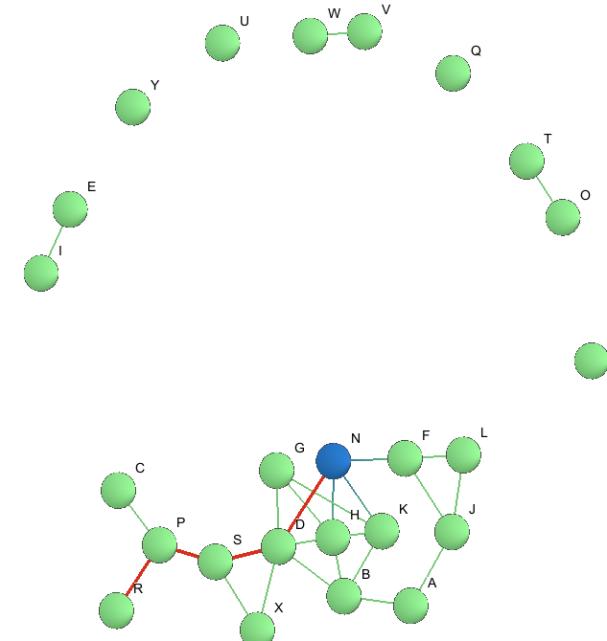
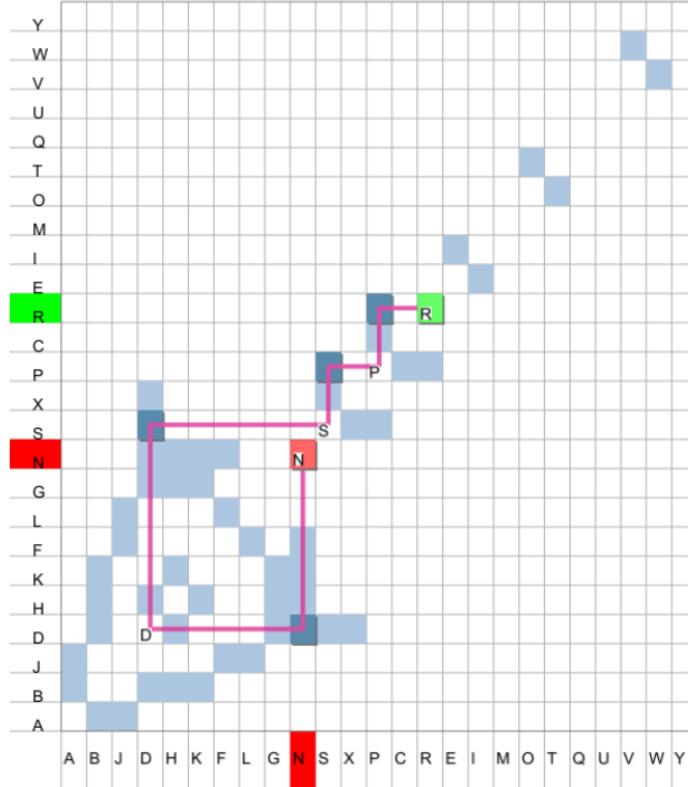
# Good Ordering of Adjacency Matrix



<http://hci.stanford.edu/jheer/files/zoo>

立志成才报国裕民

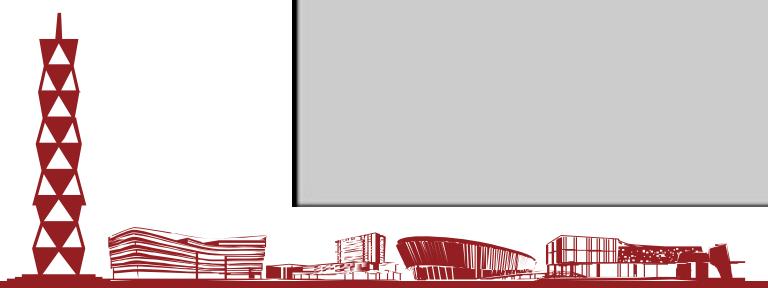
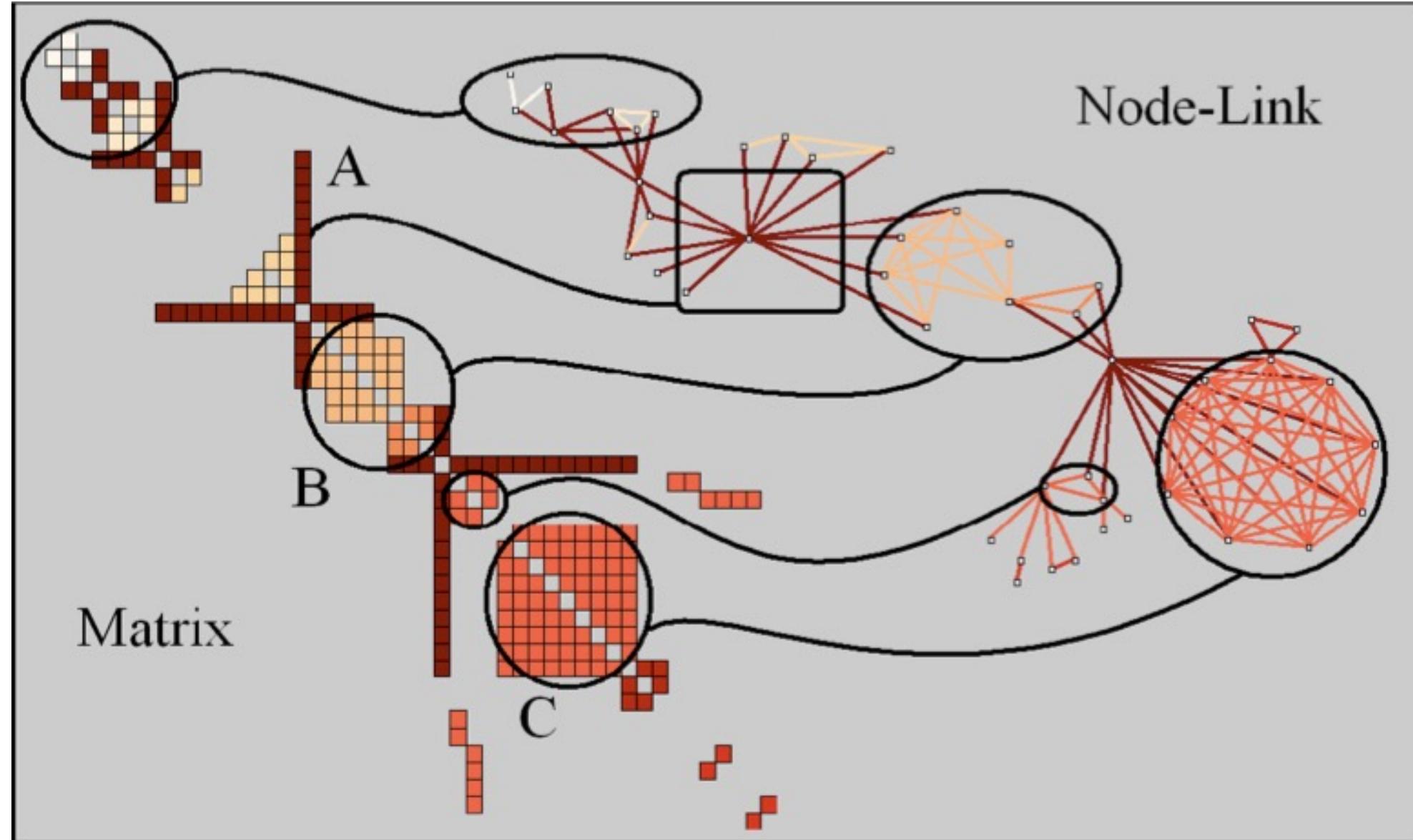
# Path Visualization for Adjacency Matrix



Left: Connecting matrix entries and the diagonal.  
Right: Path  $\{N;D;S;P;R\}$  is Highlighted in the Node-link Diagram.



# Recognizing the Patterns of Matrix





# Adjacency Matrix Summary

- No edge crossing, good for edge cluttered graph
- Good visual scalability
- Good presentation of graph pattern
- Visualization is too abstract to understand
- Difficult to follow a transitive relation path



# Hybrid Layout

# Hybrid Layout

- Complex edge relations—adjacency matrix
- Too many nodes—node-link diagram
- What if there are so many nodes, and some of them with complex edge relations?

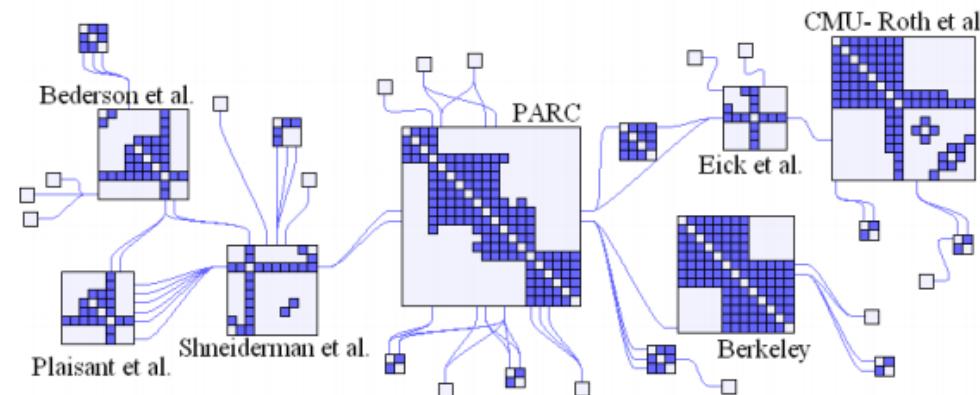
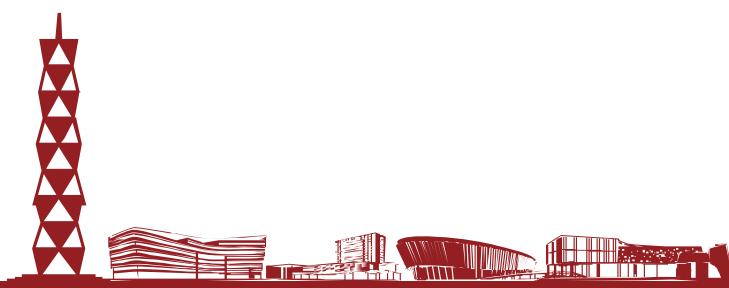
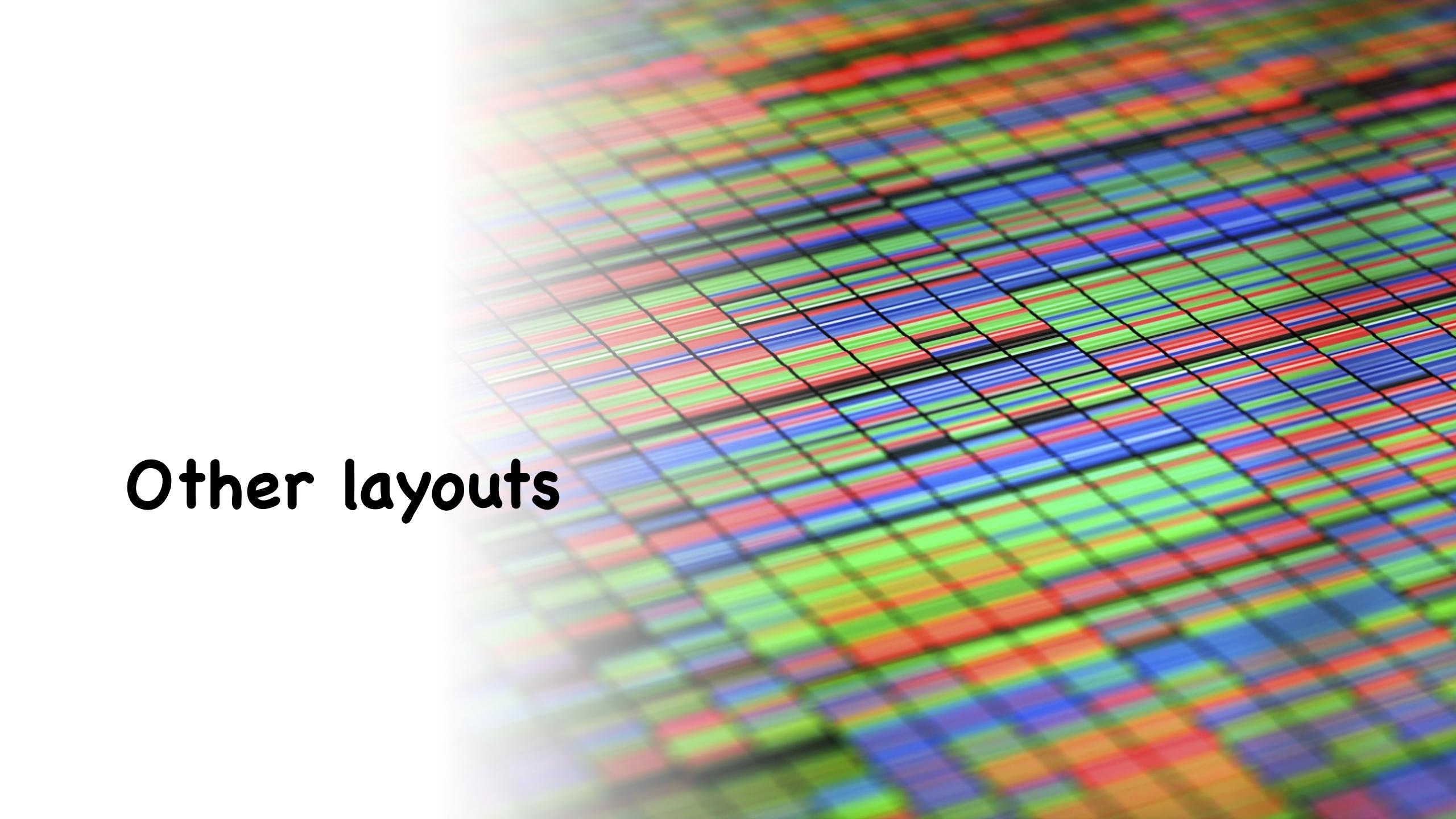


Fig. 1: NodeTrix Representation of the largest component of the InfoVis Co-authorship Network

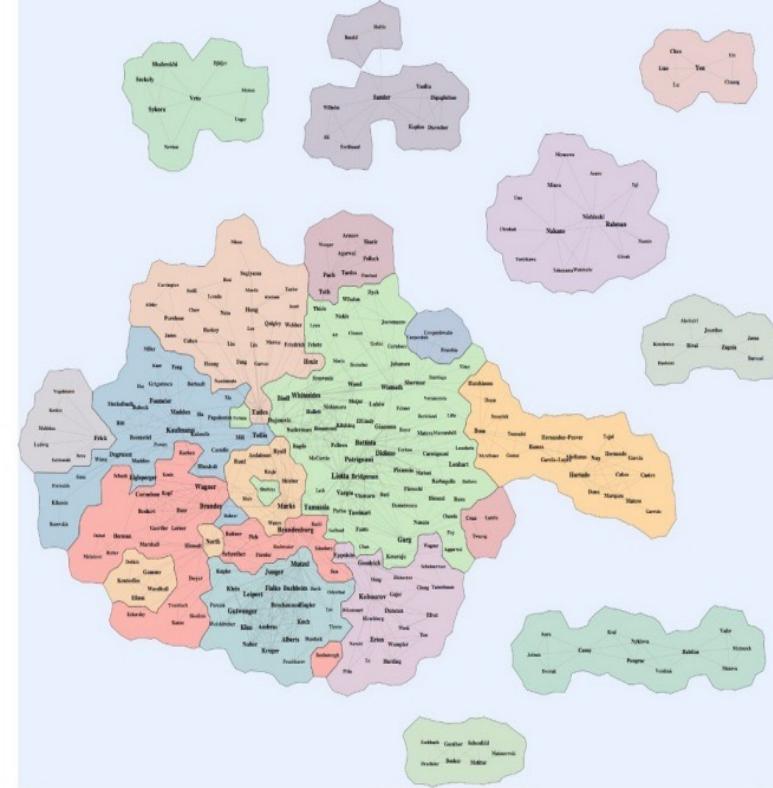
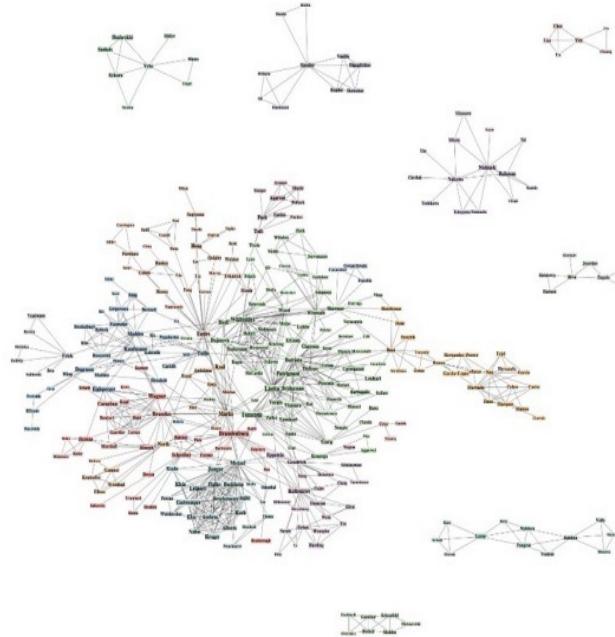
Nathalie Henry, et al. NodeTrix: A Hybrid Visualization of Social Networks. TVCG 2007.





# Other layouts

- Gmap represents object clusters with plane areas.





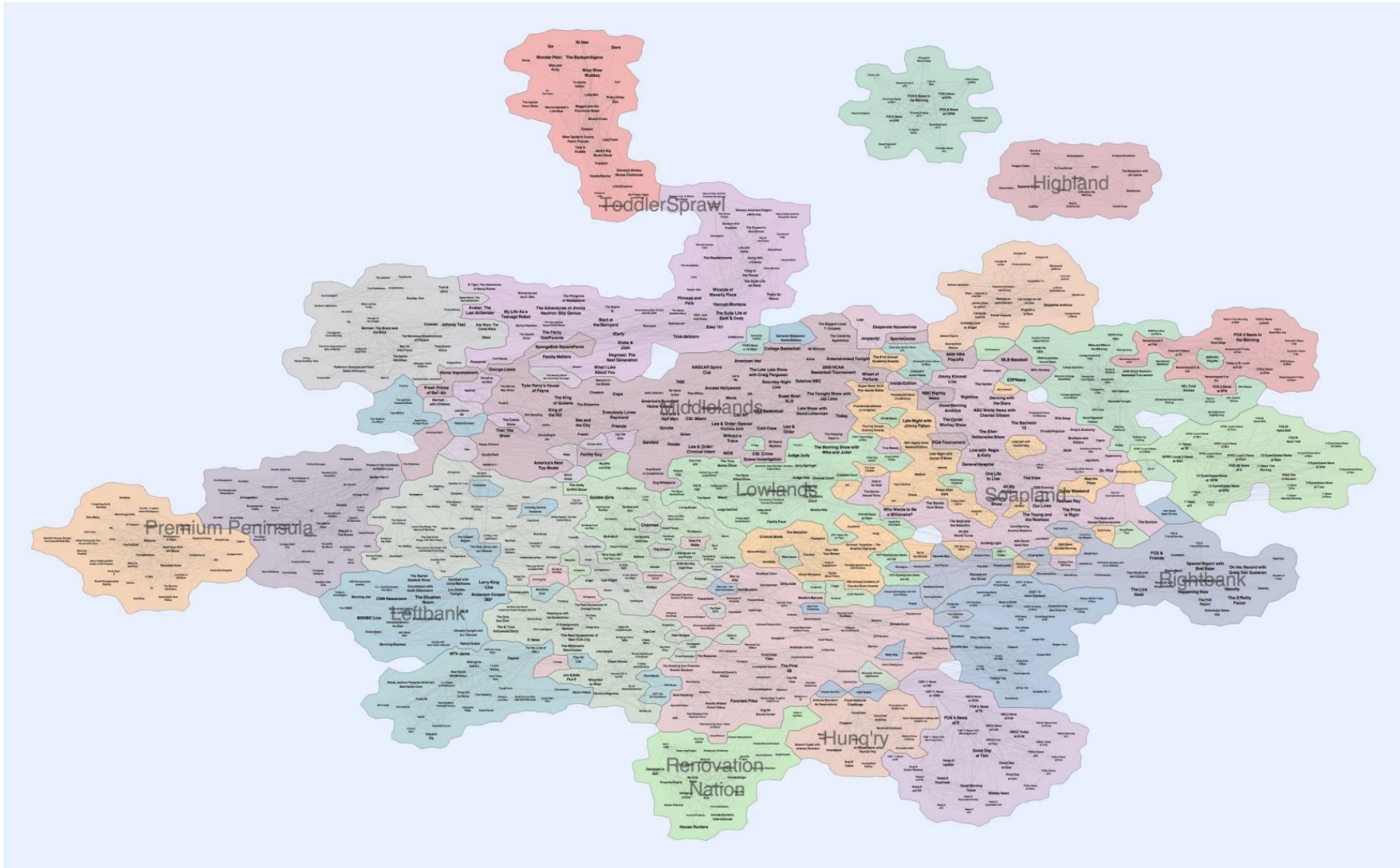
## Gmap Procedures\*

---

1. Layout the nodes on the graph
2. Clustering
3. Constructing voronoi diagram for every cluster
4. Coloring



# Gmap Procedures\*





# Graph Simplification

---



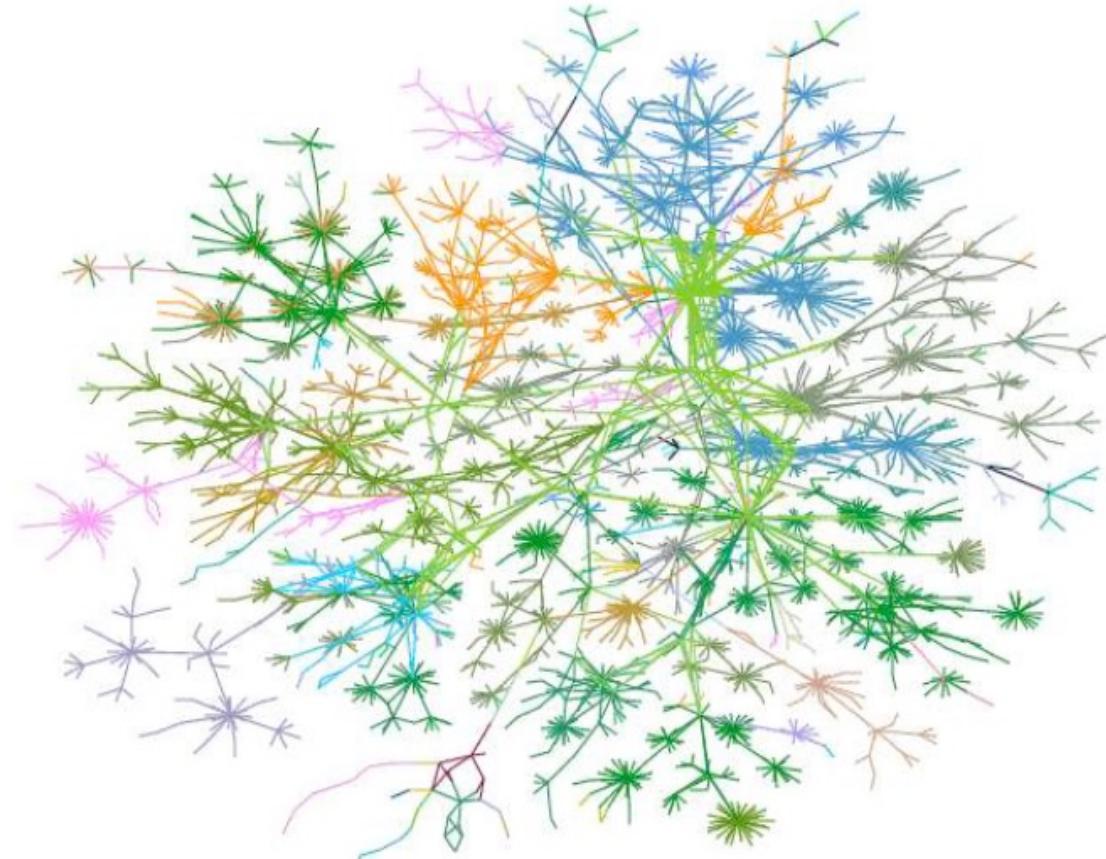
# Topology Simplification

- Data reduction
  - Nodes reduction
    - Clustering
  - Edge reduction
    - Minimal spanning tree



