Extracting Knowledge from Social Sites

Yiannis Kompatsiaris

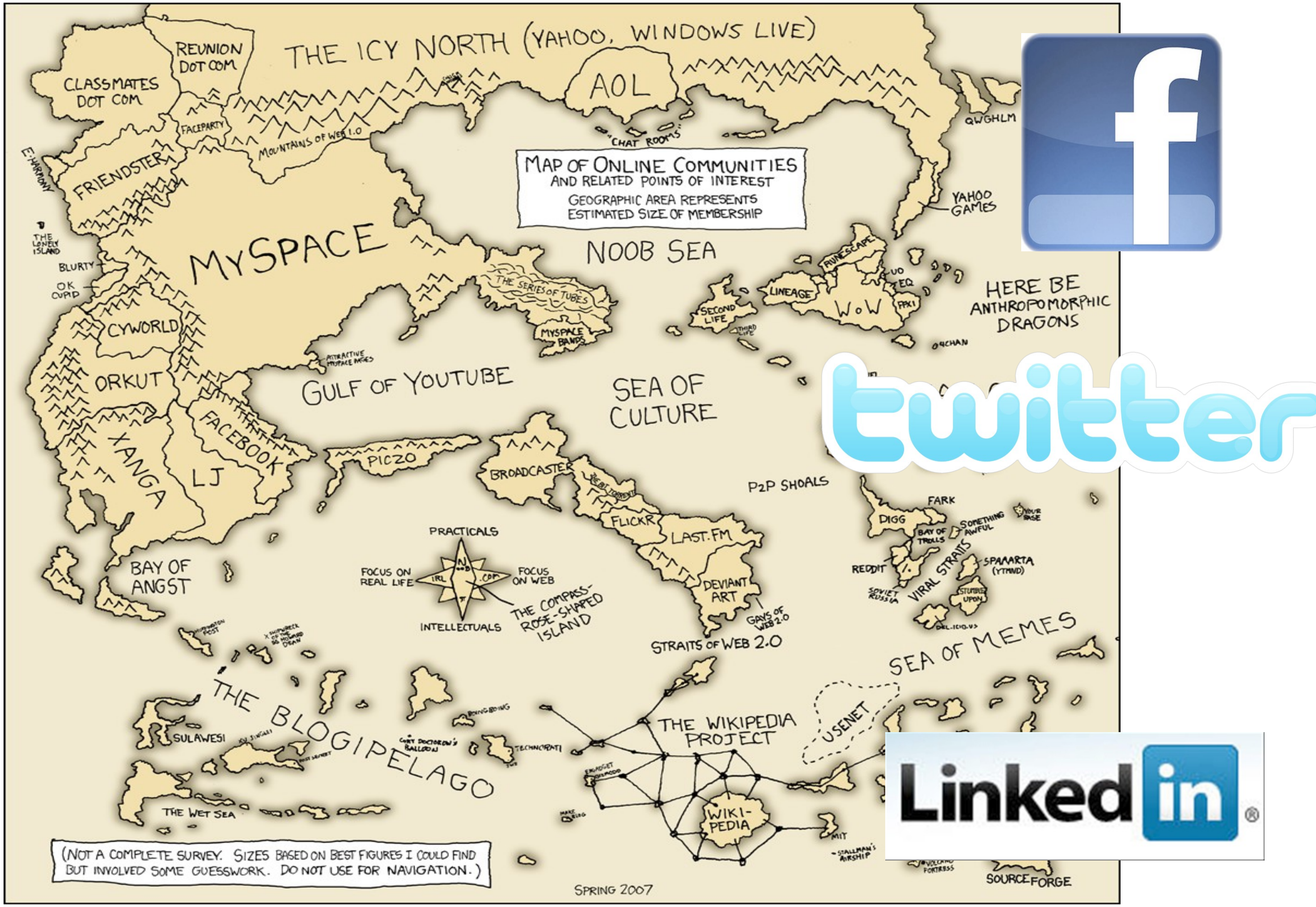
Eirini Yannakidou, Symeon Papadopoulos,
Spyros Nikolopoulos, Elisavet Chatzilari

Athena Vakali, AUTh

Contents

- Introduction
- Clustering in Social Media
- Social media “teacher” of the machine
- Community detection in Social Media
- WeKnowIt project
- Conclusions - Issues

Our world today (already old)

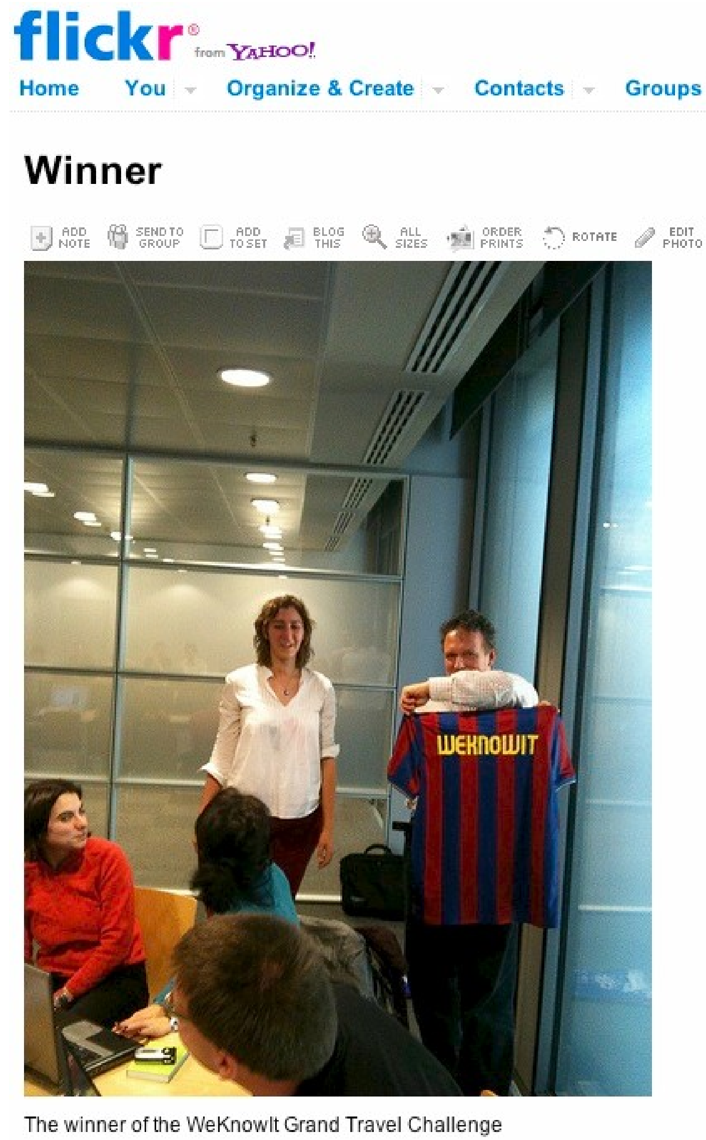


Web 2.0 content

- 20h of video content uploaded every minute at YouTube (2009)
- 3,024,780,142 photos in Flickr @ 11:52, 12 Nov 2008
- 2 million geotagged photos uploaded each month (2008)

Facebook:

- More than 250 million active users
- More than 120 million users log on to Facebook at least once each day
- More than 1 billion photos uploaded to the site each month



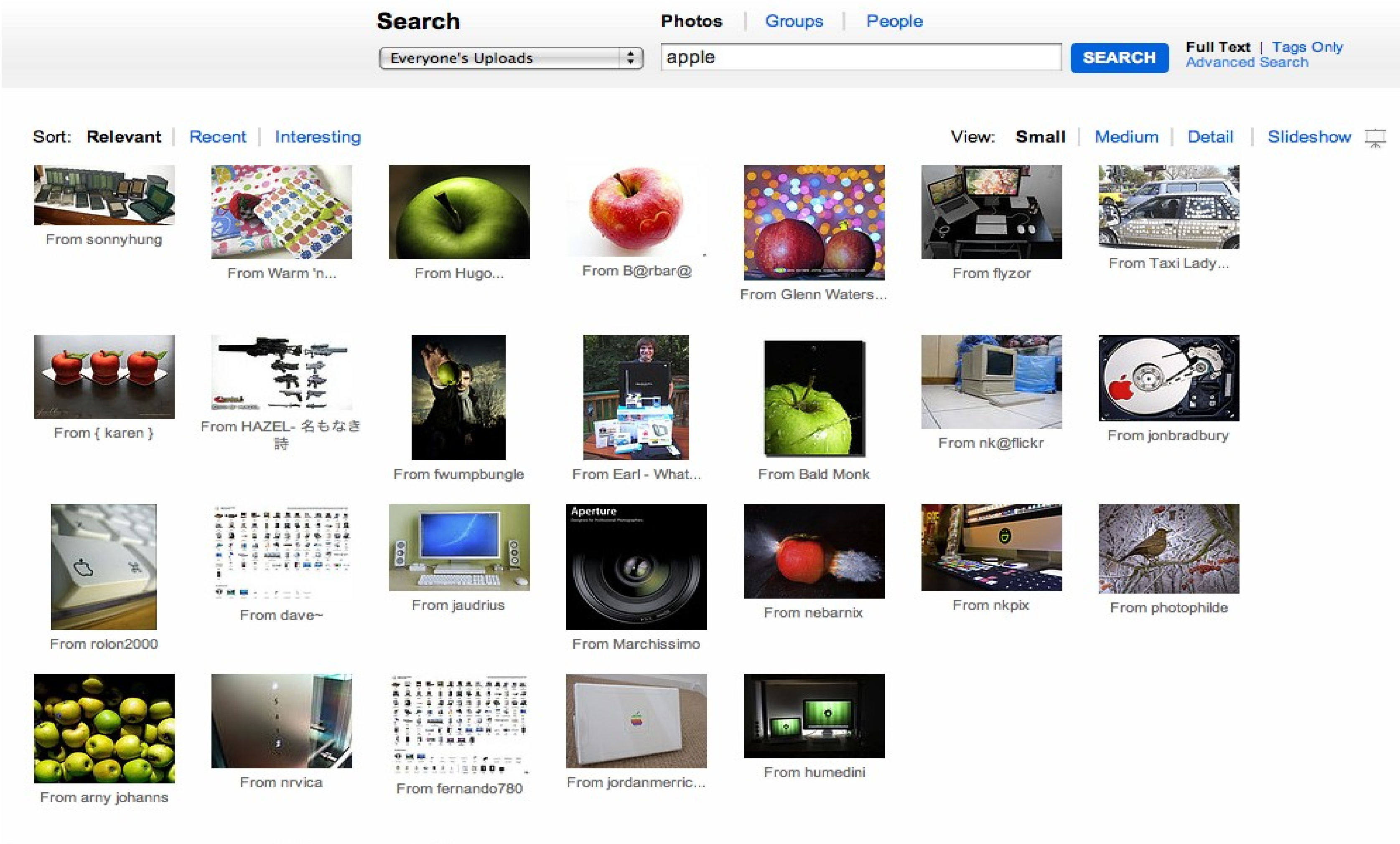
Tags everywhere

tag cloud
Call for papers CIVR2009 Collective Intelligence Conference content popularity images Invited Talk IVUS Multimedia Retrieval Multimedia Semantics News object detection Ontologies Patents proceedings Project Semantic Multimedia Semantics social bookmarking tutorial video retrieval WeKnowIt Workshop WWW2009
more tags

Search, Describe content, Extract knowledge

amsterdam animal animals april architecture art australia baby barcelona beach berlin bird birthday black blackandwhite blue boston bridge building bw california cameraphone camping canada car cat cats chicago china christmas church city clouds color concert day dc dog dogs england europe family festival florida flower flowers food france friends fun garden geotagged germany girl graduation graffiti green hawaii holiday home hongkong house india ireland italy japan june kids lake landscape light london losangeles macro march may me mexico moblog mountains museum music nature new newyork newyorkcity newzealand night nyc ocean old orange oregon paris park party people phone photo pink portrait red reflection river roadtrip rock rome sanfrancisco school scotland sea seattle sign sky snow spain spring street summer sun sunset taiwan texas thailand tokyo toronto travel tree trees trip uk unfound urban usa vacation vancouver washington water wedding white window winter work yellow zoo

Very low precision

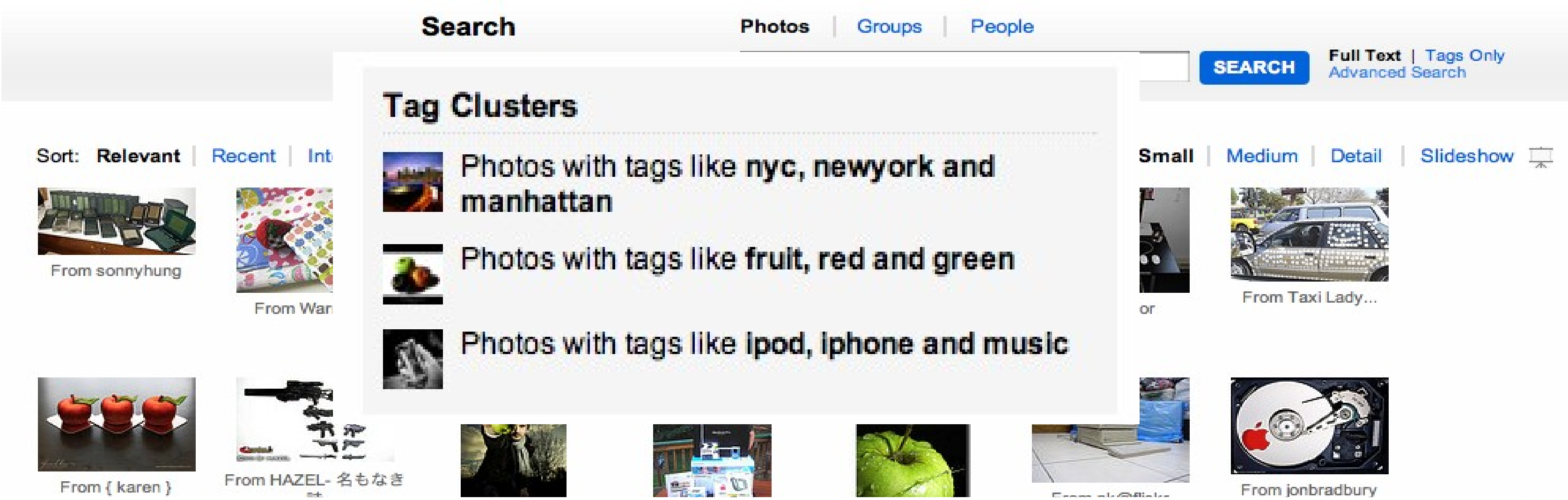


Very low recall

Tags

- Property#1
- Canada
- photo
- image
- digital
- urban
- Halifax
- park
- morning
- afternoon
- night
- Pentax K20D
- Sigma 70-300
- early
- Sackville

Can we improve things?



