Leveraging Metropolis-Hastings Algorithm on Graph-based Model for Multimodal IR

Serwah Sabetghadam, Mihai Lupu, Andreas Rauber

Vienna University of Technology

Vienna, Austria

Abstract

The velocity of multimodal information shared on web has increased significantly.

Many reranking approaches try to improve the performance of multimodal retrieval, however not in the direction of true relevancy of a multimodal object.

Metropolis-Hastings (MH) is a method based on Monte Carlo Markov Chain (MCMC) for sampling from a probability distribution when traditional sampling methods such as transformation or inversion fail.

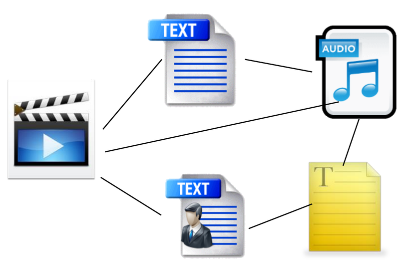
If we assume this probability distribution as true relevancy of documents for an information need, in this paper we explore how leveraging our model with Metropolis-Hastings algorithm, may help towards true relevancy in multimodal IR (MMIR).

Graph of Information obj.

There are many related work on MMIR, particularly in combination of image and text, and video and text.

The generated multimodal data today are not isolate, specially through social network platforms.

We present a graph based model for MMIR name Astera.



Probability Ranking Principle in IR

Relevancy of a doc (d) to a query (q)

$$p(d|q) = \frac{p(q|d)p(d)}{p(q)}$$

The probabilities of p(d) & p(q) are not known.

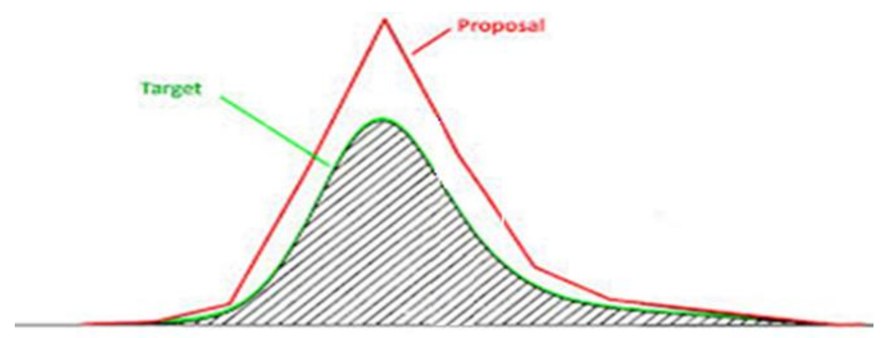
Ranking models e.g. TF.IDF, BM25, LM probe the true ranking by modelling p(q/d).

Having a **graph-based model for MMIR,**
How we can approach
true relevancy?

Methodology

We assume that the true relevancy of info. objects is the final probability distribution. But we cannot sample from this distribution of π(x).

We can use Monte Carlo Markov Chain (MCMC) methods like Metropolis-Hastings to make a Markov chain resulting in the same distribution.



What do we need:

1- An adjacency matrix e.g. *W* to get *y* from *W(x,y)* with ergodic property : if the graph is connected and not bipartite.

2- A function $\tilde{\pi}(x)$, which is proportional to the desired probability distribution $\pi(x) = \tilde{\pi}(x)/k$

Metropolis-Hastings

1- Start from a random node e.g. *x*

2- Get the next candidate *y* from matrix *W(x,y)*

3- Decide to accept or reject this new candidate based on parameter $\lambda = \tilde{\pi}(y)/\tilde{\pi}(x)$. If the density increases, we choose *y*, otherwise we accept it with a probability.

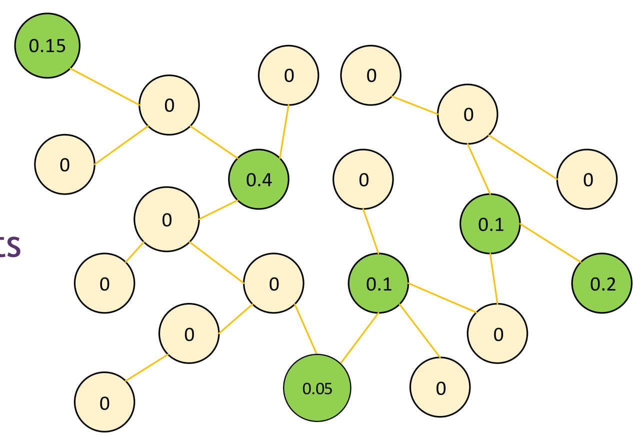
Mapping to IR problem

1- We assume the target probability distribution of true relevancy as $\pi(x)$.

2- The proportional function of $\tilde{\pi}(x)$ could be BM25, TF.IDF or LM .

3- Our graph of multimodal objects is the adjacency matrix

The ideal graph of true relevancy of information objects to the query



Experiment Design

Test Collection: ImageCLEF 2011 Wikipedia

About 400,000 documents and images

Each image has metadata of its comment, caption and description.

We make a graph of documents with their images.

We use Lucene TF.IDF standard result on documents and image metadata to find top 20 result from each. We start the propagation in the graph from top results.