# Minimal Spanning Tree

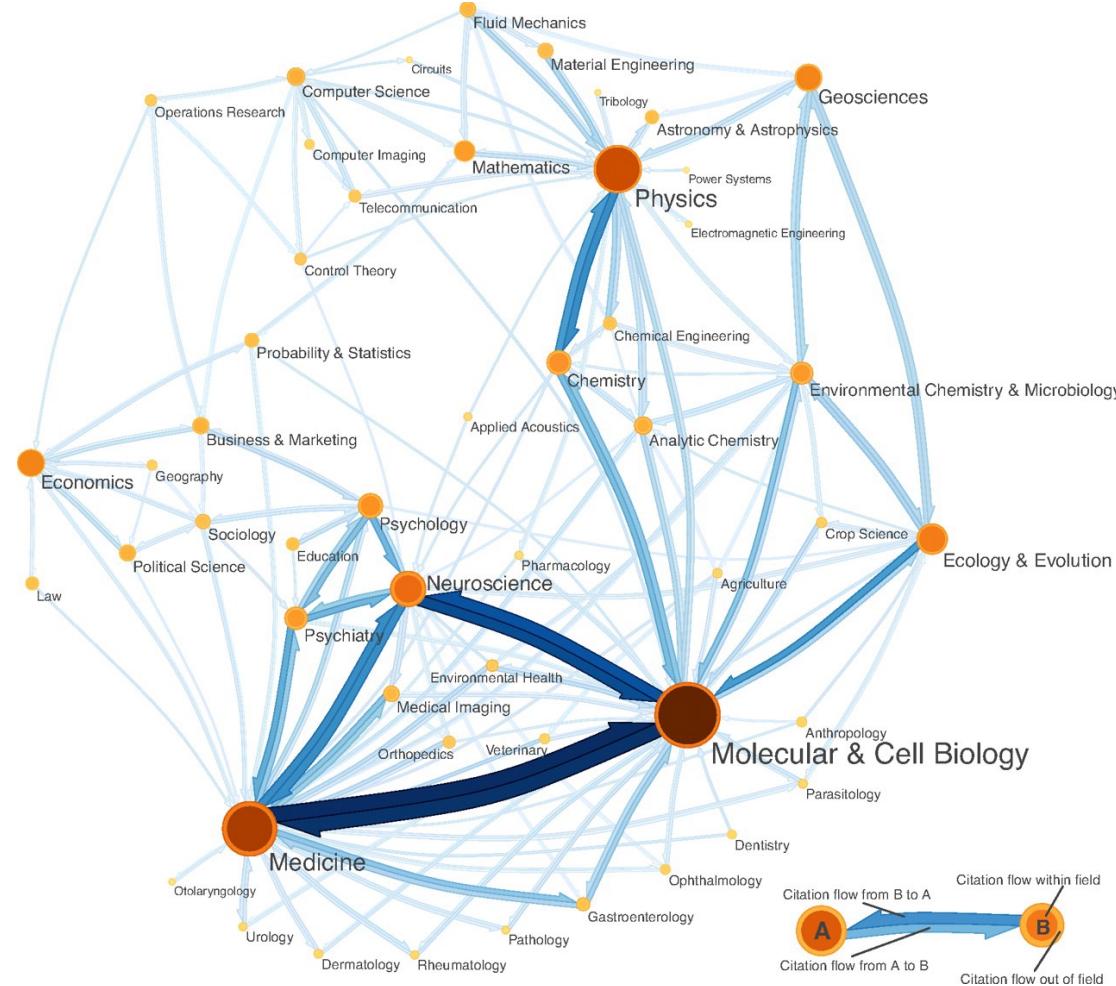
- Edge deletion
- Prim algorithm



Hal Burch. Measuring an IP Network in situ



# Visualization of Clusters

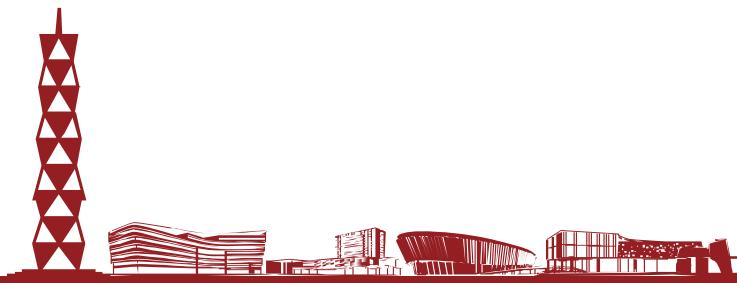
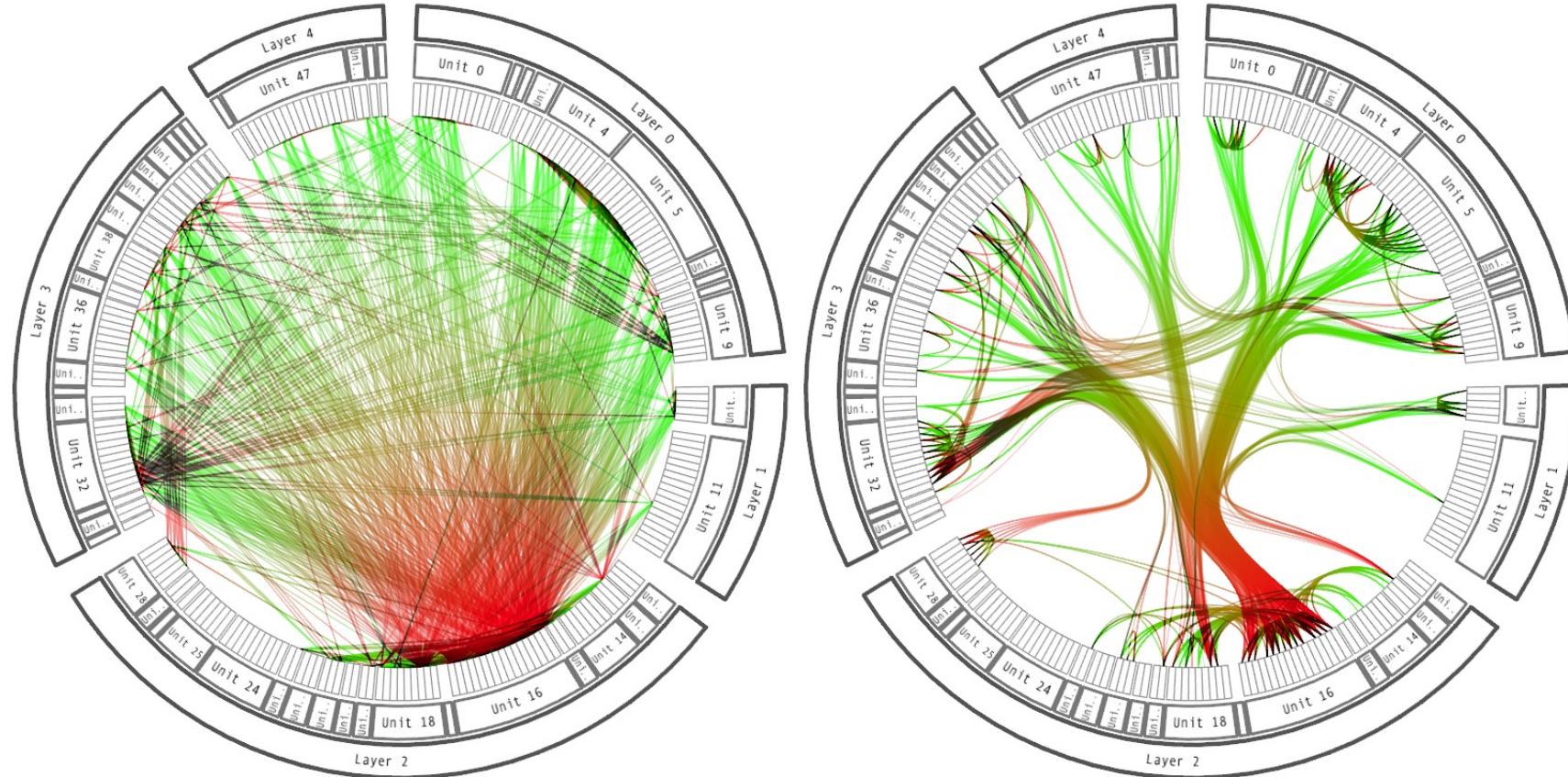


M. Rosvall and C.T. Bergstrom. Maps of Random Walks on Complex Networks Reveal Community Structure

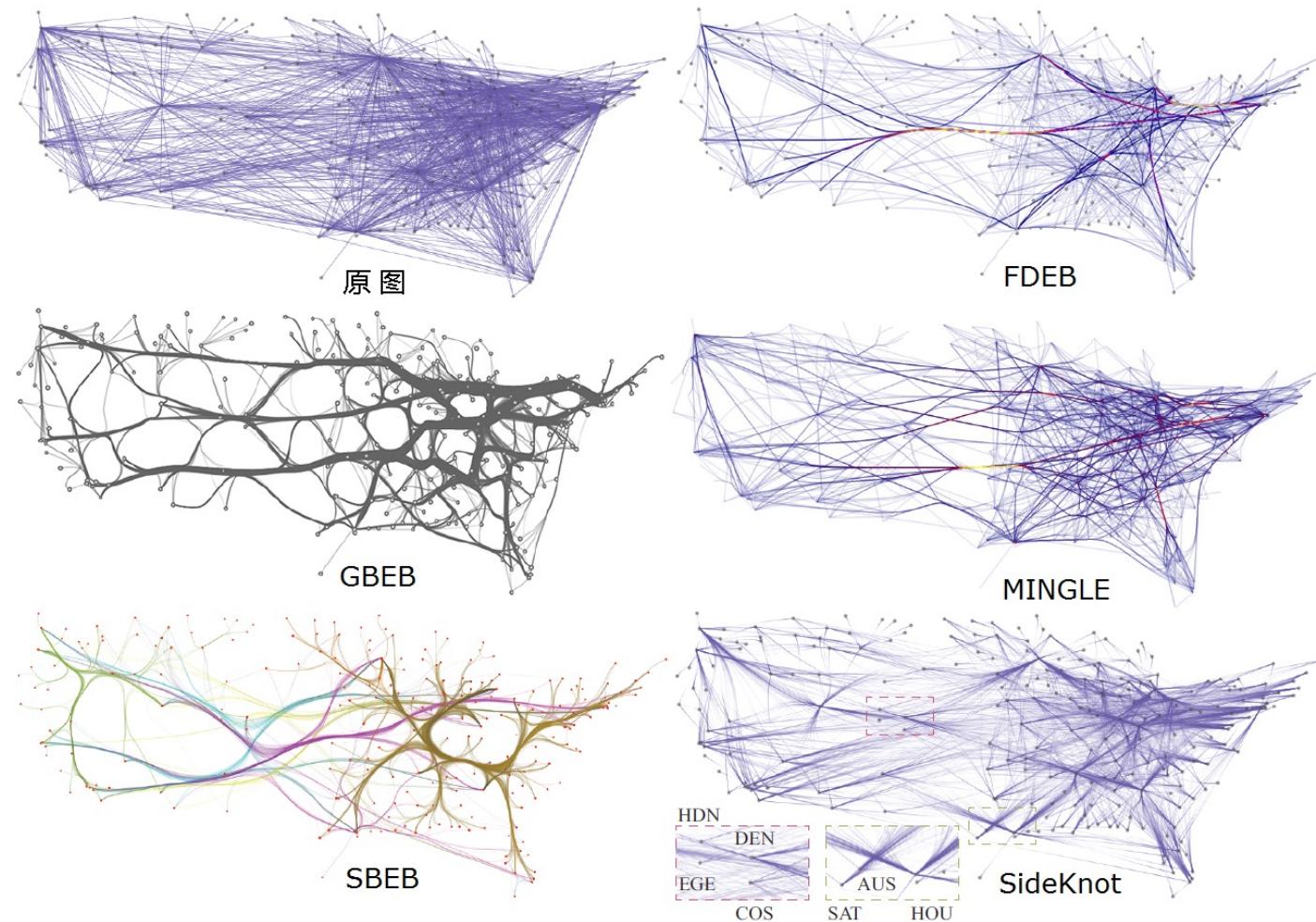


立志成才报国裕民

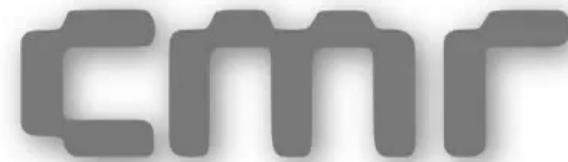
# Edge Bundling



# Edge Bundling



# KDE-based Graph Simplification



## Interactive Visualization of Streaming Data with Kernel Density Estimation

Ove Daae Lampe, CMR and UiB \*  
Helwig Hauser, UiB +

\* odl@cmr.no  
<http://cmr.no>

+ Helwig.Hauser@UiB.no  
<http://www.ii.uib.no/vis/>

O. D. Lampe and H. Hauser. Interactive visualization of streaming data with kernel density estimation.

Ove Daae Lampe, et al. Interactive visualization of stream data with kernel density estimation. PacificVis 2011



立志成才报国裕民

# KDE-based Graph Simplification



上海科技大学  
ShanghaiTech University



Michael Zinsmaier, et al. Interactive Level-of-Detail Rendering of Large Graphs. TVCG 2012

立志成才报国裕民

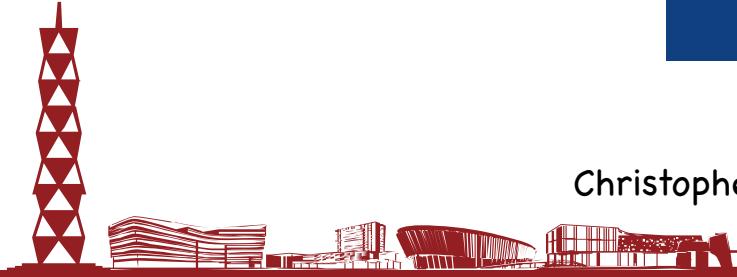
# KDE-based Graph Simplification



上海科技大学  
ShanghaiTech University



Christophe Hurter, et al. Graph Bundling by Kernel Density Estimation. EuroVis 2012



立志成才报国裕民

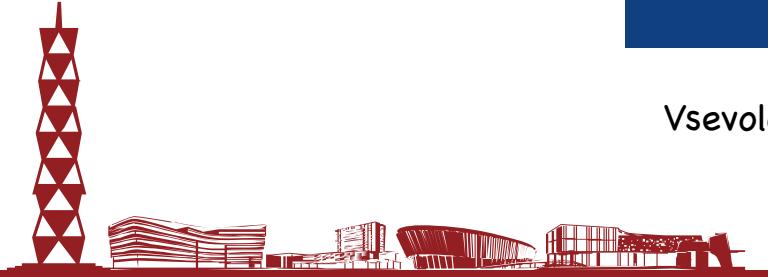
# Attribute-Driven Edge Bundling



上海科技大学  
ShanghaiTech University



Vsevolod Peysakhovich, et al. Attribute-Driven Edge Bundling for General  
Graphs with Applications in Trail Analysis. PacificVis 2015



立志成才报国裕民



# Edge bundling of huge graphs



上海科技大学  
ShanghaiTech University

## FFTEB: Edge Bundling of Huge Graphs by the Fast Fourier Transform



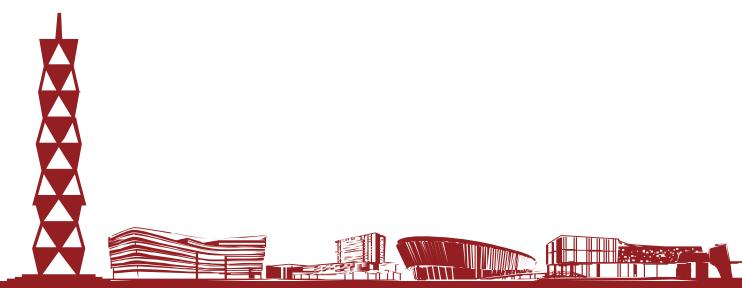
Antoine Lhuillier, et al. FFTEB: Edge bundling of huge graphs by the Fast Fourier Transform. PacificVis 2017

立志成才报国裕民



上海科技大学  
ShanghaiTech University

# Interaction



立志成才报国裕民

# Labeling

## Non-text labels

- Color, size, shape, thickness
- Too many labels lead to misinterpretation

## Labels of nodes

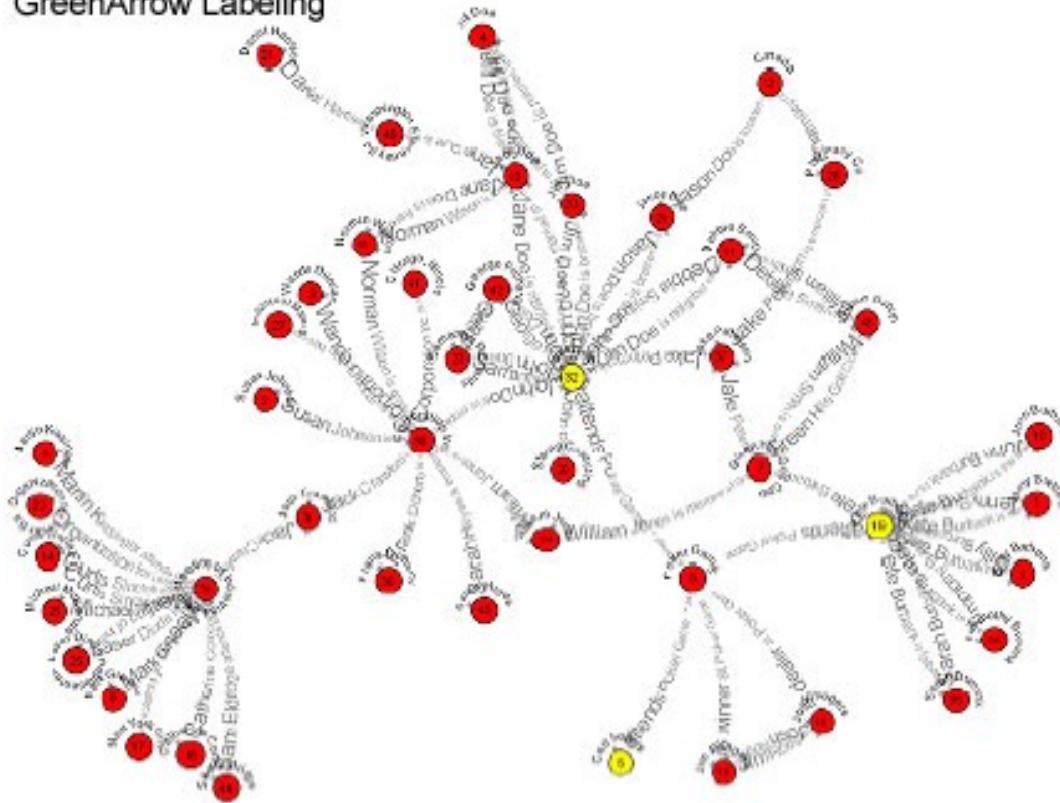
- Inside the node object

## Labels of edge

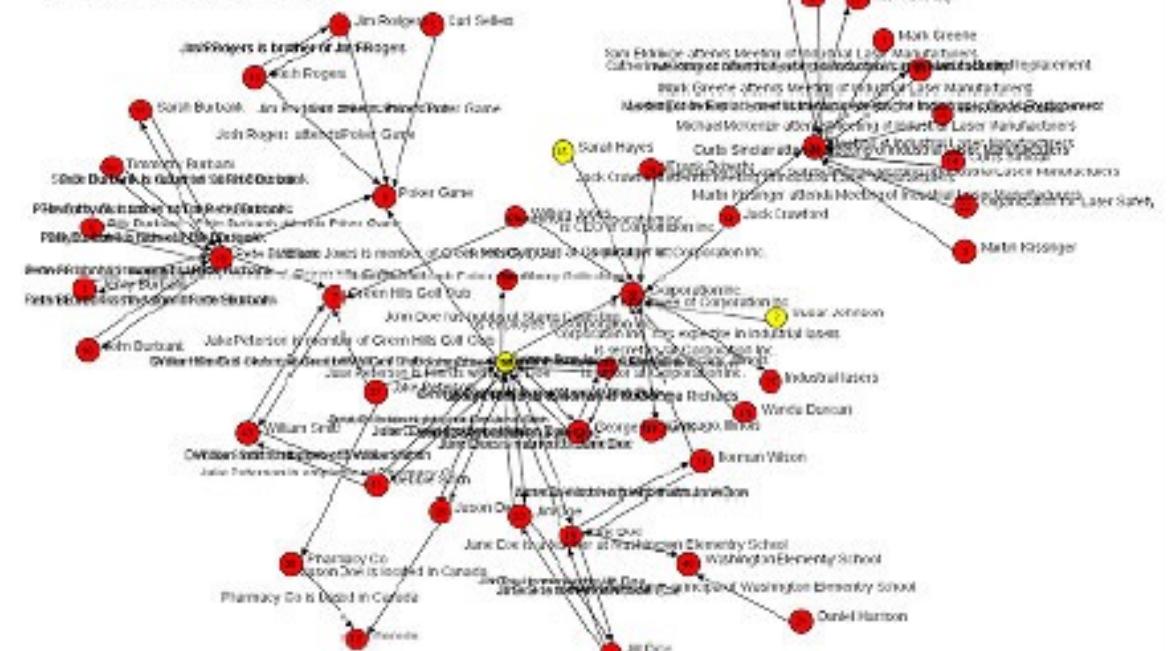
- Along the edge

# Labeling

GreenArrow Labeling



Traditional Labeling



Labels are also nodes and edges

立志成才报国裕民

# Interactions

## View-based interaction

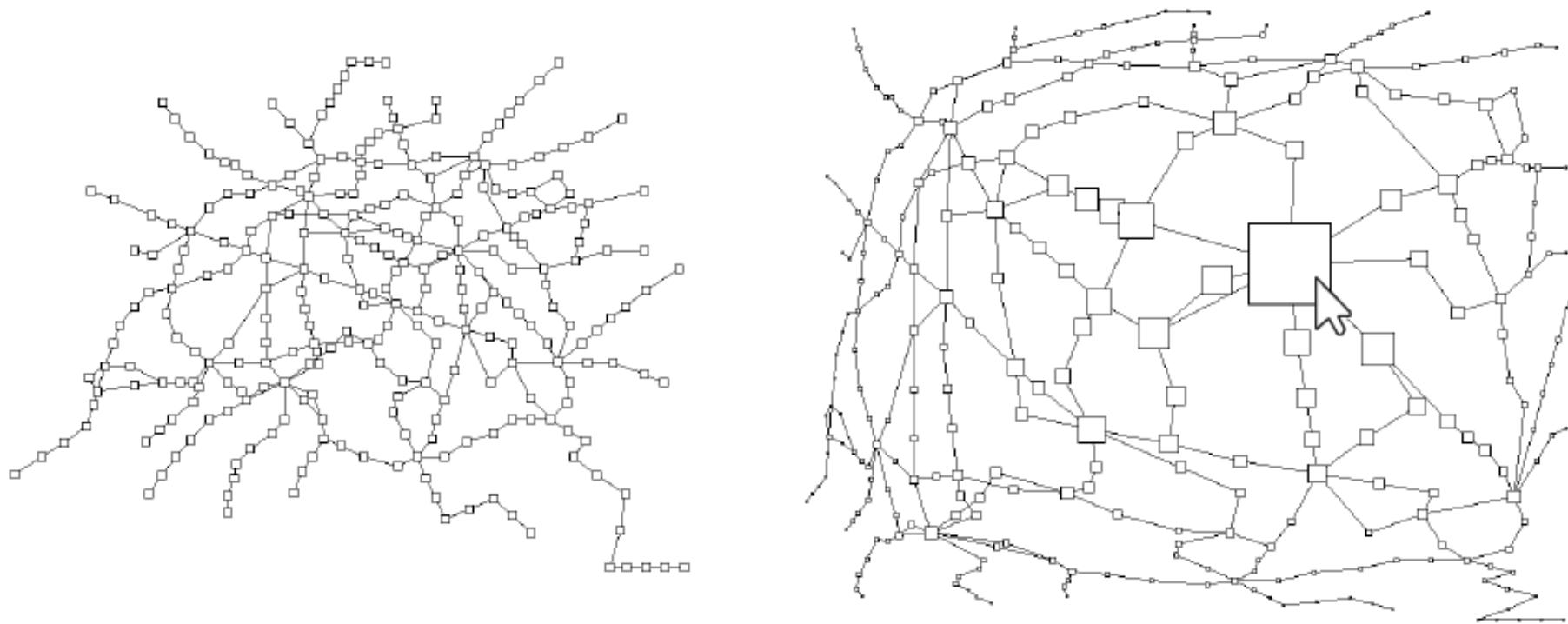
- Panning, zooming, rotating

## Object-based interaction

- Selecting, highlighting, deleting, dragging

## Structure-based interaction

- Reorder, re-layout
- Focus + context
- Roll-up, drill-down



# Fisheye

- Originate in wide-angle lens in photography.
- Popping out objects in the center and also covering its context.