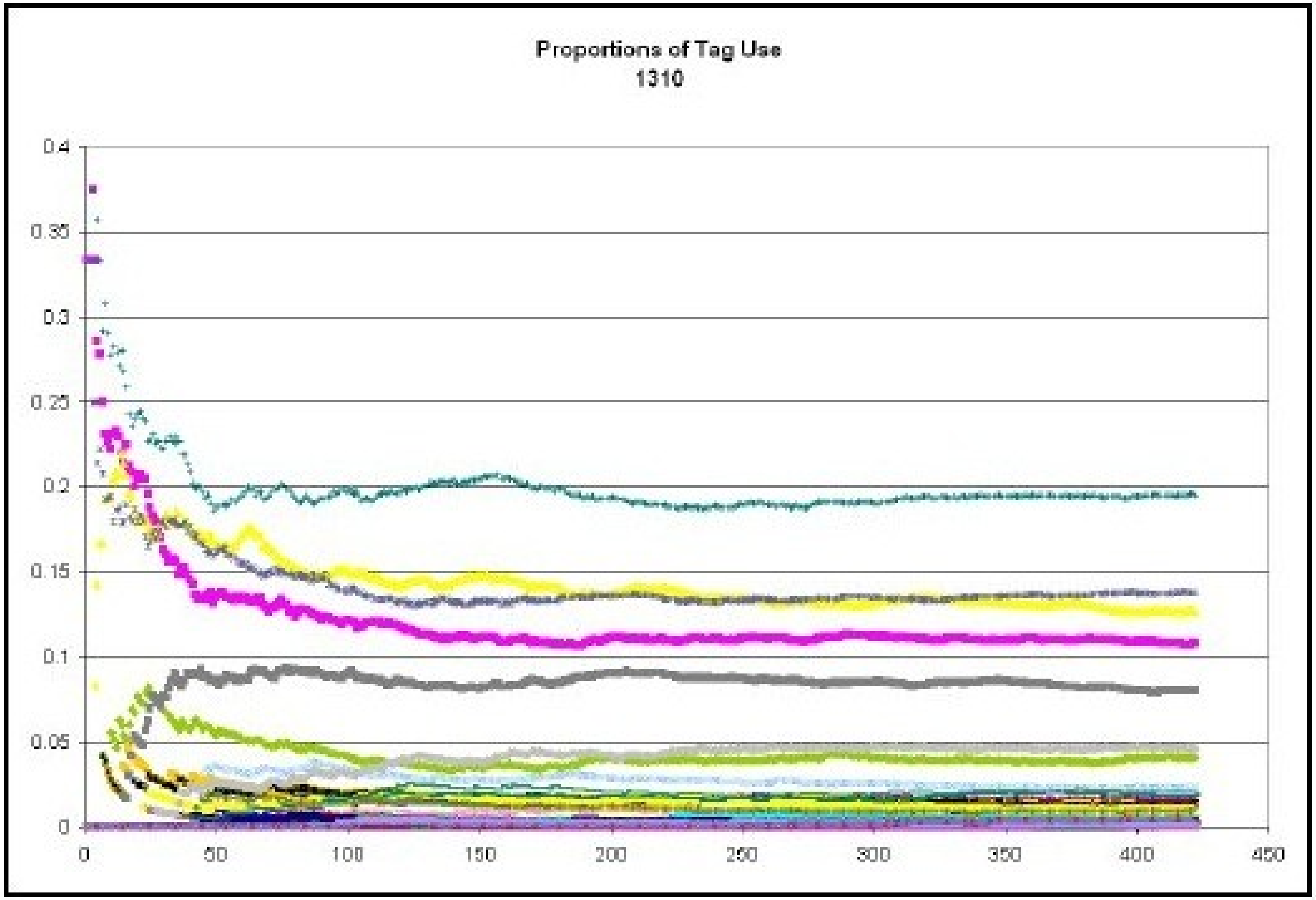
By combining information from many photos - tags, it seems that we can

Stable patterns

in tagging systems over time

From arny johanns From nrvica From fernando780 From jordanmerric... From humedini

Stable tagging patterns

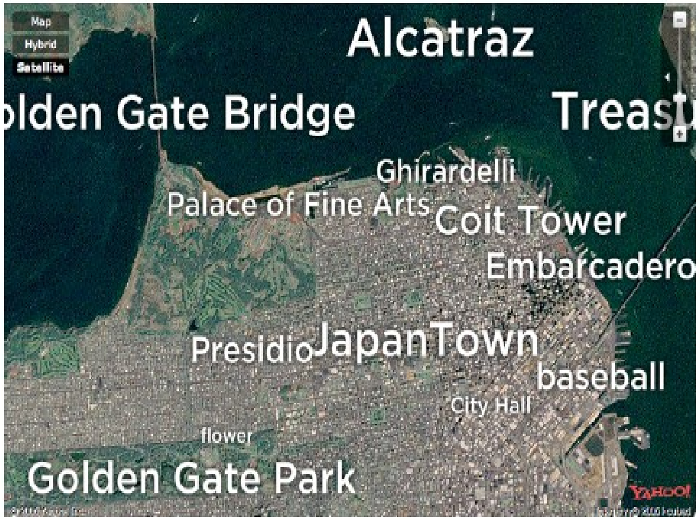


What else we can do?

Contribute to our understanding of the world

Tags that are "representative" for a geographical area

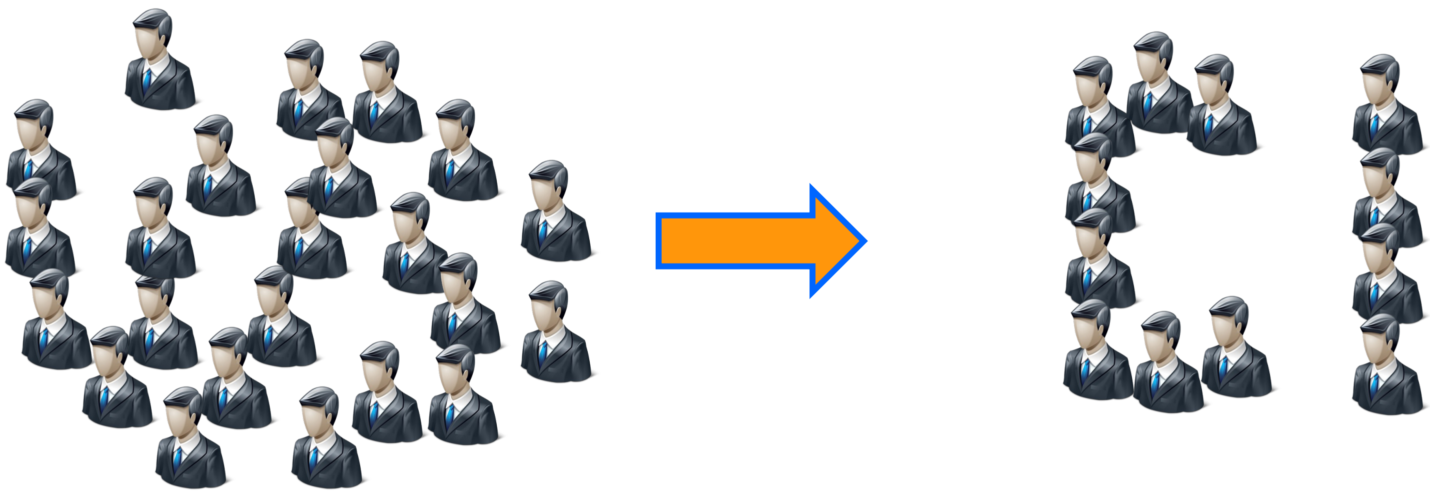
- 1. Clustering of photos
  - K-means, based on their location [Kennedy07]
- 2. Rank each cluster's tags
- 3. Get tags above a certain threshold



Representative tags for San Francisco [Kennedy07]

Collective Intelligence, PeopleWeb, Croudsourcing, Wisdom of crouds …

Collective Intelligence is the Intelligence which emerges from the collaboration, competition and coordination among individuals.



…an Intelligence greater than the sum of the individuals' intelligence

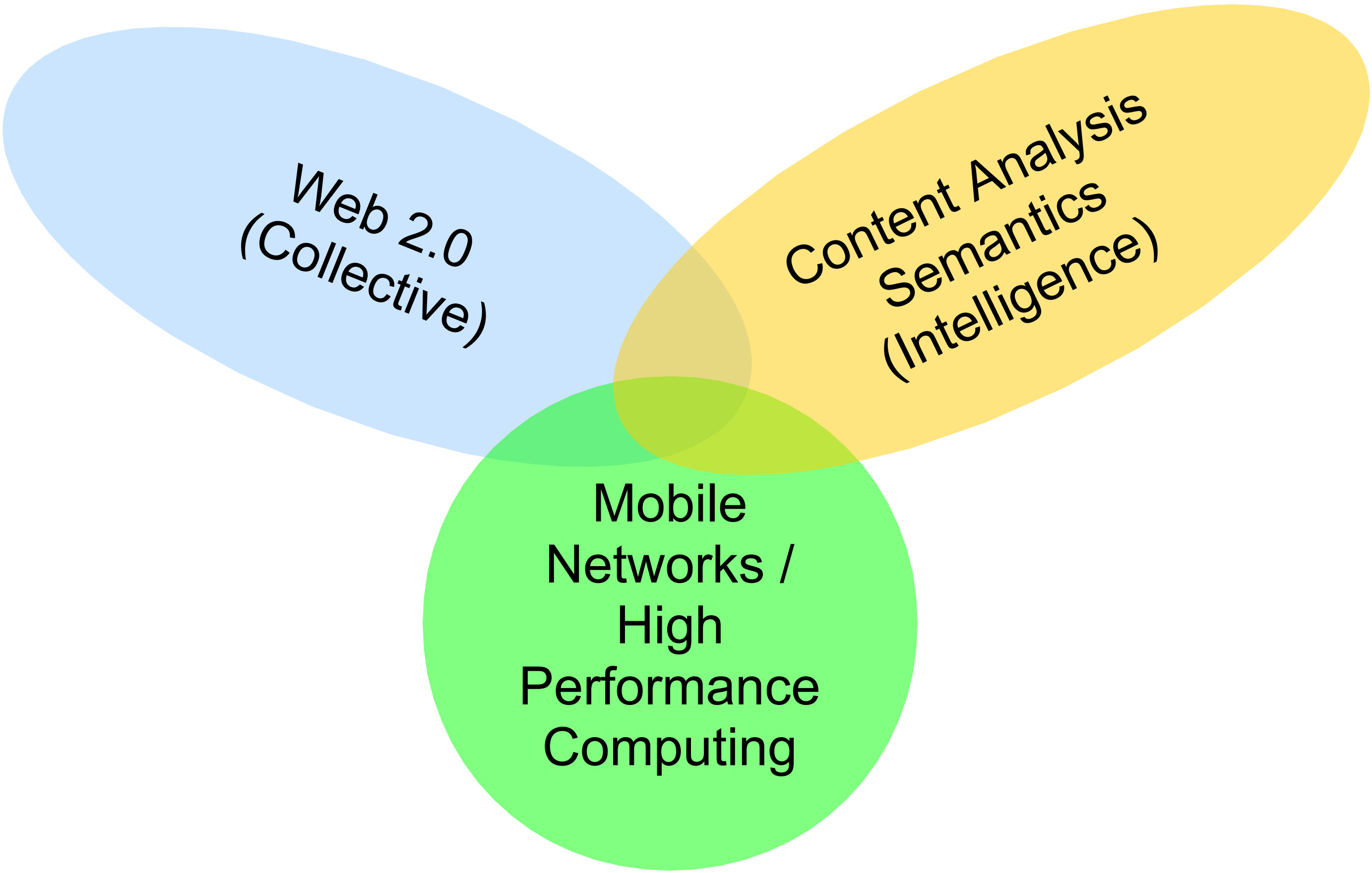
CI and Web 2.0?

- Analyze user-generated content, such as tags that are manually assigned to photos, and its relation to context over time, space and social connectivity
- Sources
  - Tags
  - Content
  - Social info
  - Time, Location
  - Other sources (e.g. Wikipedia)

http://www.iyouit.eu


Low-level
Comms
flickr
Deutsches Eck from Ehrenbreitstein Fortress, Koblenz, Germany
Favs
Time
by schaengel
121 comments
69 faves
Tagged with koblenz, ehrenbreitstein ...
Tags
Taken on November 15, 2009, uploaded November 17, 2009
See more of schaengel photos, or visit his profile.
When you're high up on the hill above Koblenz at Ehrenbreitstein Fortress you can get a great panoramic view of the city and the surrounding area.
Caption
Geo
loungefly
User Profile
Groups
Social network

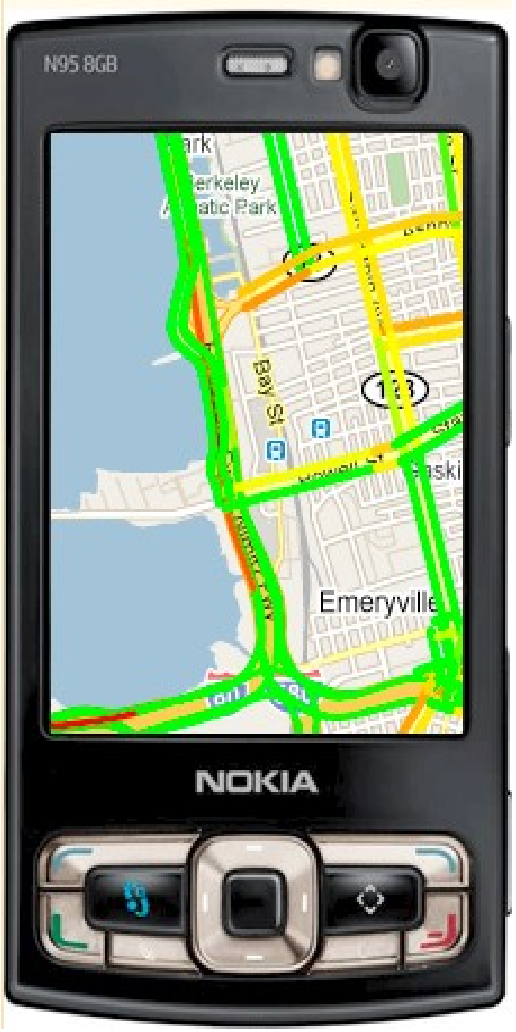
Why today?


Web 2.0
(Collective)
Content Analysis
Semantics
(Intelligence)
Mobile
Networks /
High
Performance
Computing

A “simple” example

Uses the GPS in cellular phones to gather traffic information, process it, and distribute it back to the phones in real time

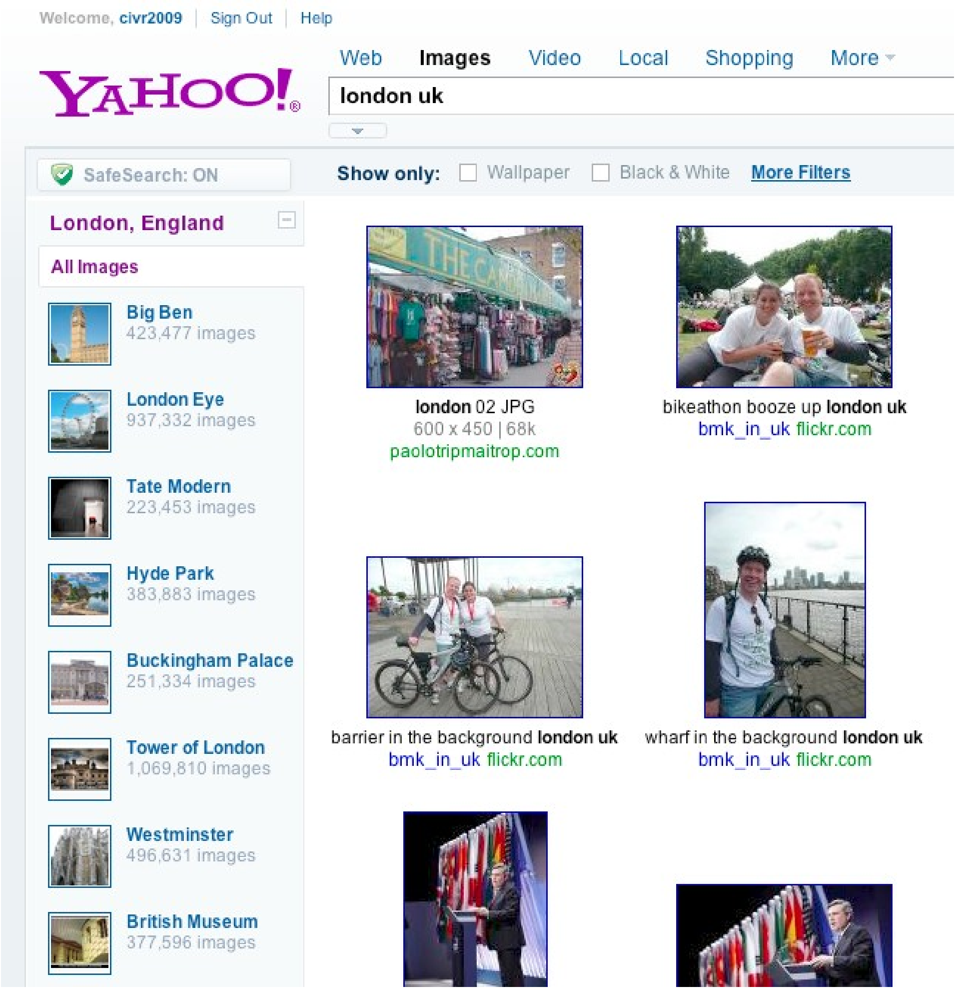
- online, real-time data processing
- privacy-preservation
- data efficiency, i.e. not requiring excessive cellular network



