

Regression Relevance Models for Data Fusion

Shengli Wu

School of Computing and Mathematics

University of Ulster, Northern Ireland, UK

Data fusion with scoring information

- Data fusion: to search the same collection of documents with different information retrieval systems, then to merge those results from different systems for effectiveness improvement.
- Sometimes scores, indicating the estimated probability of relevance, or the estimated degree of relevance, are associated with each document in the result, a few methods such as CombSum, CombMNZ, the linear combination methods can be used.

Data fusion with ranking information

- Sometimes no scores are available, only a ranked list of documents are given. For example, Web documents searched from Web search engines do not have scores associated.
- How to use data fusion methods such as CombSum, and others?
- Estimating relevance probabilities at each rank position, then CombSum can be used.

Modeling the rank-probability of relevance relationship

- For a slightly different purpose (distributed information retrieval), Calve and Savoy used the logistic model for this:

$$p = \frac{1}{\frac{1}{u} + a_2 * e^{\ln(rank) * \ln(b_2)}}$$

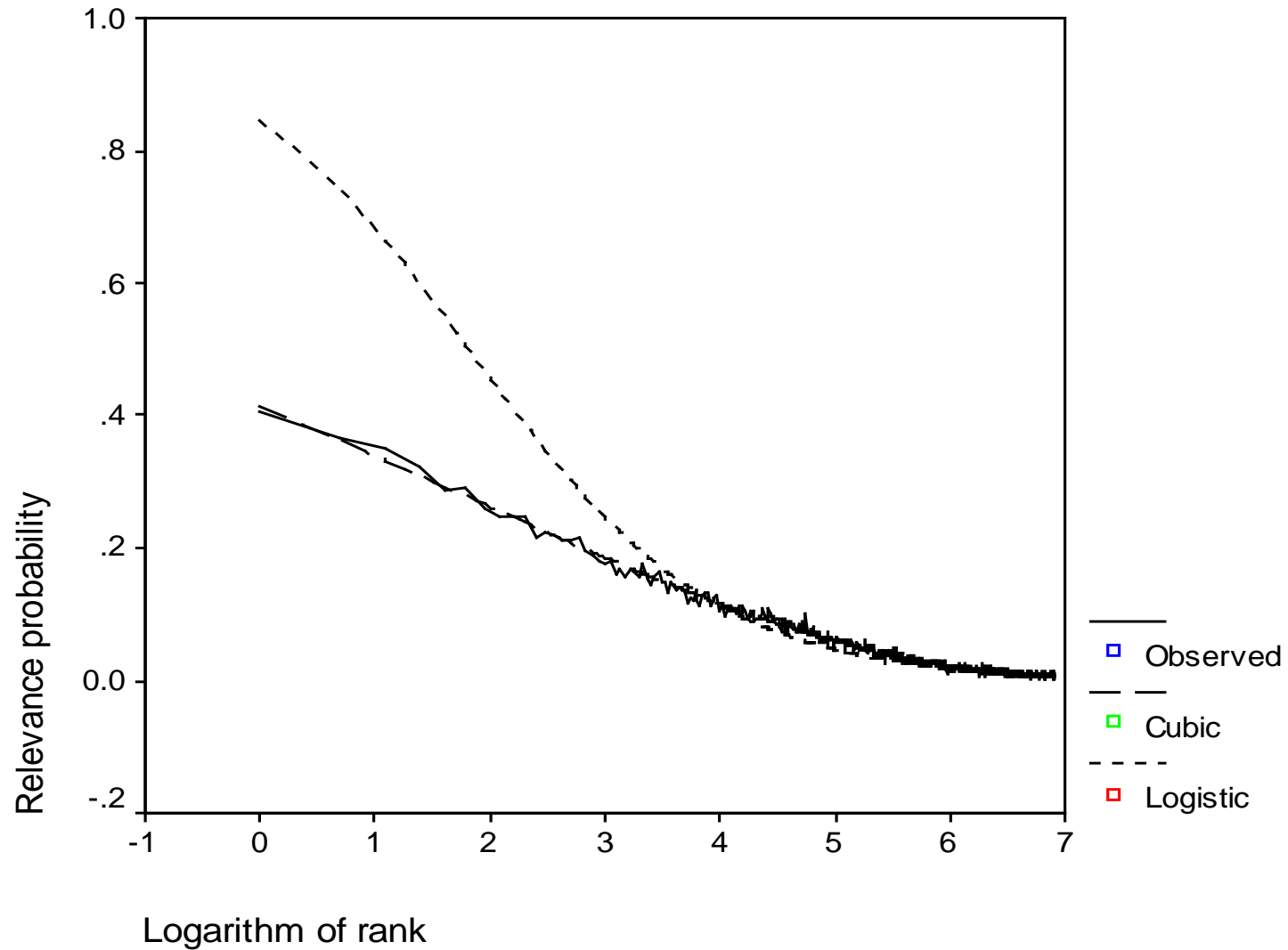
- We tried several different functions for this, and found that cubic function is a good option.