# Large Scale Graph Interaction

TVCG PAPER

## CUBu: Universal real-time bundling for large graphs

Matthew van der Zwan, Valeriu Codreanu, Alexandru Telea



23-28 October 2016  
Baltimore, Maryland, USA

[ieeevis.org](http://ieeevis.org)

Matthew van der Zwan, et al. CUBu: Universal real-time bundling  
for large graphs. IEEE VIS 2016



立志成才报国裕民

# Tools and Applications

---





上海科技大学  
ShanghaiTech University

the **prefuse**  
interactive visualization toolkit  
  
selected applications

<http://prefuse.org/>

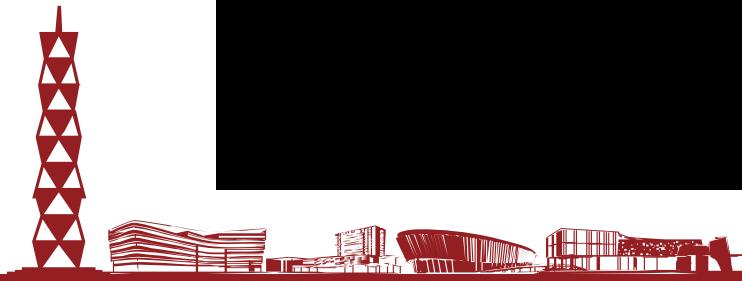
立志成才报国裕民



上海科技大学  
ShanghaiTech University

# Gephi

The Open Graph Viz Platform



<https://gephi.org/>

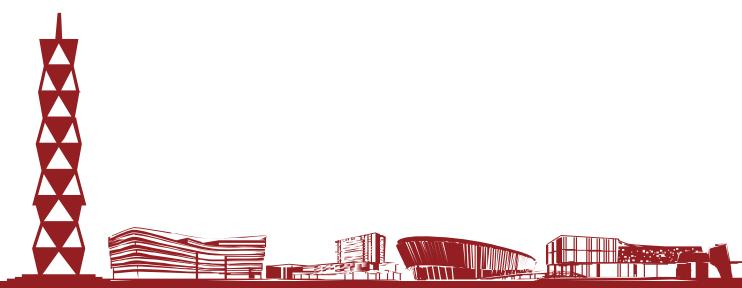
立志成才报国裕民





上海科技大学  
ShanghaiTech University

# Social Media Visual Analytics



立志成才报国裕民

# Outline



Social Media  
Data Category



Social Media  
Visualization



Social Media  
Visual Analytics



Application and  
System

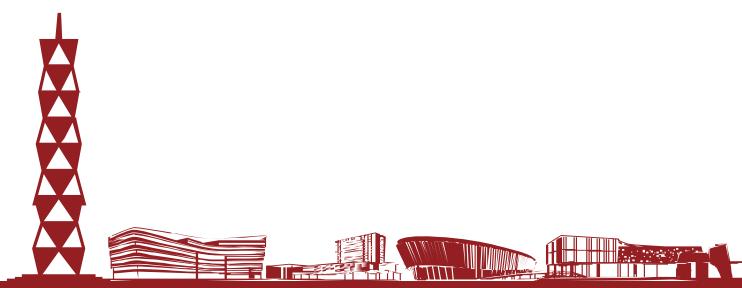
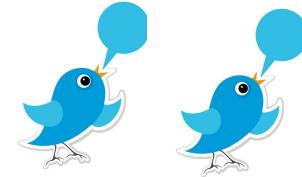


Discussion and  
Summary



# Social Media

- Entity
  - Users
  - Messages
- User behaviors
  - Posting messages (Tweet)
  - Commenting and Re-posting (Re-tweet)



# Social Media Data



Marian Dörk @nrchtct · 8h great winning projects of this year's DH award by @ifDHberlin:  
1st: [edition-humboldt.de](http://edition-humboldt.de)  
2nd: [kunstgeschichte.hu-berlin.de/institut/media...](http://kunstgeschichte.hu-berlin.de/institut/media...)  
#DHPreis17

Robert Kosara @eagereyes · 6h The TV in my hotel room here in Barcelona has every analog video connector known to mankind: component, composite, S-Video, VGA, even SCART

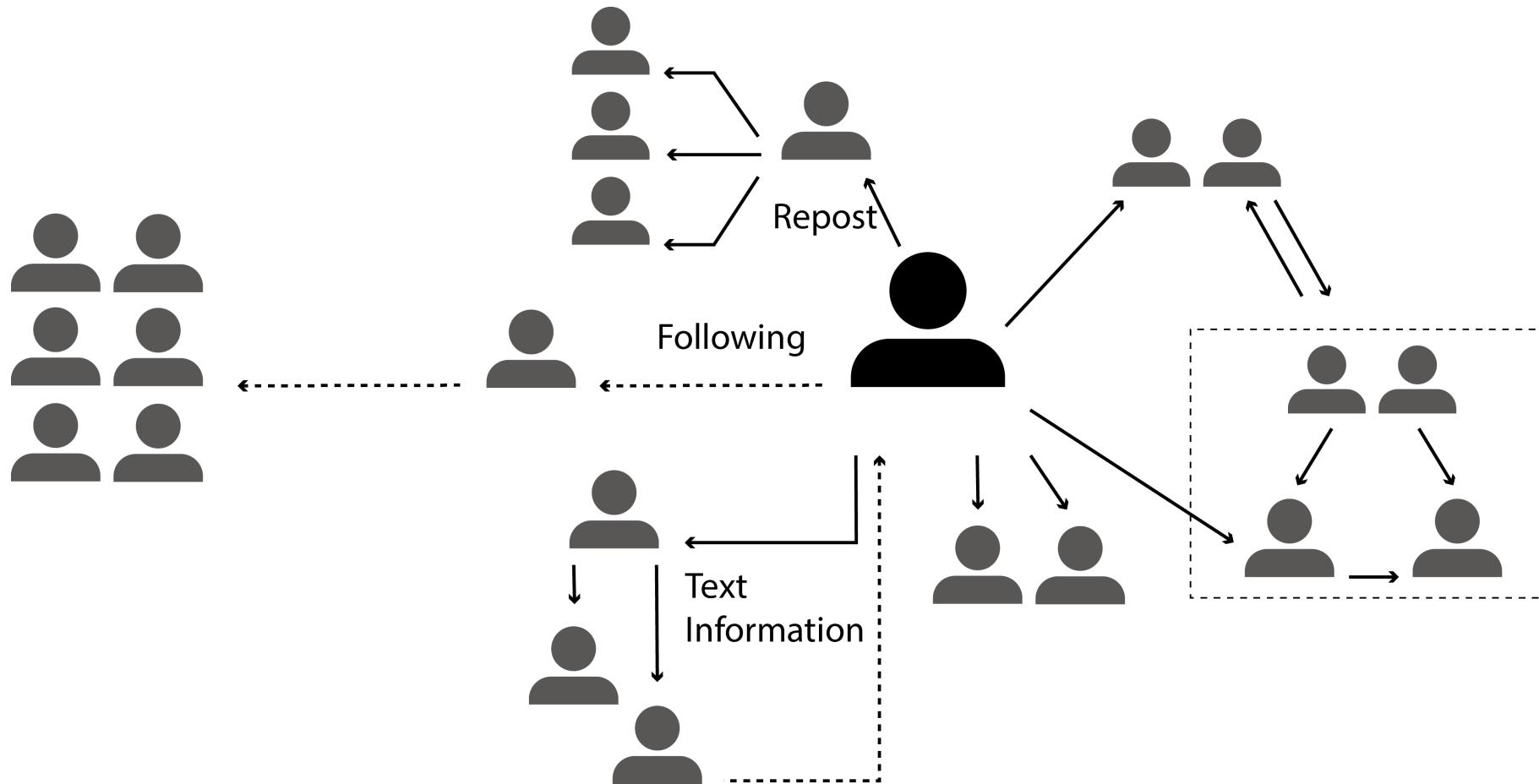


IBM Security @IBMSecurity · Apr 7 Discover all endpoints on the network and get advanced behavioral analytics to detect evasive attacks.

Complete visibility into your endpoint landscape  
IBM BigFix Detect

- Users
- Messages
  - Time
  - Location
  - Content
  - Multi Media
- Reposting Behaviors

# Social Network in Social Media



# Geo-tagged Social Media

垦丁的风景确实很漂亮。  台湾



6-17 21:23 come from iPad客户端

每个城市的夜市都人满为患，摊位上吆喝着各自的黑暗料理。  台湾·花莲火车站



6-19 22:44 come from iPad客户端

看完后很大的感触是，有钱人真会玩。  台湾·台北故宫博...

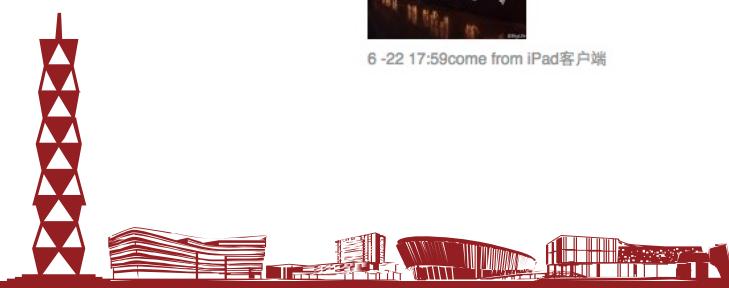
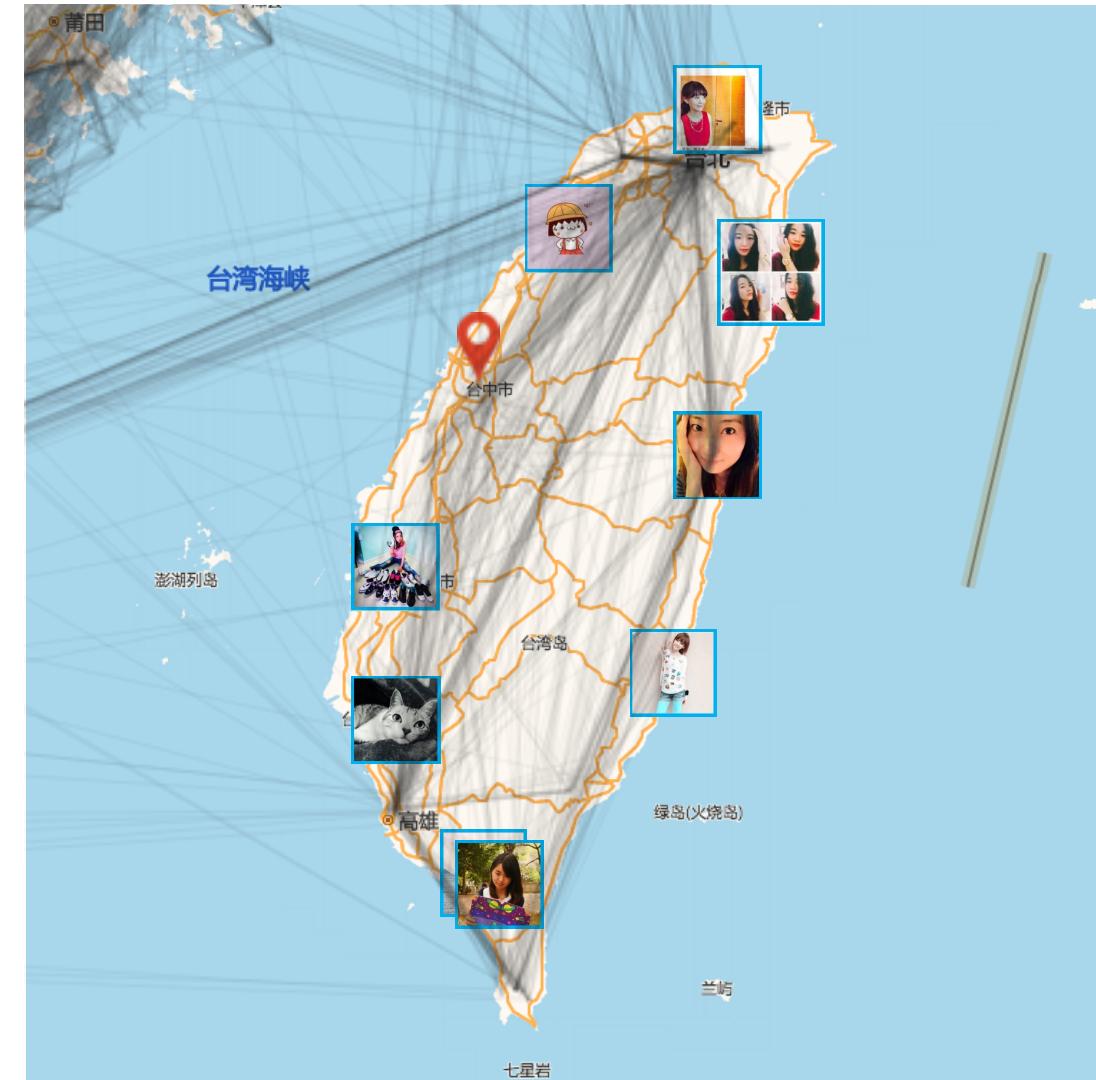


6-22 08:55 come from iPad客户端

这里有我们丢失的一些东西。  台湾



6-22 17:59 come from iPad客户端



# Classification of Social Media Entities and Related Visualization Techniques

Entities in Social Media			
Network	Geographic Information	Text and Content	
People's Follower Network	Geographic Information Diffusion	Keywords	
Messages' Diffusion Network	Spatial Temporal Event Distribution	Topic	
People's Reposting Network	Movement Trajectory	Sentiment	

**Network** 

**Geographic Information** 

**Text and Content**

**Keywords**

**Topic**

**Sentiment**

+ Time  
↓  
Dynamic Network

Geographic Information Diffusion

Spatial Temporal Event Distribution

Movement Trajectory

+ Time  
↓  
Spatial Temporal Scenes

+ Time  
↓  
Dynamic Semantic Flow

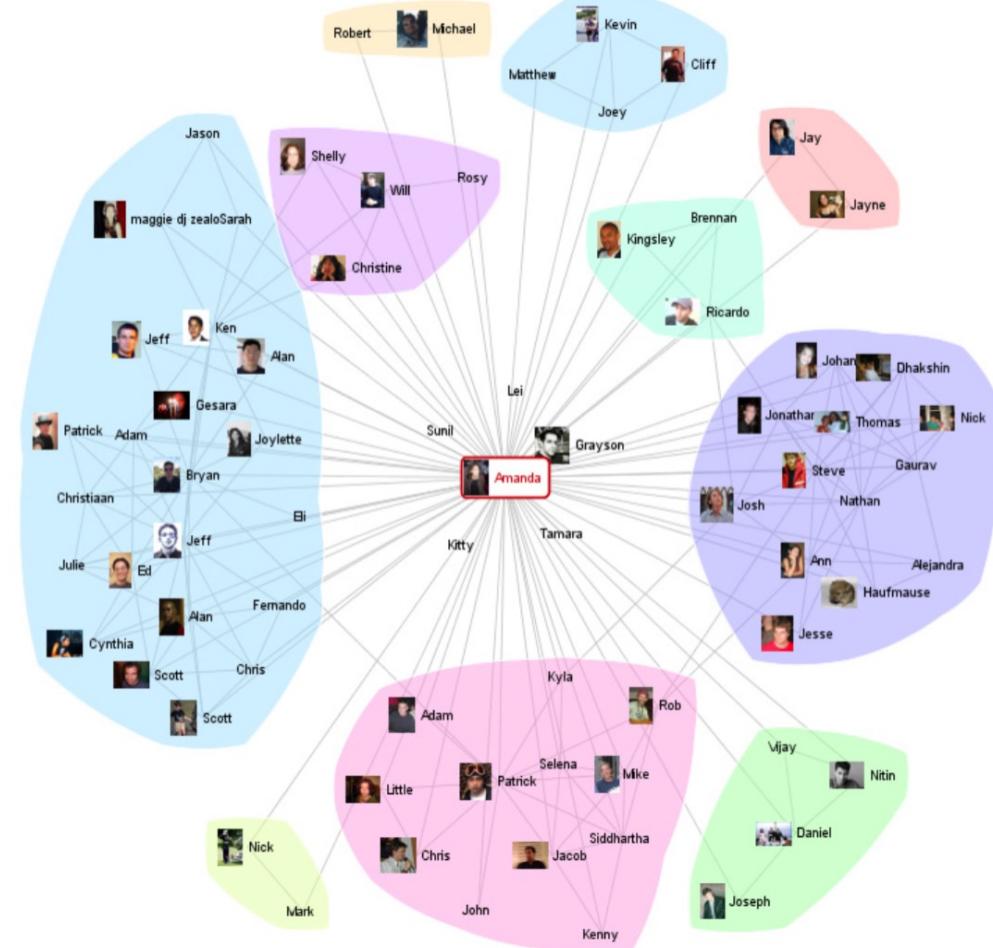
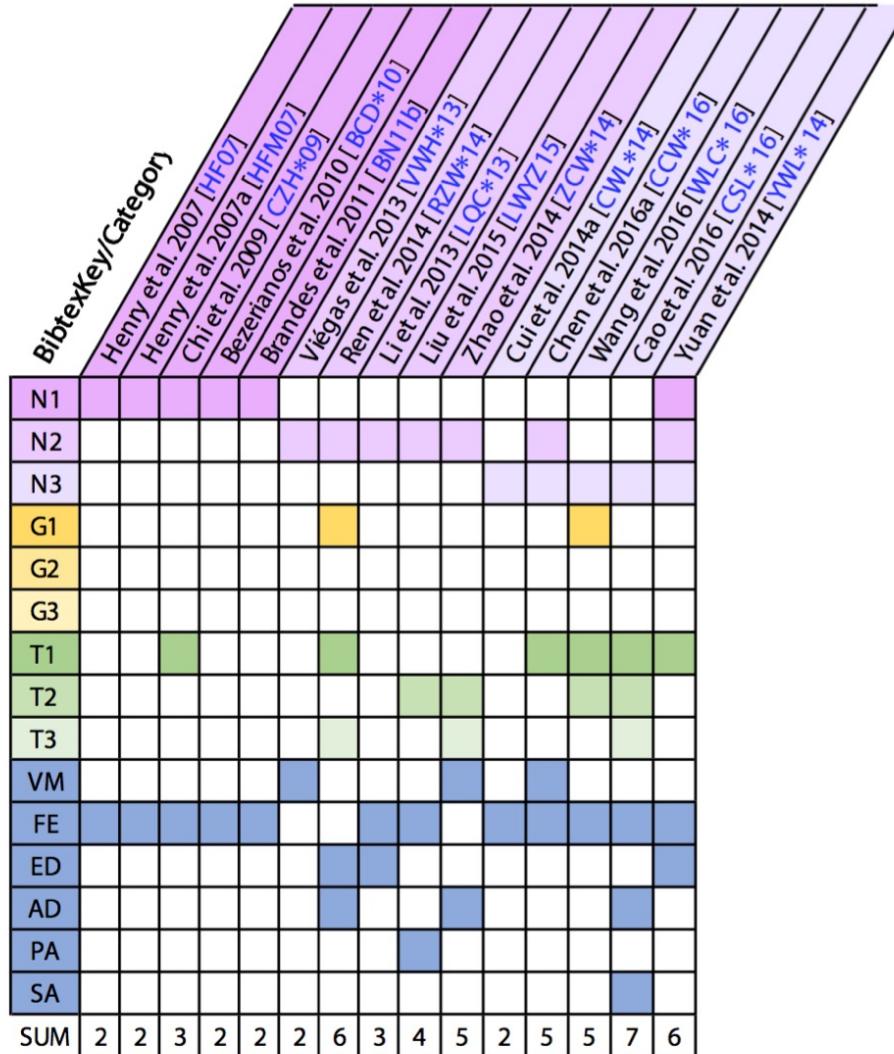
# Survey Methodology

BibTeX Key/Category	Henry et al. 2007 [HFH07]	Henry et al. 2007 [HFH07]	Chi et al. 2007 [CHF07]	Bezerianos et al. 2009 [CBH09]	Brändes et al. 2010 [BCD*10]	Riegas et al. 2011 [BRI11b]	Li et al. 2013 [LWH*13]	Liu et al. 2013 [RZW*13]	Zhao et al. 2015 [LOC13]	Cui et al. 2014 [WZL15]	Chen et al. 2014 [ZCM*14]	Wang et al. 2016 [CCW*14]	Yuan et al. 2016 [WLC*16]	Cao et al. 2014 [CS*16]	Croitoru et al. 2012 [WZL*14]	Zhang et al. 2013 [CLS*12]	MacEachren et al. 2013 [CCB613]	MacEachren et al. 2013 [ZW13]	Kraft et al. 2012 [MFR*11]	Chae et al. 2012 [CTB*11]	Thom et al. 2013 [KWD*12]	Bosch et al. 2014 [CTU*13]	Thom et al. 2012 [TBK*12]	Prieto et al. 2015 [TKE*13]	McKenzie et al. 2015 [PHE*15]	Xia et al. 2014 [XSK*14]	Chen et al. 2014 [MG*14]	Chua et al. 2016 [CYW*16]	Wu et al. 2014 [CSV*14]	Chae et al. 2014 [LSKG*16]	Andrienko et al. 2015 [WZS14]	Andrienko et al. 2012 [AAS*12]	Dörk et al. 2013 [AAF*12]	Dialloopoulos et al. 2010 [DGWC10]	Fischer et al. 2011 [DNYS11]	Archambault et al. 2010 [DNK10]	Besler et al. 2014 [FS14]	Hu et al. 2012 [BBG*12]	Liu et al. 2017 [HWS17]	Wanner 2016 [LLZ*16]	Dou et al. 2016 [WJS*16]	Dou et al. 2012 [DW*16]	Gansner et al. 2015a [DCS*15]	Liu et al. 2013 [GHW13]	Ribarsky et al. 2013 [LWW*13]	Wang et al. 2014 [HYZ*14]	Dou et al. 2014 [RWD14]	Xu et al. 2013 [CLWW14]	Wu et al. 2013 [XWW*13]	Sun et al. 2014 [WLY*13]	Hu et al. 2014 [DW*12]	Zhao et al. 2013 [HYZ*14]	Steed et al. 2014 [ZGWZ14]	Hu et al. 2013 [HYZ*14]	Kucher et al. 2015 [LKT*14]	Lu et al. 2014 [SBWZ14]	Li et al. 2014 [HSB*15]	Bollen et al. 2014 [LWM14]	Marcus et al. 2011 [BHZ11]	Cao et al. 2011 [BHZ11]	Rohrdantz 2014 [GLL*14]	Liu et al. 2012 [RHD*12]	SUM
N1																													7																																		
N2																													14																																		
N3																													11																																		
G1																													7																																		
G2																													23																																		
G3																													8																																		
T1																													48																																		
T2																													30																																		
T3																													20																																		
VM																													30																																		
FE																													43																																		
ED																													22																																		
AD																													12																																		
PA																													7																																		
SA																													5																																		
SUM	2	2	3	2	2	2	6	3	4	5	2	5	5	7	6	4	5	4	5	6	6	2	5	5	3	3	4	3	5	5	5																																

- N1-N3: Following Network, Information Diffusion Network and Reposting Network;
- G1-G3: Geographic Information Diffusion, Spatial Temporal Event Distribution and Movement Trajectories;
- T1-T3: Keywords, Topic and Sentiment

VM: Visual Monitoring;  
 FE: Feature Extraction;  
 ED: Event Detection;  
 AD: Anomaly Detection;  
 PA: Predictive Analysis;  
 SA: Situation Awareness

# Social Networks

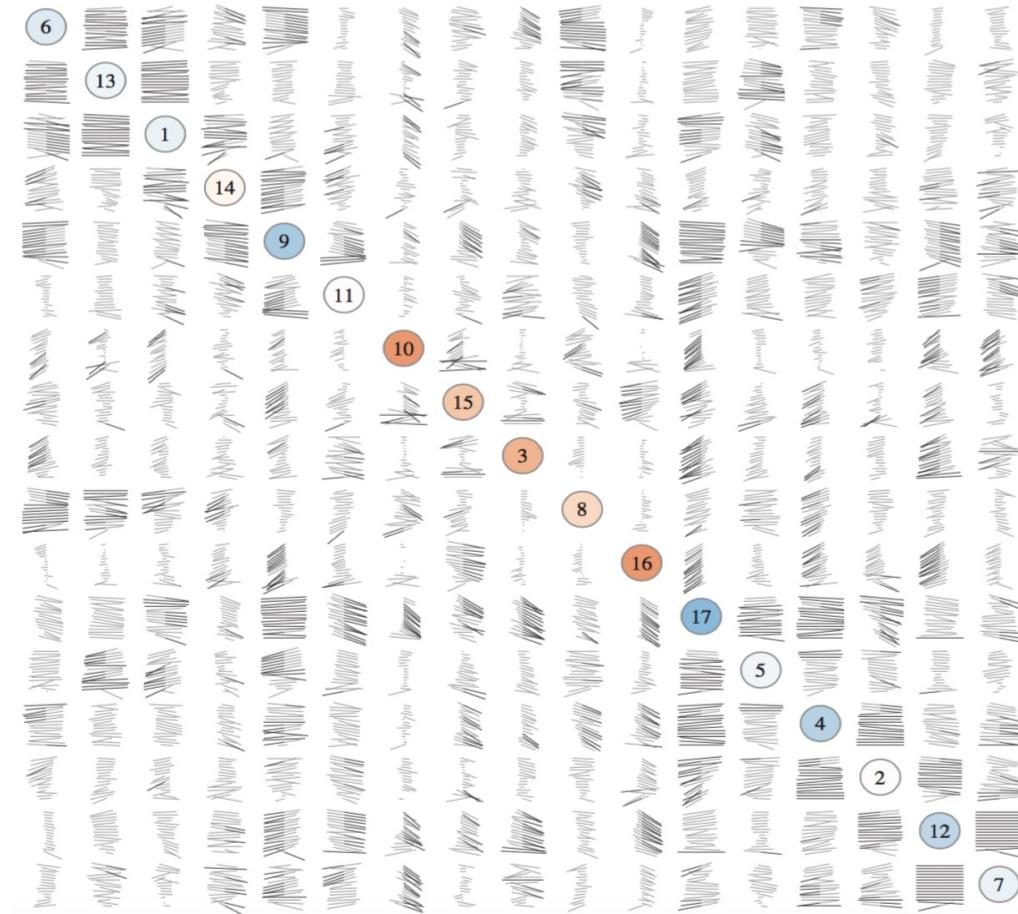
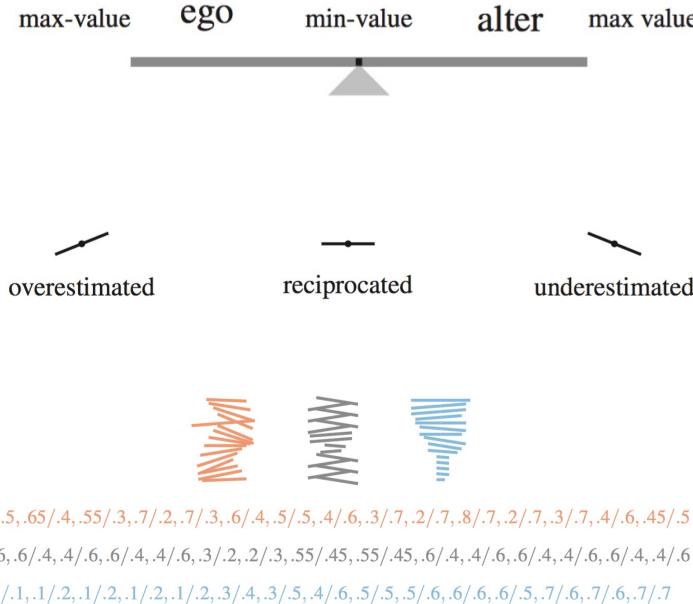