Mobile Century Project: http://traffic.berkeley.edu/mobilecentury.html

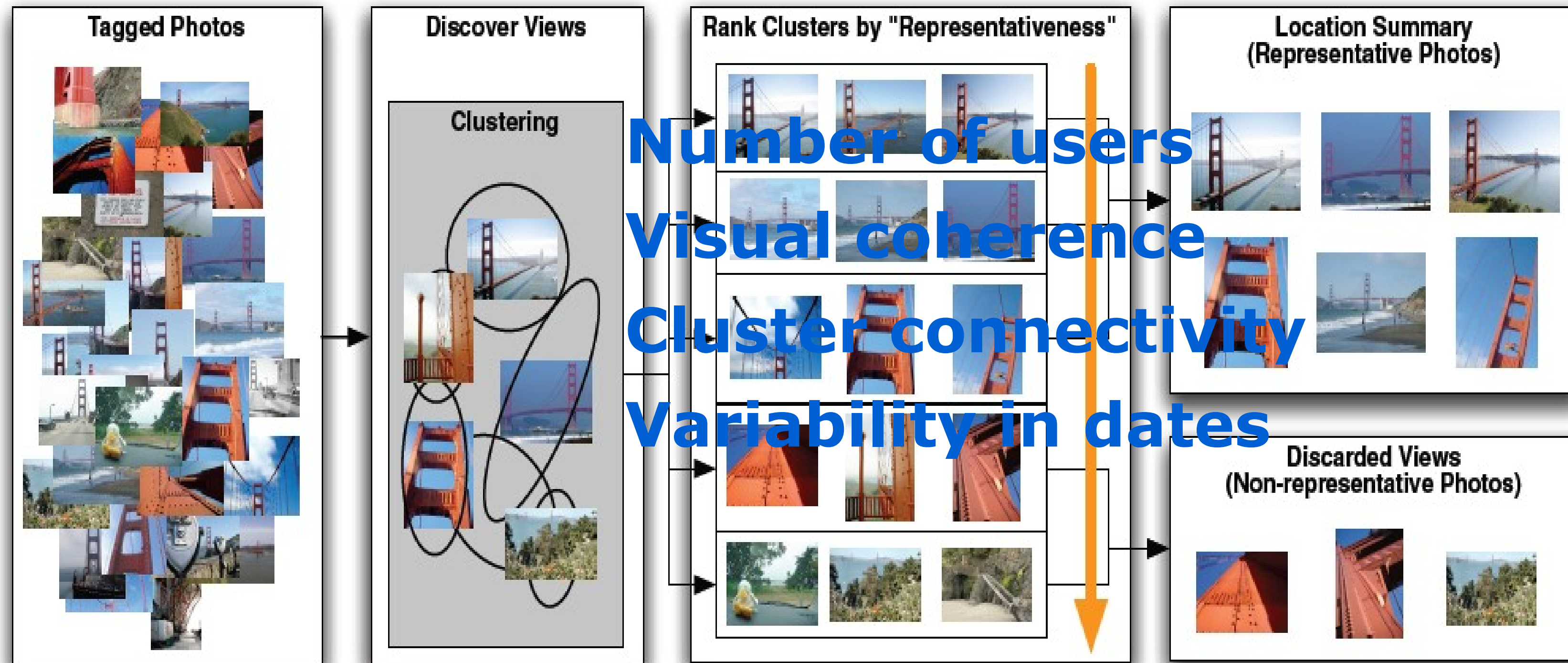
Image search - Tourism

- Linguistic processing of semi-structured sources
  - Wikipedia, Geoplanet
- Statistical analysis for ranking
  - User Queries
  - Flickr tags



Generating photo summaries

- **Problem formulation**: Having identified a tag x as representative of a cluster, compute a set of photos that are representative for that tag



Number of users
Visual coherence
Cluster connectivity
Variability in dates

Generating photo summaries for geographic objects in [Kennedy07]

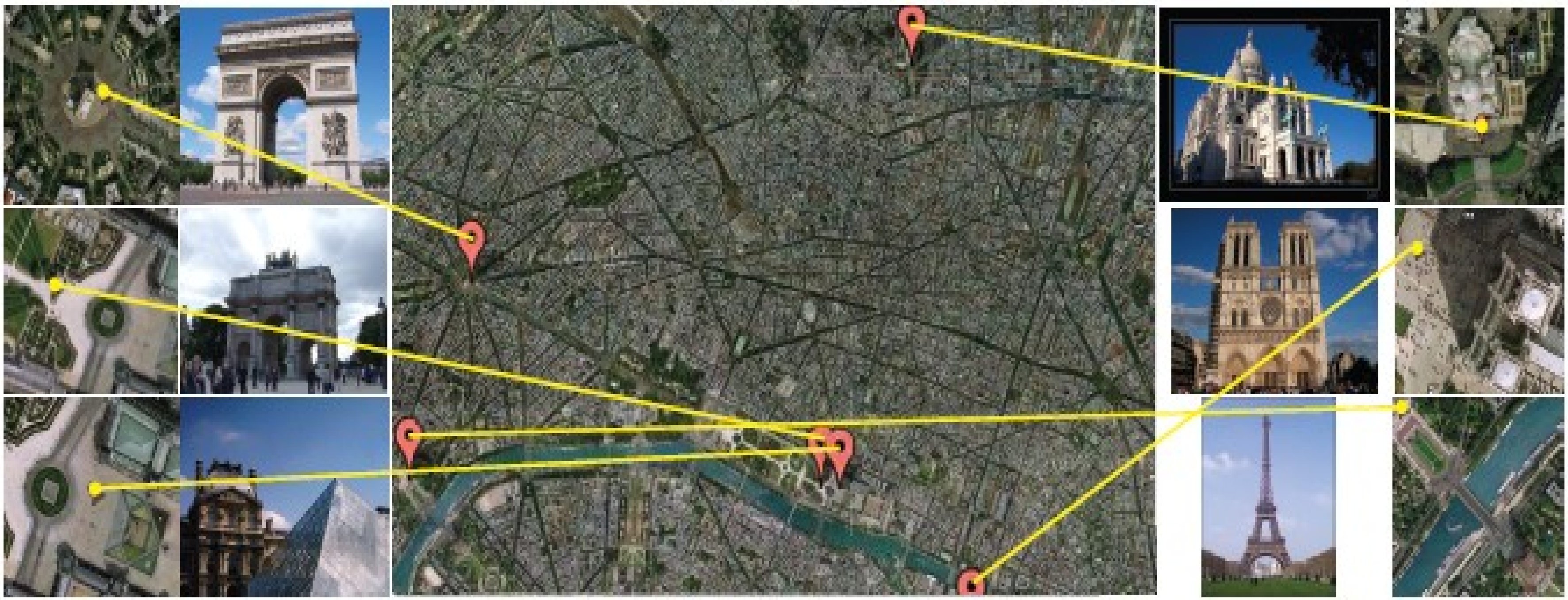
Sample photo summaries of events [Quack08]

DATASET: Divide the earth's surface into square tiles of 200m2
70000 geographic tiles
220000 geotagged photos from Flickr
After preprocessing, 73000 photos were assigned to clusters
Manually labeling of 700 clusters

"Oxford Geek nights"
"Movie premiere Italy"
"Exhibition gallery paris"

The most commonly identified event (single day covered by a single photographer)

Auto annotation & geo-location



# Images	222'757
Size Metadata	1.1 GB
Size Features	111 GB
# Images assigned to clusters	73'236
# Similarities computed	217'330'144
# Similarities > 0	751'457

[Quack08]

EpiCollect: Science - epidemiology example

A scientist or member of the public collects and records data, photos and videos then sends this information to a central web-based database

- e.g. to document the presence of an animal or plant species that are “representative” for a geographical area
- Location information – maps
- Citizen scientists

EpiCollect: Linking Smartphones to Web Applications for Epidemiology, Ecology and Community Data Collection, David M. Aanensen, Derek M. Huntley, Edward J. Feil, Fada'a al-Own, Brian G. Spratt

Research Fields and Issues

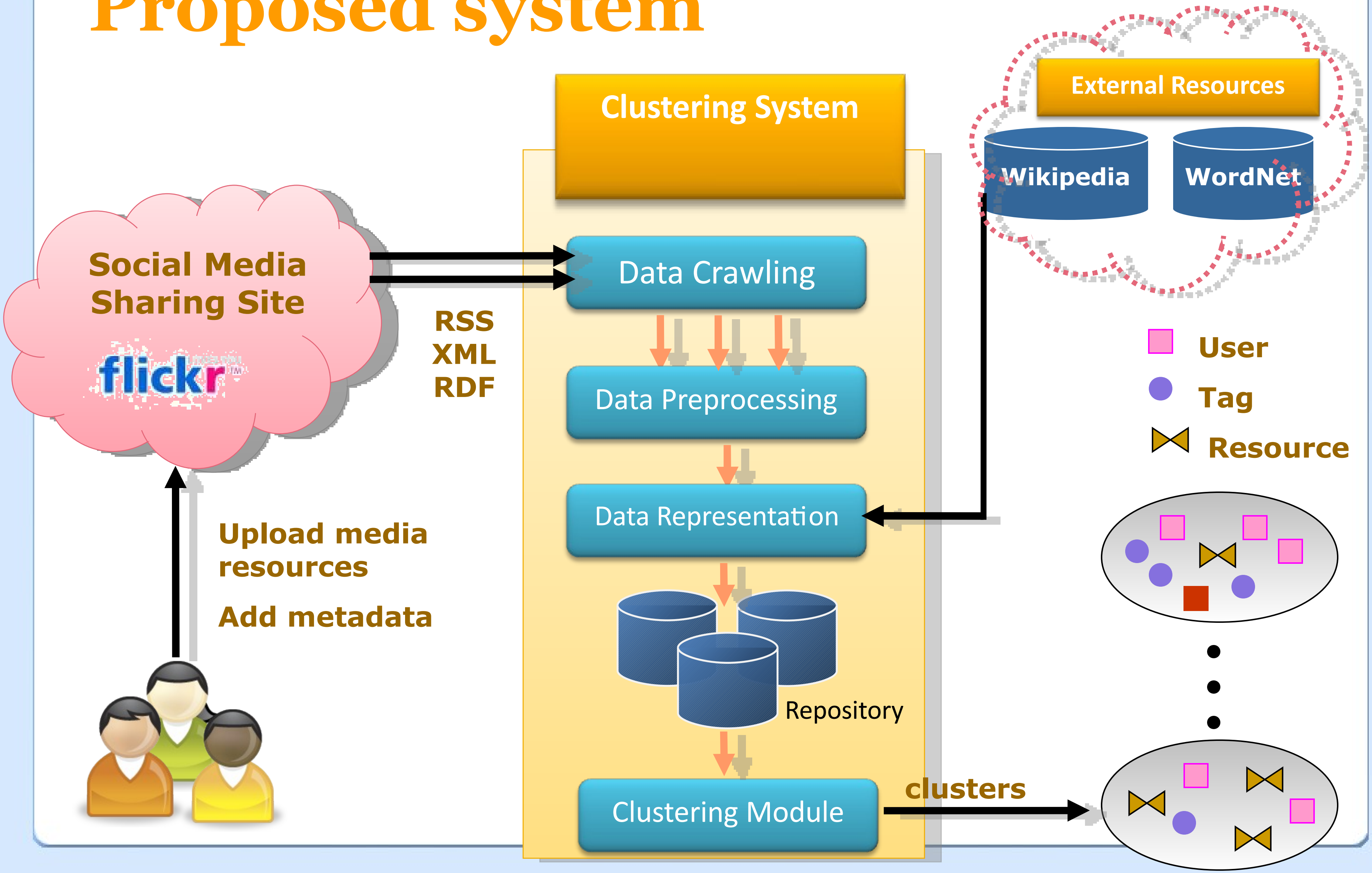
- Statistical analysis, machine learning, data mining, pattern recognition, social network analysis
- Clustering
- Graph theory
- Image, text, video analysis
- Information extraction
- Fusion techniques
- Trust, security, privacy
- Performance, scalability
    - speed, storage, power, grids, clouds

Clustering for Social Media

Clustering Approaches

- Tag-Based
- Content-Based
- Time-based

Proposed system



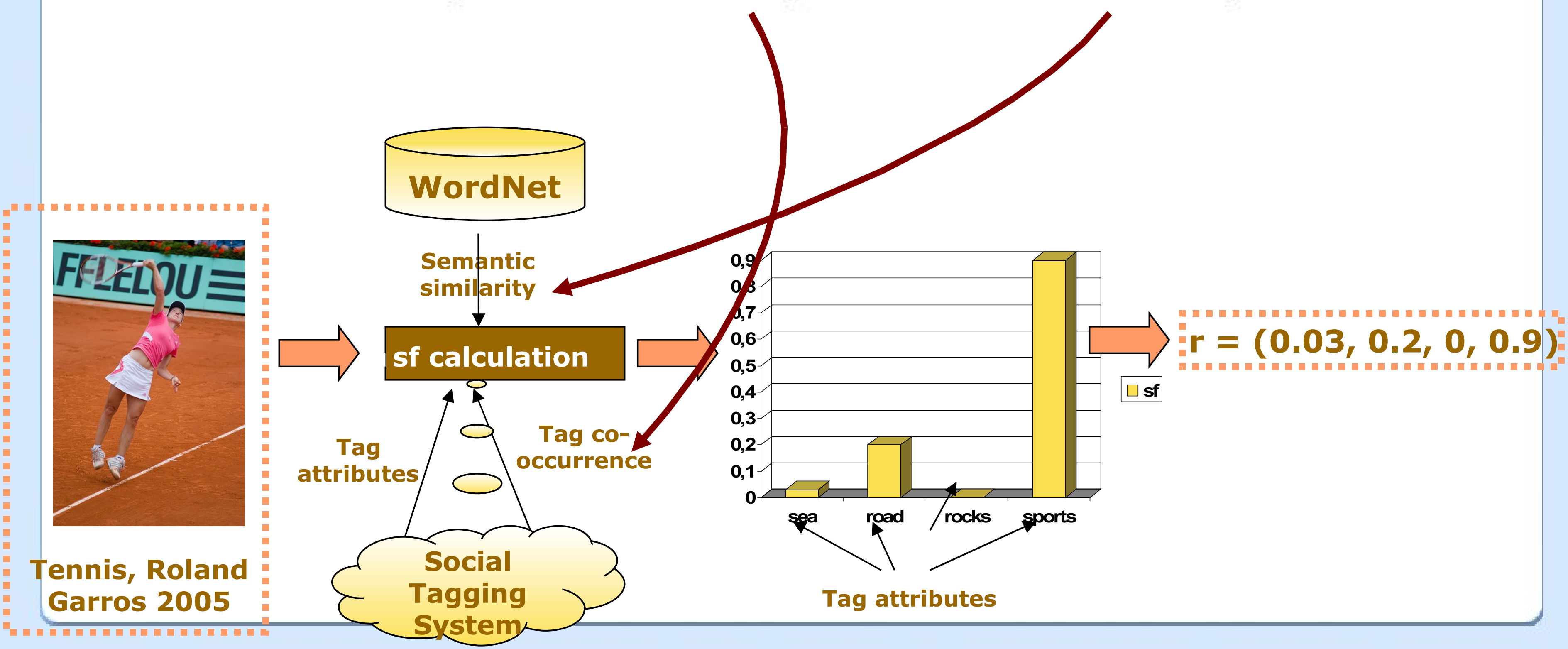
Tag-based Clustering (I)

- **1. Vector data model**
- Assume **n** resources and **d** attribute-tags
  - **d:** a representative set of tags
- A resource representation in vector space (**sf**) is based on semantic similarity and tag co-occurrence between the resource's tags and the attribute-tags
- A resource $\mathbf{r}_i$ is represented by a **d**-dimensional vector $\mathbf{r}_i = (\mathbf{sf}_1, \mathbf{sf}_2, \ldots, \mathbf{sf}_d)$
- All resources can be represented by an **n** x **d** matrix

Tag-based Clustering (II)

- **2. Clustering on n (resources, r) x d (attributes) matrix (K-means, Hierarchical, COBWEB)**

$$SS(t_x, t_y) = w * SoS(t_x, t_y) + (1 - w) * SeS(t_x, t_y)$$



Tag-based Clustering - Experimental Results

- **Dataset:** 3000 images downloaded from Flickr
- Meaningful subdomains of **roadside**:

buildings, roof, street, road

(a)

cars, vehicles, race

(b)

people, street, festival

(c)

- Different clusters for the **ambiguous tag** *wave, rock:*

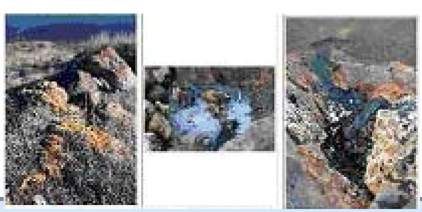
wave, sea, ocean

(a)

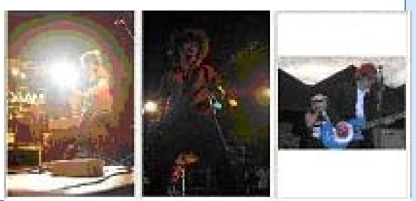
wave, person, hand

(b)

rocks, stone, rockyside



rock, music, band



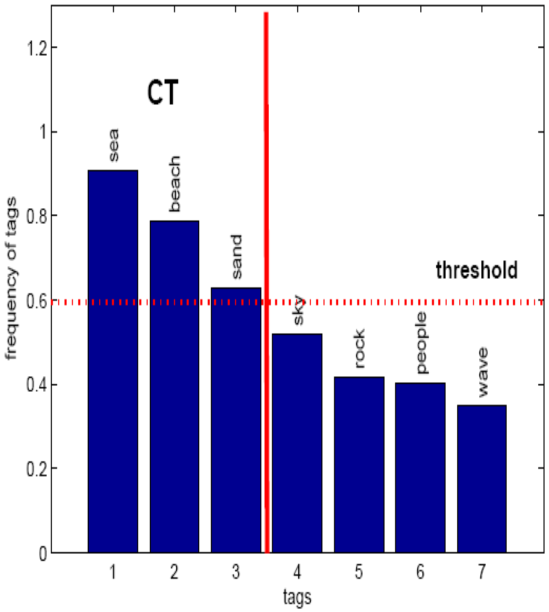
Tag & Content-based Clustering

- After performing tag-based clustering, low-level features of resources are used for cluster refinement
- Outlier Detection (mahalanobis distance)
- For each resource the following visual descriptors are extracted:
  - Scalable Color, $\mathcal{SC}$
  - Color Structure, $\mathcal{CS}$
  - Color Layout, $\mathcal{CL}$
  - Edge Histogram, $\mathcal{EH}$
  - Homogenous Texture, $\mathcal{HT}$
- A single image feature vector per each resource is produced, encompassing all descriptors normalized in [0,1]
- Feature extraction and distances between image feature vectors are according to MPEG-7 XM.