F $a^t = a^0 \cdot Pr^t$ nodes is computed based on where a^0 is the starting scores, *Pr* is the transition matrix ($Pr(x,y) = W(x,y) \cdot \lambda(x,y)$), and a^t is the final score of nodes after *t* iterations

Preliminary Results

The results with the base graph, without Metropolis-Hastings

steps	p@10	r@10	p@20	r@20
1	0.297	0.135	0.229	0.158
2	0.297	0.135	0.229	0.158
3	0.252	0.123	0.188	0.138
4	0.224	0.120	0.184	0.134
5	0.206	0.1148	0.173	0.124
6	0.182	0.1104	0.156	0.113
7	0.142	0.106	0.13	0.115

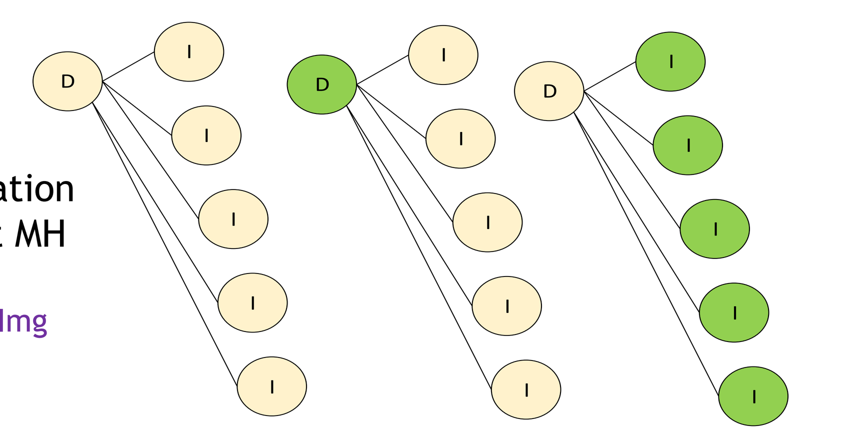
Table 1: Result for documents without image facets, self-transitivity: 0.9, links: δ, β

The results with the base graph, with Metropolis-Hastings

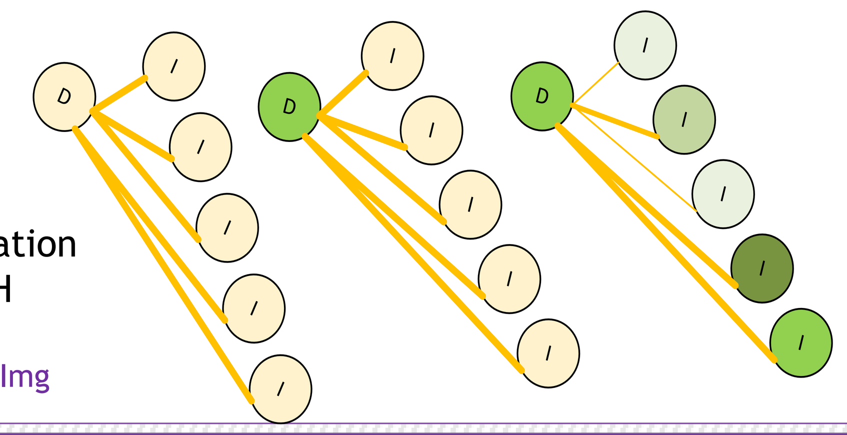
steps	p@10	r@10	p@20	r@20
1	0.27	0.125	0.151	0.135
2	0.27	0.125	0.151	0.135
3	0.23	0.113	0.148	0.1295
4	0.22	0.1097	0.133	0.1163
5	0.18	0.1091	0.113	0.1163
6	0.17	0.107	0.111	0.109
7	0.14	0.08	0.108	0.087

Table 2: Result for documents without image facets, self-transitivity: 0.9, links: δ, β

Score propagation without MH applied
D:Doc, I:Img



Score propagation with MH applied
D:Doc, I:Img



Discussion

- How much the final probability distribution is dependent to the defined RSV function?
- Is Metropolis-Hastingws algorithm on graph-based collections an opportunity to compare the effect of different ranking models?
- How much expensive is this approach? Many steps needed to burn in the matrix.
- The stochastic property in Multimodal multi-relation graph is a challenge. This property goes back to the utility of different modalities as neighbours to the user. The difficulty is whether these neighbours are equally useful to the user?

Contact

Serwah Sabetghadam
Technical University of Vienna
Email: Sabetghadam@ifs.tuwien.ac.at
Website: http://www.ifs.tuwien.ac.at/~Sabetghadam/Astera.html

References

[1] A. Awan, R. A. Ferreira, S. Jagannathan, and A. Grama. Distributed uniform sampling in unstructured peer-to-peer networks. In HICSS, 2006.
[2] T. Berber, A. H. Vahid, O. Ozturkmenoglu, R. G. Hamed, and A. Alpkocak. Demir at imageclefwiki 2011: Evaluating dierent weighting schemes in information retrieval. In CLEF, 2011.
[3] R. A. Ferreira, M. Krishna Ramanathan, A. Awan, A. Grama, and S. Jagannathan. Search with probabilistic guarantees in unstructured peer-to-peer networks. In P2P, 2005.
[4] M. Hlynka and M. Cylwa. Observations on the Metropolis-Hastings Algorithm. University of Windsor, Department of Mathematics and Statistics, 2009.
[5] W. H. Hsu, L. S. Kennedy, and S.-F. Chang. Video search reranking through random walk over document-level context graph. MULTIMEDIA, 2007
[6] Y. Jing and S. Baluja. Visualrank: Applying pagerank to large-scale image search. IEEE Trans. Pattern Anal. Mach. Intell., 2008.
[7] J. Martinet and S. Satoh. An information theoretic approach for automatic document annotation from intermodal analysis. In Workshop on Multimodal Information Retrieval, 2007.
[8] T. Mei, Y. Rui, S. Li, and Q. Tian. Multimedia search reranking: A literature survey. ACM Computing Surveys (CSUR), 2014.