Vizster: Visualizing Online Social Networks

Jeffrey Heer, danah boyd  
IEEE Information Visualization (InfoVis), 32–39, 2005

# Social Networks - Follower Network Visualization

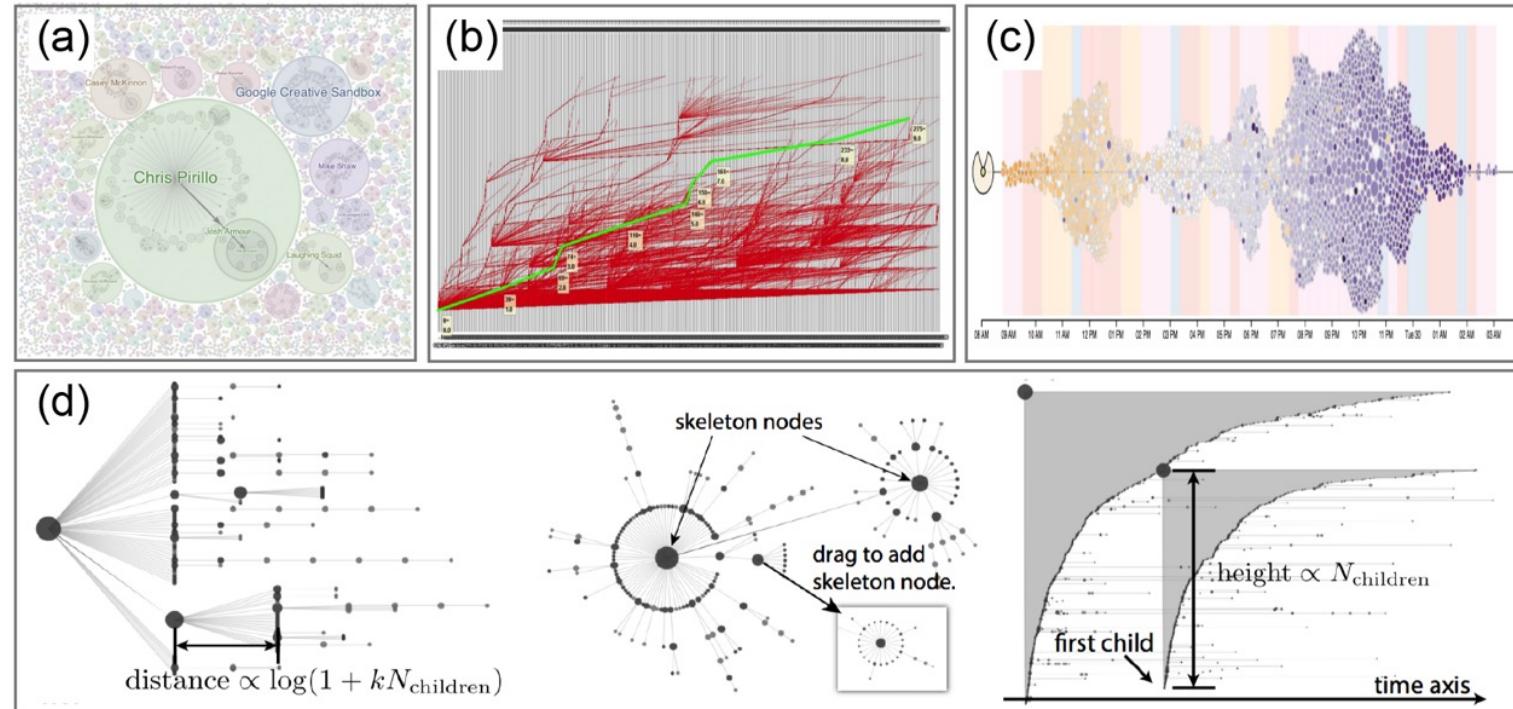
- Social Relationship
  - Multiple Attributes
  - Dynamic Features



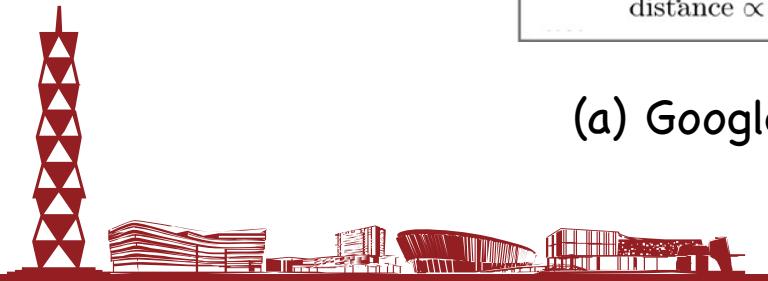
Gestaltmatrix [BN 11]

# Social Networks – Diffusion Network

- Entity - Messages
- How information diffuse along time
  - Topic
  - Relationship

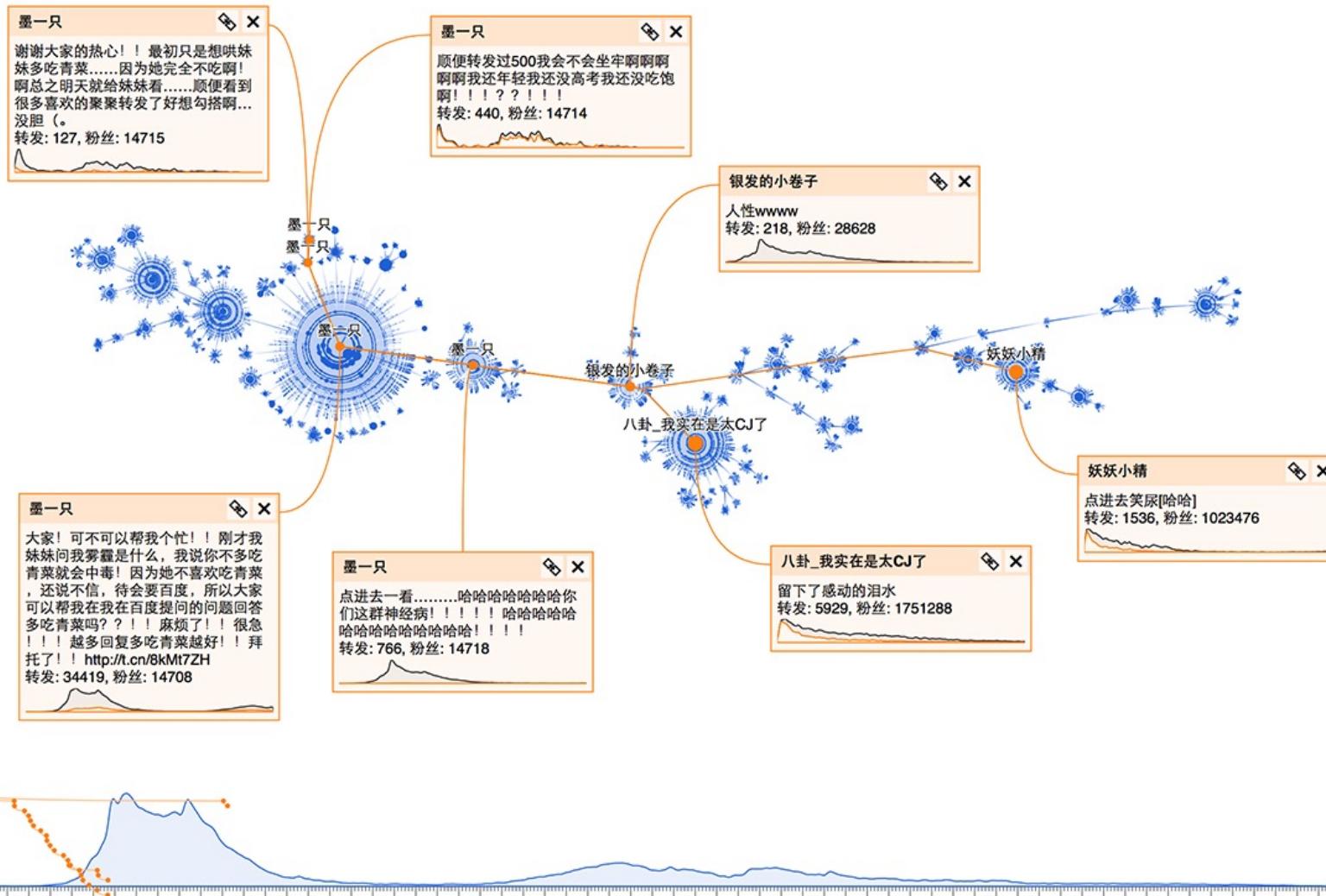


(a) Google+ Ripples [VWH\*13], (b) Li et al. [LQC\*13], (c) FluxFlow [ZCW\*14], (d) WeiboEvents [RZW\*14]



# Diffusion Network Visualization: WeiboEvents

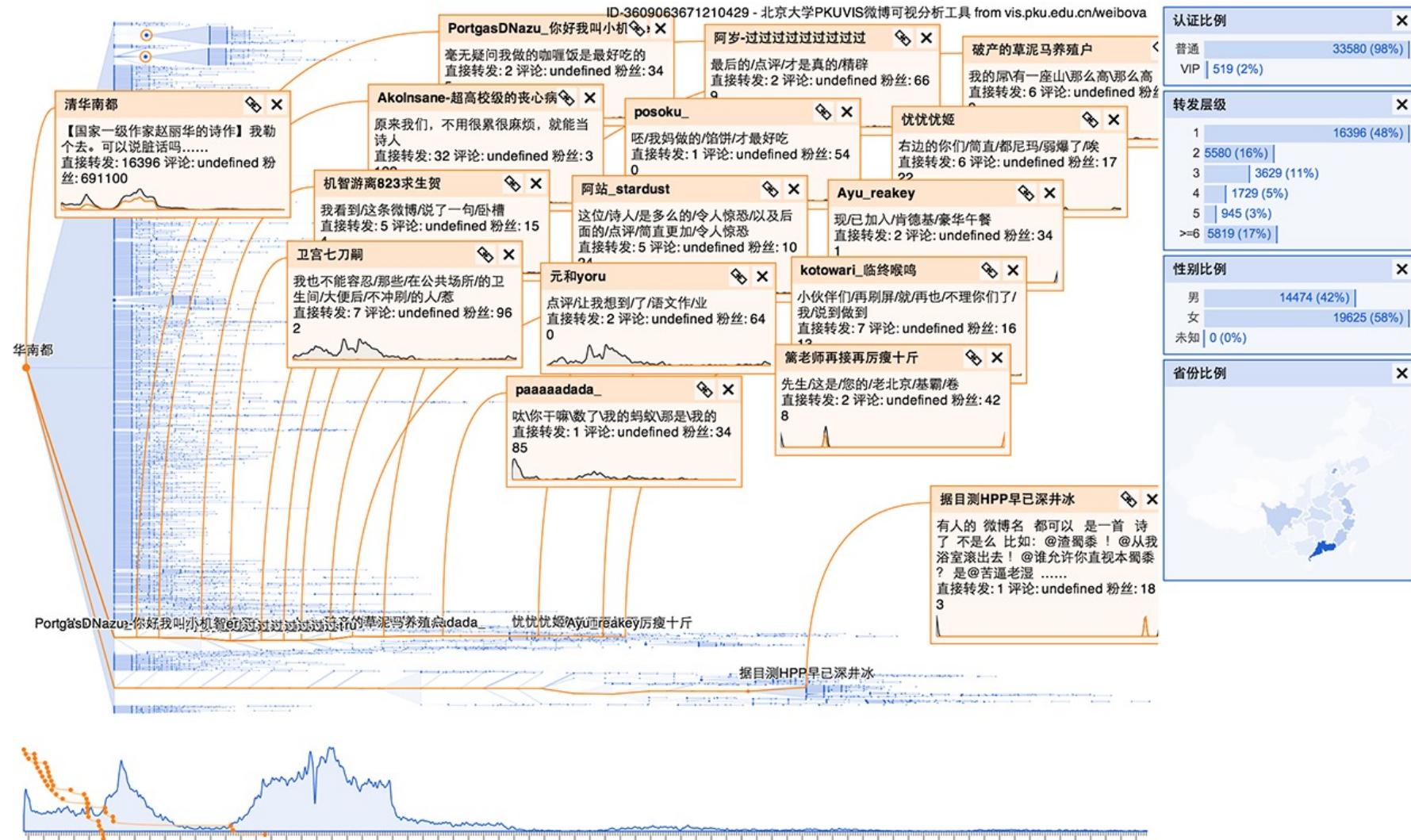
## - Normal Diffusion Patterns



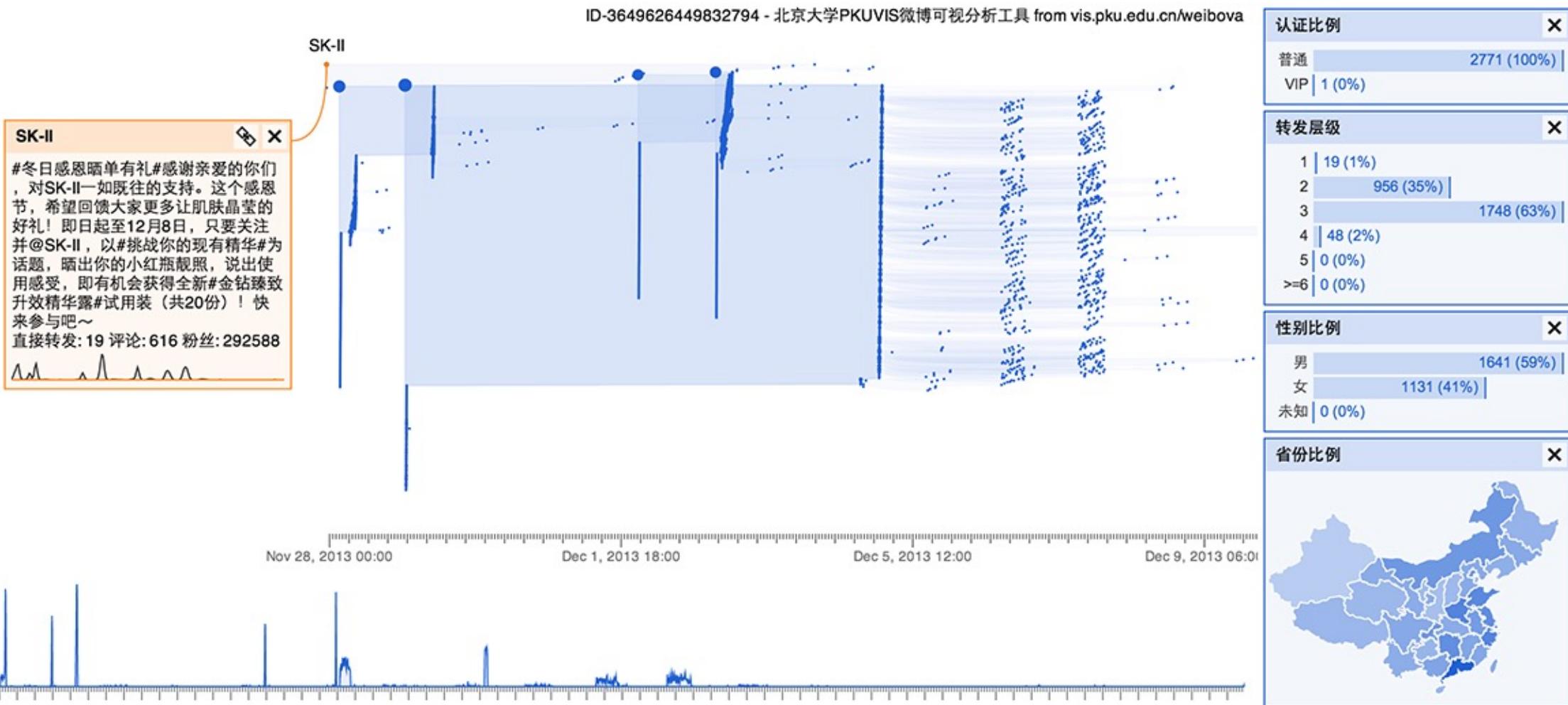
Ren, Donghao, Xin Zhang, Zhenhuang Wang, Jing Li, and Xiaoru Yuan. "WeiboEvents: A Crowd Sourcing Weibo Visual Analytic System." In *Pacific Visualization Symposium (PacificVis) Notes, 2014 IEEE*, pp. 330-334. IEEE, 2014.

# Diffusion Network Visualization: WeiboEvents

## - In-depth Discussions



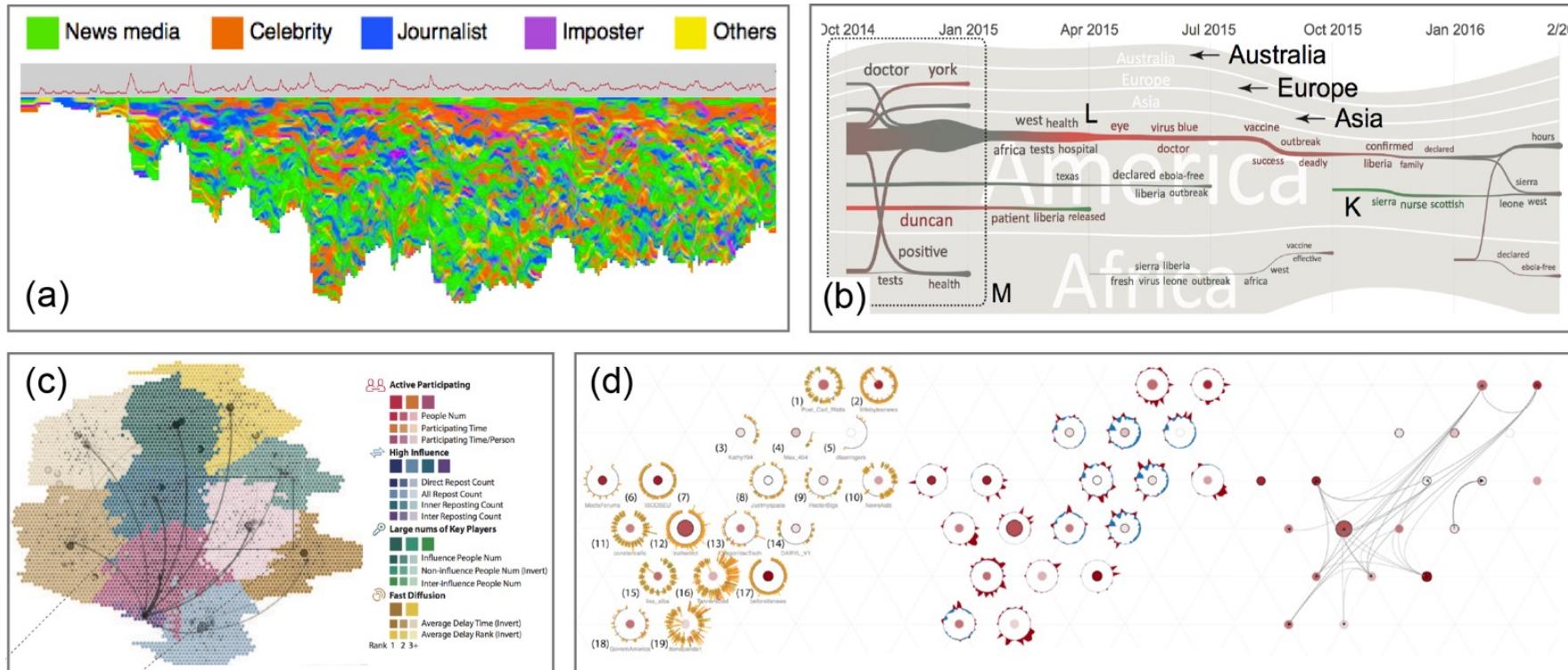
# Diffusion Network Visualization: WeiboEvents: Machine Account Behavior



Ren, Donghao, Xin Zhang, Zhenhuang Wang, Jing Li, and Xiaoru Yuan. "WeiboEvents: A Crowd Sourcing Weibo Visual Analytic System." In *Pacific Visualization Symposium (PacificVis) Notes, 2014 IEEE*, pp. 330–334. IEEE, 2014.

# Social Network: Reposter Network

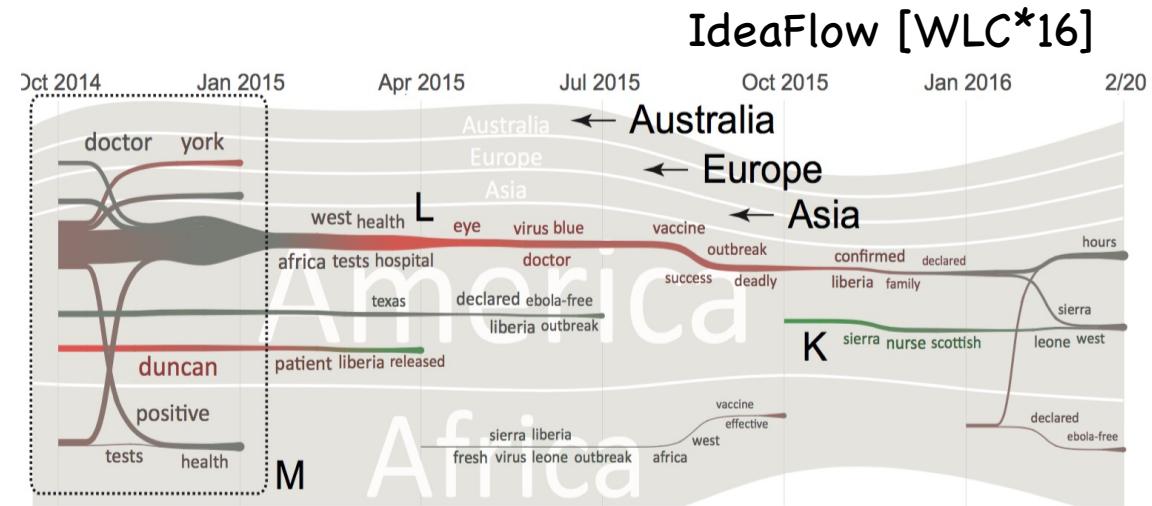
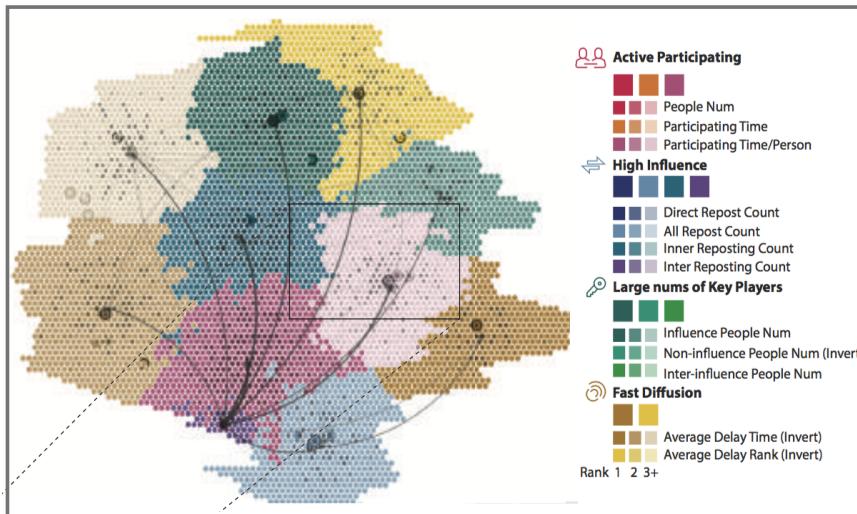
- Entities – Social media users
- Reposting behaviors leading to information diffusion
  - User behavior / type / relationship
  - Dynamic scenarios



(a) GraphFlow [CWL\*14], (b) IdeaFlow [WLC\*16], (c) D-Map [CCW\*16], (d) TargetVue [CSL\*16]

# Social Network: Reposter Network

- Dynamic Relationship
- Community Behaviors
- Topic Changes



- River-based Visual Metaphor
- Map-based Visual Metaphor

**D-Map [CCW\*16]**

# D-Map: Visual Analysis of Ego-centric Information Diffusion Patterns in Social Media

Siming Chen<sup>1</sup>, Shuai Chen<sup>1</sup>, Zhenhuang Wang<sup>1</sup>, Jie Liang<sup>1,2</sup>,  
Xiaoru Yuan<sup>1</sup>, Nan Cao<sup>3</sup>, Yadong Wu<sup>4</sup>

<sup>1</sup>Key Laboratory of Machine Perception (Ministry of Education), and School of EECS, Peking University, China

<sup>2</sup>Faculty of Engineering and Information Technology, The University of Technology, Sydney, Australia