Evaluation Method

- **<u>Definition:</u>** Cluster Topic, CT, are the tags that have frequency in cluster's resources annotation over a threshold `T`.

- **Evaluation Metrics**
- Precision $$Pr\left(C_j\right) = \frac{\left|C_j \cap RR(C_j)\right|}{\left|C_j\right|}$$
- Recall $$R\left(C_j\right) = \frac{\left|RR(C_j) \cap C_j\right|}{\left|RR(C_j)\right|}$$
- F-Measure $$F\left(C_j\right) = \frac{2 * Pr\left(C_j\right) * R\left(C_j\right)}{Pr\left(C_j\right) + R\left(C_j\right)}$$

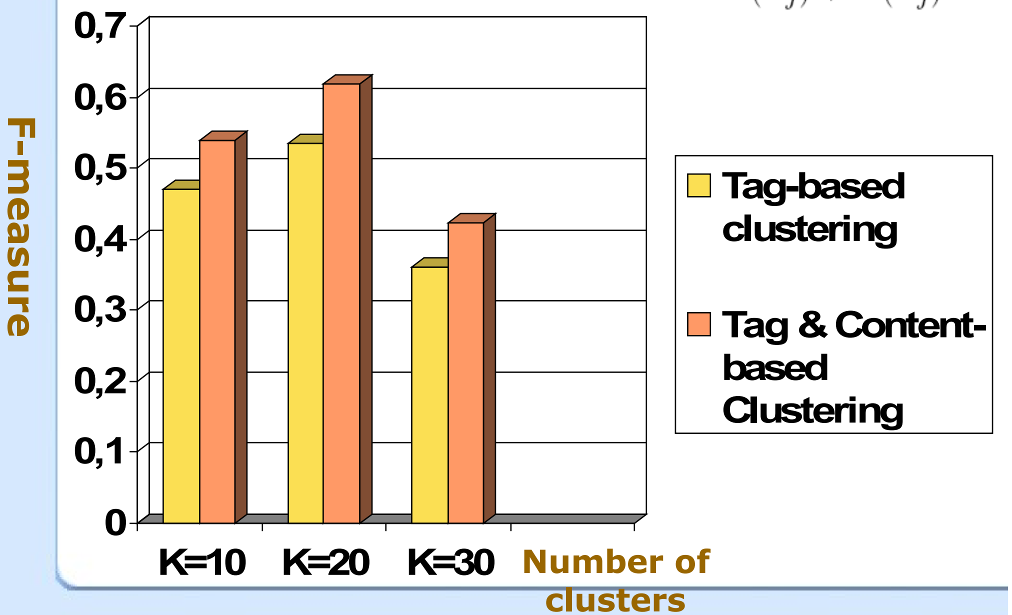


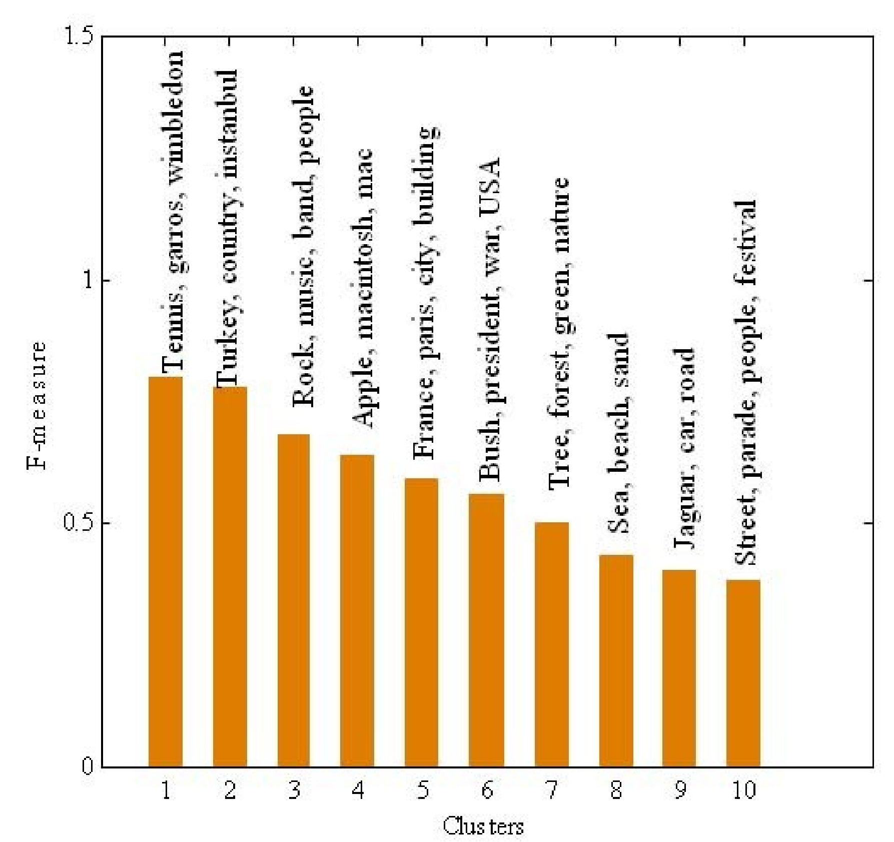
Tag & Content-based Clustering – Experimental Results

Dataset: 10000 images (with their tags) downloaded from Flickr

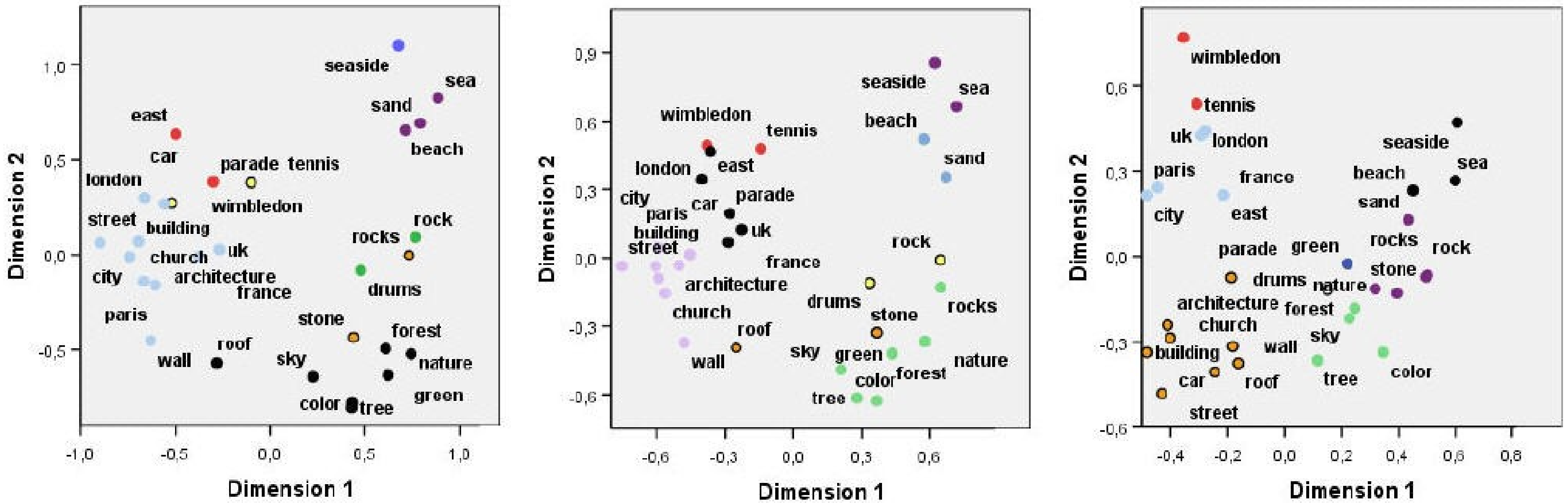
Evaluation: Manual annotation and use of F-Measure.

$$F(C_j) = \frac{2 * Pr(C_j) * R(C_j)}{Pr(C_j) + R(C_j)}$$





Experimental Results (II)



(a) $w = 0.2$ (b) $w = 0.5$ (c) $w = 0.8$

Attributes Assignment to k=8 clusters,

$\mathcal{W}$: weighting factor of semantic similarity against similarity derived from tag co-occurrence

Why consider time?

- Motivation

Events, Trends, Changing of user interests

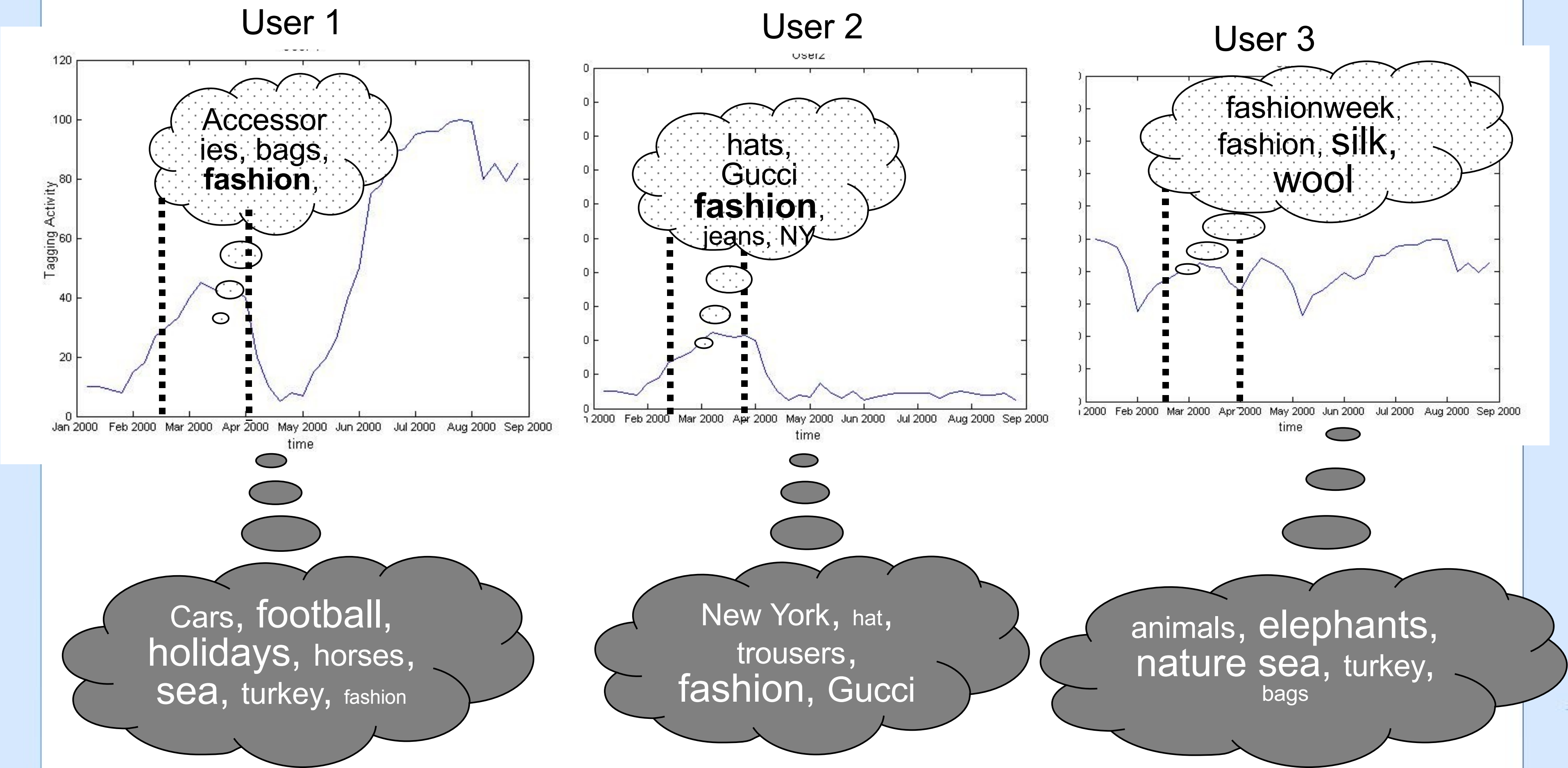
⇩

Users Tagging Behavior changes over time

⇩

Time is a fundamental dimension in analysis of users and tags in a social tagging system

Many times, a user's targeted interest is hidden in the general tagging activity….



The basic idea

$$UTF = \begin{array}{c} \\ u_1 \\ u_2 \\ \vdots \\ u_N \end{array} \begin{array}{c} \begin{array}{cccc} i_1 & i_2 & \cdots & i_D \end{array} \\ \begin{bmatrix} ut_{11} & ut_{12} & \cdots & ut_{1D} \\ ut_{21} & ut_{22} & \cdots & ut_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ ut_{N1} & ut_{N2} & \cdots & ut_{ND} \end{bmatrix} \end{array},$$

$$TTF = \begin{array}{c} \\ t_1 \\ t_2 \\ \vdots \\ t_S \end{array} \begin{array}{c} \begin{array}{cccc} i_1 & i_2 & \cdots & i_D \end{array} \\ \begin{bmatrix} tt_{11} & tt_{12} & \cdots & tt_{1D} \\ tt_{21} & tt_{22} & \cdots & tt_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ tt_{N1} & tt_{N2} & \cdots & tt_{SD} \end{bmatrix} \end{array},$$

Step 1: Representation

Step 3: Focus on time locality

$$< u_j, t_k > = \frac{u_j * t_k}{\sqrt{\sum_{j=1}^{N} u_i^2 * \sum_{k=1}^{S} t_k^2}}$$

Step 4: Combination of semantic and time information

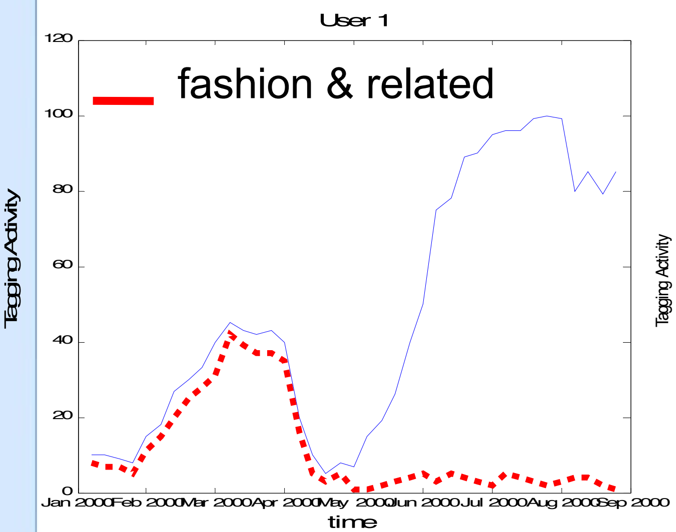
(u_i,t_j)= *SemSim*(u_i,t_j) * *InnerProduct*(u_i,t_j)

