Picture Tags and World Knowledge

learning tag relations from visual semantic sources

Lexing Xie, Xuming He

Australian National University
NICTA

Oct 2013 @ ACM MM

An image is worth 1000 words …

Which 1000? How are they related?

#barcelona #tibidabo #montjuïc #catalunya #catalonia #hdr #nikon #nikond90 #d90 #18200mmf3556gvr #clouds #nuvols #cityscape #people #gent #view #sight

http://www.flickr.com/photos/bcnbits/3663297224/

Two related goals

- Which 1000?

  Why we tag [Ames, Naaman, 2007]

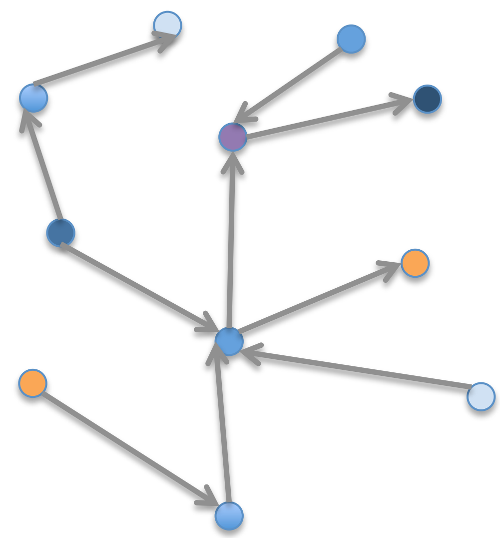
  Events and places from Flickr tags

  [Rattenbury, Good, Naaman, 2007]

  ClassTag: Tag→ Wikipedia → Wordnet

  [Overell, Sigurbjornsson, van Zwol 2009]

  Can all tags be used for search? [Bischoff et al 2008]



- How are they related?

  Knowledge graph 1.0:
  [Cyc 1984-] [WordNet mid-1980s- now] [ConceptNet 2003-now]
  Data-driven knowledge graphs:
  NELL @CMU1010-, Google Knowledge Graph, Yahoo! Clickstream Graph
  Vispedia, Perona 2010-

Plan for this talk

- A vision for multimedia knowledge graph

- A novel connection across three data sources
- Quantifying visual tags from co-ocurrence
- A network inference model on tag relationships

- Evaluation of relation learning and tagging

How to find 1000s of tags?

flickr

Millions of photos with tags

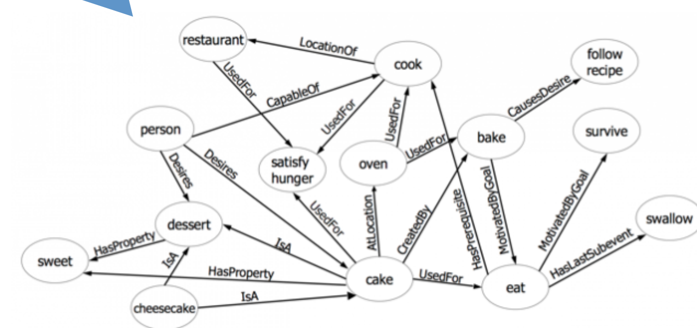
What can help quantify the "visual"-ness of tags?

How are tags related?

IMAGENET

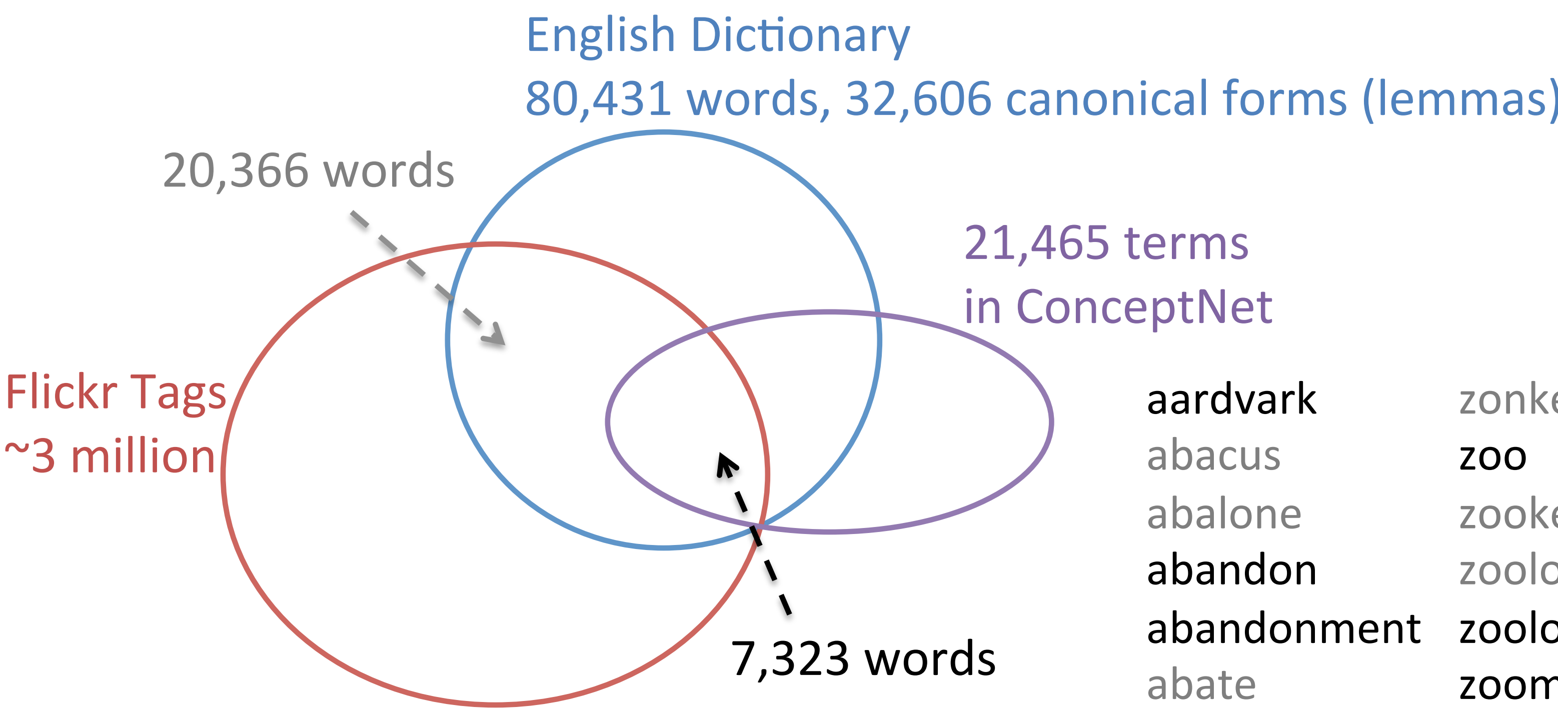
14 M images, 21K illustrated synsets
5.1M with Flickr tags, 13,288 synsets