<sup>3</sup>New York University, Shanghai, China

<sup>4</sup>Southwest University of Science and Technology, China

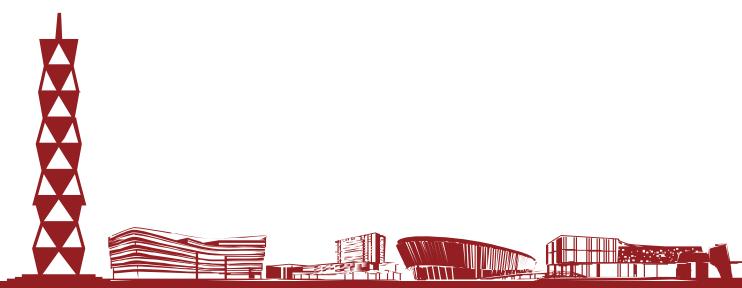
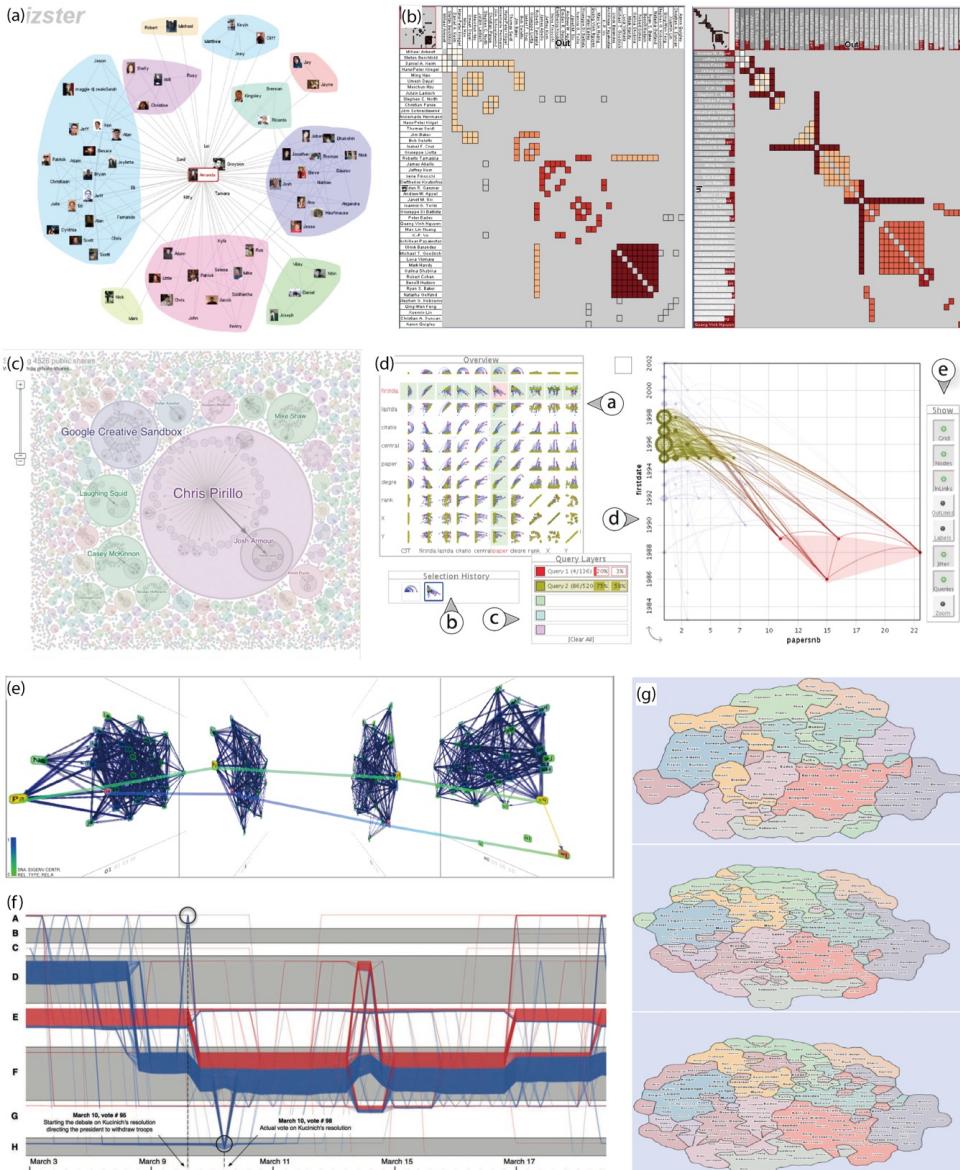
# Social Network Visualization Techniques: Summary

- Graph-based visualization

- Network [Heer2005]
- Matrix [Henry2006]
- Space filling techniques [Viegas2013]
- Multi-variate exploration [Bezerianos2010]

- Dynamic graph visualization

- 2.5D representation [Federico2011]
- Storyline [Reda2011]
- Map-like visual metaphor [Hu2012]



# Classification of Social Media Entities and Related Visualization Techniques

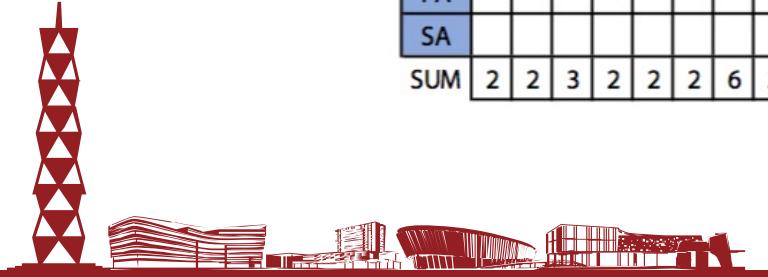
Entities in Social Media					
Network			Geographic Information	Text and Content	
People's Follower Network	 - 	+ Time ↓ Dynamic Network	Geographic Information Diffusion  -   - 	Spatial Temporal Event Distribution  	Keywords Topic Sentiment    
Messages' Diffusion Network	 - 		Movement Trajectory  	Spatial Temporal Scenes 	+ Time ↓ Dynamic Semantic Flow
People's Reposting Network	 - 				



# Paper Collections

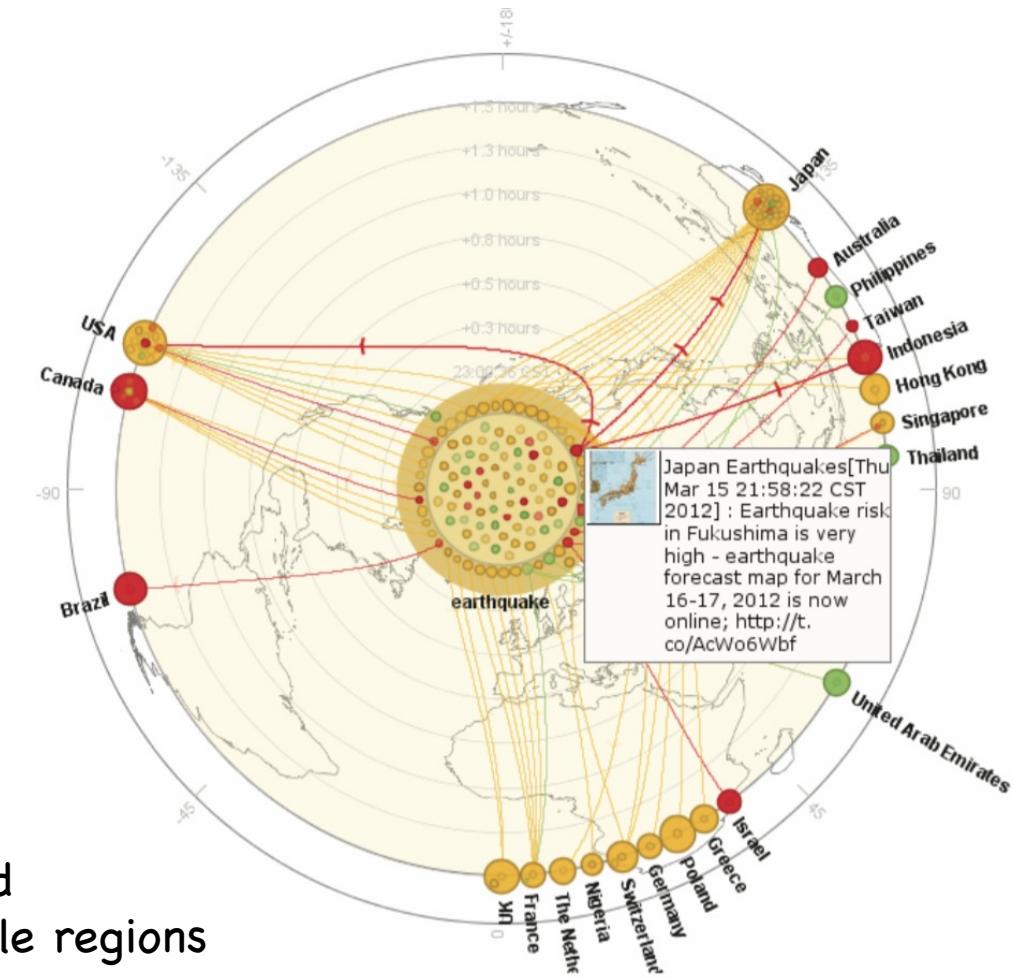


		BibTeX Key/Categories																																				
		N1		N2		N3		G1		G2		G3		T1		T2		T3		VM		FE		ED		AD		PA		SA								
		Henry et al. 2007 [HF07]	Henry et al. 2009 [HFM09]	Chi et al. 2007a [HFM07]	Bezerianos et al. 2009 [CZH*09]	Vilégas et al. 2010 [BCD*10]	Ren et al. 2011 [BN11b]	Li et al. 2013 [VWH*13]	Liu et al. 2013 [RZW*13]	Zhao et al. 2015 [LQC*13]	Cui et al. 2014 [WYZ15]	Chen et al. 2014a [ZCW*14]	Wang et al. 2016a [CCW*16]	Cao et al. 2016 [MLC*16]	Yuan et al. 2016 [CSL*16]	Cao et al. 2014 [YWL*16]	Croitoru et al. 2012 [CLS*14]	Zhang et al. 2013 [CCRS13]	MacEachren et al. [ZLW13]	Chae et al. 2011 [MJR*11]	Kraft et al. 2012 [CTB*12]	Chae et al. 2013 [KWD*13]	Thom et al. 2014 [CTJ*14]	Bosch et al. 2012 [TBK*12]	Prieto et al. 2015 [BTH*13]	McKenzie et al. 2015 [TKE*15]	Xia et al. 2016 [CNW*16]	Chen et al. 2014 [CMSVM14]	Liu et al. 2016 [LSKG14]	Chae 2015 [WZSL14]	Andrienko et al. 2015 [CCJ*15]	An et al. 2015 [An*15]						
SUM		2	2	3	2	2	2	6	3	4	5	2	5	5	7	6	7	6	4	5	4	5	6	4	5	6	6	2	5	5	4	3	3	4	4	4	3	4



# Geographic Information: Geo-tagged Information Diffusion Visualization

- Two sources of geo information
  - Where the social media users come from (Relative static)
  - Where the social media users post the messages (Dynamic)



Integrating the spatial temporal information and representing the diffusion process across multiple regions

Whisper [CLS\*12]

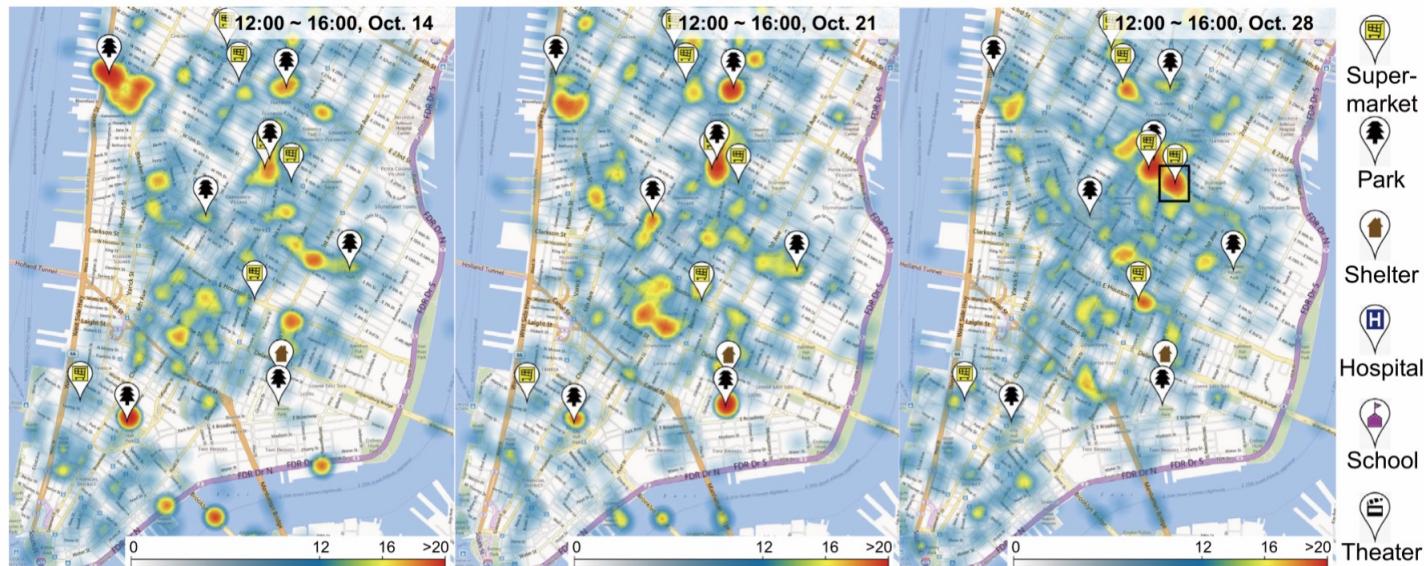
# Whisper:Tracing the Spatiotemporal Process of Information Diffusion in Real Time

InfoVis 2012

Nan Cao, Yu-Ru Lin, Xiaohua Sun,  
David Lazer, Shixia Liu, and Huamin Qu

# Geographic Information: Spatial Temporal Event Distribution Visualization

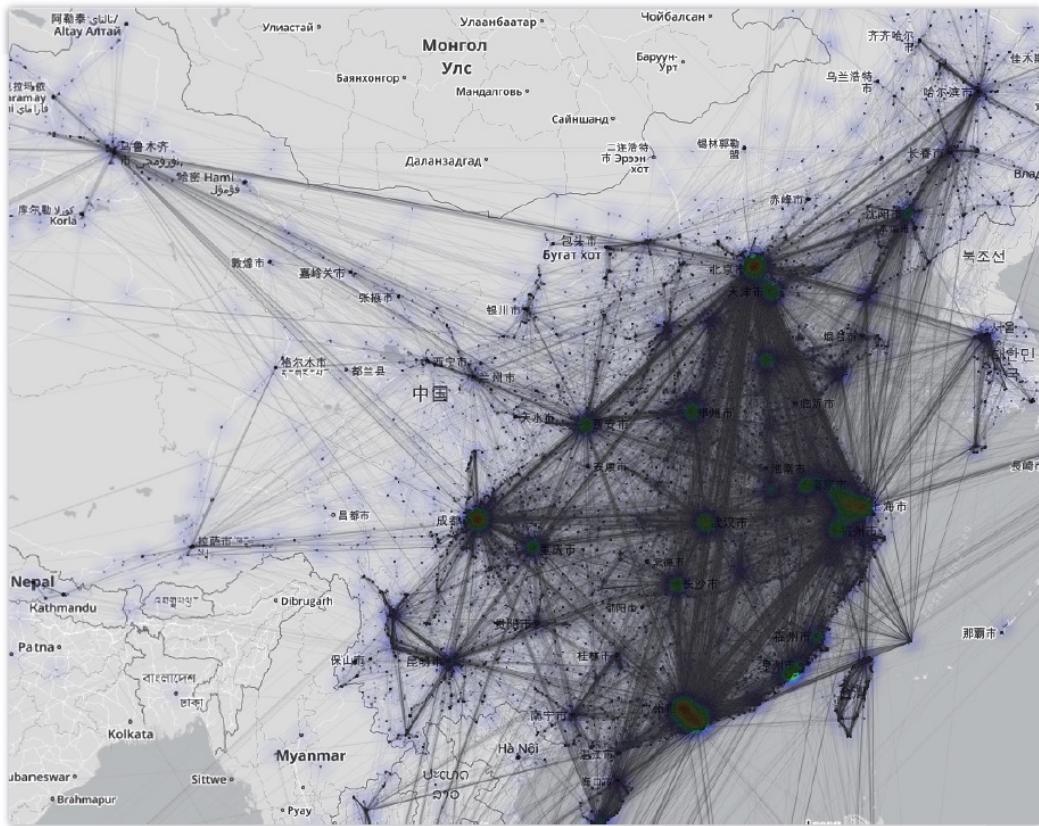
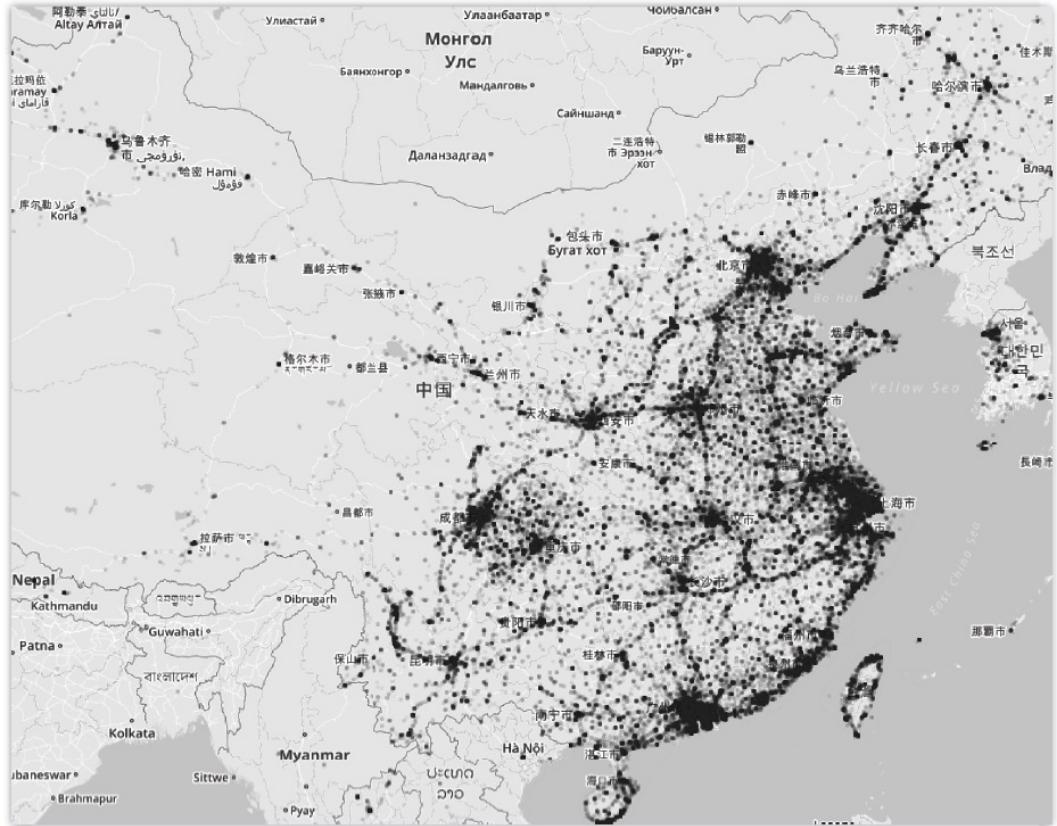
- Two sources of geo information
  - Where the social media users come from (Relative static)
  - Where the social media users post the messages (Dynamic)



A density map design with the aggregation to show the distribution of the social events [CTJ\*14]



# Geographic Information: Movement Trajectories Visualization



- Two weeks Weibo distribution in China and its neighborhood

[CYW\*16]

立志成才报国裕民



# Interactive Visual Discovering of Movement Patterns from Sparse Sampling Geo-tagged Socail Media

Siming Chen<sup>1</sup>, Xiaoru Yuan<sup>1</sup>, Zhenhuang Wang<sup>1</sup>, Cong Guo<sup>1</sup>, Jie Liang<sup>1</sup>,  
Zuchao Wang<sup>1</sup>, Xiaolong (Luke) Zhang<sup>2</sup>, Jiawan Zhang<sup>3</sup>

<sup>1</sup>Key Laboratory of Machine Perception (Ministry of Education), School of EECS, Peking University

<sup>2</sup>Pennsylvania State University

<sup>3</sup>Tianjin University

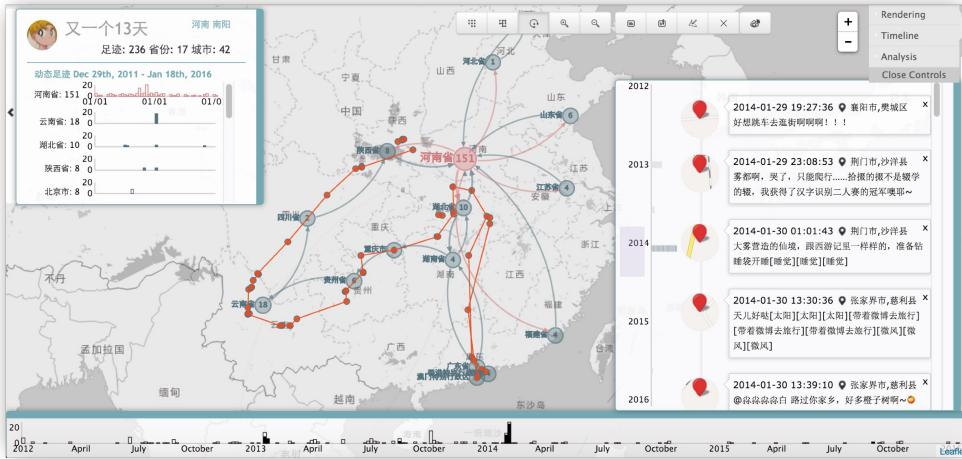
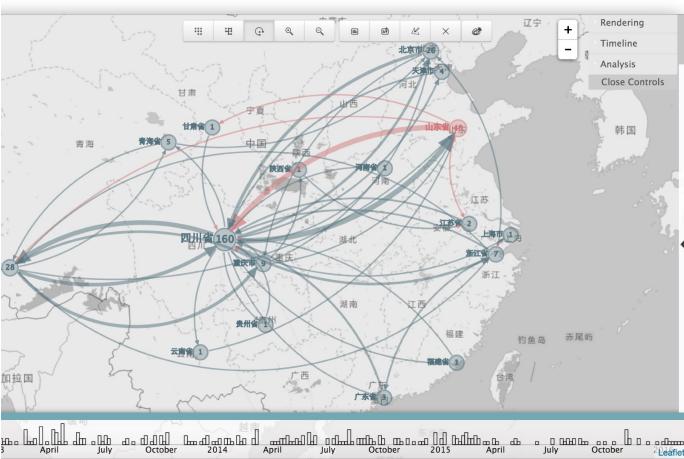
Contact:  
xiaoru.yuan@pku.edu.cn | <http://vis.pku.edu.cn>

**Vis**  **PKU**

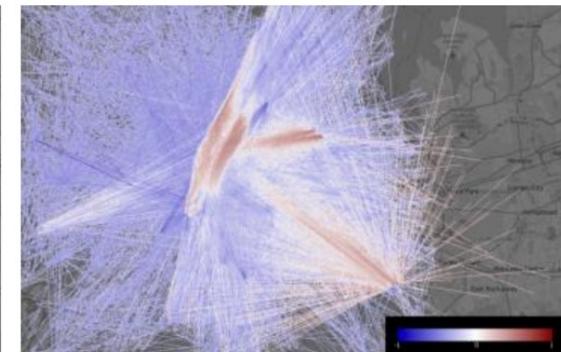
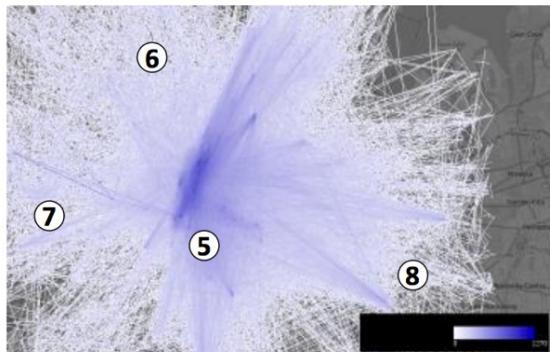
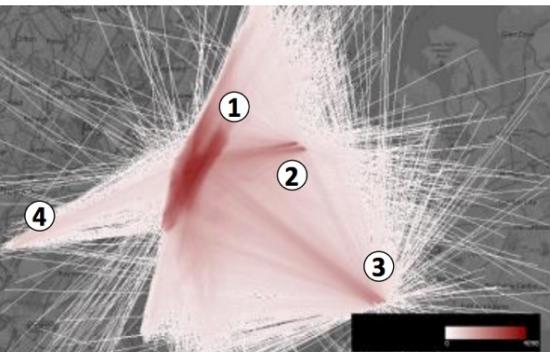
Contact: [xiaoru.yuan@pku.edu.cn](mailto:xiaoru.yuan@pku.edu.cn)  
<http://vis.pku.edu.cn>

# Geographic Information: Movement Trajectories Visualization

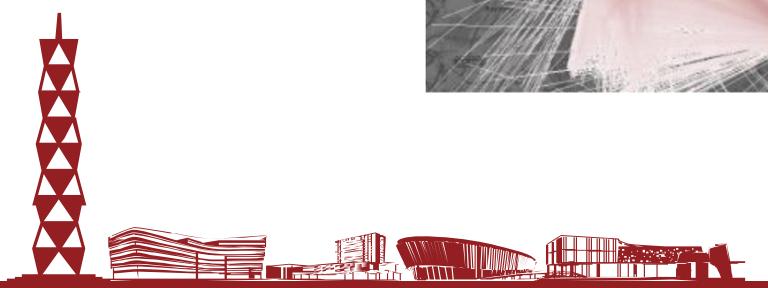
- Semantic Movement Patterns



Weibo Footprint [CWLY16]



