In: Bullinger, Hans-Jörg; Ziegler, Jürgen (eds.): Human-Computer Interaction: Communication, Cooperation and Application Design. Proceedings of the HCI International ´99 (8th International Conference on Human-Computer Interaction), Munich, Germany, 22-27. August 1999. Lawrence Erlbaum Associates: Mahwah, NJ; London. Vol. 2. S. 604-608.

Adapting Meta Information Retrieval to User Preferences and Document Features

Christa Womser-Hacker & Thomas Mandl
Information Science - University of Hildesheim

1 Introduction

Human and contextual factors of information processes need to be integrated into Information Retrieval (IR) models, although often ignored. This article introduces an approach to adapting of an IR system by optimizing the choice of retrieval functions based on user preferences and document features. The framework is provided by MIMOR (Multiple Indexing and Method-Object Relations), a novel approach in Information Retrieval which exploits users' relevance feedback information to model method-object relations. By processing the knowledge from feedback, MIMOR automatically chooses the optimal retrieval methods for a specific situation. In this paper, MIMOR is presented and a formal model introduced. In the last section, the model is extended to incorporate user preferences in order to allow users to fine tune his/her own Meta IR System. All users contribute to a public model providing relevance feedback information. They are also to control the weighting of the private versus the public model during retrieval.

2 The MIMOR Approach

MIMOR (cf. Womser-Hacker 1997) is based on empirical results from IR research and especially from TREC (Text Retrieval Conference; cf. Voorhees/Harman 1998), a large testbed for IR experiments granted by US government institutions. TREC is organized by the National Institute for Standards and Technology (NIST) which provides a large real world collection of documents, queries/topics and relevance judgements, and evaluates the results. Research groups apply their IR systems to this corpus and train it with the results from previous years. The final results are sent to NIST, where they are analyzed, visualized and compared to one another following a well defined evaluation methodology (cf. Harman 1993). At an annual conference the systems are demonstrated. IR systems consist of different indexing, weight assignment and matching functions. Two major results of TREC have influenced MIMOR:

- User relevance feedback proves to be very effective in improving retrieval performance. This means that the natural interactivity of the retrieval process can be exploited using additional knowledge about the distribution of relevant and non-relevant information objects. This knowledge can be used to modify query terms as well as their weights in order to place the relevant material at the beginning of the ranking list (cf. Belkin 1993).
- Many systems have led to comparable quality, however the overlap of their result sets has been rather small. Some researchers have attempted to exploit this fact by introducing poly-representation techniques of retrieval objects and achieved good results. The so called data fusion is based on the idea that different object representation corresponds better to individual users' interests and perspectives (cf. Fox & Shaw 1994).

MIMOR integrates these two approaches. It can be seen as an additional layer within an IR system managing the poly-representation of queries and documents by selecting appropriate methods for indexing and matching. Through learning from relevance feedback information, the model adapts itself during an initial phase by assigning weights to the different method–object combinations. MIMOR is not limited to text documents as many common IR applications, but rather it is open to data types like structured data and multimedia objects.

3 Formalization

Each Information Retrieval System calculates a retrieval status value (RSV) for each document by comparing it to the query.

Table 1: MIMOR combines the results of different retrieval systems

	Doc. 1	Doc. 2	Doc. 3	Doc. 4	...
System A	RSV_{A1}	RSV_{A2}	RSV_{A3}	RSV_{A4}	
System B	RSV_{B1}	RSV_{B2}	RSV_{B3}	RSV_{B4}	
System C	RSV_{C1}	RSV_{C2}	RSV_{C3}	RSV_{C4}	
...					...

MIMOR	RSV_1	RSV_2	RSV_3	RSV_4	...

A standard IR system presents a ranked list containing the documents with the highest RSV as result. In data fusion IR systems, the entire matrix of documents and systems is considered. In order to present a single ranked list to the user, MIMOR needs to summarize the matrix into a single vector. Each system influences the final result according to a weight ω_I which is changed through learning. The final RSV for document j is calculated using the following formula:

(1) $RSV_j = (\omega_A\, RSV_{Aj} + \omega_B\, RSV_{Bj} + \omega_C\, RSV_{Cj} + ... + \omega_N\, RSV_{Nj}) / N$

Initially, all weights are equal. While MIMOR is used, the weights are changed according to users' relevance feedback R_U, which can be positive or negative:

(2) $\omega_I = \omega_I + (e\, RSV_{Ij}\, R_{Uj})$

e learning rate
R_{Uj} relevance judgement of user U for document j

The formula shows that the user does not judge the retrieval systems overall, but instead s/he evaluates single documents, a strategy which is more appropriate. The weight of retrieval systems assigning high RSV to relevant documents will increase over time. Thus, the influence of a successful system can increase and the weights can be observed by a supervisor.

The MIMOR model is further refined so that information on document features can be included. Criteria building document types, such as length or topic difficulty, can also be found in TREC but should be further developed.

The features are modeled as exclusive clusters of documents. Instead of one single weight, each system has a weight for each cluster which can be adapted independently. Thus, a system can perform well for one type of documents and therefore have a large impact on the final result whereas on the other hand it can have a small weight for another type of documents.