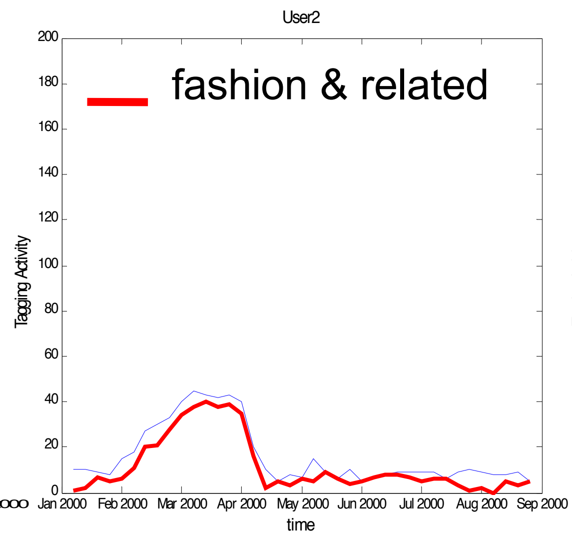
WordNet

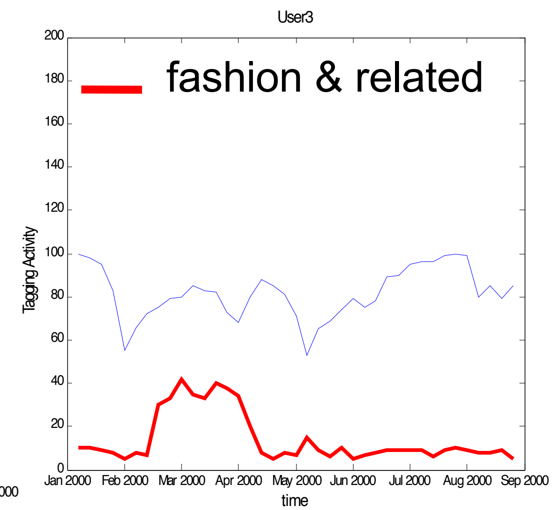
Step 2: Focus on contents (tags semantics)

$$\begin{array}{c} \\ u_1 \\ u_2 \\ \vdots \\ u_S \end{array} \begin{array}{c} \begin{array}{cccc} t_1 & t_2 & \cdots & t_D \end{array} \\ \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1D} \\ w_{21} & w_{22} & \cdots & w_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & w_{SD} \end{bmatrix} \end{array}$$

An example



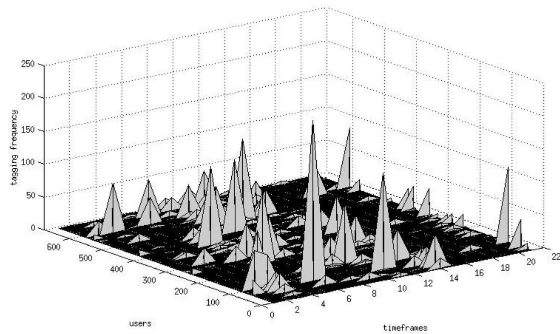


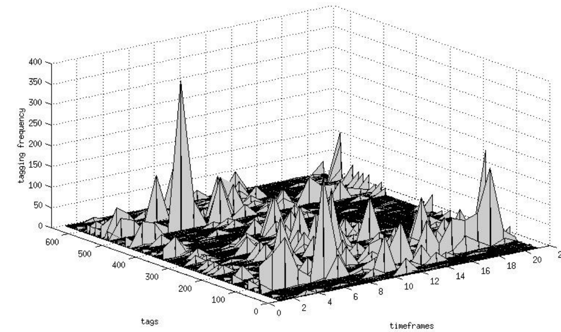


$$(u_i,t_j)= SemSim(u_i,t_j) * InnerProduct(u_i,t_j)$$

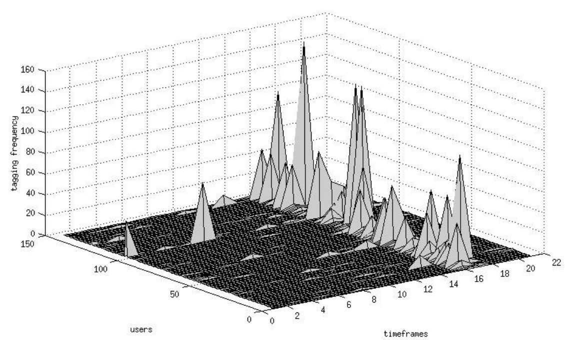
Time-aware user/tags clusters on Flickr (I)

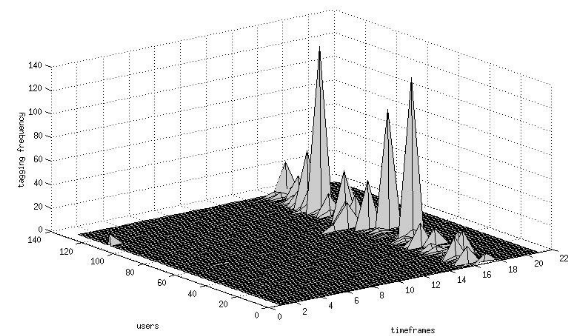
Cluster of users interested regularly in weddings and related tags



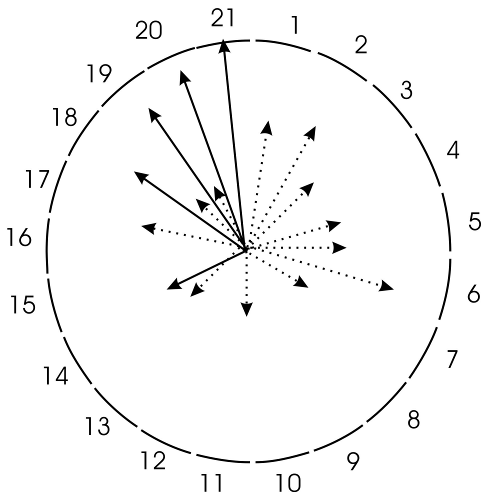


Cluster of users interested in Olympics and related tags

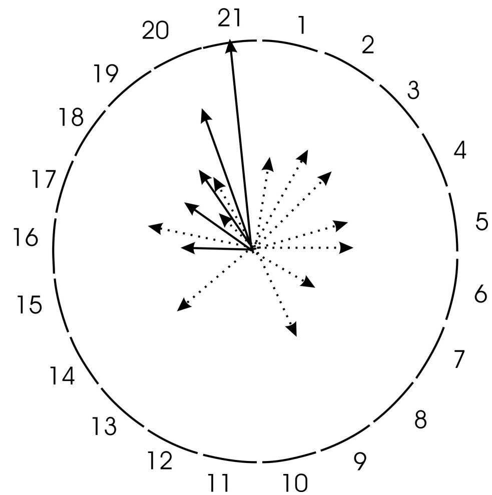




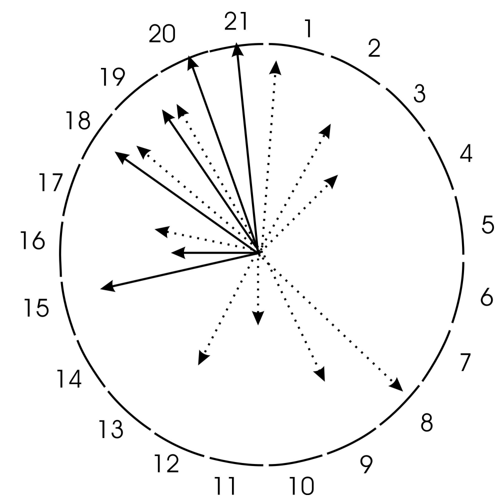
Time-aware user/tags clusters on Flickr (II)



Tags distribution in a cluster



User1' s tags distribution

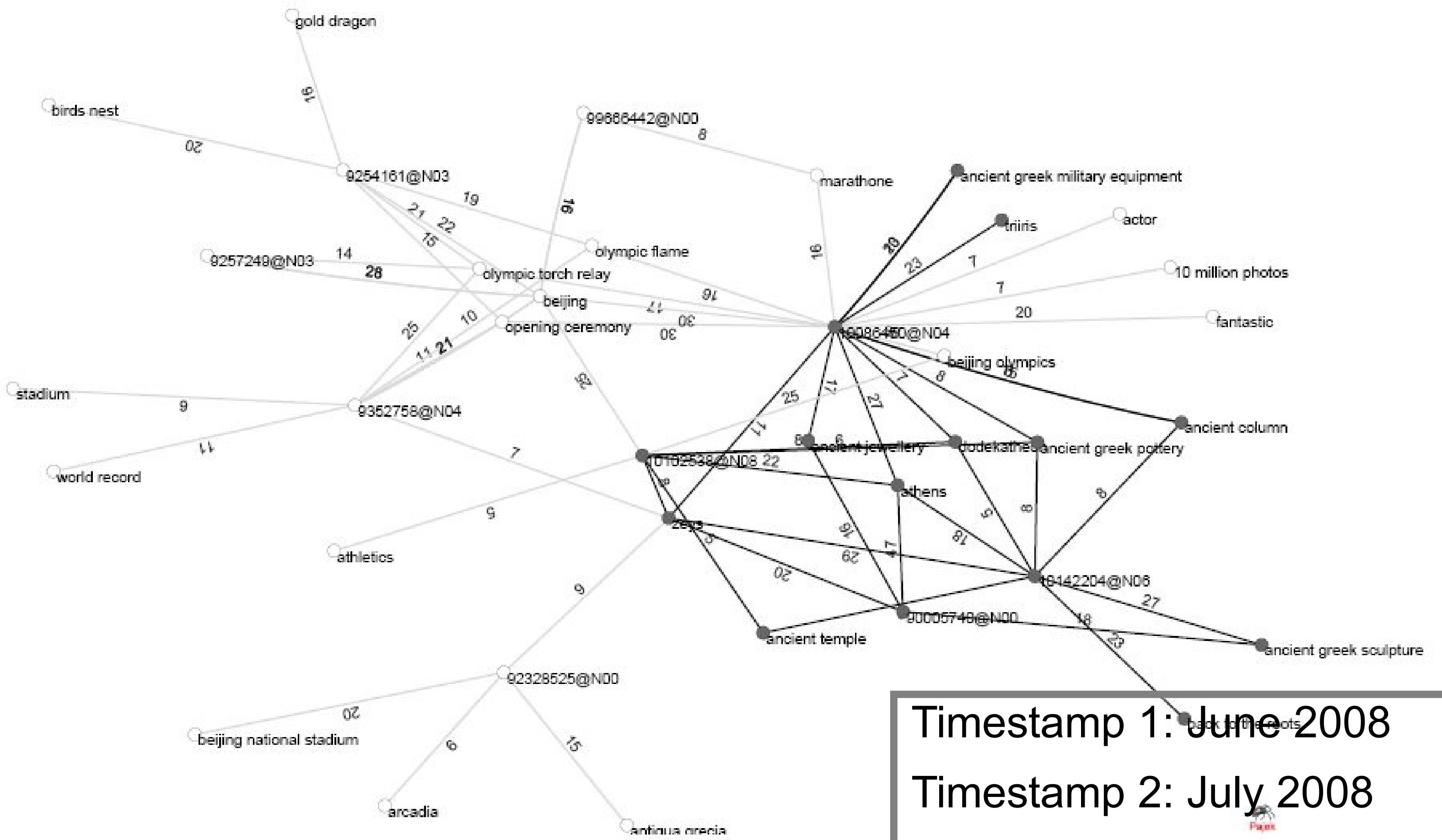


User2' s tags distribution

Olympics –related tags

Ancient Greece –related tags

Cluster evolution (timestamp 1)

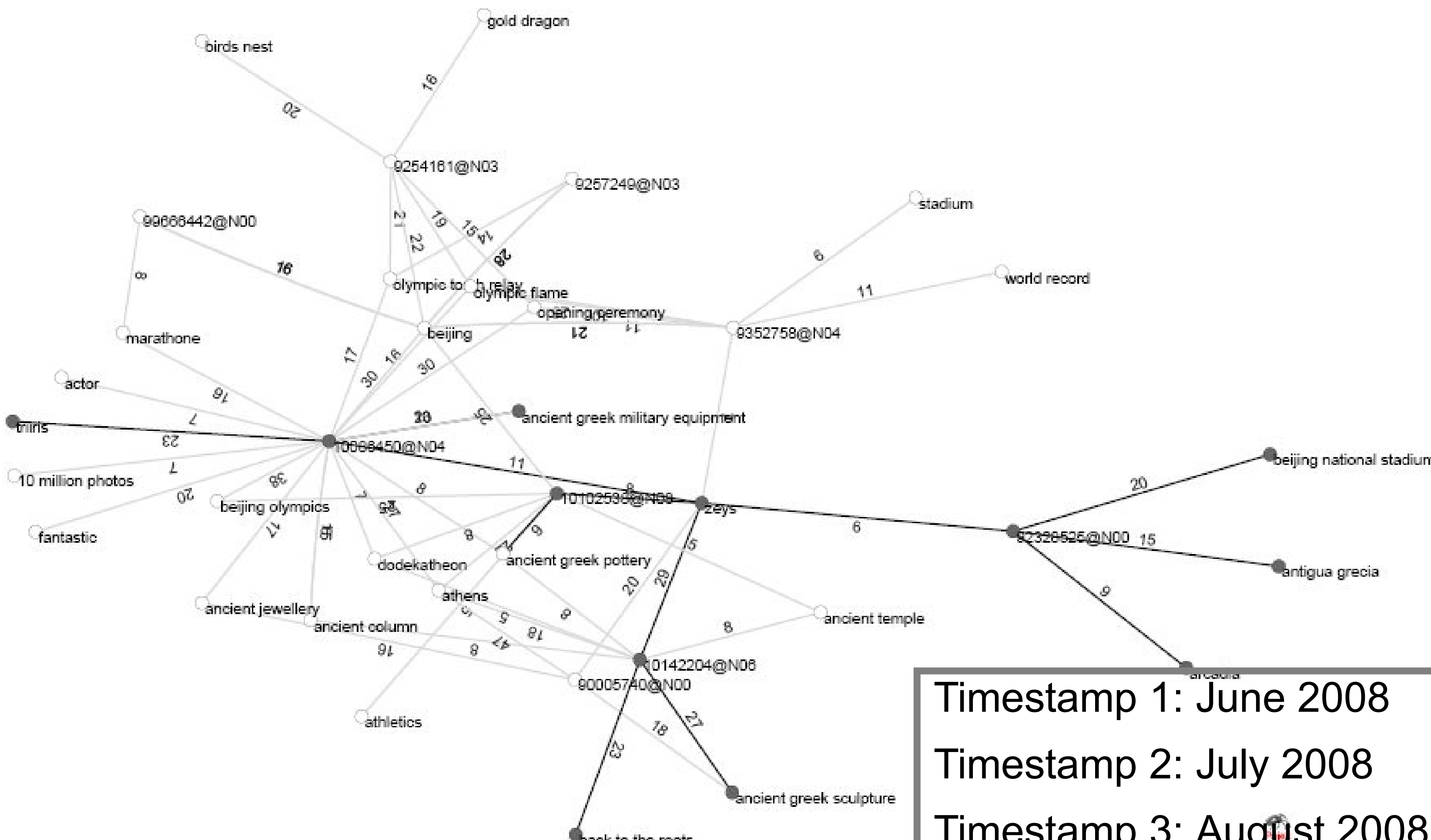


Timestamp 1: June 2008

Timestamp 2: July 2008

Timestamp 3: August 2008

Cluster evolution (timestamp 2)

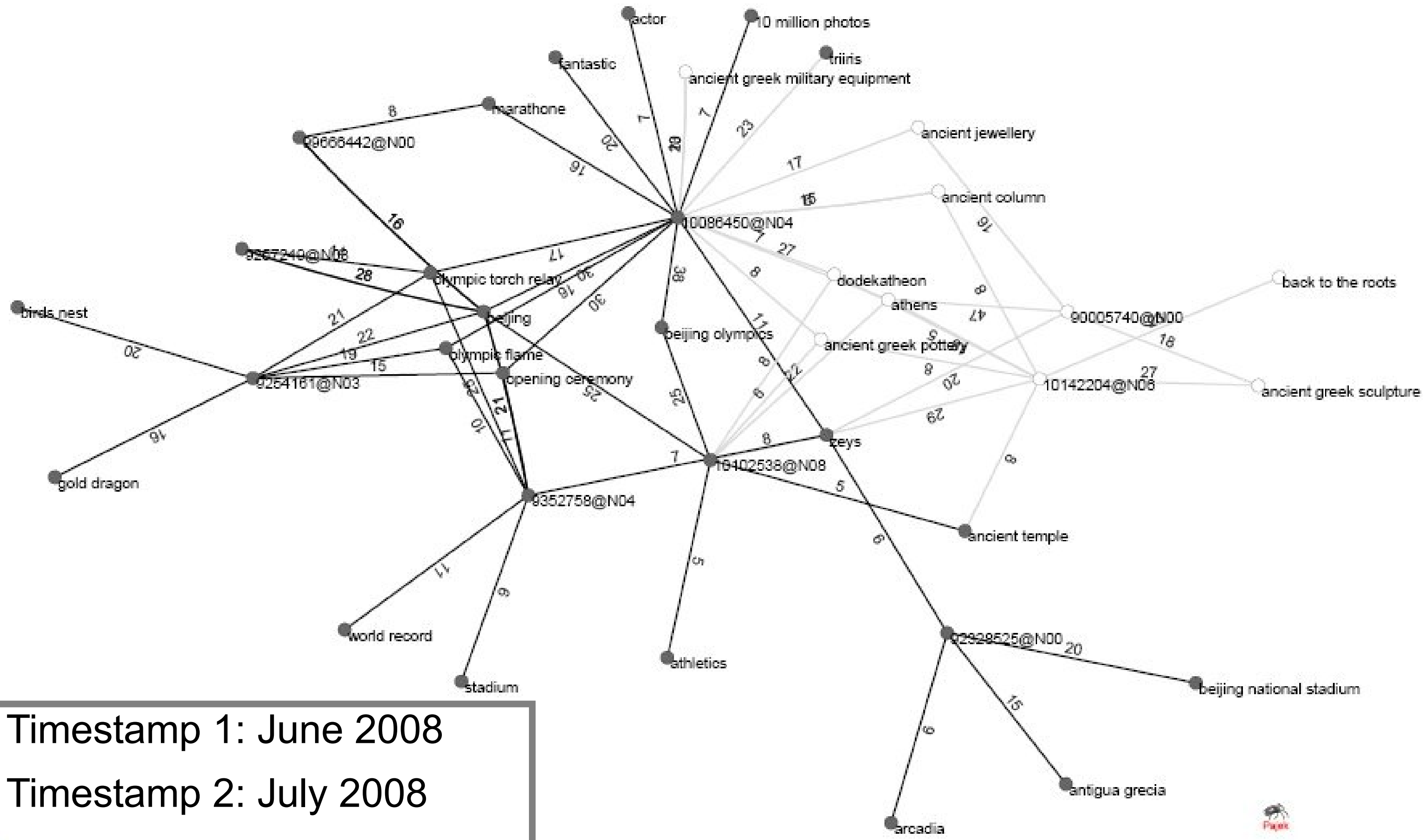


Timestamp 1: June 2008

Timestamp 2: July 2008

Timestamp 3: August 2008

Cluster evolution (timestamp 3)