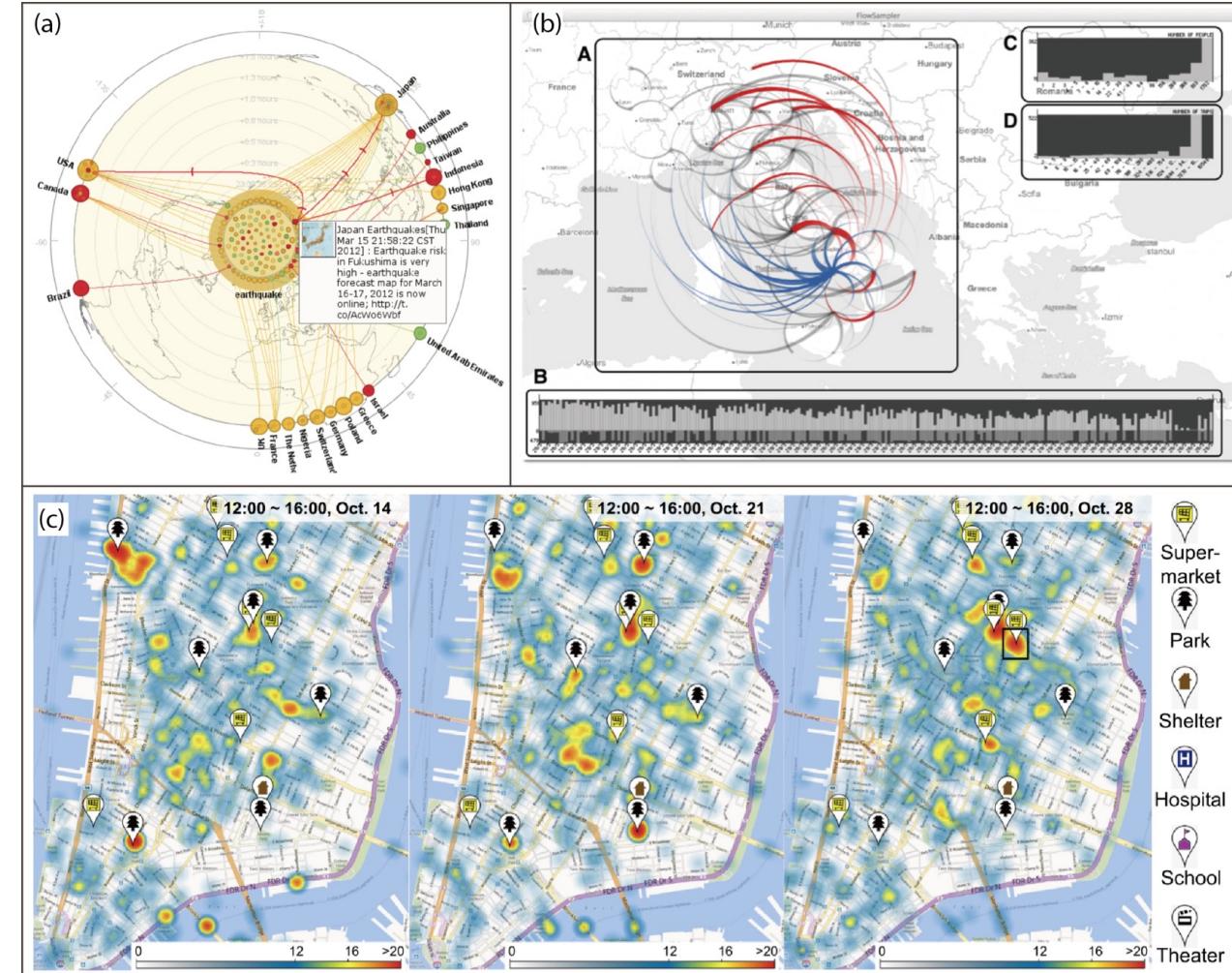
TravelDiff [KSB\*16]



立志成才报国裕民

# Geographic Information Visualization Summary: Visualization on the Map

- Density-based Visualization
- Flow-based Visualization



(a) Whisper [CLS\*12], (b) Flowsampler [CMSVM14], (c) Density Map [CTJ\*14]



# Classification of Social Media Entities and Related Visualization Techniques

Entities in Social Media					
Network			Geographic Information	Text and Content	
People's Follower Network	  	+ Time ↓ Dynamic Network	Geographic Information Diffusion  Spatial Temporal Event Distribution  Movement Trajectory 	+ Time ↓ Spatial Temporal Scenes	Keywords Topic Sentiment  
Messages' Diffusion Network					+ Time ↓ Dynamic Semantic Flow
People's Reposting Network					



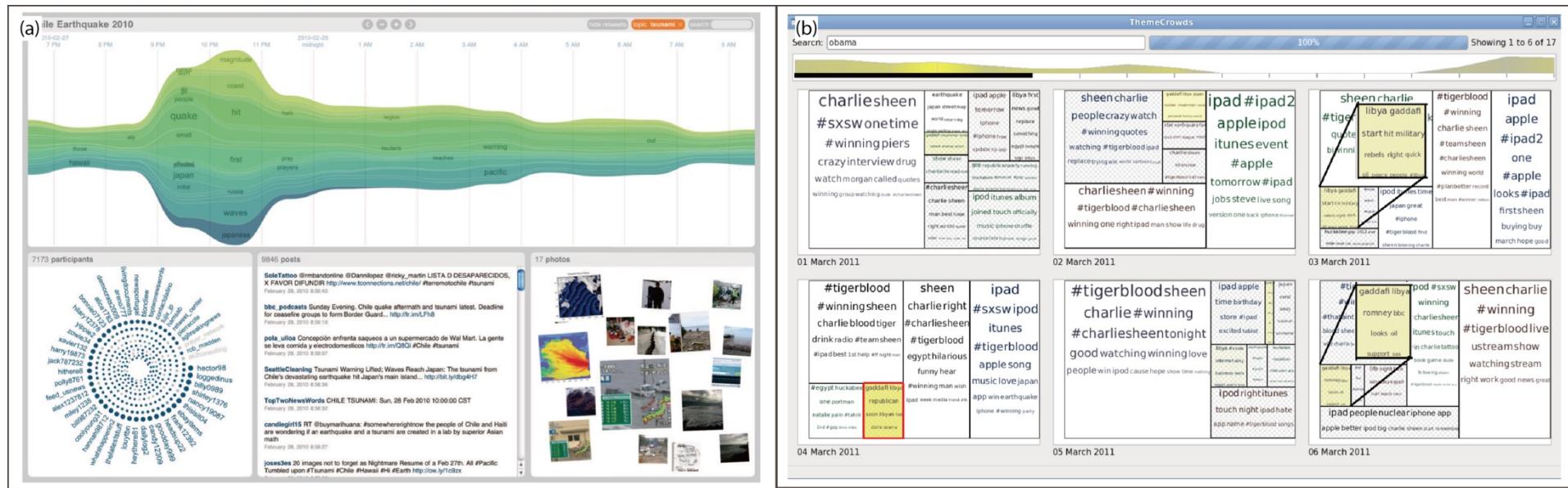


# Paper Collections



# Text and Content: Keywords Visualization

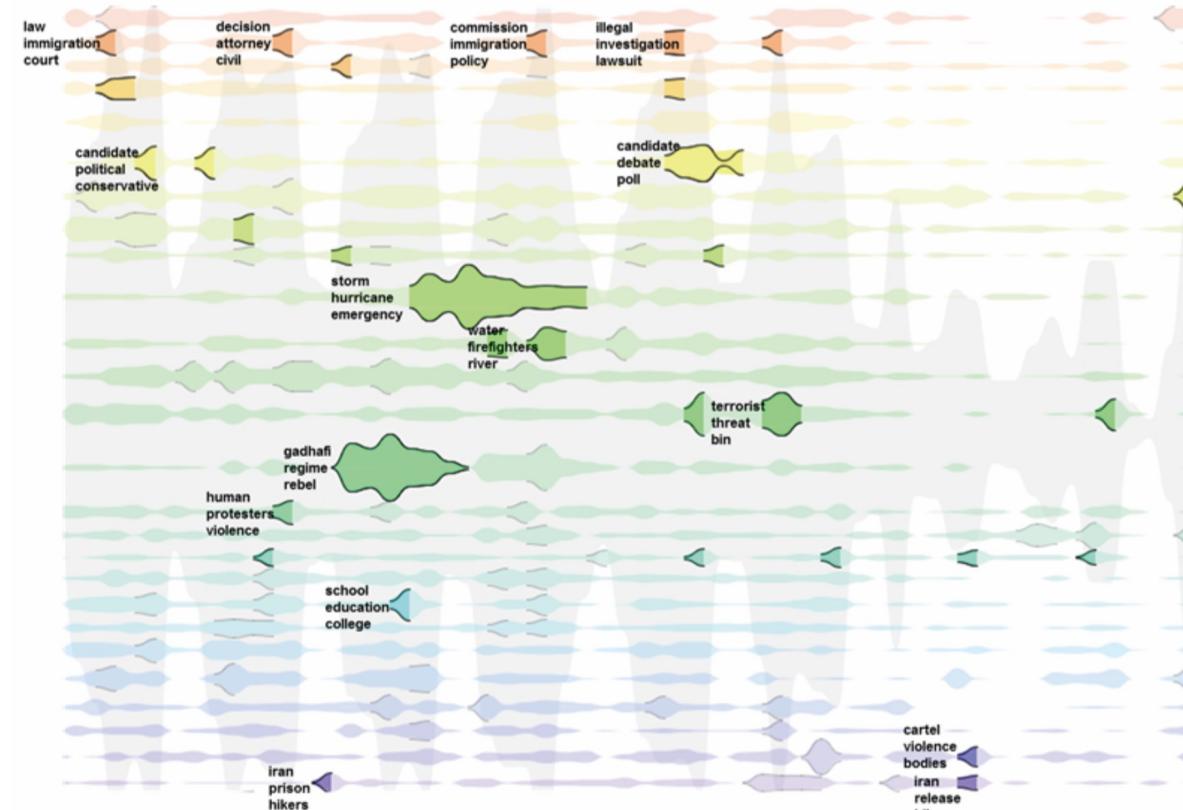
- Extracted keywords from content
  - Dynamic keywords along the time
  - Multi-level aggregation of keywords



(a) Visual BackChannel [DGWC10], (b) ThemeCrowds [AGCH11]

# Text and Content: Topic Visualization

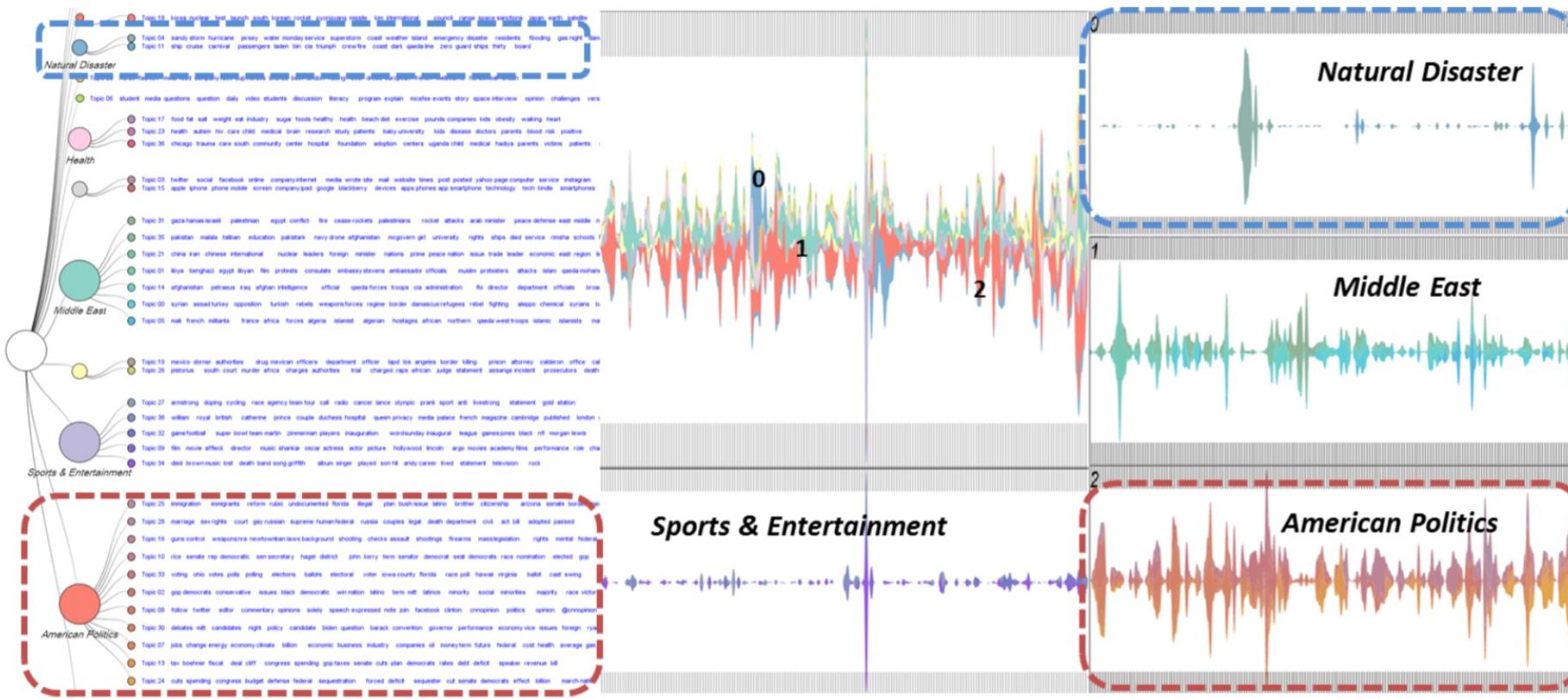
- Topic Structure
  - Dynamic change along the time



LeadLine [DWS\*12]

# Text and Content: Topic Visualization

- Topic Structure
  - Dynamic change along the time
  - Hierarchical Topics



Hierarchical Topics [DYW\*13]

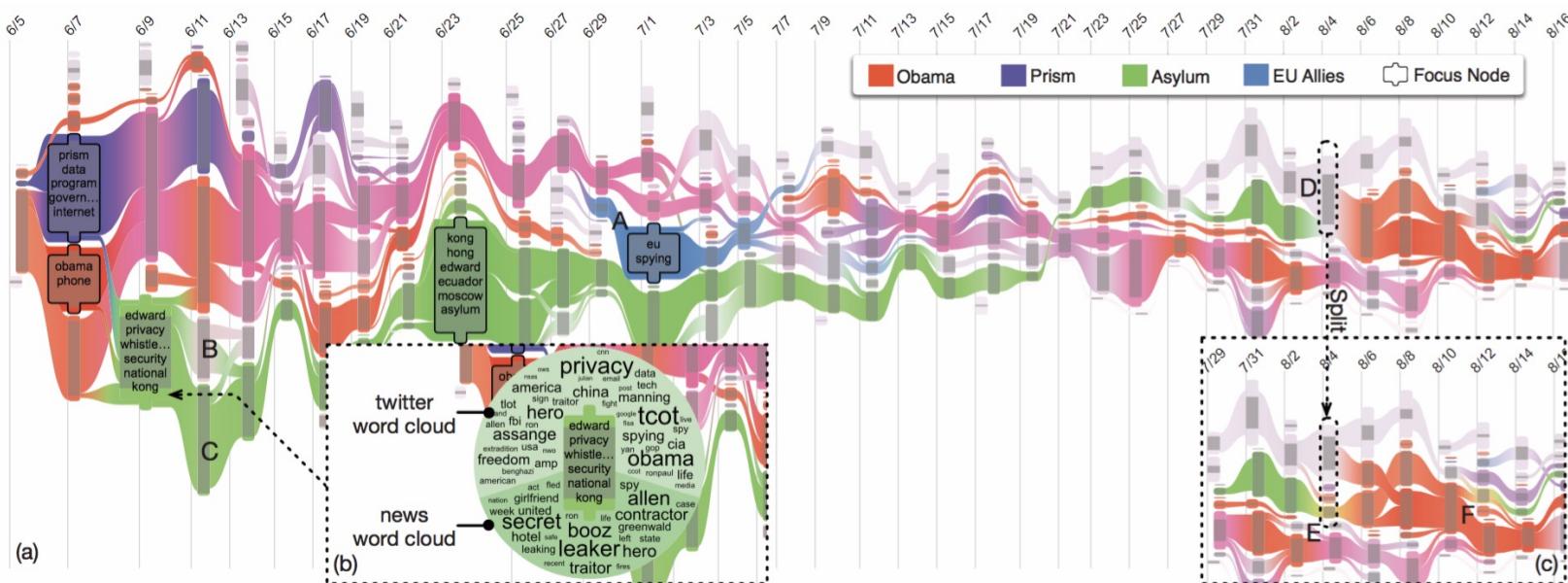
# HierarchicalTopics: Visually Exploring Large Text Collections Using Topic Hierarchies

Wenwen Dou, Li Yu, Xiaoyu Wang, Zhiqiang Ma, and William Ribarsky  
[wdou1@uncc.edu](mailto:wdou1@uncc.edu)  
Charlotte Visualization Center  
UNC Charlotte

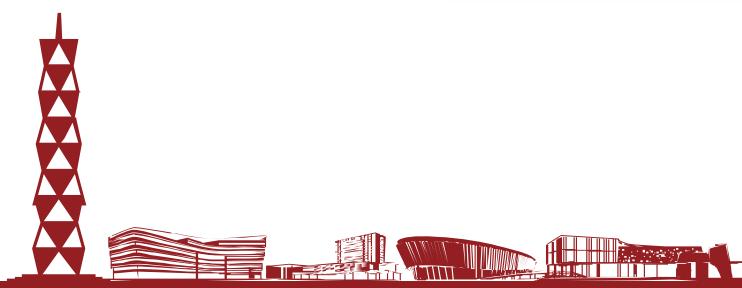
# Text and Content: Topic Visualization

- Topic Structure

- Dynamic change along the time
- Hierarchical Topics
- Dynamic Hierarchical Topics



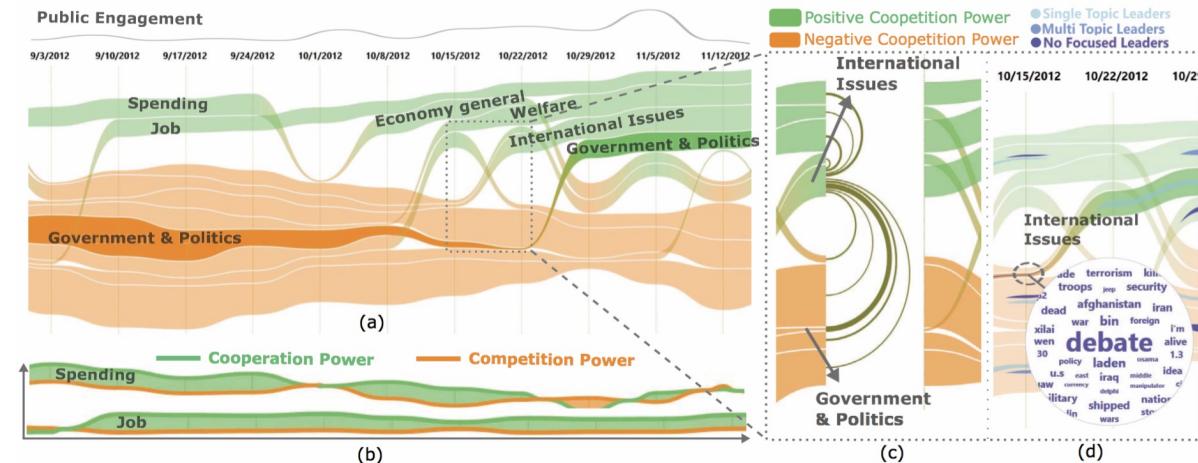
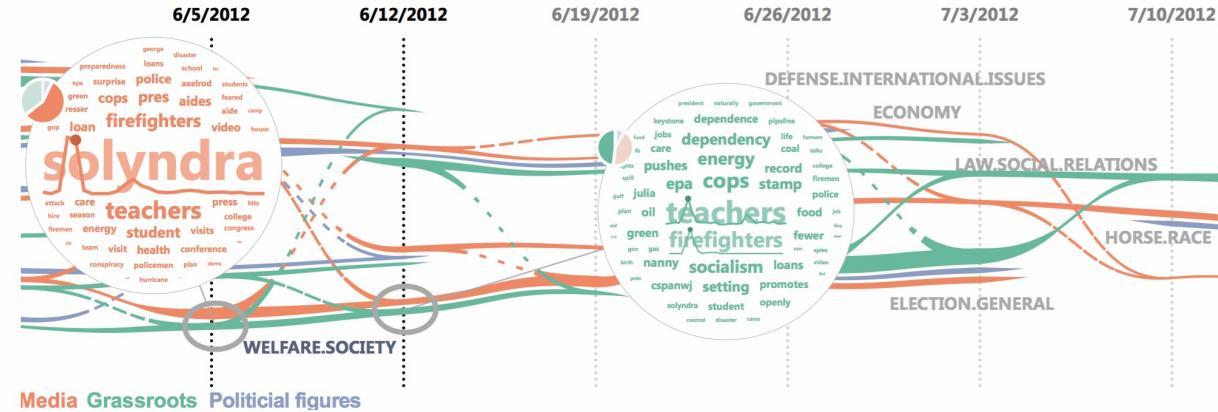
RoseRiver [CLWW14]



# Text and Content: Topic Visualization

- Topic Interactions
  - Competition
  - Collaboration

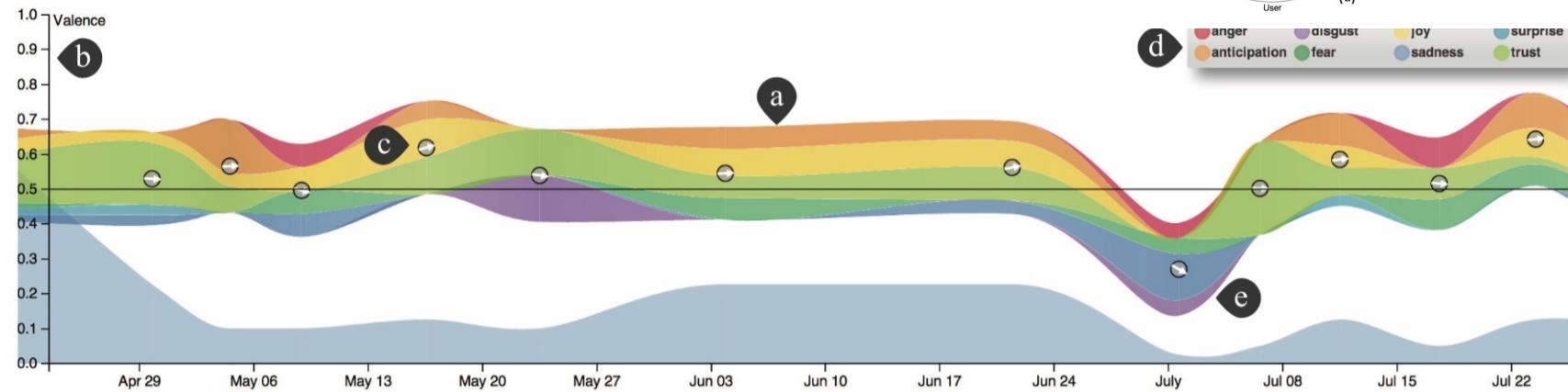
Topic competition [XWW\*13].  
 EvoRiver, visualizing the competition and collaboration relationships of topics [SWL\*14].