ConceptNet



Crowd-sourced commonse knowledge
22K Concepts, 450K relations

Mapping tags to words

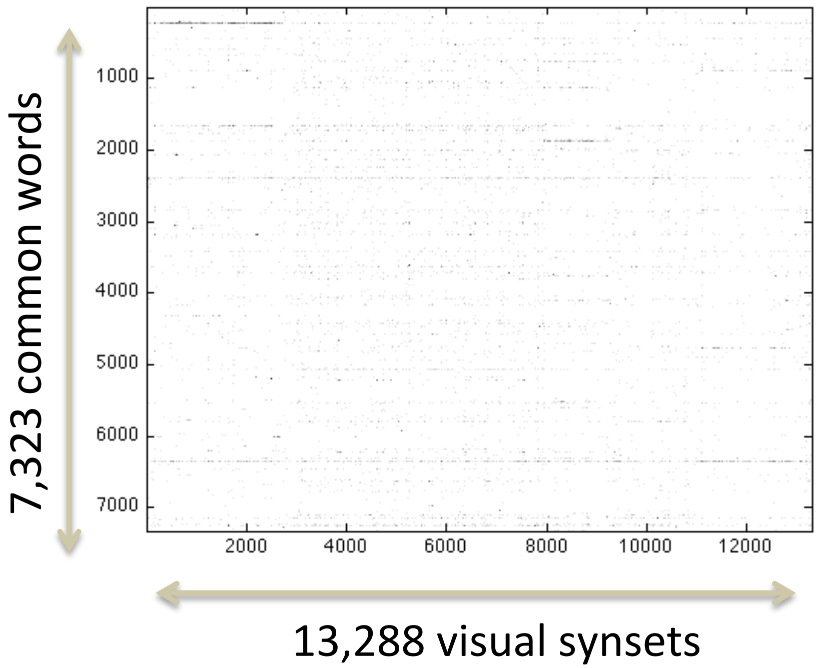


aardvark	zonked
abacus	zoo
abalone	zookeeper
abandon	zoological
abandonment	zoology
abate	zoom
abatement	zucchini
abattoir	zwieback
abbe	zydeco
abbess	zymurgy

…

Observation: Visual Synsets vs Tags

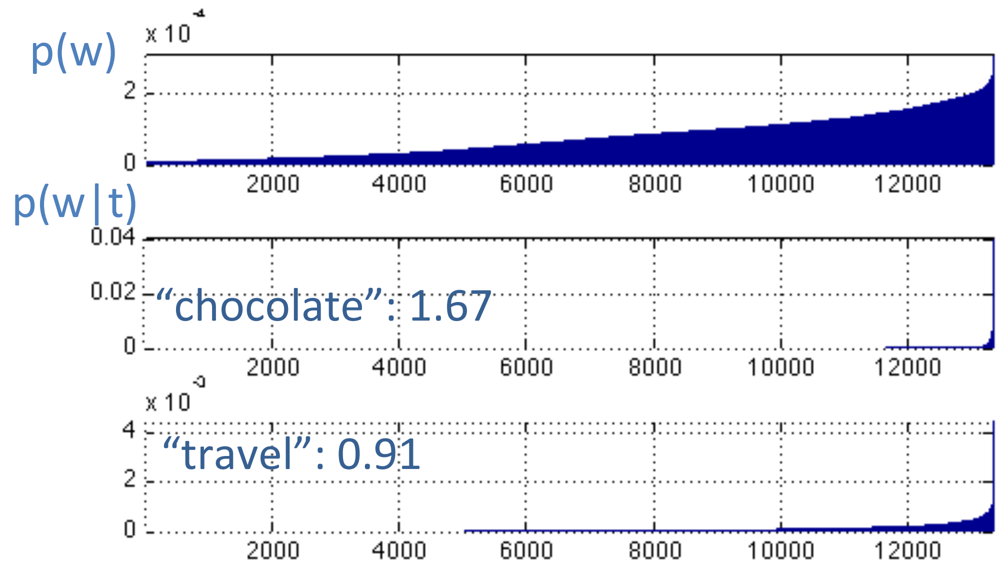
Y log(coocurrence) over 5.1M images



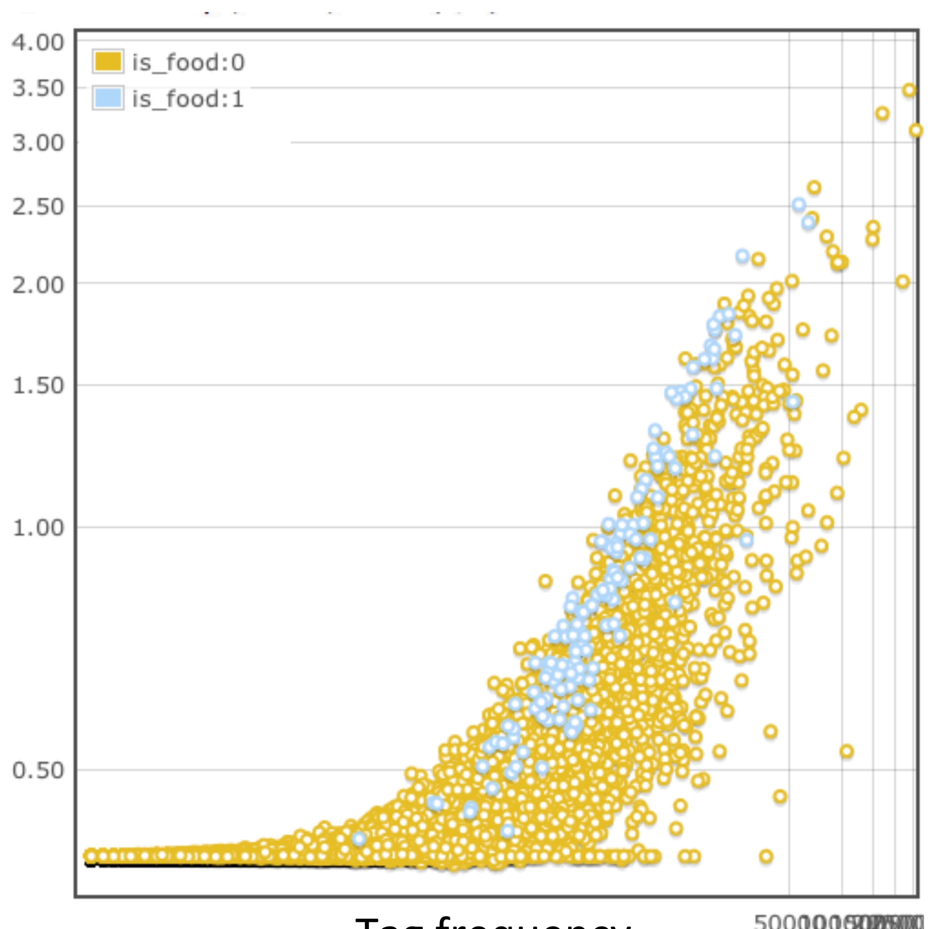
Which tags are visually descriptive?