Table 2: Clusters of documents, each cluster having a different weight

	Cluster 1		Cluster 2		...
	Doc. 1	Doc. 2	Doc. 3	Doc. 4	...
System A	ω_{A1}		ω_{A2}		
System B	ω_{B1}		ω_{B2}		
System C	ω_{C1}		ω_{C2}		
...					...

The calculation of the final RSV as presented in Formula (1) has to be adapted for a document j in cluster x:

(3) $RSV_{jx} = (\omega_{Ax}\ RSV_{Aj} + \omega_{Bx}\ RSV_{Bj} + \omega_{C}\ RSV_{Cj} + ... + \omega_{Nx}\ RSV_{Nj}) / N$

The learning formula is changed accordingly. Only the weights of clusters containing the document evaluated by user relevance feedback are adapted.

4 Extension of MIMOR to User Preferences

In many IR experiments researchers found that users do not agree on all relevance judgements as a result of their different perspective or background. To achieve optimal retrieval quality, each user needs to train his own MIMOR model. However, this approach is not applicable until a significant number of relevance judgements have been made. Although relevance feedback proved to be one of the best methods to improve retrieval results (cf. Over 1998), users do not give many relevance judgements as it is considered time consuming. MIMOR can collect relevance feedback information from many users and train an optimal system. However, such a general MIMOR does not account for differences between users. A tradeoff between degree of personalization and the available number of user judgements needs to be found.

In order to manage this tradeoff, an extension of MIMOR introduces a private and a public model. The public model is formed by all user relevance feedback information available to the system while the private model is restricted to the judgements of one user. Users who start to work with the system have not presented many relevance judgements yet. Because of that the influence p of their private model for the final result should be low. The weight of the private model can grow with the number of relevance judgements but can also be set by the user directly. The formula for the calculation of the final result has to be further adapted considering that each weight ω_{Ix} for system I and cluster x is split into $\omega_{Ix(pub)}$ and $\omega_{Ix(pri)}$ for the public and private model respectively:

(4) $RSV_{jx} = ((p\ \omega_{Ax(pri)}\ RSV_{Aj} + (1\text{-}p)\ \omega_{Ax(pub)}\ RSV_{Aj}) + (p\ \omega_{Bx(pri)}\ RSV_{Bj} + (1\text{-}p)\ \omega_{Bx(pub)}\ RSV_{Bj}) + ... + (p\ \omega_{Nx(pri)}\ RSV_{Nj} + (1\text{-}p)\ \omega_{Nx(pub)}\ RSV_{Nj})) / N$

A similar technique of additive components will be used for the document type clusters. The final weight may be a weighted combination of a weight for the cluster and a general weight for the system. The influence of the cluster-model would grow with the number of relevance judgements available for the cluster. Clearly, the presented model requires a considerable amount of training data. The number of adaptable parameters increases with the complexity of the model. Each additional IR system, the introduction of document clusters and again the consideration of user preferences lead to a growing number of parameters. Especially the training of the private model demands many relevance judgements and may only be feasible for a dedicated, professional searcher. However, each single relevance judgement supports the construction of the public model and thus all users can profit. One source for relevance judgements is the TREC corpus, where a great amount of has been accumulated over the years. This knowledge can serve as a starting point for the implementation of MIMOR and its extension.

5 Summary

The presented extension of MIMOR is a promising approach to model human computer interaction in IR systems. Derived from the results of the largest empirical evaluation initiative in IR, the TREC conferences, it has a sound basis. It combines two of the most prominent approaches to improve IR systems, user relevance feedback and data fusion based on poly-representation. Thus, the experience of IR systems development can be exploited in a single model to account for well known shortcomings of current systems. The extension of MIMOR regards different users as well as document types.

6 References

Belkin, N. (1993). Interaction with Texts: Information Retrieval as Information Seeking Behavior. In Knorz, G.; Krause, J. & Womser-Hacker, C. (Eds.). *Information Retrieval '93. Von der Modellierung zur Anwendung*. Proc. d. 1. Tagung Information Retrieval. Konstanz. pp. 55-66.

Fox, E. & Shaw, J. (1994). Combination of Multiple Searches. In Harman, D. (Ed.): *The Second Text Retrieval Conference (TREC-2). NIST Special Publication 500-215.* pp. 243-252.

Harman, D (1993). Overview of the First Text Retrieval Conference (TREC-1). In Harman, D (Ed.). *The First Text Retrieval Conference (TREC-1).* NIST Special Publication 500-207. pp. 1-20

Over, P. (1998). TREC-6 Interactive Track Report. In Voorhees, E. & Harman, D. (Eds.). *The Sixth Text Retrieval Conference (TREC-6).* NIST Special Publication 500-240. August 1998, pp. 73-81.

Voorhees, E. & Harman, D. (1998). *The Sixth Text Retrieval Conference (TREC-6).* NIST Special Publication 500-240. August 1998.

Womser-Hacker, C. (1997). *Das MIMOR-Modell. Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval.* Habilitationschrift. Universität Regensburg.