



Application and Evaluation of Multi-Dimensional Diversity

Teerapong Leelanupab, Martin Halvey, Joemon M. Jose
Multimedia Information Retrieval Group,
University of Glasgow, UK



Outline

- Background
- Related Works
- Proposed Methods
 - Application of Diversity
 - Evaluation of Diversity
- Conclusion

Background

- Traditional IR system
 - Employ independent ranking approach to rank documents in order of relevance to queries
 - Very successful, such as Google, Bing, Yahoo etc.
 - Suitable where relevant documents are very few, and high recall is required
- But**
- Ignore contents of documents already ranked in the search results.

Background (Continue)

- In some situations, inappropriate
 - Many relevant documents containing similar information
 - Results might be very similar to each other
 - User queries
 - Related to broad search topics
 - Ex: topic: London
 - Weather, Transport, People, Travel, Big Ben
 - Have multiple distinct meaning
 - Ex: query: Chelsea
 - Chelsea UK, Chelsea Clinton, Chelsea FC

Background (Continue)

Search results need to be diverse?

- Novelty in ranking
 - Highly duplicate information within document in ranking
- Ambiguity of Search Query
 - Broad topics given query
 - Equivocal query

Objective

- Retrieve as diverse results as possible

Charles L.A. Clarke, Maheedhar Kolla, Gordon V. Cormack, Olga Vechtomova, Azin Ashkan, Stefan B. Utcher, and Ian MacKinnon. Novelty and diversity in information retrieval evaluation. In SIGIR '08

Outline

- Background
- Related Works
- Proposed Methods
 - Application of Diversity
 - Evaluation of Diversity
- Conclusion

Related works

- Explicit re-ranking functions with tunable parameter
 - Maximal Marginal Relevance (MMR), Harmonic measure, Risk Minimisation, Portfolio theory

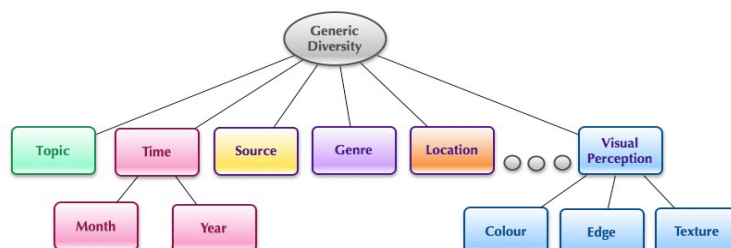
$$MMR_{J+1} \equiv \operatorname{argmax}_{x_i \in I \setminus J} [\lambda S(x_i; q) + (1 - \lambda) D(x_i; (x_1, \dots, x_J))]$$

- Subtopic retrieval measures
 - S-recall, S-precision, S-MRR, α -nDCG etc.

$$CR@k = \frac{|\cup_{i=1}^k \text{subtopics}(d_i)|}{n_Q}$$

Related work: dimensions of Diversity

- Other research works raise importance of specific dimensions of diversity
 - Ex: Topic, Source (supplier), Genre (products and accessories), Temporal, Spatial, Visual Diversification, and so on



Related works: ImageClef 2009, Photo Retrieval task

- Create a test collection and define clusters based on analysing the distribution of query variations.
- More accurately specify diversity based on user information needs.



- The clusters of topic, "Beckham", defined according to ImageClef 2009.
- Four dimensions, i.e. anchor persons (topic), genre, location, time
- Some documents falls into clusters overlapping two dimensions, i.e. "David Beckham 2009"

Related Work (Continue)

- How to effectively develop diversity algorithms?
- How to evaluate the results from the combination of varied dimensions in diversity?
- Current widely used evaluation measures account for subtopics in a single dimension of diversity.
- Diversity should depends on application, user context, information needs. Ex: product search, sport journalist, magazine editor

Outline

- Background
- Related Works
- Proposed Methods
 - Application of Diversity
 - Evaluation of Diversity
- Conclusion

Application of Diversity

- Intuition
 - Separately promote diversity based on predefined dimensions
 - Cover as many dimensions as possible
 - User centred approach, taking into account the context and information needs of users
- Solutions
 - Visually present results, separated into different viewpoint according to dimensions
 - Fuse results by considering which dimensions are important to users and show in a single ranking.

Application of Diversity (Continue)

Documents	α_1 -Topic (2)		α_2 -Location (3)			α_3 -Genre (1)		Total Scores
	$C_{\alpha_1,1}$	$C_{\alpha_1,2}$	$C_{\alpha_2,1}$	$C_{\alpha_2,2}$	$C_{\alpha_2,3}$	$C_{\alpha_3,1}$	$C_{\alpha_3,2}$	
x_1	X				X		X	6
x_2		X		X				5
x_3			X					3
x_4					X	X		1
x_5	X							0

- Assume that documents which fit into multiple subtopics from different dimensions are more important
- Documents fall into a particular subtopic where no other document exists.
- Weight of each dimensions is pre-defined according to different search domains

Application of Diversity (Continue)

- Dimension coverage scores, for example, can be treated as a graded diversity value that will be added to dissimilarity value, such as in MMR function.

$$MMR_{J+1} \equiv \operatorname{argmax}_{x_i \in I \setminus J} [\lambda S(x_i; q) + (1 - \lambda)(D(x_i; (x_1, \dots, x_J)) + \alpha(x_i; q))]$$

Outline

- Background
- Related Works
- Proposed Methods
 - Application of Diversity
 - Evaluation of Diversity
- Conclusion

Evaluation of Diversity

- Suggest to evaluate systems by separately considering clusters from the same dimensions

$$CR_{\alpha_a; q}@k = \frac{|\cup_{i=1}^k \text{subtopics}(d_i)|}{n_Q}$$

- Then, evaluate overall performance of systems by average sum of S-recall from possible dimensions related to query, including weight specified by search domain or user context.

$$CR_{total; q}@k = \frac{1}{A} \sum_{a=1}^A w_{\alpha_a} \cdot (CR_{\alpha_a; q}@k)$$

Outline

- Background
- Related Works
- Proposed Methods
 - Application of Diversity
 - Evaluation of Diversity
- Conclusion

Conclusion

- Propose important aspects needed to be considered in produce document diversity in ranking
- Many research challenges needed to be solved in different dimensions of diversity
- Need to specify the desirable level of granularity of dimensions in different context.

Questions?



Teerapong Leelanupab (Kim)
Information Retrieval Group
University of Glasgow
kim@dcs.gla.ac.uk