$$KL(p(w)||p(w|t_j))$$

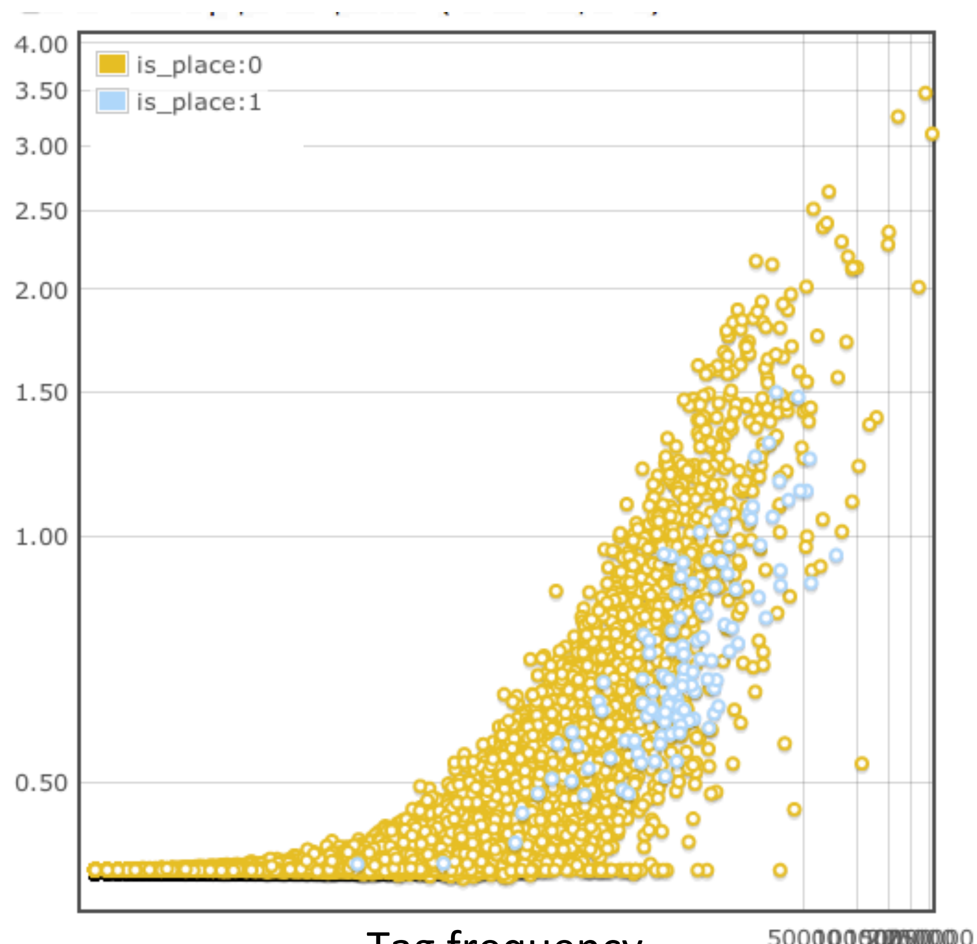
$$= \Sigma_i p_i(w) log \left(\frac{p_i(w)}{p_i(w|t_j)} \right)$$