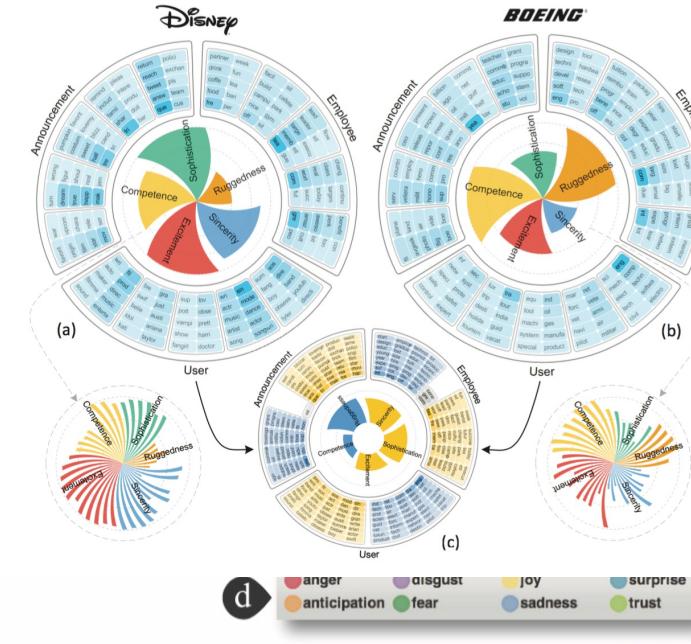
# Text and Content: Sentiment Visualization

- Sentiment
    - Attitude Distribution
    - Dynamic Sentiment Change

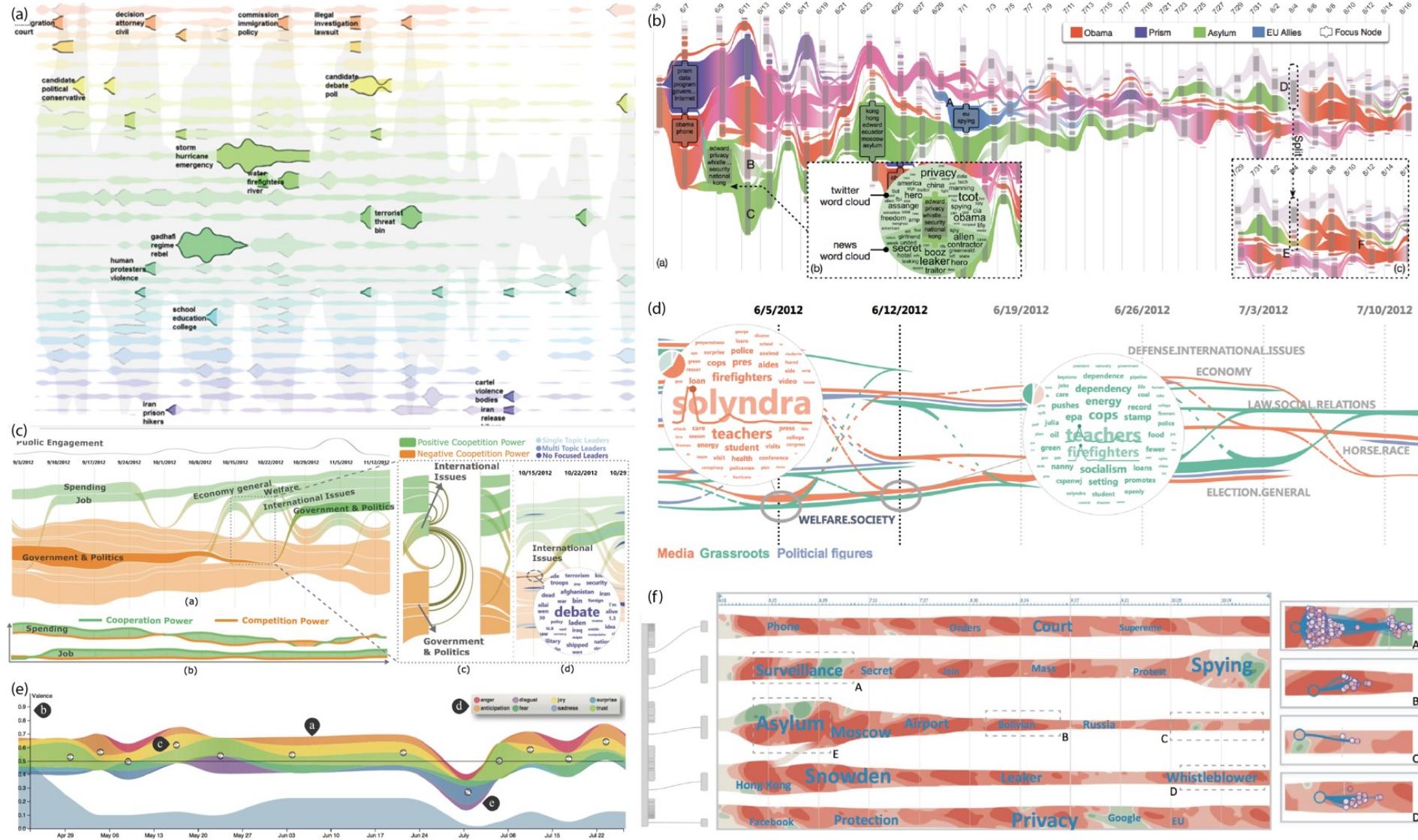
# Sentiment visualization [ZGWZ14]



SocialBrands [LXG\*16]

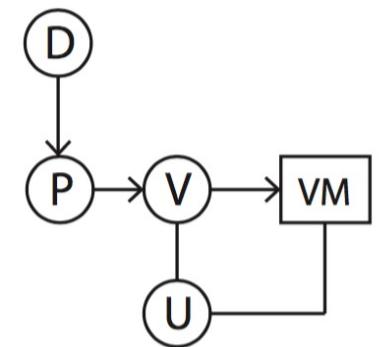


# Text and Content Visualization Summary: River-like Visual Metaphor



# Social Media Visual Analytics (1) – Visual Monitoring

- Monitoring (e.g., Information diffusion process, reposting behaviors, moving behaviors etc.)
    - Overview design
    - Streaming Data
    - Animation



## D Data      V Visualization

P Processing

V Visualizatio

Users

VM Visual Monitoring

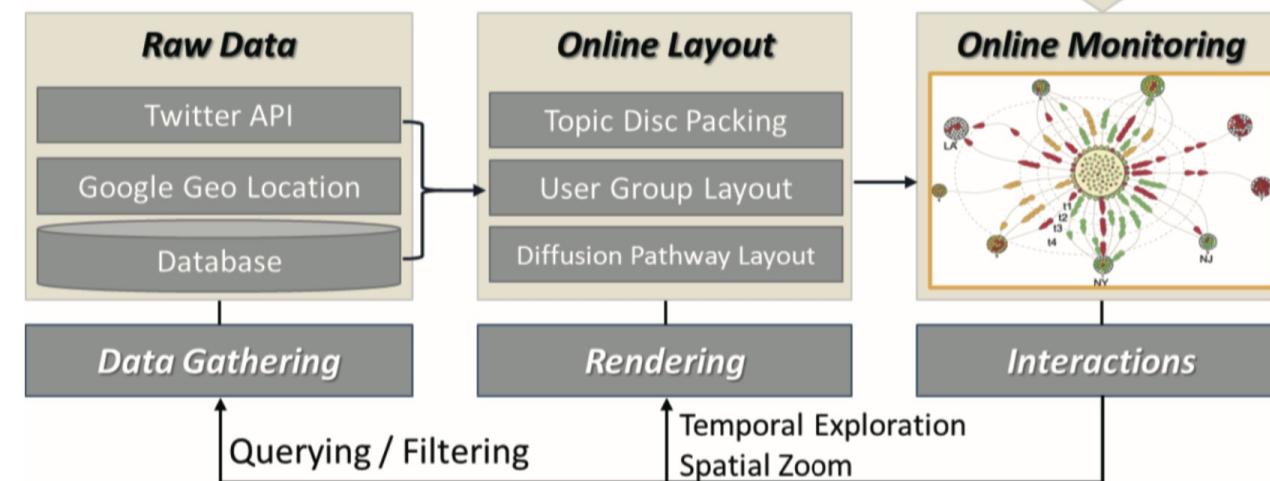
AD Anomaly Detection

## FE Feature Extraction

**PA** Predictive Analysis

ED Event Detection

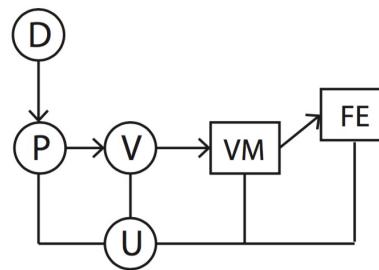
SA Situation Awareness



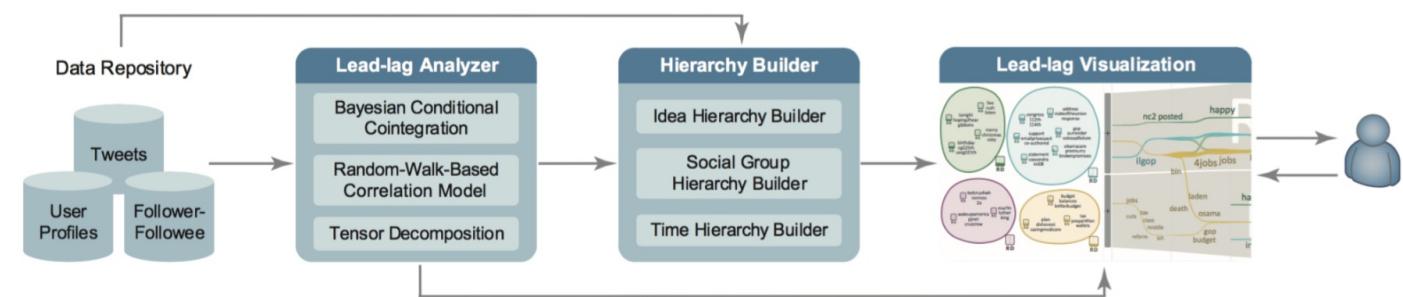
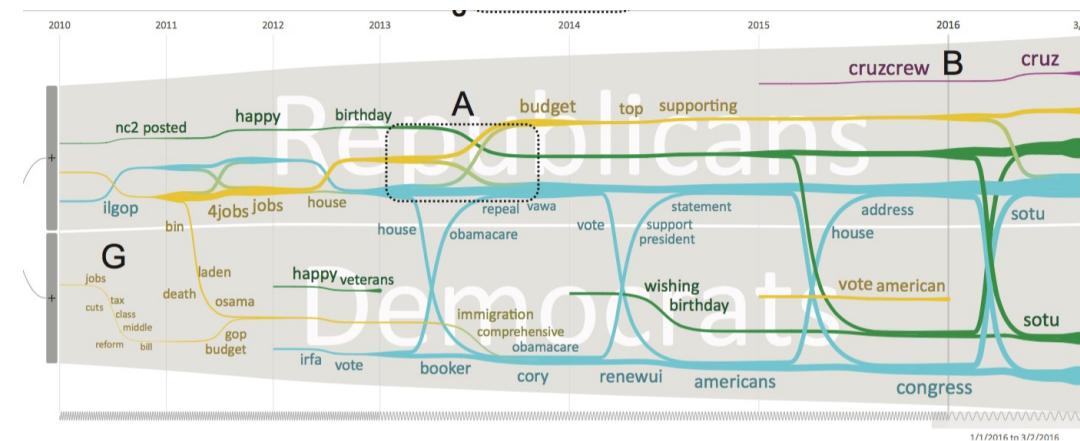
## Whisper [CLS\*12]

# Social Media Visual Analytics (2) – Feature Extraction

- Feature Extraction (e.g. summarized diffusion patterns, movement trajectories with semantics, etc.)
  - Significant characteristics of specific data attributes
  - Combining data mining algorithms
  - Preprocessing



	Data		Visualization
	Processing		Users
	Visual Monitoring		Anomaly Detection
	Feature Extraction		Predictive Analysis
	Event Detection		Situation Awareness

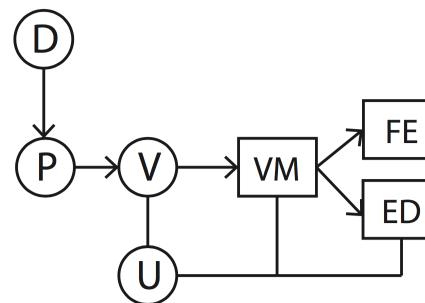


IdeaFlow [WLC\*16]

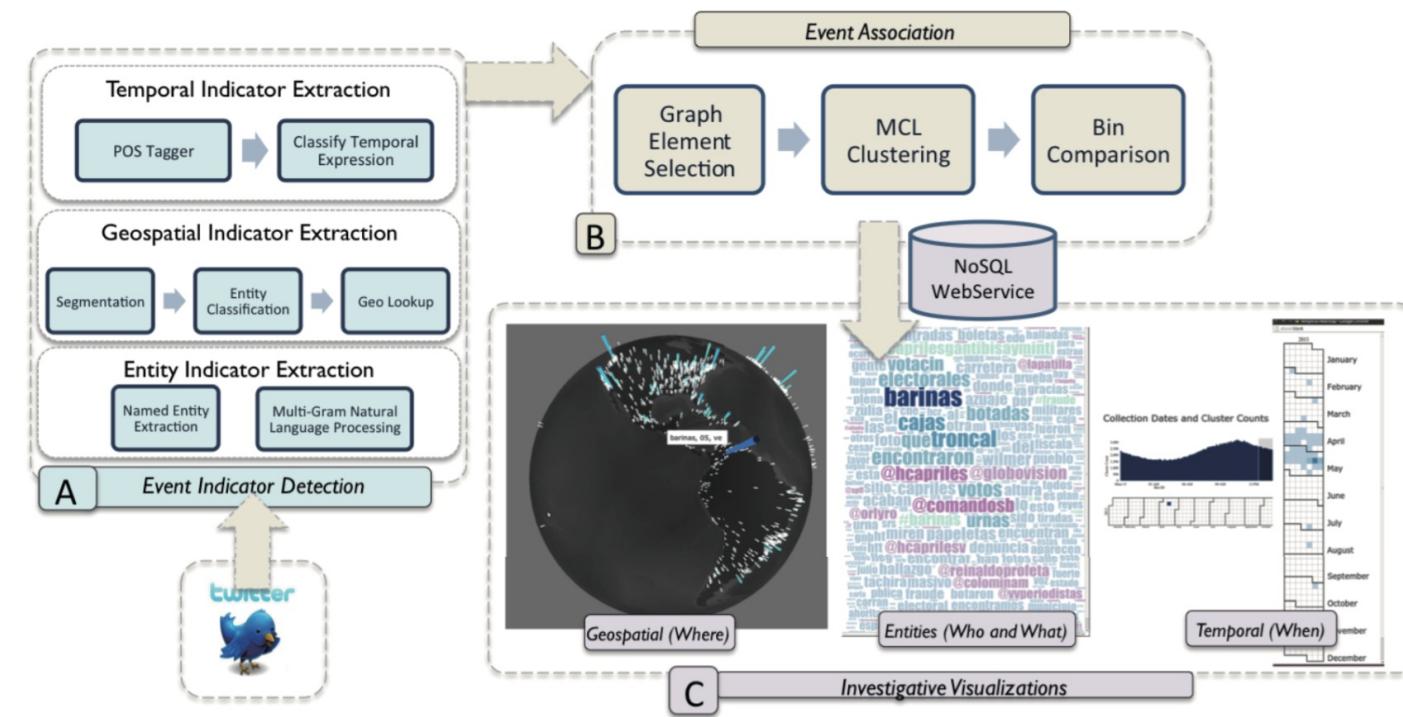
立志成才报国裕民

# Social Media Visual Analytics (3) – Event Detection

- Event: <People, Messages, Time, [Reposting, Location, Themes]>
  - Event extraction
  - Event visualization
  - Event analysis



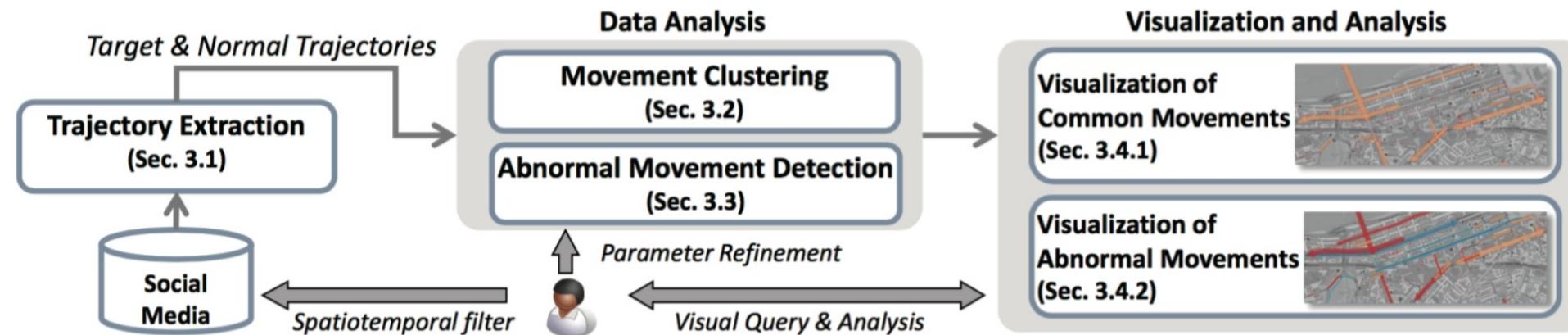
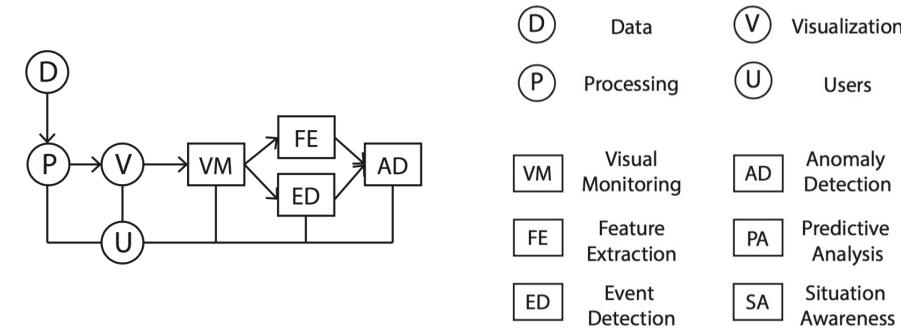
(D)	Data	(V)	Visualization
(P)	Processing	(U)	Users
[VM]	Visual Monitoring	[AD]	Anomaly Detection
[FE]	Feature Extraction	[PA]	Predictive Analysis
[ED]	Event Detection	[SA]	Situation Awareness



GCAT [KWD\*13]

# Social Media Visual Analytics (4) – Anomaly Detection

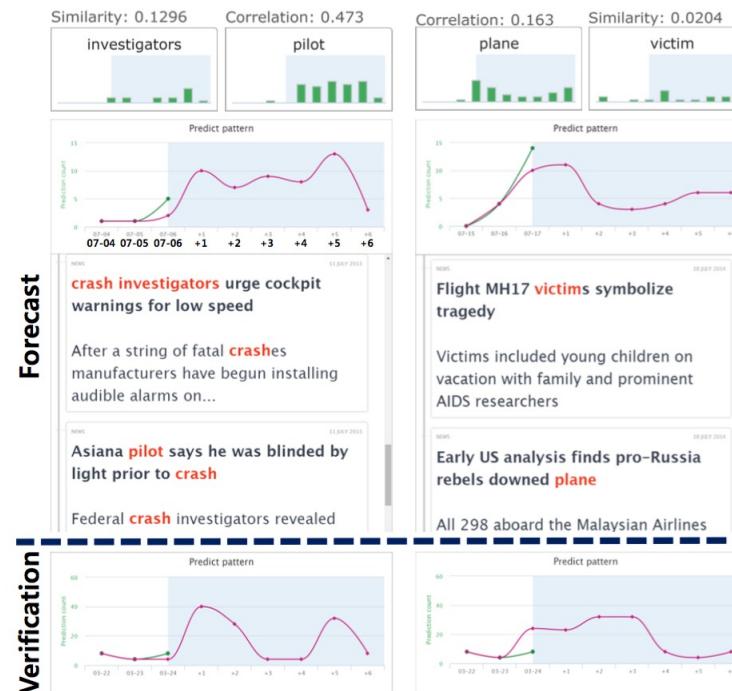
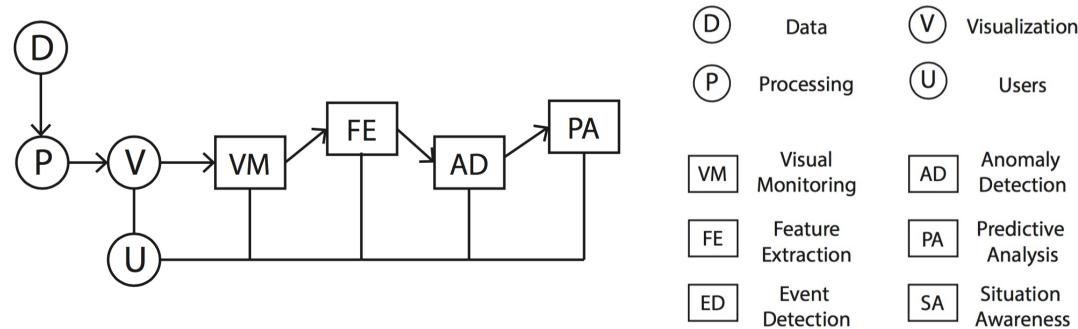
- Normal patterns: data is around average / common ranges of its attributes
- Abnormal patterns
  - Classification model
  - Visual query and analysis



Public Response Analysis [CCJ\*15]

# Social Media Visual Analytics (5) – Predictive Analysis

- Predictive Analysis (Topic trend, keywords rating, etc.)
  - Event detection
  - Abnormal event detection
  - Similarity analysis
  - Prediction and verification

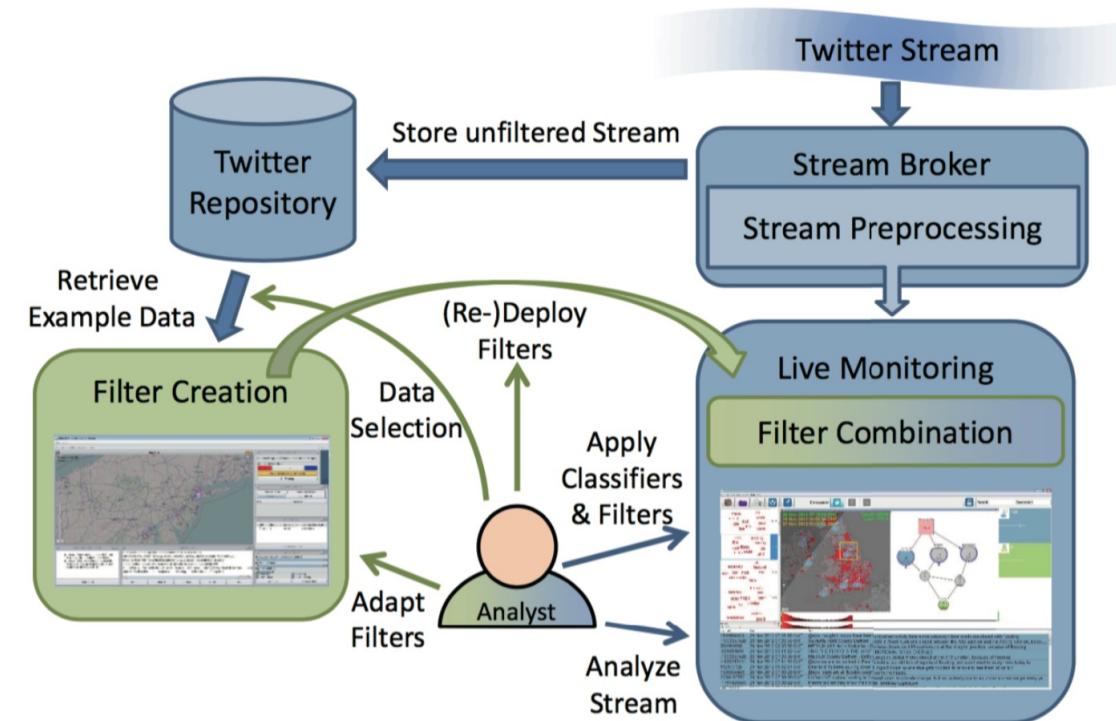
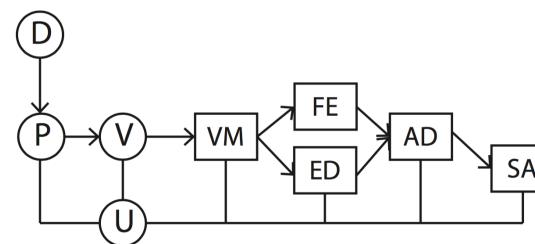


Topic Decomposition [YJ15]

立志成才报国裕民 95

# Social Media Visual Analytics (6) – Situation Awareness

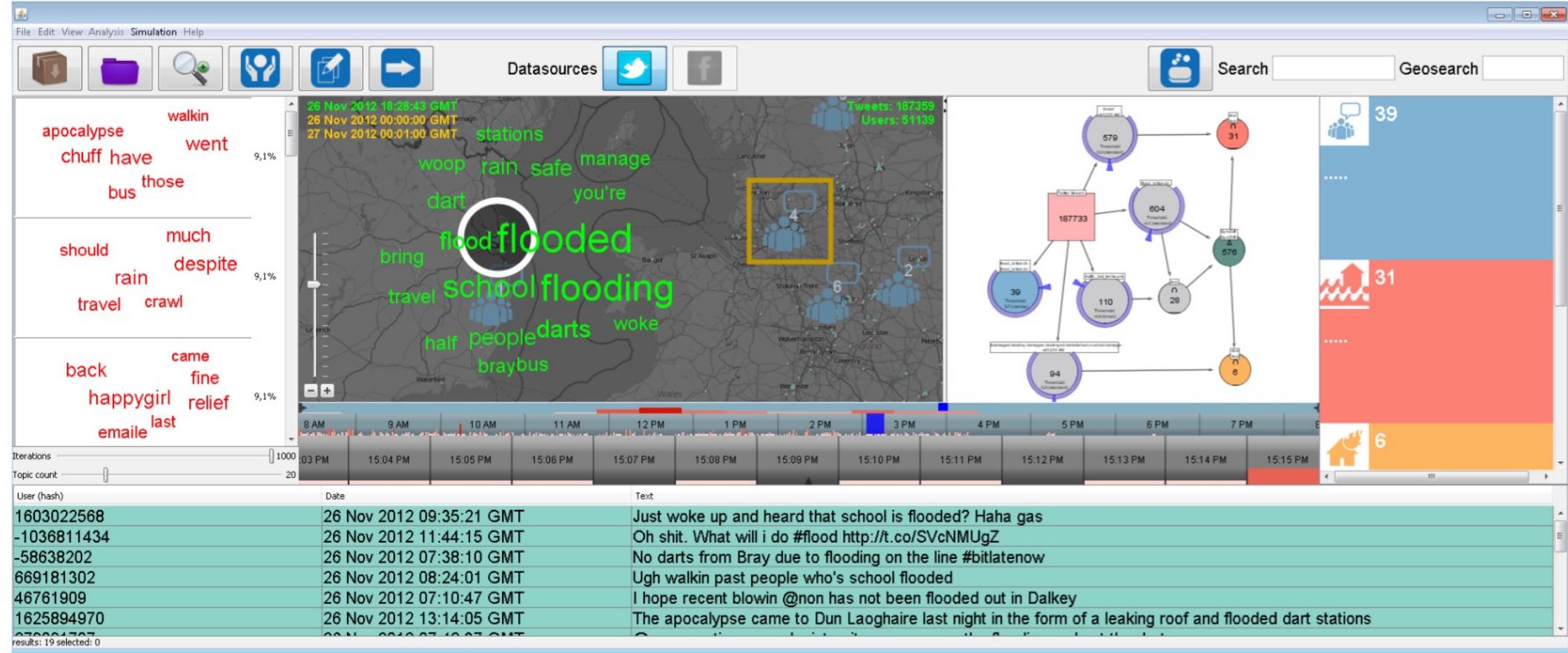
- Situation Awareness (Combination of spatial temporal events, topic and social networks) – for Decision Making
  - Visual monitoring
  - Feature extraction / Event detection
  - Anomaly detection



ScatterBlogs2 [BTH\*13]

立志成才报国裕民

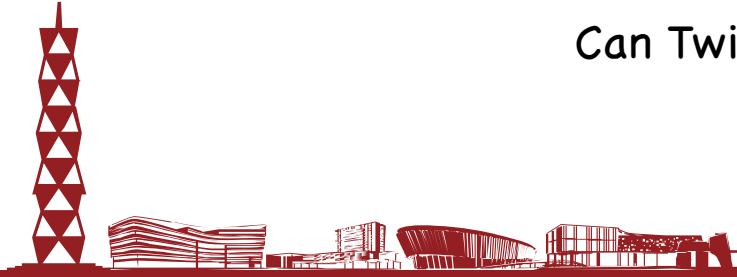
# Social Media Visual Analytics (6) – Situation Awareness