Timestamp 1: June 2008

Timestamp 2: July 2008

Timestamp 3: August 2008

Social Media "teacher" of the machine

Exploiting clustering for machine learning

Objective: Develop a framework able to create strongly annotated training samples from weakly annotated images

Problems:

- Object detection schemes require region-detail annotations
- Manual annotation is laborious and time consuming

Tagged images

sand, wave, rock, sky

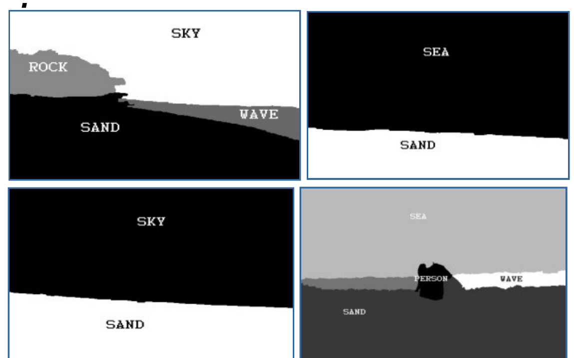
sea, sand

sand, sky

person, sand, wave, see

Social information + Image analysis

Region-detail annotated



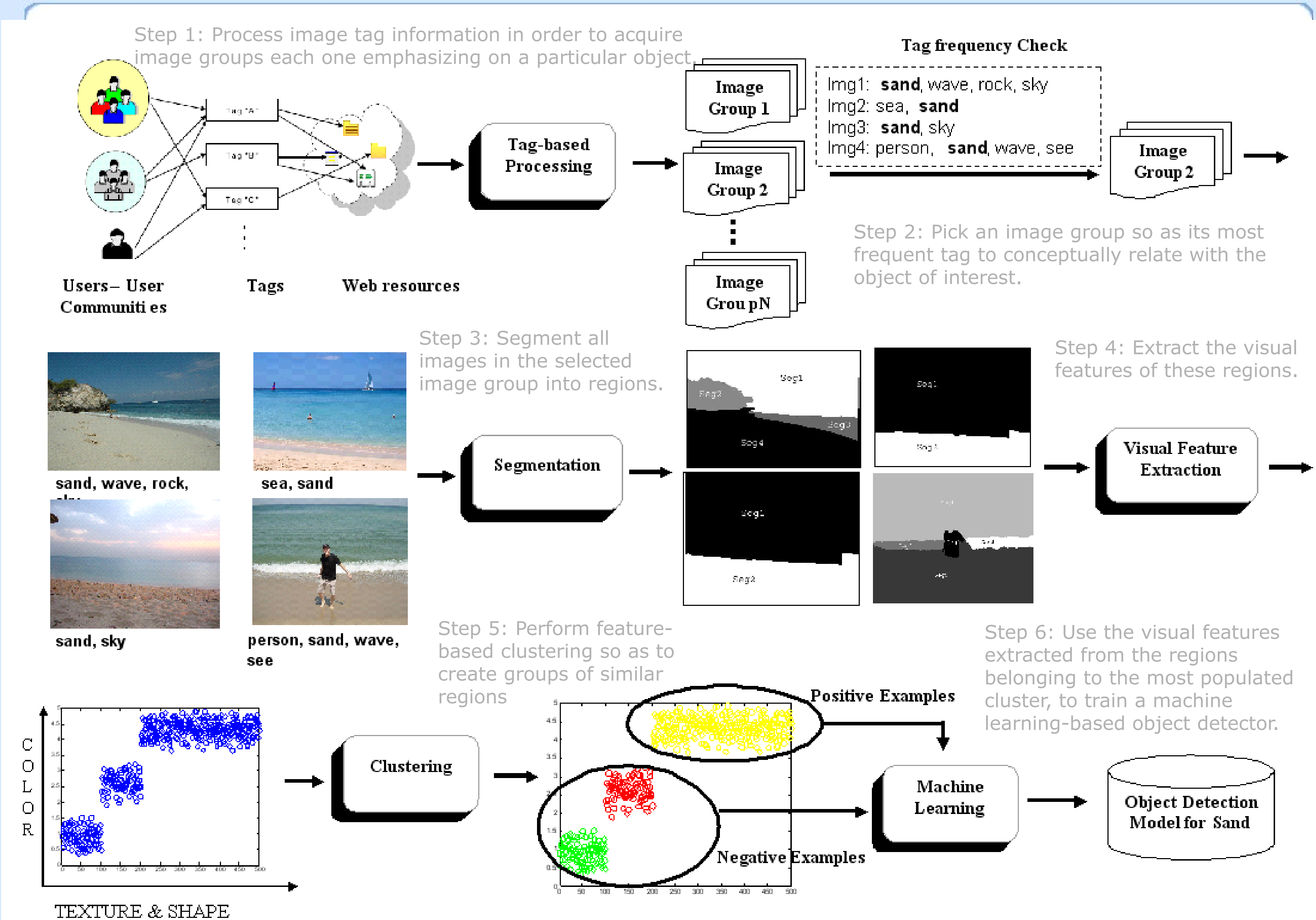
[Chatzilari09]

Machine Learning

Object Detectors

Solutions:

- Exploit user tagged images from social sites like flickr
- Combine techniques operating on tag and visual information space


Step 1: Process image tag information in order to acquire image groups each one emphasizing on a particular object.
Users – User Communities
Tags
Web resources
Tag-based Processing
Image Group 1
Image Group 2
Image Group pN
Tag frequency Check
Img1: sand, wave, rock, sky
Img2: sea, sand
Img3: sand, sky
Img4: person, sand, wave, see
Image Group 2
Step 2: Pick an image group so as its most frequent tag to conceptually relate with the object of interest.
Step 3: Segment all images in the selected image group into regions.
sand, wave, rock, sky
sea, sand
sand, sky
person, sand, wave, see
Segmentation
Step 4: Extract the visual features of these regions.
Visual Feature Extraction
Step 5: Perform feature-based clustering so as to create groups of similar regions
COLOR
TEXTURE & SHAPE
Clustering
Positive Examples
Negative Examples
Step 6: Use the visual features extracted from the regions belonging to the most populated cluster, to train a machine learning-based object detector.
Machine Learning
Object Detection Model for Sand

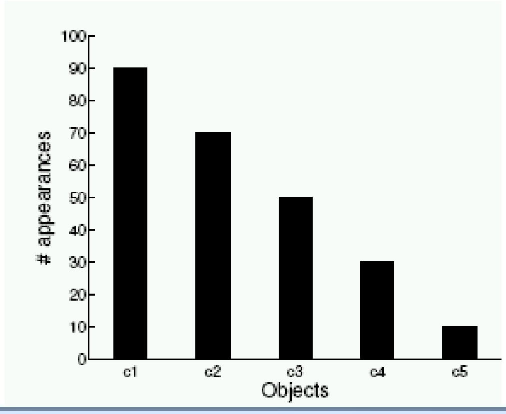
Tag-based processing

[Giannakidou08]

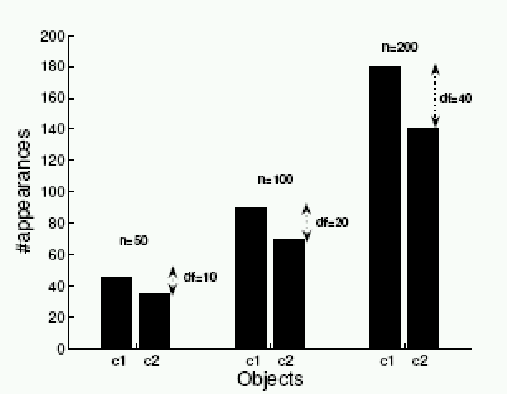
SEMSOC, vector space model where each image is projected onto a space defined by the most prominent tags

SEMSOC output example

Distribution of objects based on their frequency rank



Absolute difference between 1st and 2nd most highly ranked objects increases as n increases



Segmentation & Visual Descriptors

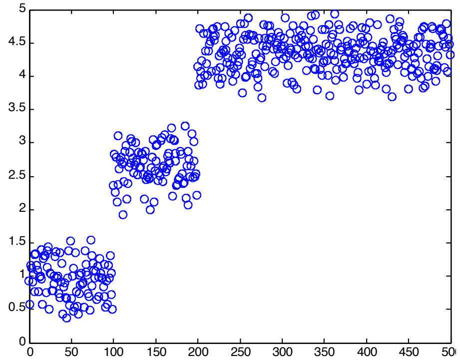
- Segmentation
  - K-means with connectivity constraint (KMCC)

[Mezaris et al., 2004]

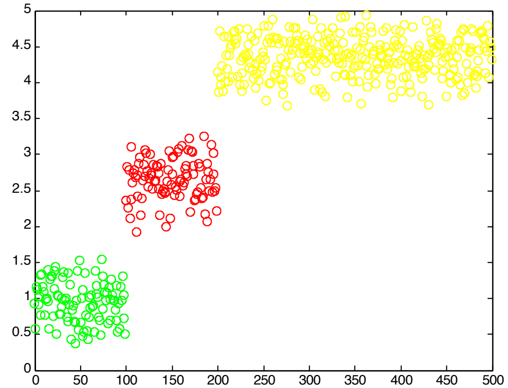
- Visual Descriptors
  - MPEG-7 standard
    - *Dominant Color , Color Layout, Color Structure, Scalable Color, Edge Histogram, Homogeneous Texture, Region Shape.*

[Bober et al., 2001], [Manjunath et al., 2001].

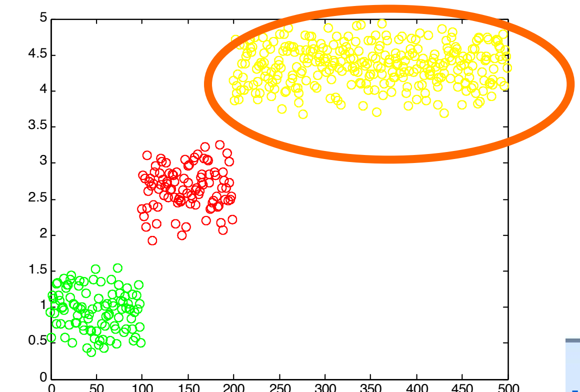
Region-based Clustering & Cluster Selection Region clustering



- Perform segmentation and visual feature extraction from all images in an image group (Identified by SEMSOC)

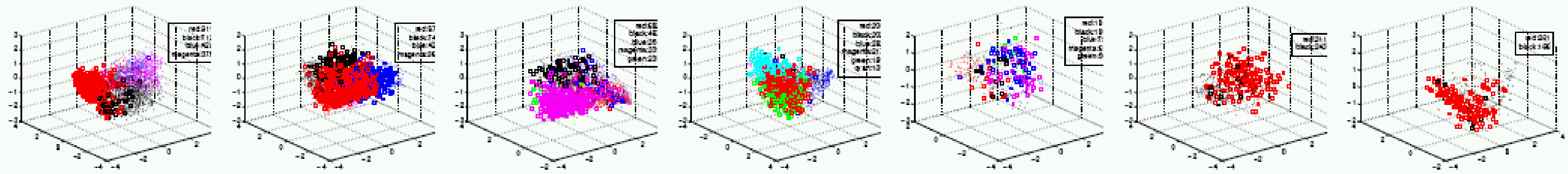


- Perform clustering based on visual features to gather together regions depicting the same object



- Pick the most populated cluster as the one representing the most frequently appearing tag of the group

Experimental Results – Cluster Selection



(a) Sky (b) Sea (c) sand (d) Person (e) Boat (f) Vegetation (g) Rock

Setting:

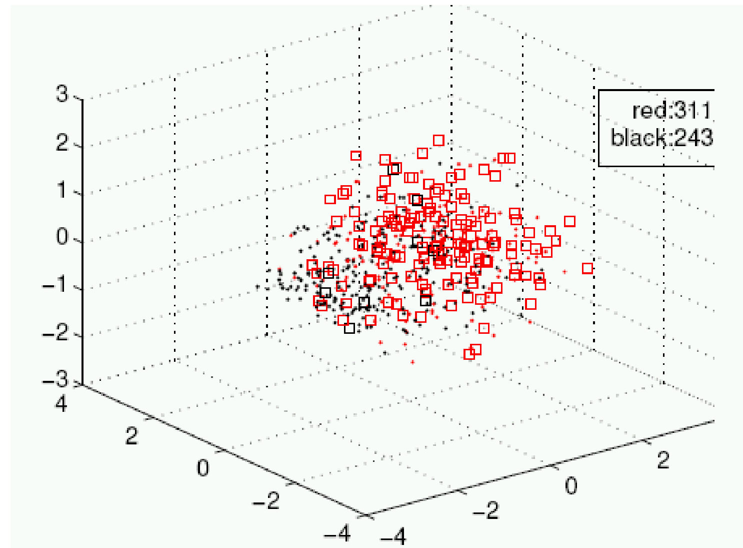
- Visualise the way regions are distributed among clusters
- Use shape-code (squares) to indicate the regions of interest and color-code to indicate a cluster's rank (largest cluster: red)
- Ideally all squares should be painted red and all dots should be painted differently

Goal:

- Validate our theoretical claim that the most populated cluster contains the majority of regions depicting the object of interest

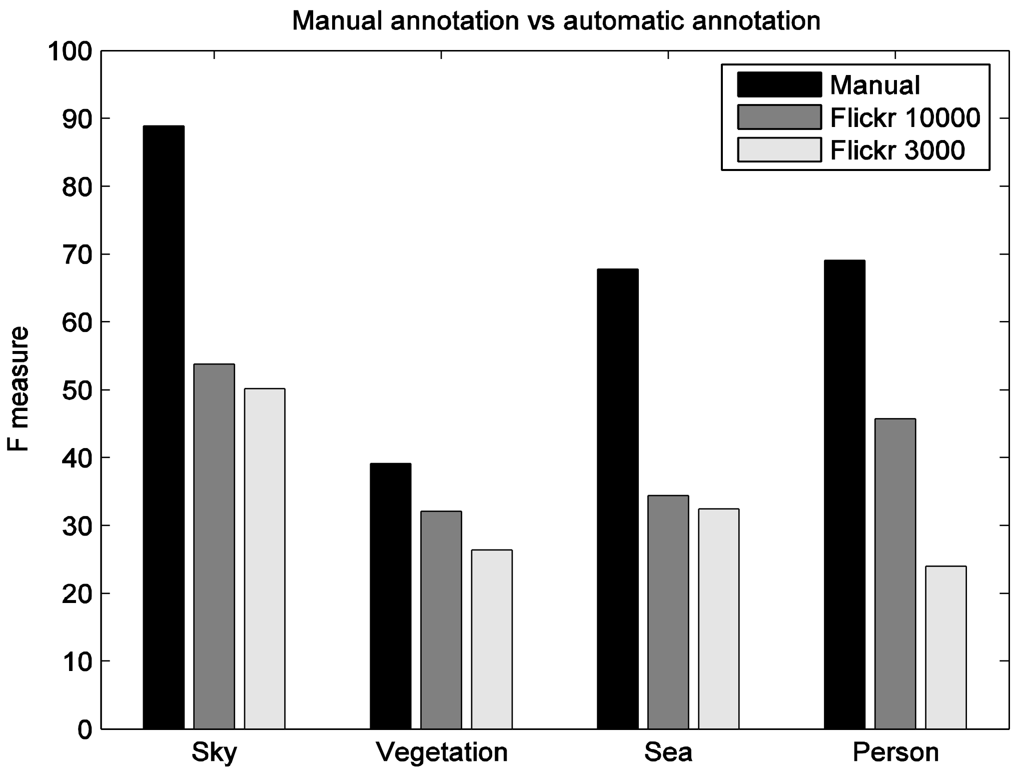
Conclusions:

- Our claim is valid in 5 (i.e., sky, sea, person, vegetation, rock) and not valid in 2 (i.e., boat, sand) cases



Vegetation in magnification

Experimental Results - Man. vs Autom. trained object detectors



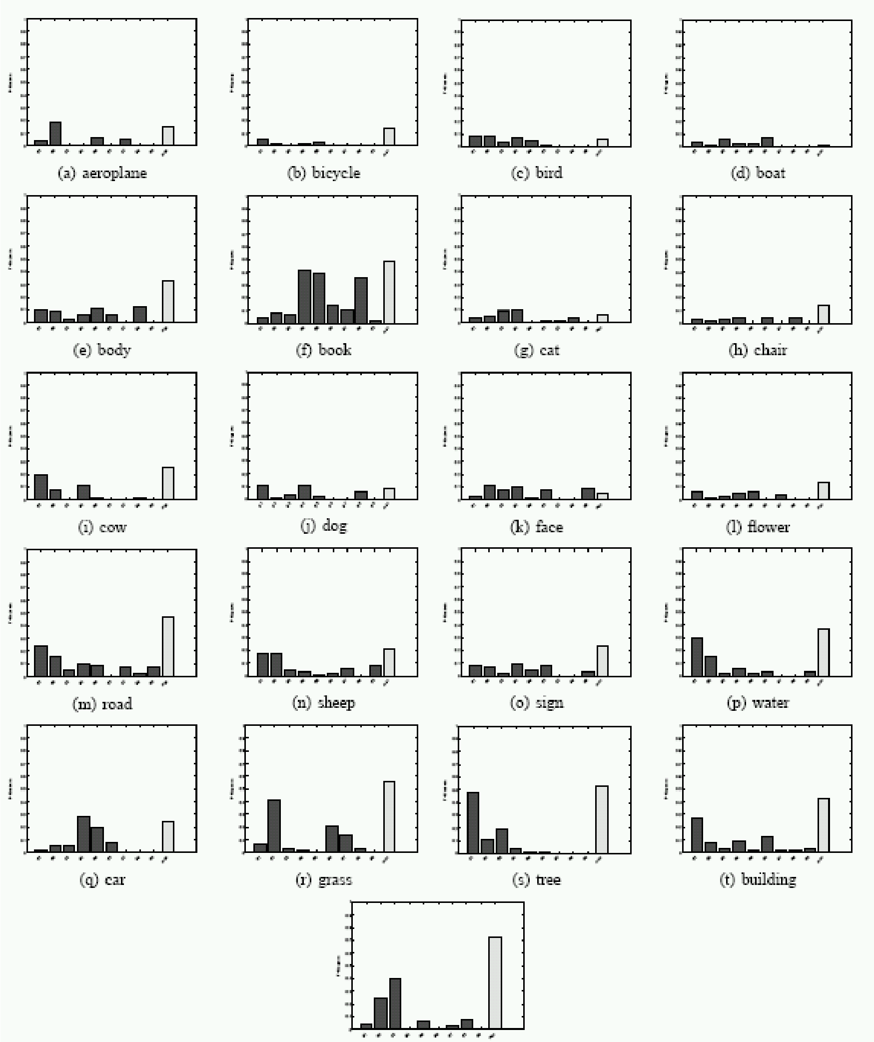
<u>Observations:</u>

- Performance lower than manually trained detectors
- Consistent performance improvement as the dataset size increases

Experimental Results – MSRC Dataset (21 objects)

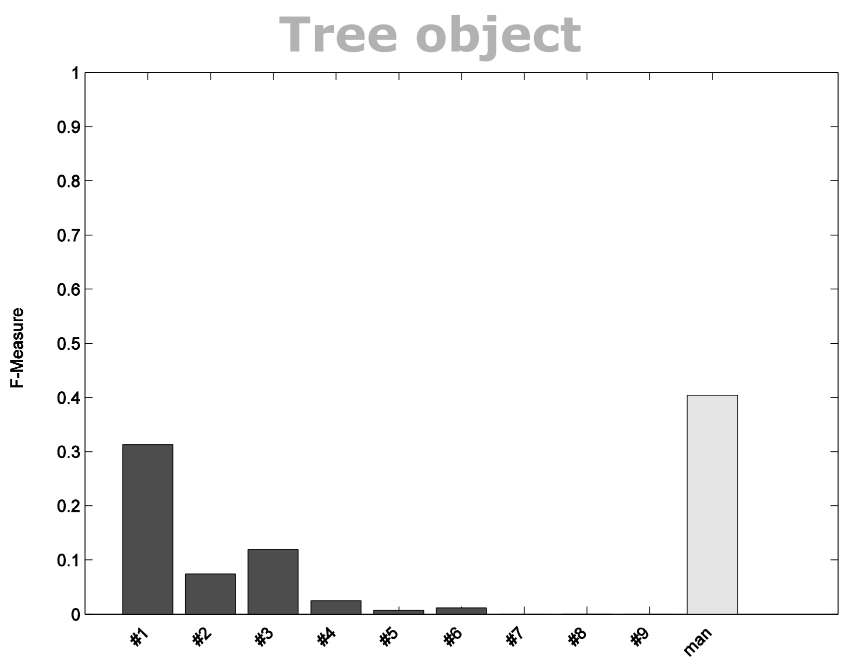
Observations:

- In 5 cases the objects were too diversiform to be described by the employed feature space (not even the manual annotations performed well)
- In 5 cases the annotation we got from Flickr groups were not appropriate
- In 6 cases, our method has failed to select the appropriate cluster
- In 5 cases our method worked well



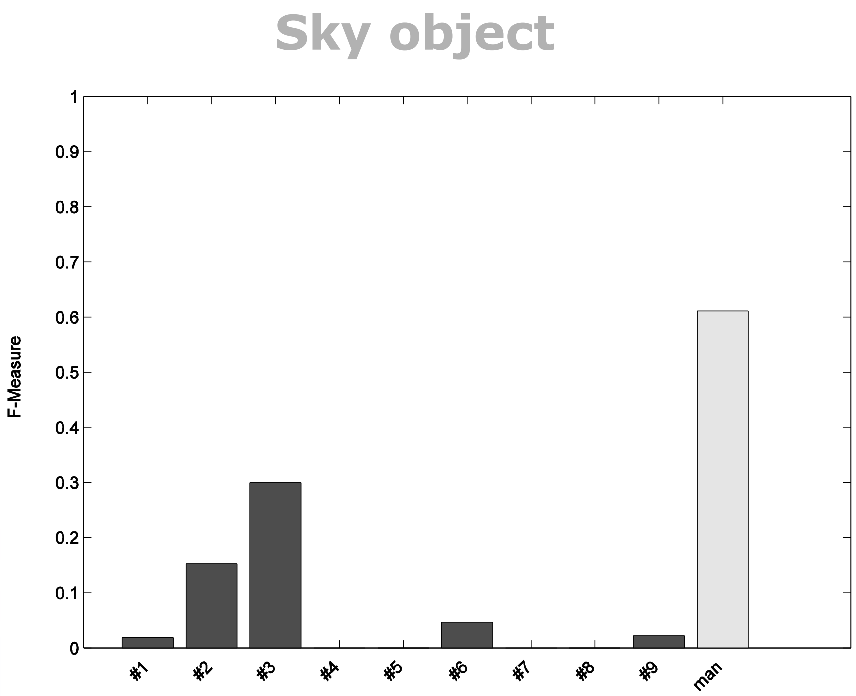
(a) aeroplane (b) bicycle (c) bird (d) boat (e) body (f) book (g) cat (h) chair (i) cow (j) dog (k) face (l) flower (m) road (n) sheep (o) sign (p) water (q) car (r) grass (s) tree (t) building (u) sky

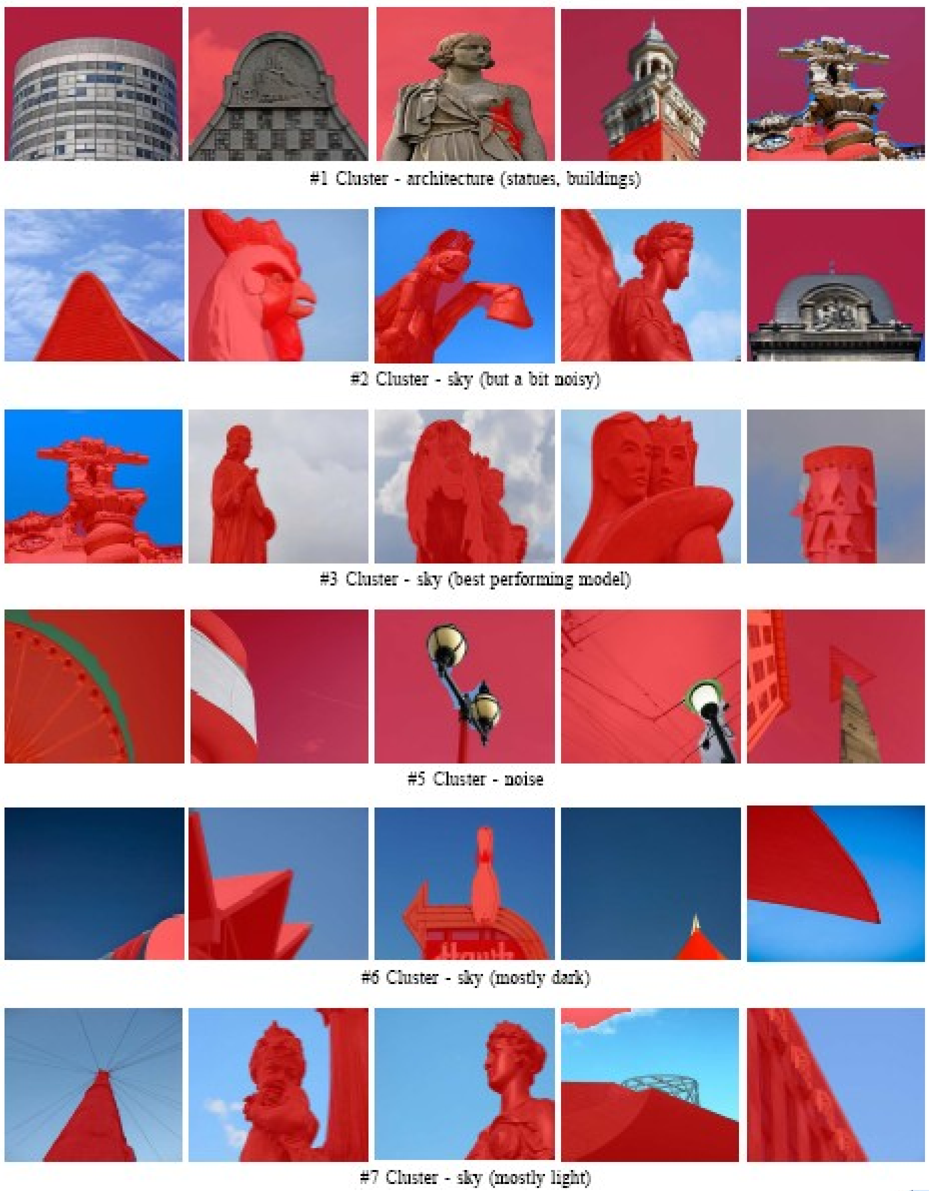
Experimental Results - MSRC vs Flickr groups
Target object: Tree



Good example: Semantic objects are correctly assigned to clusters and the most-populated cluster corresponds to the target object)

Experimental Results - MSRC vs Flickr groups Target Object: Sky



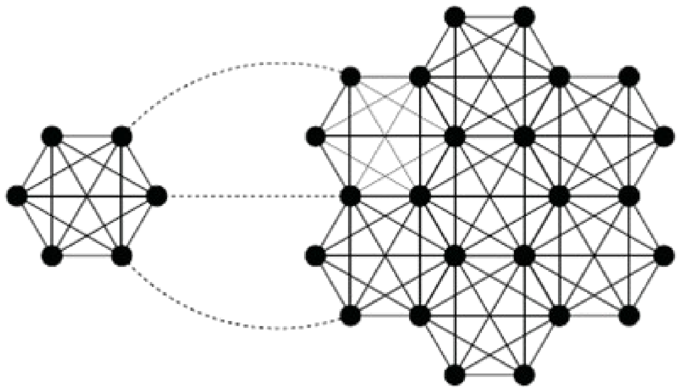


Bad example: Sky regions are split in many clusters and the most populated cluster contains noise regions

Community Detection

Community Detection in Complex Networks

- Community Detection: The Problem
- Global vs. Local Community Detection
- Bridge Bounding
- Conclusions - Future Work



communities context …

- typically … communities are defined with reference to some graph (network) which represents a set of entities / objects (nodes) and their relations (edges).

… even when there is no explicit graph, one can infer it, e.g.:

feature vectors → distances →threshold application → graph

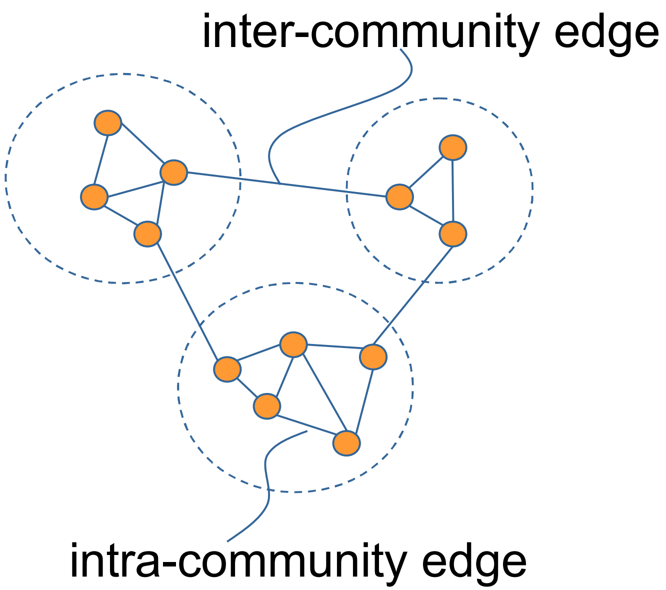
- Given a graph, a community is loosely defined as a set of nodes that are more densely connected to each other than to the rest of the graph vertices.

a simple example ...

- extremely profound community structure.

- key-concepts : within-community nodes, intra-community edges, inter-community edges.

- rarely appearing in real systems.



Definition of communities is heavily dependent on graph properties and subgraphs discovery

Global vs. Local

- **Global:** Process the whole graph to derive a partition into communities
  - \+ Abundant research
  - \+ Good results (community quality, algorithm efficiency)
  - Not practical for huge graphs or for real-time applications
- **Local:** Incremental process of the graph and output communities (streaming)
  - Relatively little research
  - Great potential for demanding applications

Bridge Bounding

Algorithm

- Start a community with a seed node
- Add neighbouring nodes as long as they are connected by edges that are not inter-community (“bridges”).
- Stop when it is not possible to add any more nodes.