Food vs Places

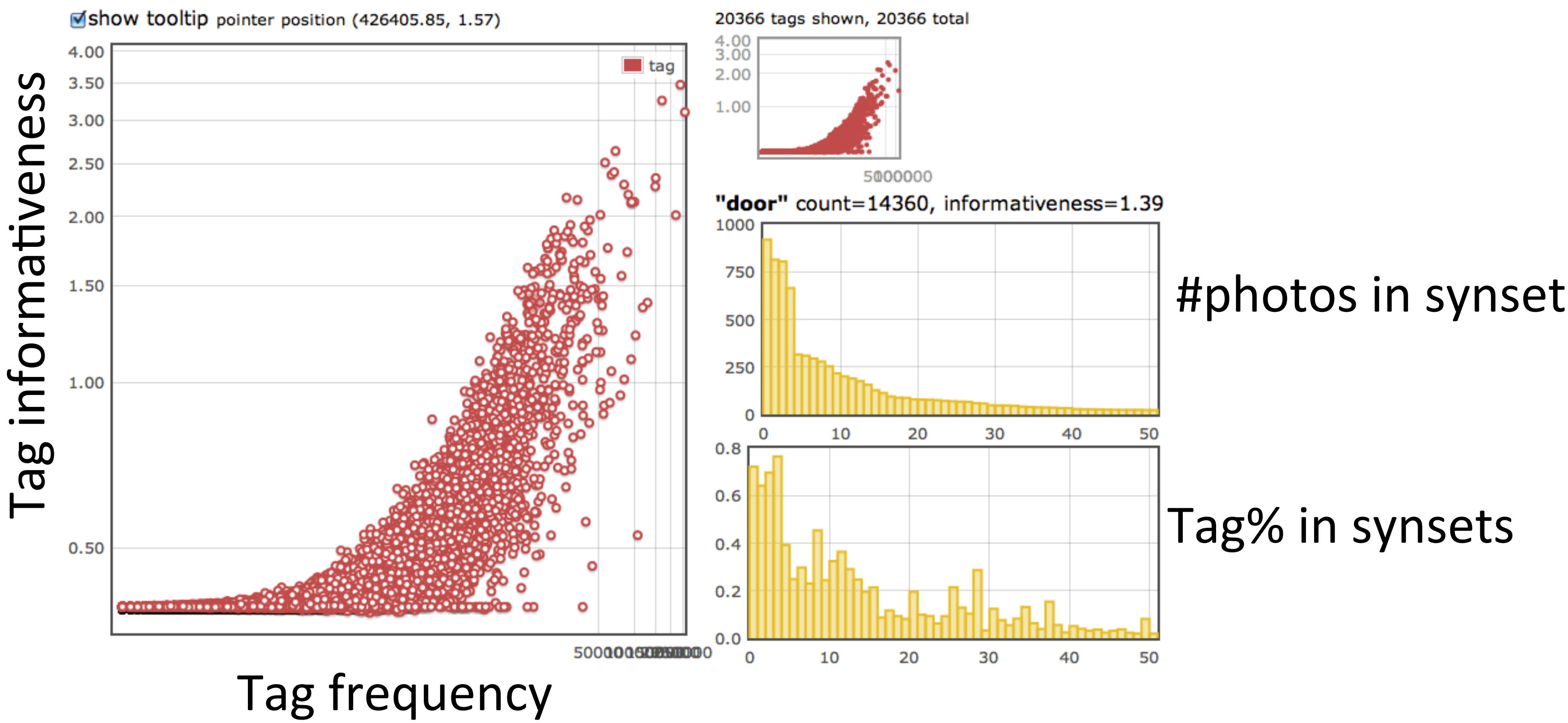


Tag frequency



Tag frequency

An interactive visualization of tag space

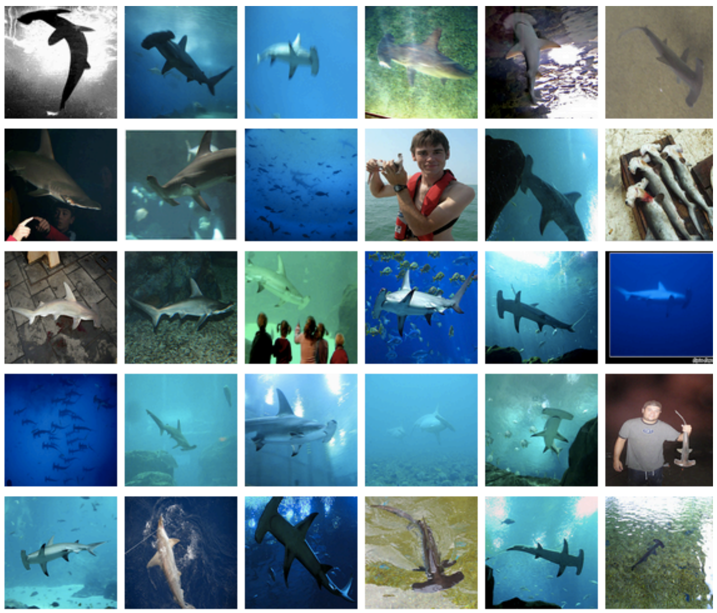


http://cecs.anu.edu.au/~xlx/proj/tag-explore

Talk Roadmap

- Goals for MM knowledge graph
- A novel connection across three data sources
- Quantifying visual tags from co-ocurrence
- Learning tag relationships
  - The Inverse Concept Rank Model
  - How good are the learned relations?
- Image tagging evaluations