Can Twitter Really Save Your Life? A Case Study of Visual Social Media Analytics for Situation Awareness

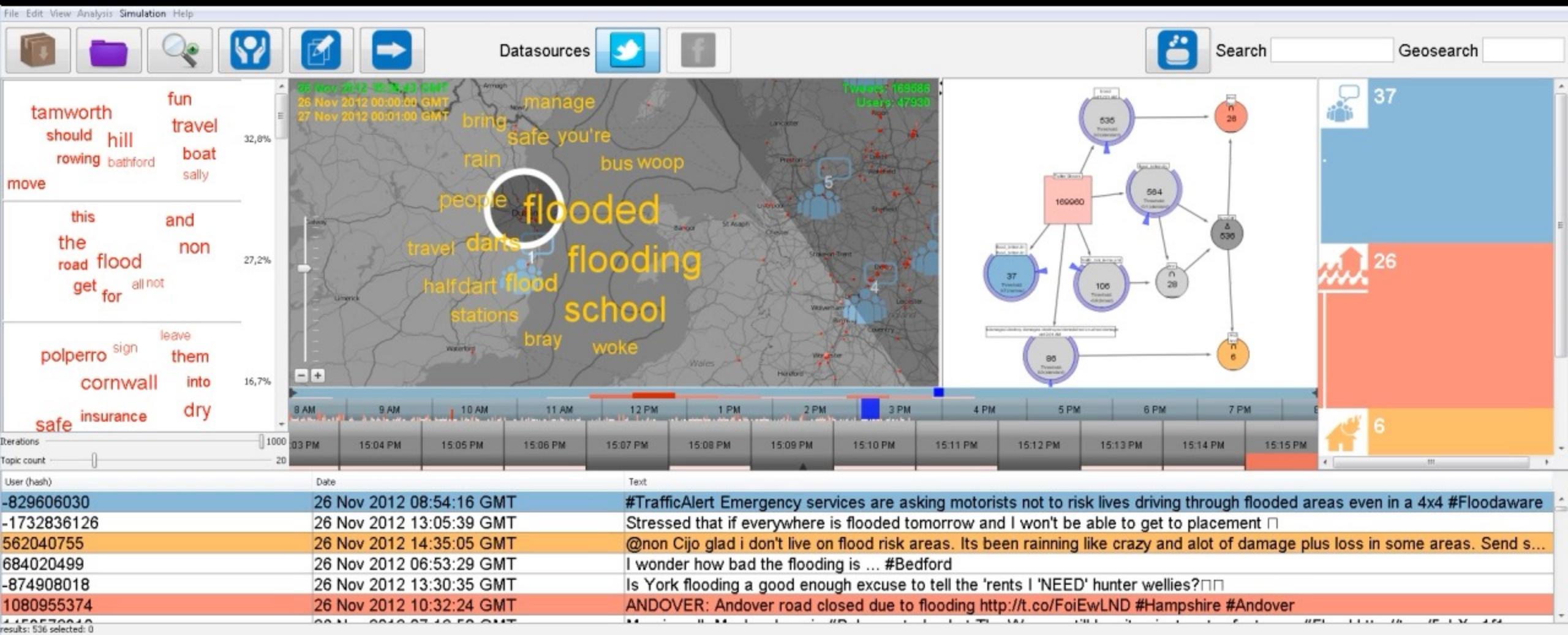
ScatterBlogs2 [BTH\*13, TKE\* 15]

立志成才报国裕民

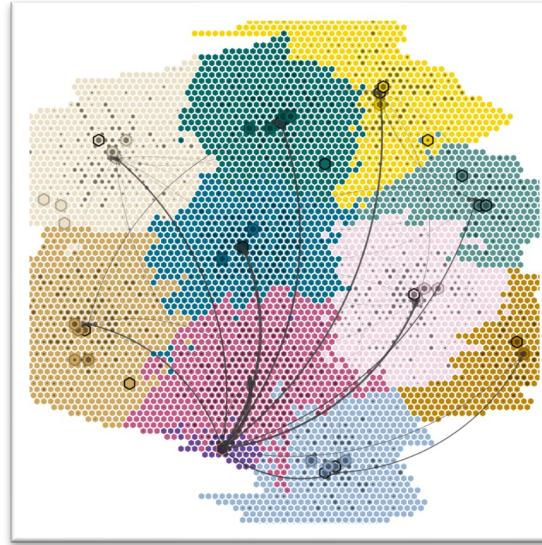


# ScatterBlogs2: Real-Time Monitoring of Microblog Messages through User-Guided Filtering

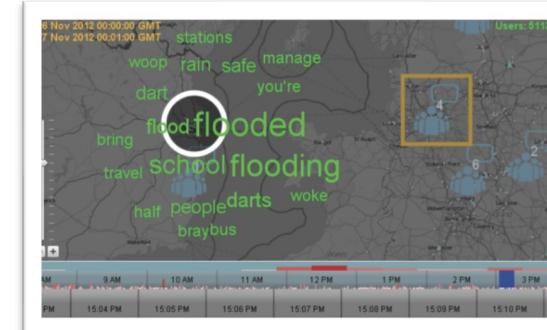
Harald Bosch, Dennis Thom, Florian Heimerl, Edwin Püttmann, Steffen Koch, Robert Krüger, Michael Wörner, Thomas Ertl  
Institute for Visualization and Interactive Systems, University of Stuttgart  
[www.vis.uni-stuttgart.de](http://www.vis.uni-stuttgart.de)



# Applications and Systems

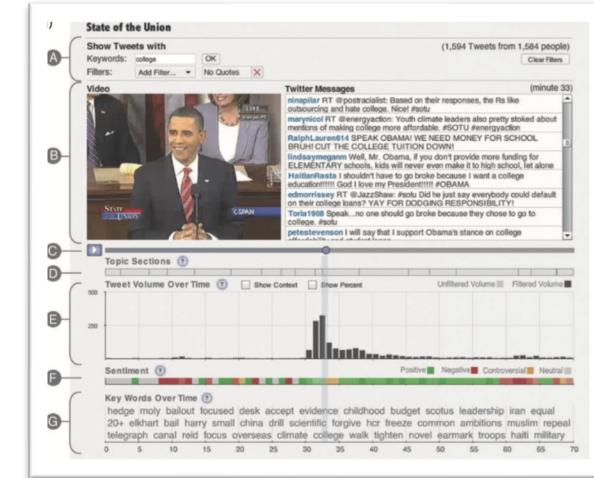


Social Science Research [CCW\*16]

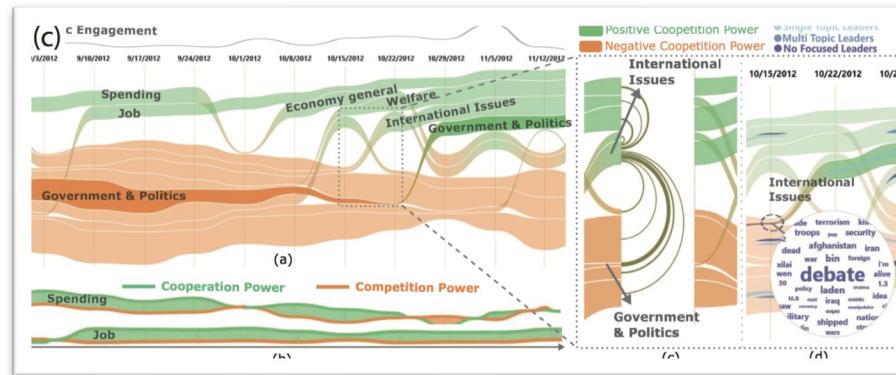


Crisis Management [DTH\*13]

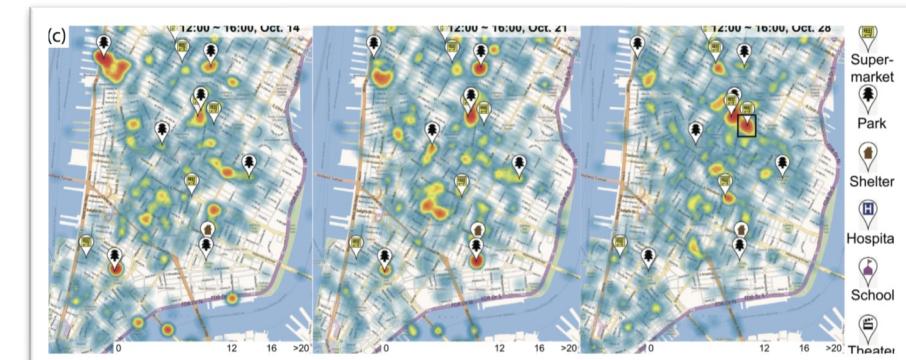
Applications  
and  
Systems



Journalism [DNYKS11]

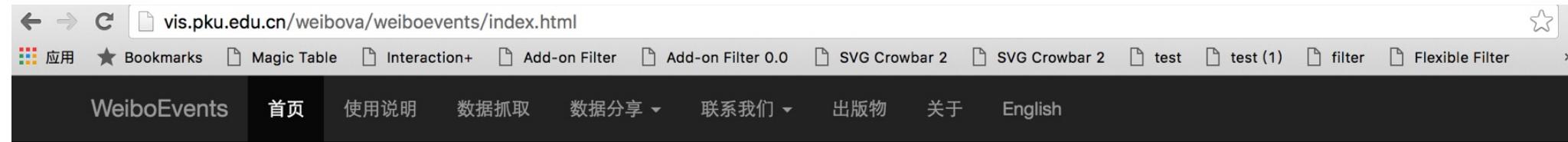


Politics [SWL\*14]



Disaster and Emergency Management [CTH\*14] 立志成才报国裕民

# Journalism Application: WeiboEvents



The screenshot shows the WeiboEvents application running in a web browser. The URL in the address bar is vis.pku.edu.cn/weibova/weiboevents/index.html. The top navigation bar includes links for 'WeiboEvents' (selected), '首页', '使用说明', '数据抓取', '数据分享', '联系我们', '出版物', '关于', and 'English'. Below the navigation is a large title '北京大学 PKUVIS 微博可视分析工具'. The main area displays several network visualizations and a timeline graph illustrating reposting patterns over time.

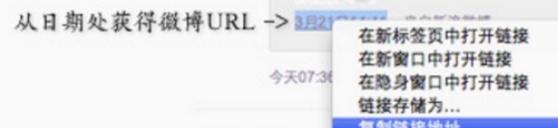
## 微博传播分析

输入微博URL:

例如: <http://weibo.com/1610356014/A3j83wqYQ>

开始分析

使用说明

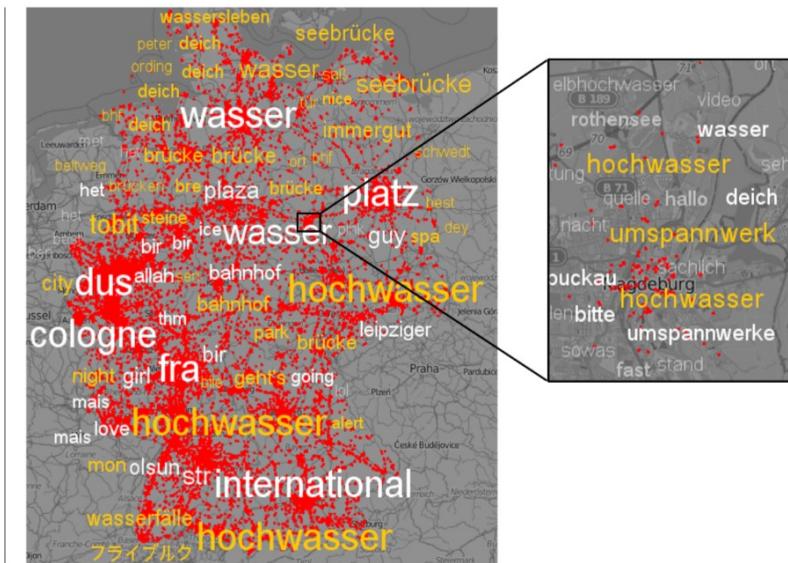
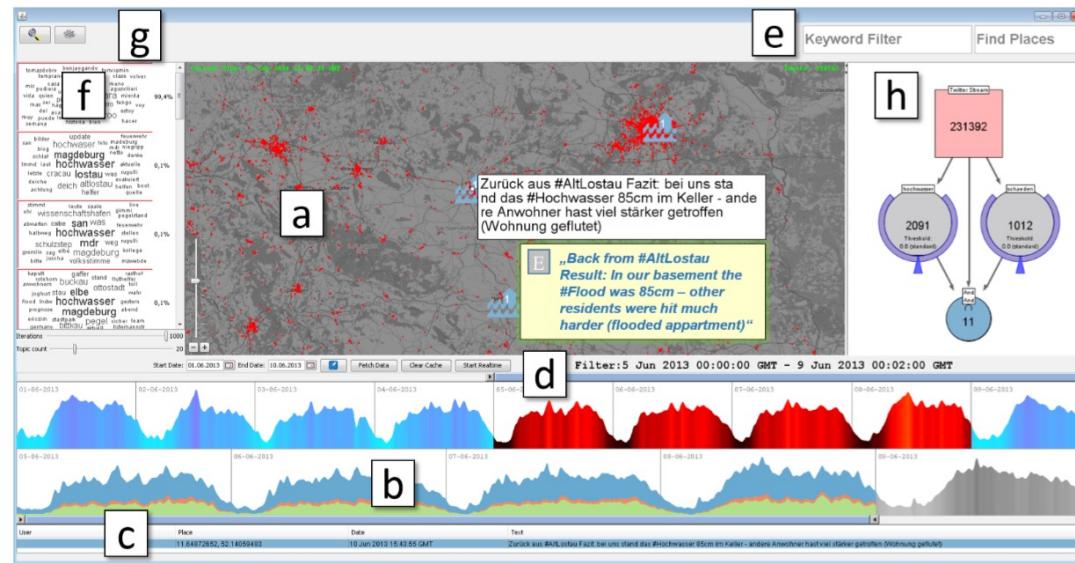


WeiboEvents [RZW\*14]

立志成才报国裕民 100

# Emergency/Disaster/Crisis Management

- Gain a general overview
- Investigate prominent topics
- Analyze message in detailed regions
- Examine where the emergency is the most severe



ScatterBlogs2 [BTH\*13, TKE\* 15]

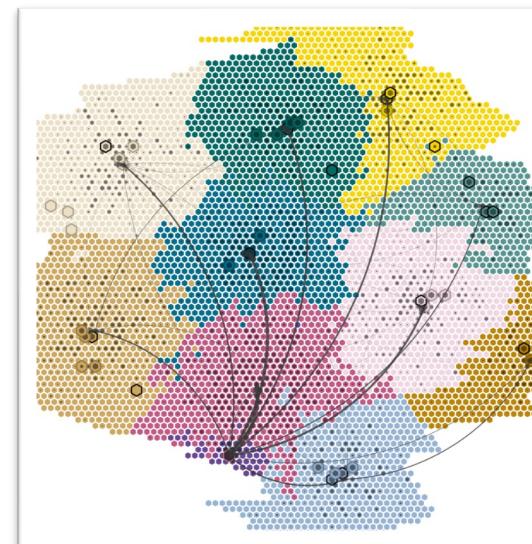
# Discussion – Challenges and Opportunities

- Social Media Data Content
  - Content analysis – Extracted keyword/topic analysis
  - Multi-media visual analytics
  - How to deal with bias / trustiness of sampling
  - How to deal with privacy issues

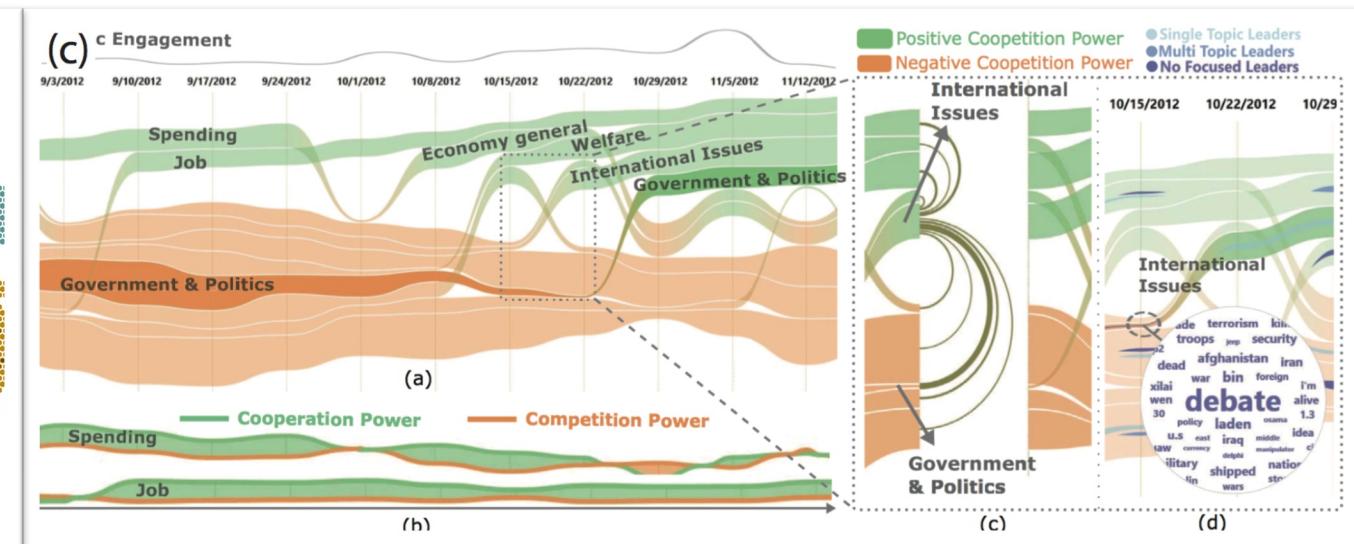


# Discussion – Challenges and Opportunities

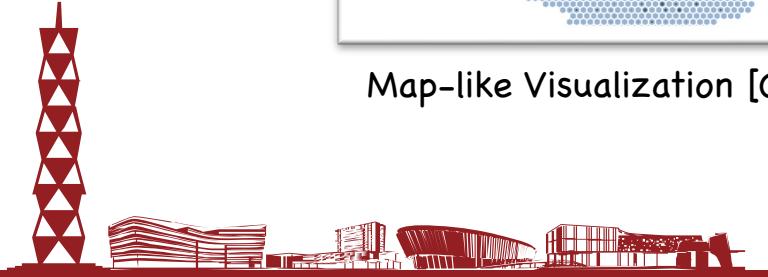
- Social Network Evolution Analysis
  - Challenge: both relationship and temporal analysis
  - Visualization: river-like and map-like visual metaphor
  - Visual analytics approach: pattern analysis
  - How to detect (subtle) anomaly? It's still in the exploration stage



Map-like Visualization [CCW\*16]



River-like Visualization [SWL\*14]



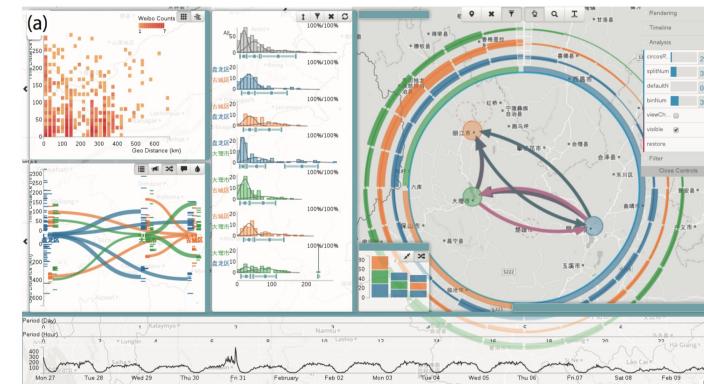
# Discussion – Challenges and Opportunities

- Movement and Trajectory Analysis

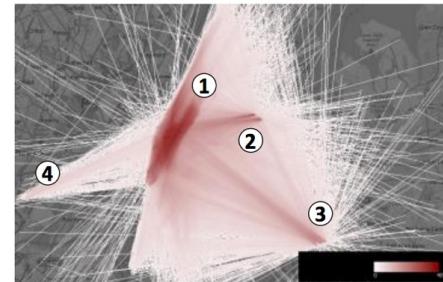
- Challenge: Understand trajectory patterns with semantics
- Detecting movement patterns considering periodical behaviors / uncertainty
- Comparison with other data sources



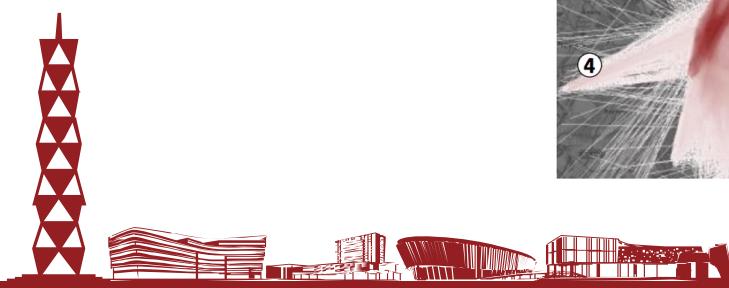
Visits [TBC13]



WeiboGeo [CYW\*16]



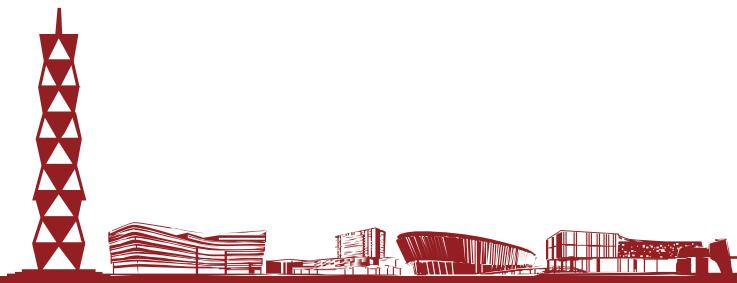
TravelDiff [KSB\*16]





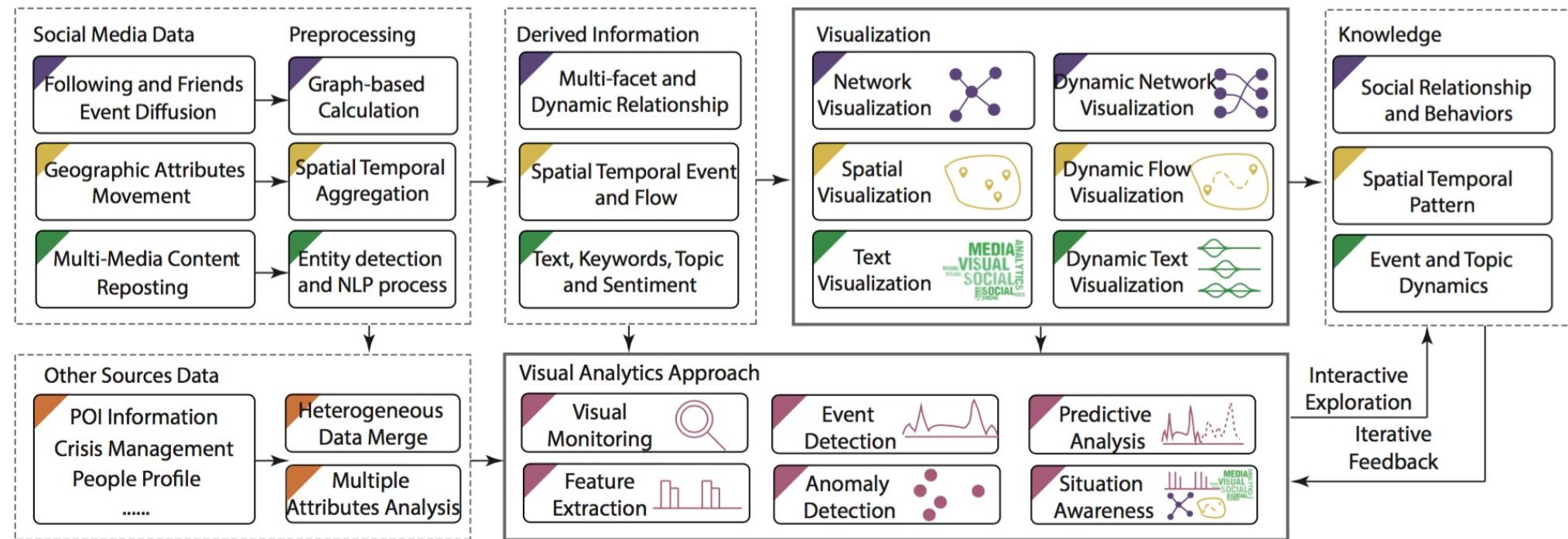
## Discussion – Challenges and Opportunities

- Heterogeneous visual analytics
  - Attribute level – Profiling social media users
  - Data level – Enriching behavior understanding with additional data source
  - Analysis level: Social Media + other discipline's analytical pipeline
- Evaluation of Social Media Visual Analytics
  - Target users
  - Tasks
  - User study
  - Expert feedback



# Summary

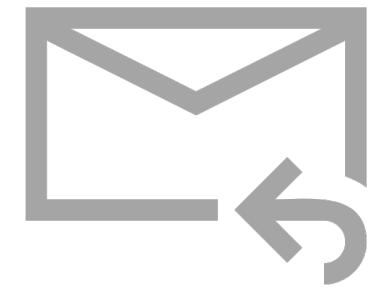
- Social Media Data Category
- Social Media Visualization
- Social Media Visual Analytics
- Application and System
- Discussion and Summary





Quan Li

Questions?  
Thank you 😊



[liquan@shanghaitech.edu.cn](mailto:liquan@shanghaitech.edu.cn)