Algorithm 1 LocalCommunityDetection

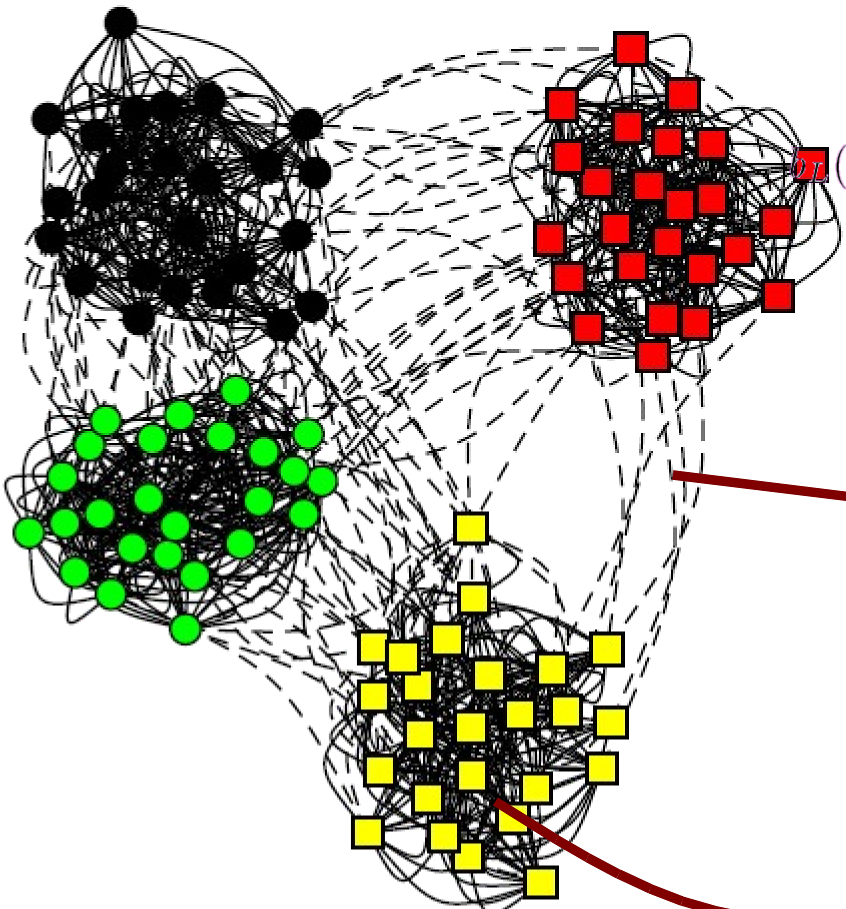
Require: Seed node $s \in G = (V, E)$
Require: Community mapping $g_C : V \rightarrow \mathbf{P}$
Require: Bridge function $b : E \rightarrow [0.0, 1.0]$

$C_s = \oslash$
Frontier set $F = \{s\}$
while $|F| > 0$ **do** $\{F$ is non-empty$\}$
  $c \leftarrow F.\text{pop}()$
  $C_s \leftarrow C_s \bigcup \{c\}$
  $C_U \leftarrow C_U \setminus \{c\}$
  for all $n \in N(c)$ such that $e_{cn} = (c, n) \in E$ **do**
    if $g_C(n) = C_U$ and $b(e_{cn}) \leq B_L$ **then**
      $F.\text{push}(n)$
    end if
  end for
end while
$\mathbf{P} \leftarrow \mathbf{P} \bigcup C_s$

Basic success factor:

Edge Bridge-ness: The property of an edge to lie between two communities.

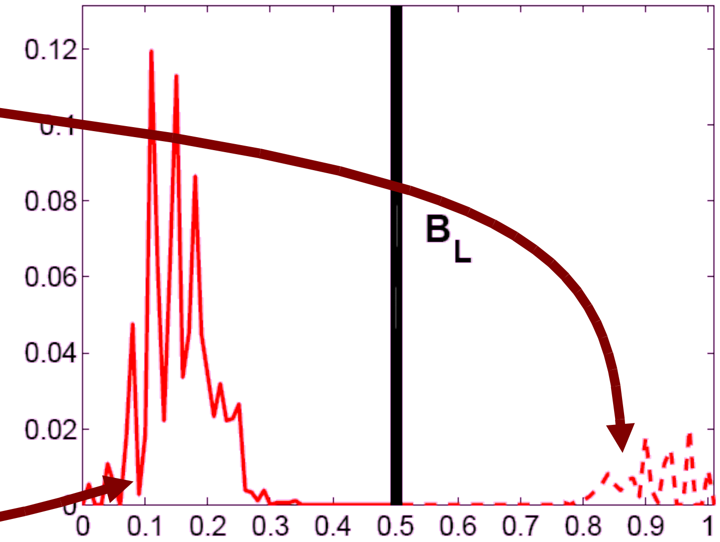
Bridge Bounding – Toy Example



Local bridging of an edge

$$b_L(e_{st}) = 1 - C^{(3)}_{st} = 1 - \frac{|N(s) \bigcap N(t)|}{min[(d(s)-1),(d(t)-1)]}$$

s, t: endpoints of edge
N(s), N(t): neighbourhoods of s, t
d(s), d(t): degrees of s, t



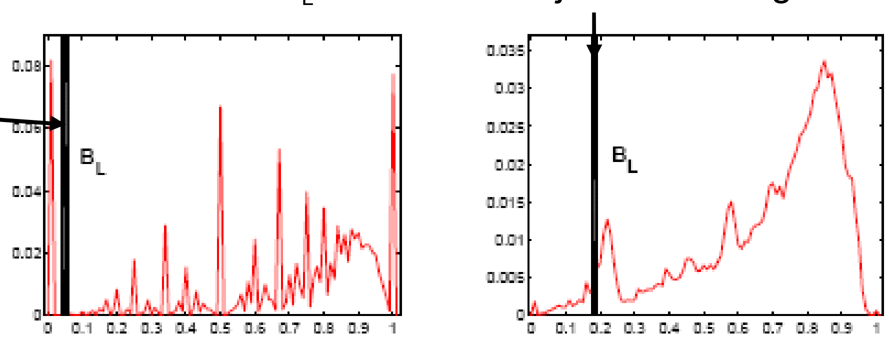
Bridge Bounding - Problems

- Local bridging not suitable for scale-free networks
- Solution (partial) 2nd order local bridging.

$$b'_L(e_{st}) = \alpha \cdot b_L(e_{st}) + (1-\alpha)\frac{1}{|N(e_{st})|}\sum_{e \in N(e_{st})} b_L(e)$$

B_L = 0.17 leaves just 1% of edges as non-bridges.

B_L as low as 0.05 leaves 8% of edges as non-bridges.

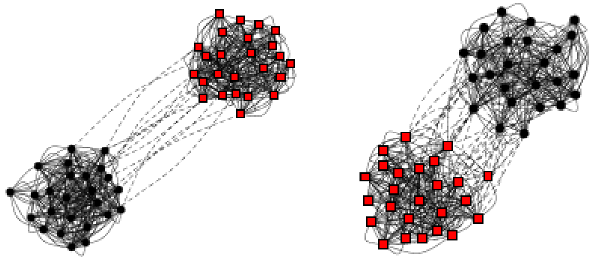


(a) b_L distribution (b) b'_L distribution, $\alpha = 0.7$

Experiments on Synthetic Community Networks

- Synthetic networks according to method of Newman and Girvan.

$$S_{PAR} = \{N, K, z_{tot}, p_{out}, s_{var}\}$$



(a) $p_{out} = 0.01$ (b) $p_{out} = 0.08$

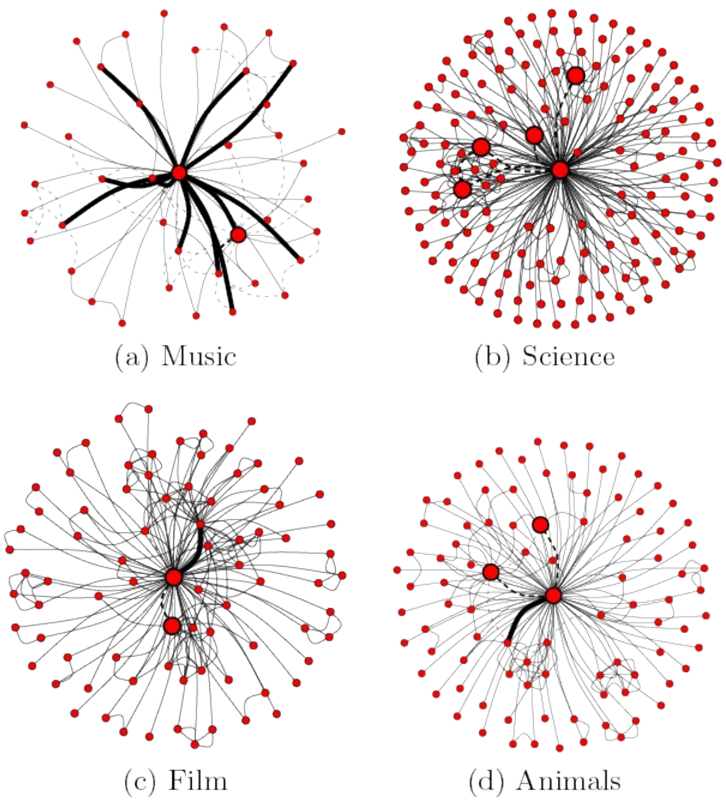
Change complexity of underlying communities.

	F_C			NMI		
p_{out}	BB	BB'	GN	BB	BB'	GN
0.01	100	100	100	1.0	1.0	1.0
0.05	100	100	100	1.0	1.0	1.0
0.1	100	100	50	1.0	1.0	0.86
0.15	100	99	50	1.0	.98	0.86
0.20	99	74	50	0.98	0.84	0.86
0.25	24	24	0	0.54	0.56	0.02

Change relative sizes of underlying communities.

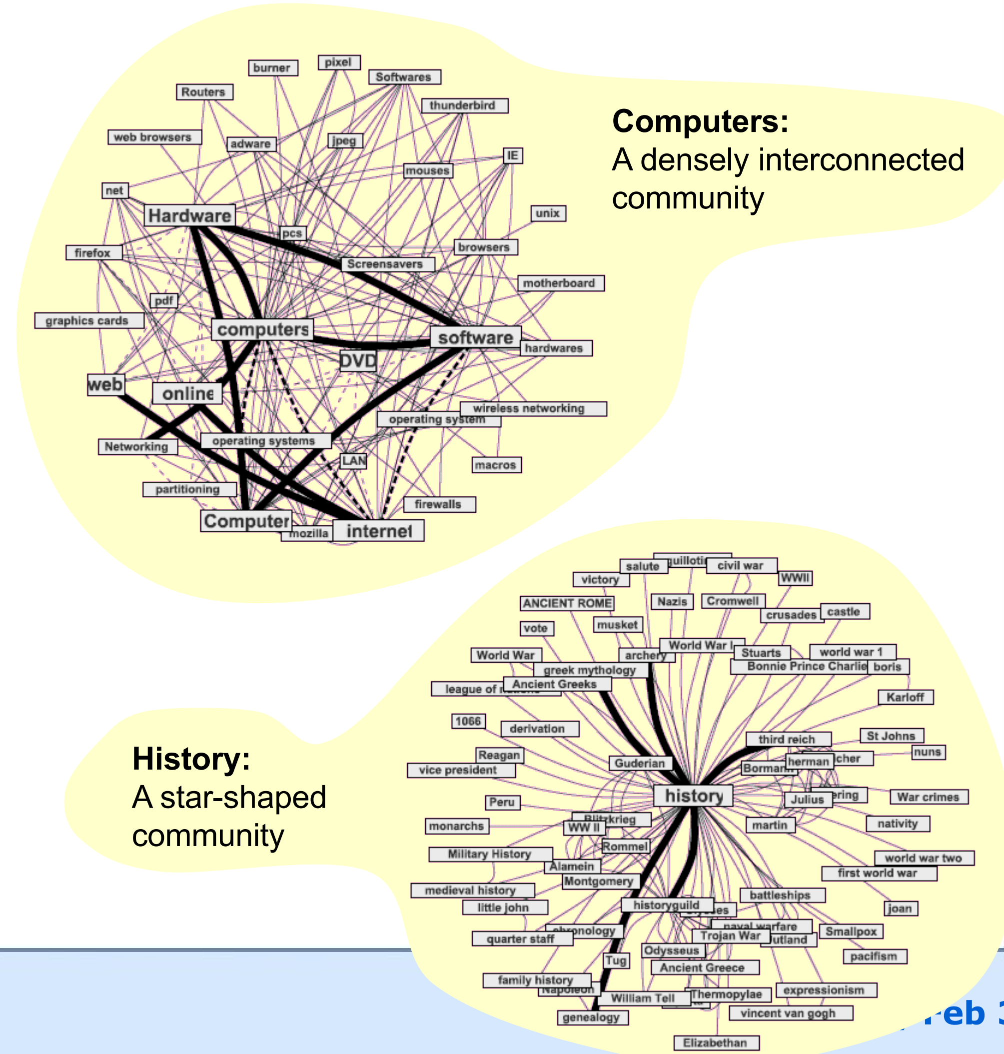
	F_C			NMI		
s_{var}	BB	BB'	GN	BB	BB'	GN
1.1	100	100	100	1.0	1.0	1.0
1.5	100	100	100	1.0	1.0	1.0
1.6	99.5	100	100	0.99	1.0	1.0
1.7	88	98	100	0.82	0.96	1.0
1.8	85.5	97	100	0.79	0.95	1.0
1.9	58.5	87	90	0.68	0.82	0.88
2.0	12.5	80	82	0.45	0.73	0.81
2.5	0	62	75	0.45	0.63	0.72

LYCOS iQ Tag Network



(a) Music (b) Science

(c) Film (d) Animals

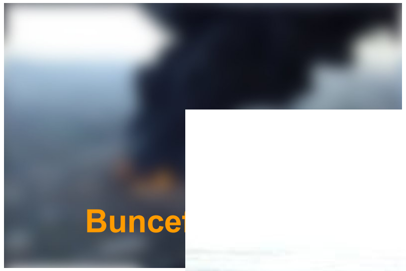


Future Work for Community Detection

- Investigate label propagation techniques.
  - Application on external memory graphs.
- Possibilities for incremental community detection.
  - Application on large dynamic networks (e.g. Social Tagging Systems)
- Applications on different domains:
  - Hybrid image clustering (use of both visual and tag features)
  - Domain-specific clustering, e.g. Points-Of-Interest in travel applications.

WeKnowIt and CI

Personal Intelligence

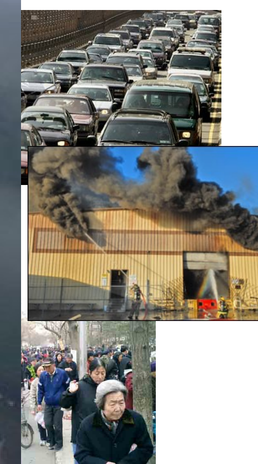


Profile of contributor

>> What to send where,
e.g. location, age

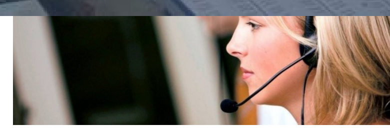
Media Intelligence

Organizati



Trust and feedback

>> Determine trustworthiness and hub-structures by SNA



Further Issues

- Not all data always available (e.g. User queries, fb)
- Long tail is forgotten (e.g. flu trends in 3rd world countries)
- “More data, less analysis”,....
- Applications and commercialization
- Efficiency of semantics and analysis
- Real integration
  - not just sum of different analysis
  - formal framework and approach
  - representation
- User interaction – Interfaces

Thank you!


sky
tower
vegetation
ground
road
</rdf:Description>
</rdf:Description>
</rdf:Description>
<rdf:Description rdf:about="http://www.iti.gr/ontologies/INSTANCES#Sky_0">

References

[Chatzilari09] Elisavet Chatzilari, Spiros Nikolopoulos, Eirini Giannakidou, Athena Vakali and Ioannis Kompatsiaris, "Leveraging Social Media For Training Object Detectors", 16th International Conference on Digital Signal Processing (DSP'09), Special Session on Social Media, 5-7 July 2009, Santorini, Greece.

[Fortunato07a] Santo Fortunato and C. Castellano, “Community structure in graphs”, arXiv:0712.2716v1, Dec 2007.

[Freeman77] L. C. Freeman : A set of measures for centrality . Resolution limit in community detection. PNAS, 104(1). pp. 36-41

[Giannakidou08] E. Giannakidou, I. Kompatsiaris, A. Vakali, "SEMSOC: SEMantic, SOcial and Content-based Clustering in Multimedia Collaborative Tagging Systems", In Proc. 2nd IEEE International Conference on Semantic Computing (ICSC' 2008), IEEE Computer Society, August 4-7, 2008 Santa Clara, CA, USA

[Kennedy07] Lyndon S. Kennedy, Mor Naaman, Shane Ahern, Rahul Nair, Tye Rattenbury: How flickr helps us make sense of the world: context and content in community-contributed media collections. ACM Multimedia 2007: 63

[Kumar99] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. Trawling the web for emerging cyber-communities. Computer Networks, 31(11-16):1481-1493, 1999.

[Lindstaedt09] S. Lindstaedt, R. Mörzinger, R. Sorschag, V. Pammer, G. Thallinger. Multimed Tools Appl (2009) 42:97–113, DOI 10.1007/s11042-008-0247-7

[Quack08] Till Quack, Bastian Leibe, Luc Van Gool. World-scale mining of objects and events from community photo collections, In Proceedings of the 2008 international conference on Content-based image and video retrieval, Jul-08

[Zhang06] Y. Zhang, J. Xu Yu, J. Hou : Web Communities: Analysis and Construction, Springer 2006.

Telematics and Informatics

[Crespoa09]Angel García-Crespoa, Javier Chamizoa, Ismael Riverab, Myriam Menckea, Ricardo Colomo-Palaciosa and Juan Miguel Gómez-Berbísa, “SPETA: Social pervasive e-Tourism advisor”, Volume 26, Issue 3, August 2009, Pages 306-315, Mobile and wireless communications: Technologies, applications, business models and diffusion.