Which tags are related?



n01494475
hammerhead.n.03

aquarium, shark
aquarium, hammerhead
aquarium, georgia
aquarium, fish
aquarium, atlanta
atlanta, shark
atlanta, georgia
atlanta, hammerhead
shark, underwater
fish, water
…

A random walk model for tagging

green

apple

tomato

market

red

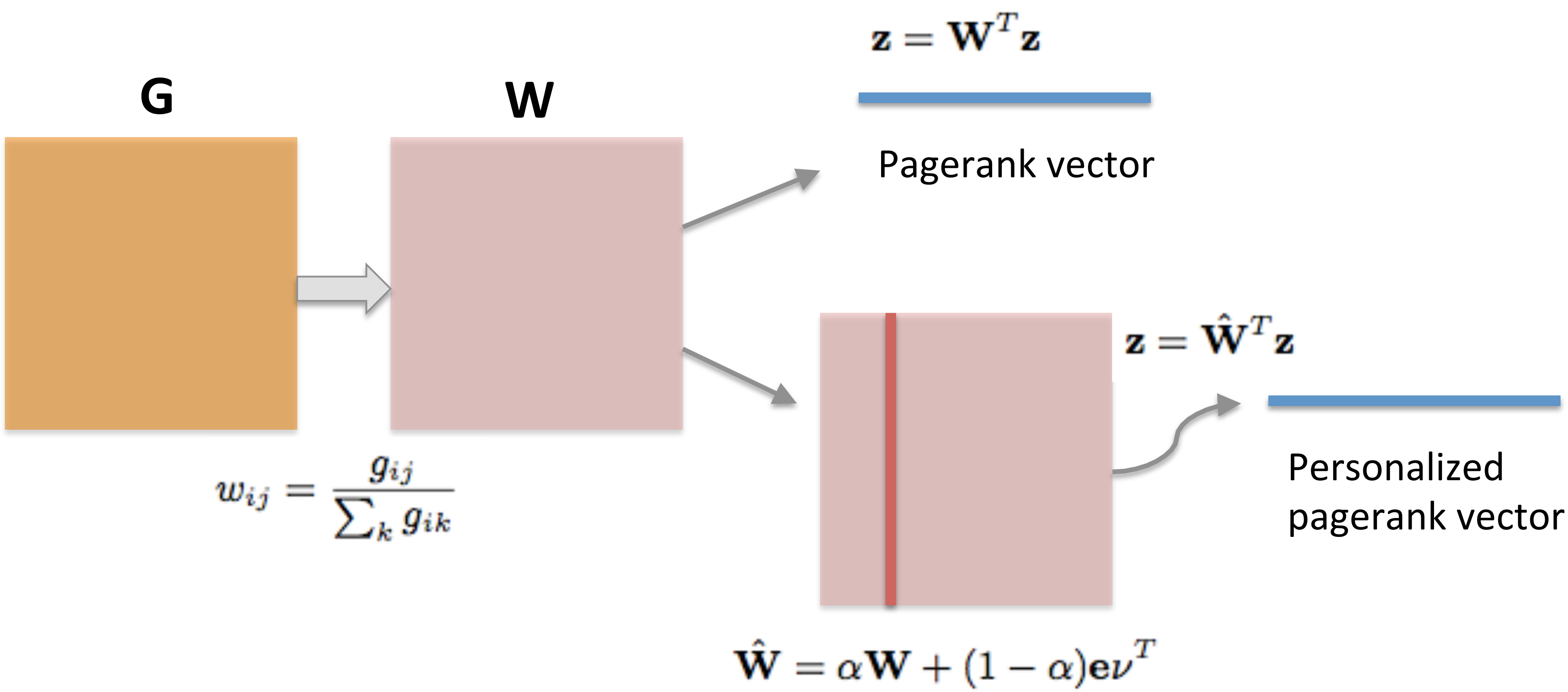
leaves

basket

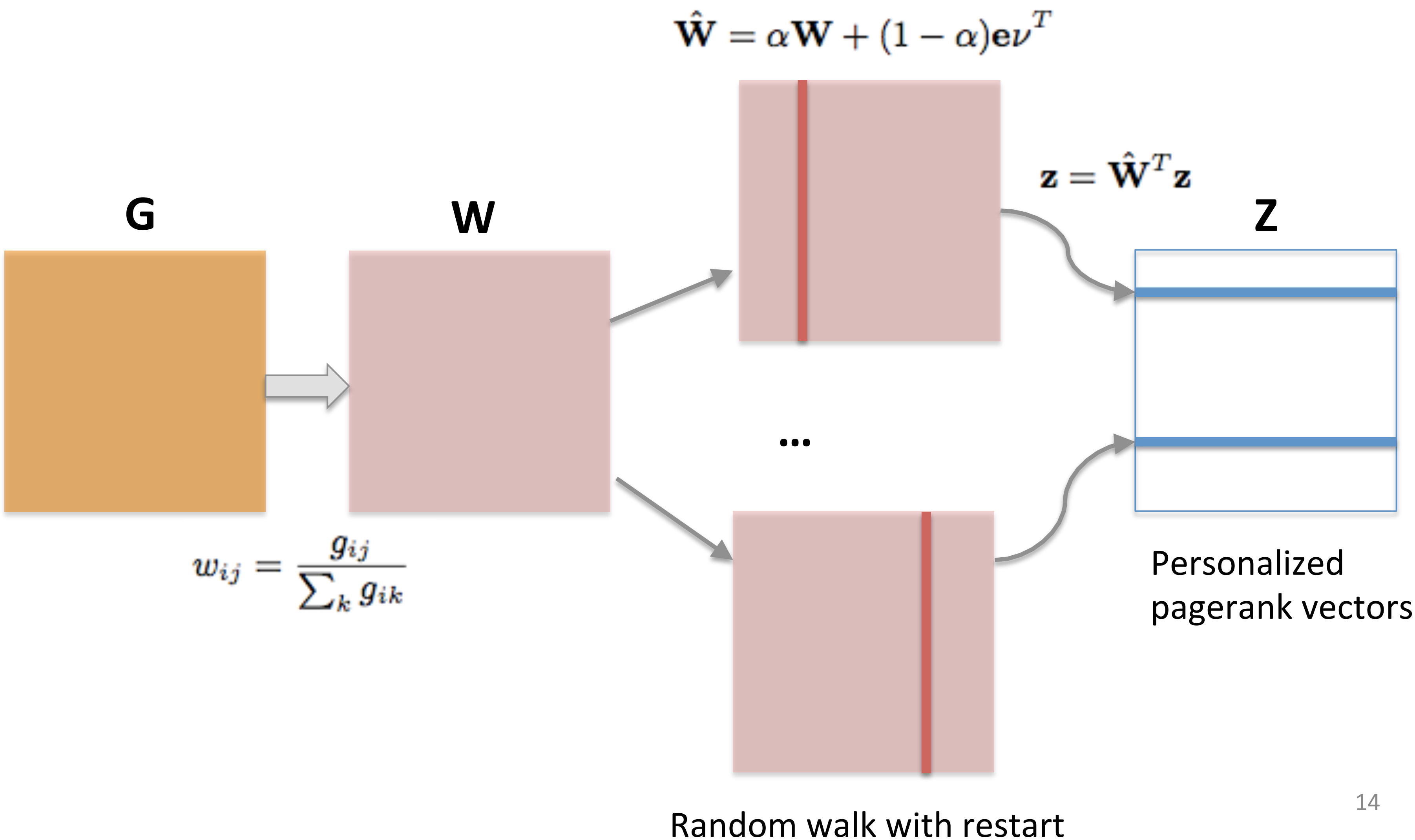
[Griffiths et al. “Google and the mind”, 2004]
[Abbott, Austerweil and Griffiths, NIPS2012]

