

## Multilingual Retrieval Experiments with MIMOR at the University of Hildesheim

René Hackl, Ralph Kölle, Thomas Mandl,  
Alexandra Ploedt, Jan-Hendrik Scheufen,  
Christa Womser-Hacker

Information Science

Universität Hildesheim  
Marienburger Platz 22 - 31131  
Hildesheim

Germany

{mandl,womser}@uni-hildesheim.de

## Overview



- Fusion in Information Retrieval
- The MIMOR Model
- Participation 2003
  - Blind Relevance Feedback
  - Fusion Parameters
  - Submitted Runs
- Outlook

## Fusion in ad-hoc Retrieval

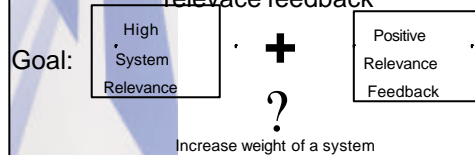
Many studies especially within TREC show

- that the quality of the best retrieval systems is similar
- that the overlap between the results is not very large
- ✎ that fusion of the results of several systems can improve the overall performance

## The MIMOR Model

Combines fusion and Relevance Feedback

- linear combination
- each individual system has a weight
- weights are adapted based on relevance feedback



## Calculation of RSV in MIMOR

Weighted sum of single RSV

$$RSV_j = \frac{\sum_{i=1}^N (w_i \cdot RSV_{ij})}{N}$$

$RSV_i$  Retrieval Status Value of System  $i$  for Document  $j$

## Participation in CLEF

- 2002: GIRT track
  - Linguistic pre-processing: Lucene
  - Basic IR systems: irf
- 2003: multilingual-4 track
  - Linguistic pre-processing: Snowball
  - Basic IR systems: Lucene and MySQL
  - Internet translation services

## Participation in CLEF 2003

- Snowball for linguistic pre-processing
- MIMOR in JAVA
  - Lucene and MySQL
  - static optimization
- Internet translation services
  - „Fusion" of translation services: use all terms found by any service
- Blind relevance feedback

## Blind Relevance Feedback

We experimented with:

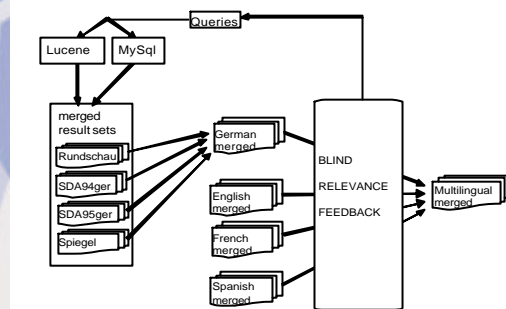
- Kullback-Leibler (KL) divergence measure
- Robertson selection value (RSV)

	Average precision	Average document precision	Documents retrieved
Lucene BRF RSV 5 10	0.3017	0.3107	5059
Lucene BRF KL 5 10	0.3138	0.3277	5216
7:1 BRF RSV 5 10	0.3049	0.3174	5157
7:1 BRF KL 5 10	0.3127	0.3264	5227

## Blind Relevance Feedback

- Blind Relevance Feedback improved performance over 10% in all tests
- Parameters could not be optimized
- Submitted runs use top five documents and extract 20 terms (5 / 20) based on Kullback-Leibler

## Sequence of operations



## Test Runs

	Number of retrieved multilingual documents	Average precision	Average document precision
<b>Data from 2001</b>			
Lucene	5167 / 6892	0.2880	0.3248
MySQL	2873	0.1037	0.1359
1:1 merged	3975	0.1856	0.2206
4:1 merged	4984	0.2673	0.3094
<b>9:1 merged</b>	5101	<b>0.2830</b>	<b>0.3248</b>
17:3 merged	5056	0.2764	0.3189
<b>Data from 2002</b>			
Lucene	4454 / 6996	0.2876	0.2769
MySQL	2446	0.0913	0.0951
9:1 merged	4543	0.2851	0.2762
17:3 merged	4533	0.2787	0.2709
<b>7:1 merged</b>	<b>4553</b>	<b>0.2822</b>	<b>0.2742</b>
33:7 merged	4511	0.2740	0.2670

## Test Runs

- Lucene performs much better than MySQL
- Still, MySQL contributes to the fusion result
- -> High weight for Lucene

## Results of submitted runs

	Average precision	Documents retrieved
UHImIt4R1 (7:1 fusion)	0.2849	3944 / 6145
UHImIt4R2 (Lucene only)	0.3057	4137
UHImnenR1 (7:1 fusion)	0.3630	951 / 1006
UHImnenR2 (Lucene only)	0.3797	945

## Submitted Runs

- Lucene performs much better than MySQL
- MySQL does not contribute to the fusion result
- Lucene only best run
- Blind Relevance Feedback leads to improvement

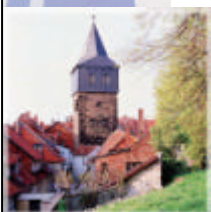
## Proper Names

- Proper Names in CLEF Topics lead to better performance of the systems (see Poster and Paper)
- Treat Topics with Proper Names differently (different system parameters) ?

## Acknowledgements

- Thanks to the developers of Lucene and MySQL
- Thanks to students in Hildesheim for implementing part of MIMOR in their course work

## Outlook



We plan to

- integrate another retrieval systems
- invest more effort in the optimization of the blind relevance feedback
- participate again in the multilingual track
- Evaluate the effect of individual relevance feedback