A model for a *forward* random walk

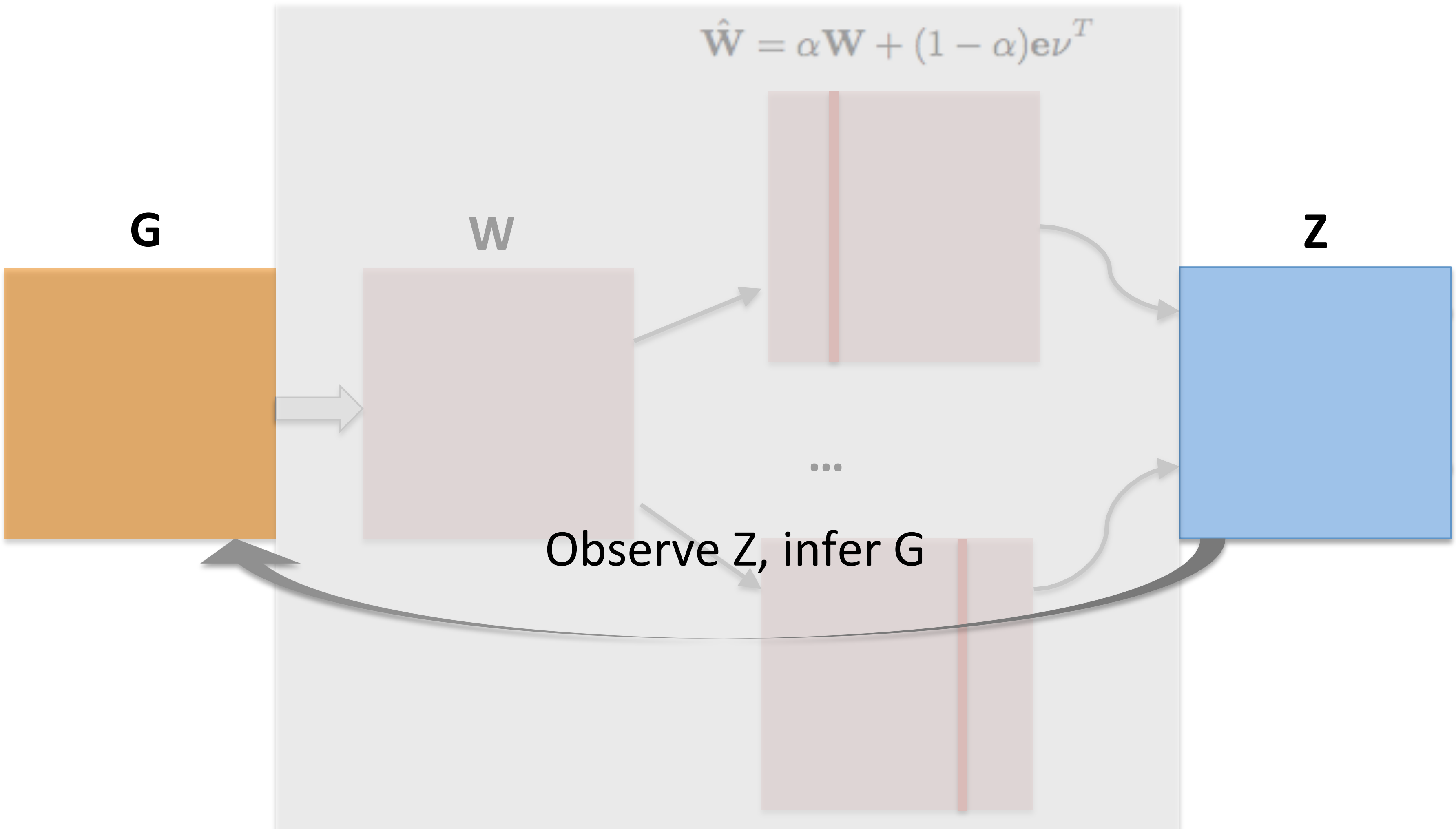
a.k.a Google v1.0



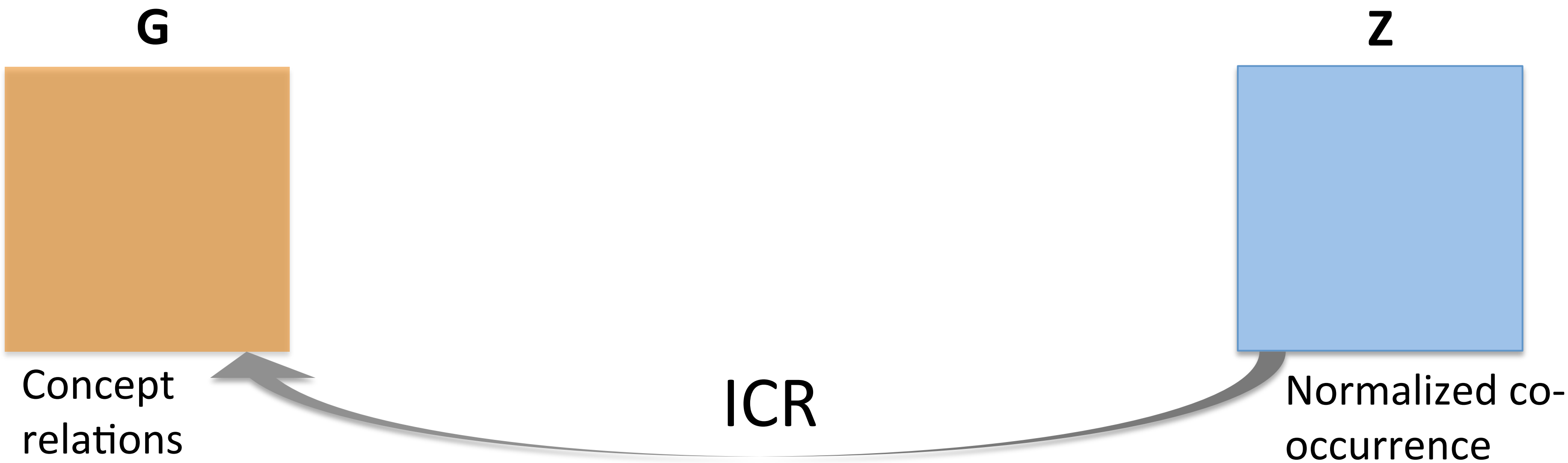
PageRank for photo tags



Network Inference for Tag Relations



The Inverse Concept Rank Problem



$$\min_{\mathbf{G}} J_R = \frac{\alpha_R}{2} tr(\mathbf{G}^T \mathbf{G}) + \sum_{i,j} h(z_{ij} - \bar{z}_{ij})$$

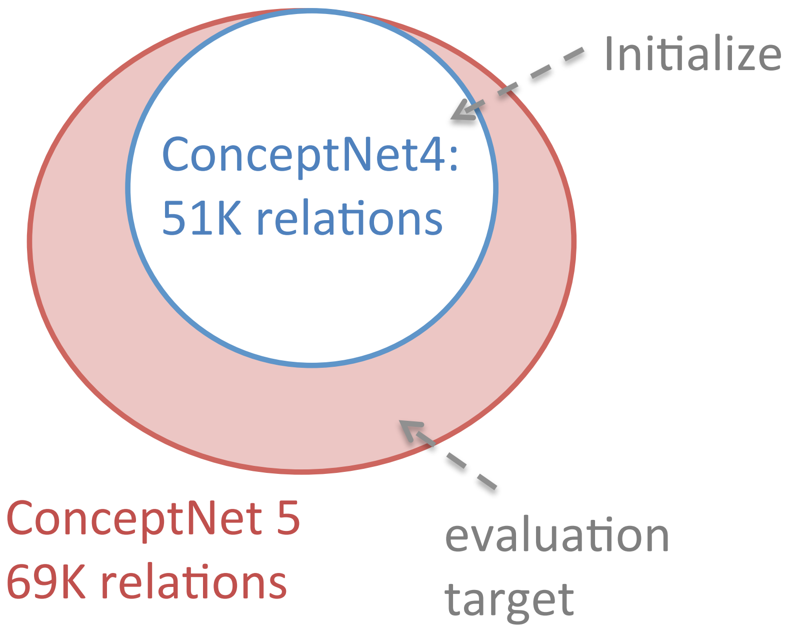
$$s.t. \quad \mathbf{G} \geq \mathbf{G}_0$$

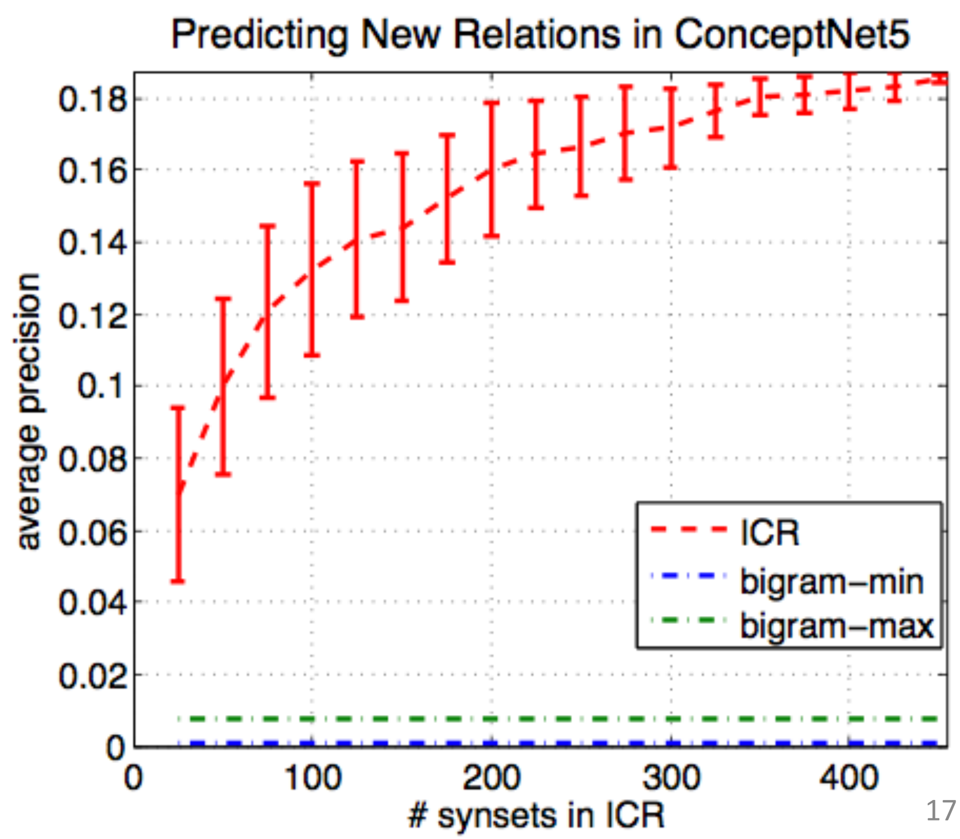
Prior knowledge

Regularizer: prefer sparse G

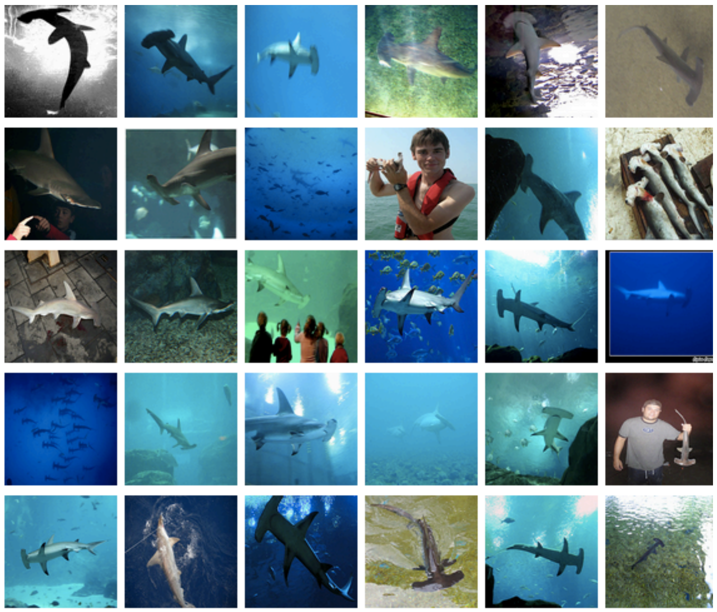
Loss function between model and observation

Evaluation: Learning New Relations





Which tags are related?

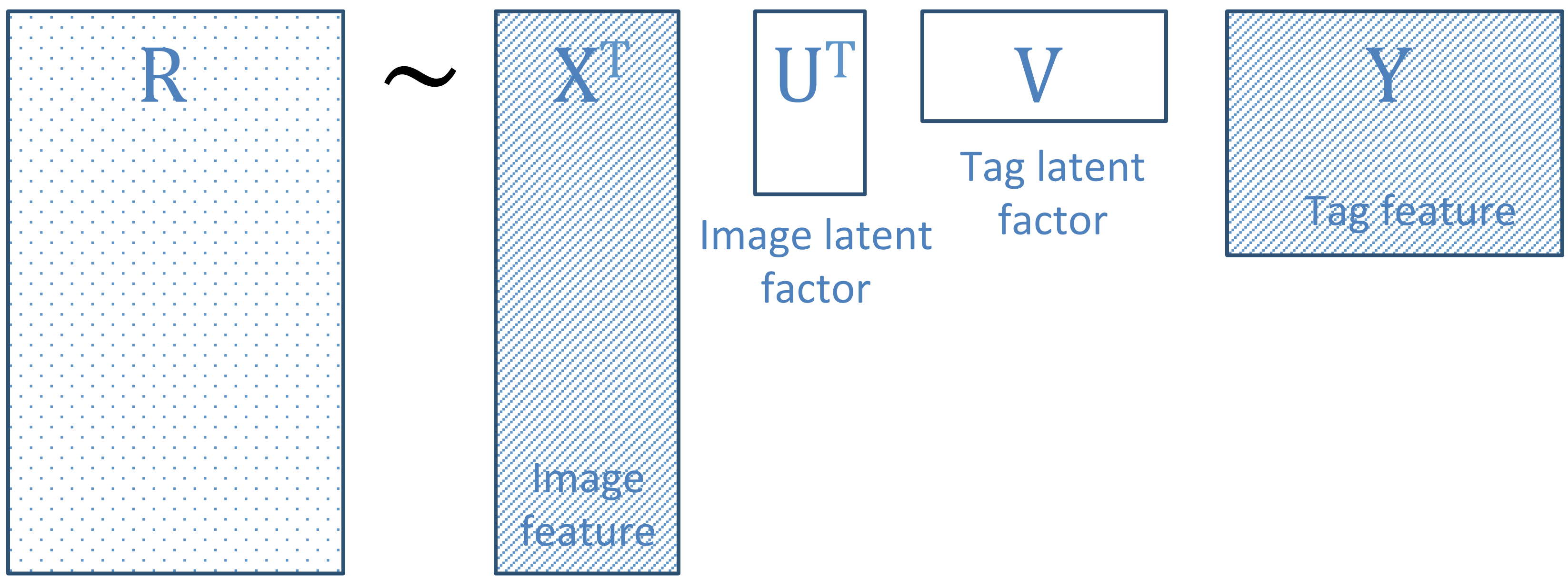


n01494475
hammerhead.n.03

aquarium, shark
aquarium, hammerhead
aquarium, georgia
aquarium, fish
aquarium, atlanta
atlanta, shark
atlanta, georgia
atlanta, hammerhead
shark, underwater
fish, water
…

Already in ConceptNet
Learned by ICR
Supressed by ICR

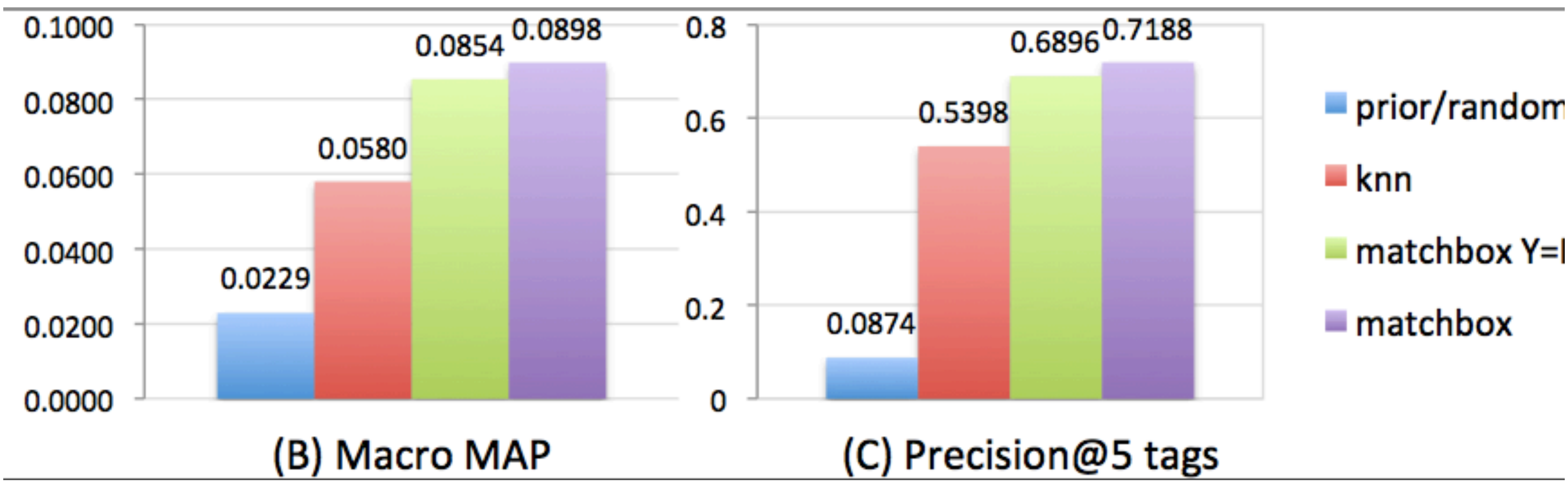
“MatchBox” model for picture tagging



- R partially known: Matrix completion
- R unknown : low-rank linear classifiers
- Y = I , low-rank linear classifiers

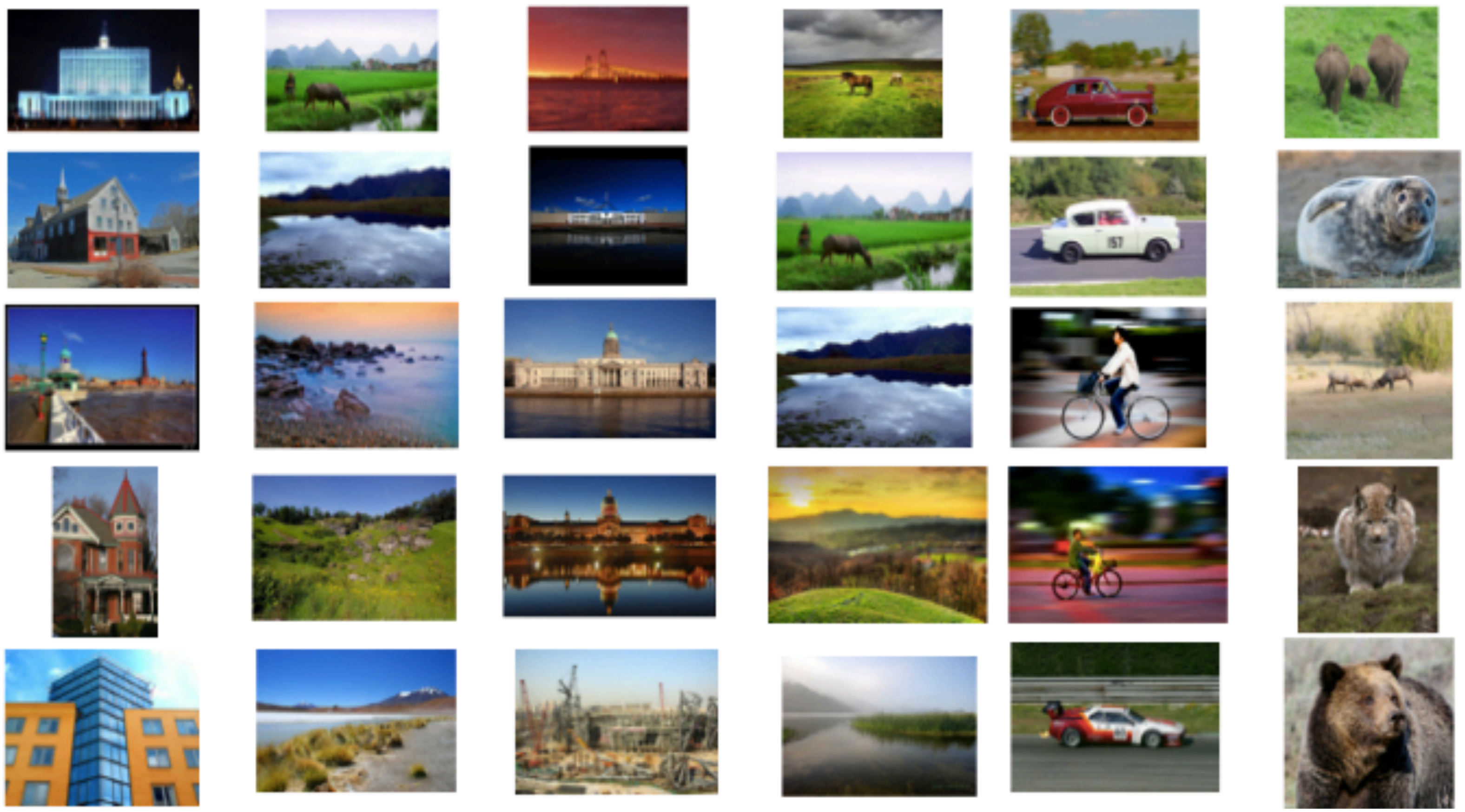
Picture tagging evaluation

- NUS-Wide Flickr dataset
  - 279K total, 63/81 tags found in ImageNet
- 177-D objectbank features
- Testing:
  - 1K images with >= 5 tags
  - fullset ~100K+ images, ~1.2 tags each

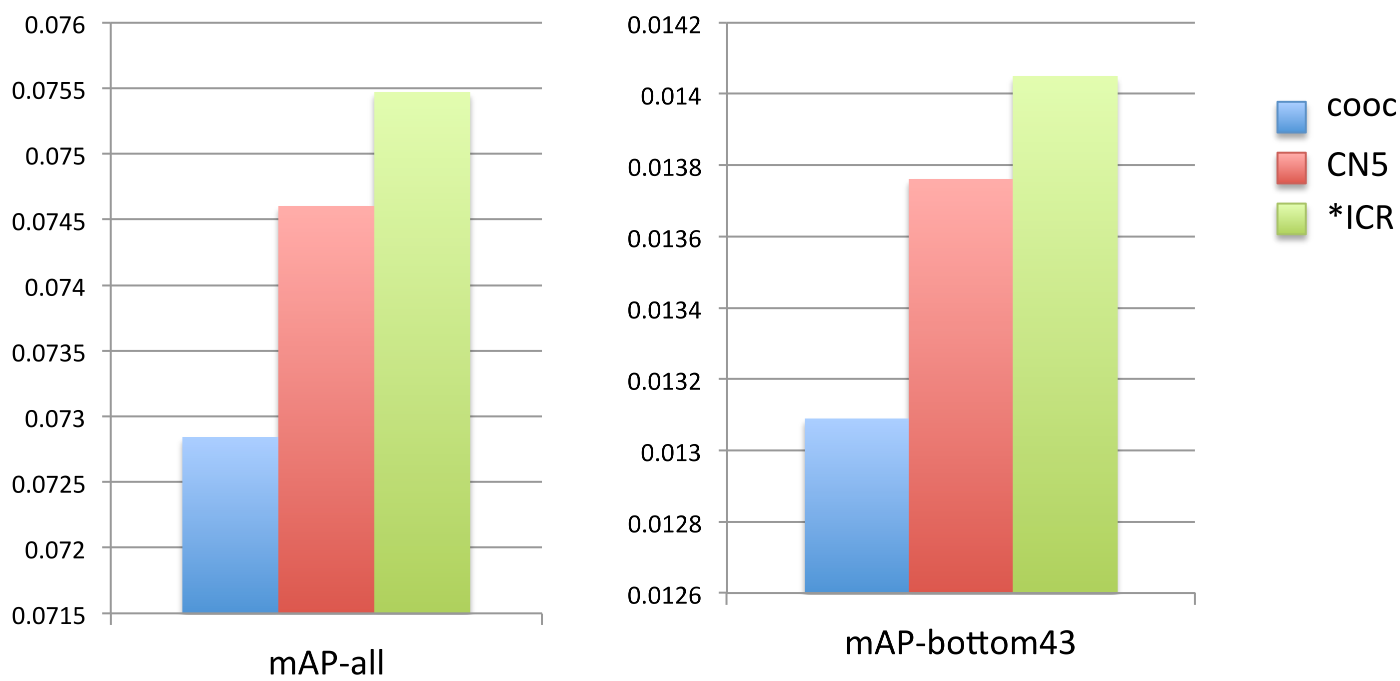


Top returns for a few 'unknown' tags

architecture | blue | city | landscape | travel | wildlife



Using ICR in picture tagging



Summary

- Statistical analysis reveals photo tag properties
- Inverse random walk algorithm for learning relations
- http://cecs.anu.edu.au/~xlx/proj/tagnet

- Also ask me about:
  - Multimedia-Hard problems [w. Ayman + Snoek]
  - Social Affinity Filtering: Rich Features > Models [COSN'13] https://code.google.com/p/social-recommendation/
  - CSS2013: Computational Social Science, Foundations and Frontiers http://cecs.anu.edu.au/~xlx/teaching/css2013/
  - Analyzing Images in Microblogs

Get in touch:
lexing.xie@anu.edu.au http://cecs.anu.edu.au/~xlx/