$$p = a_1 + b_1 * \ln(rank) + c_1 * \ln(rank)^2 + d_1 * \ln(rank)^3$$

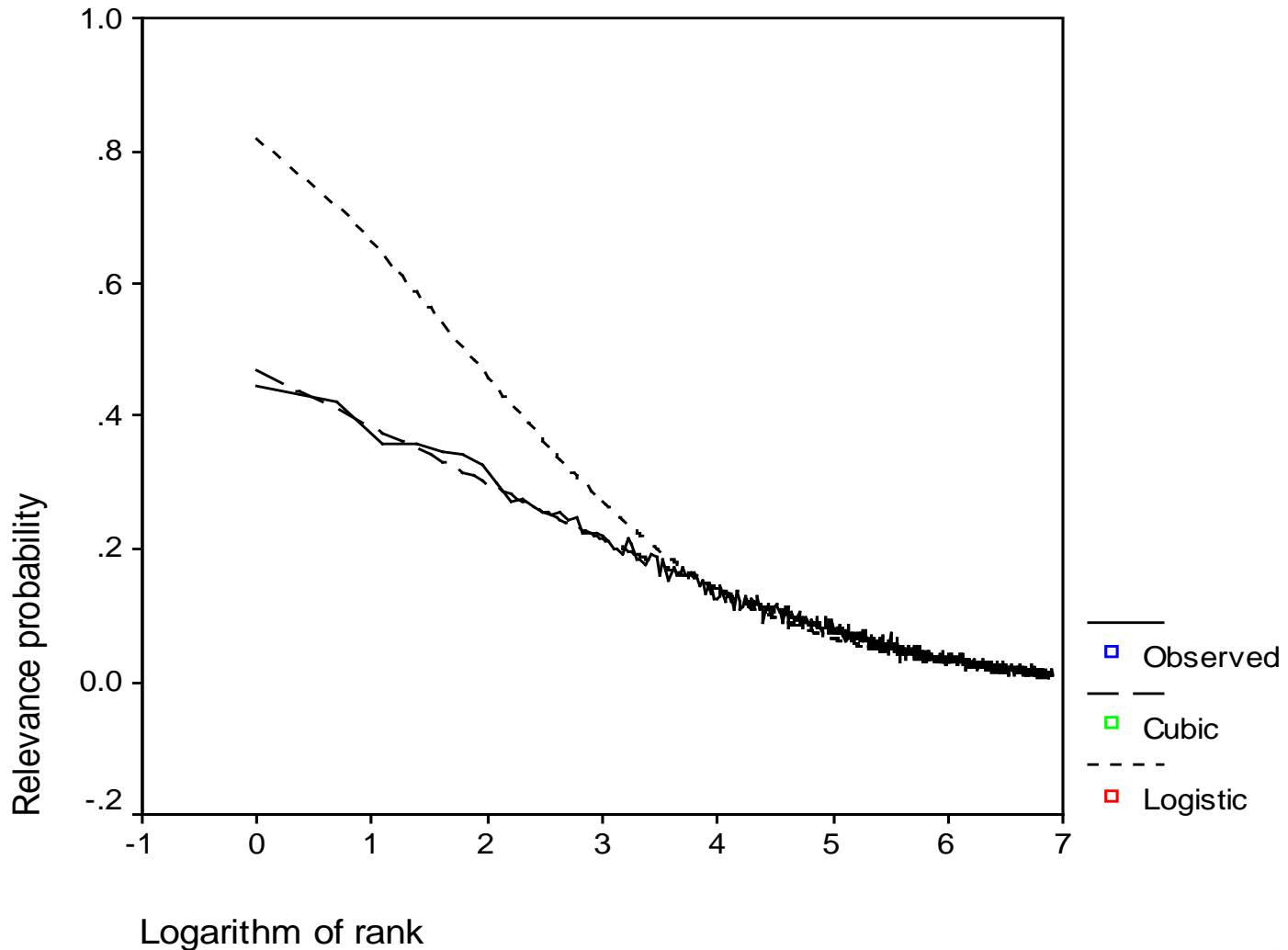
An experiment

- We used three groups of results submitted to TREC (9, 2001, and 2004)
- Regression was used to obtain the most fitting curves for those data.

Experimental results for TREC 9



Experimental results for TREC 2001



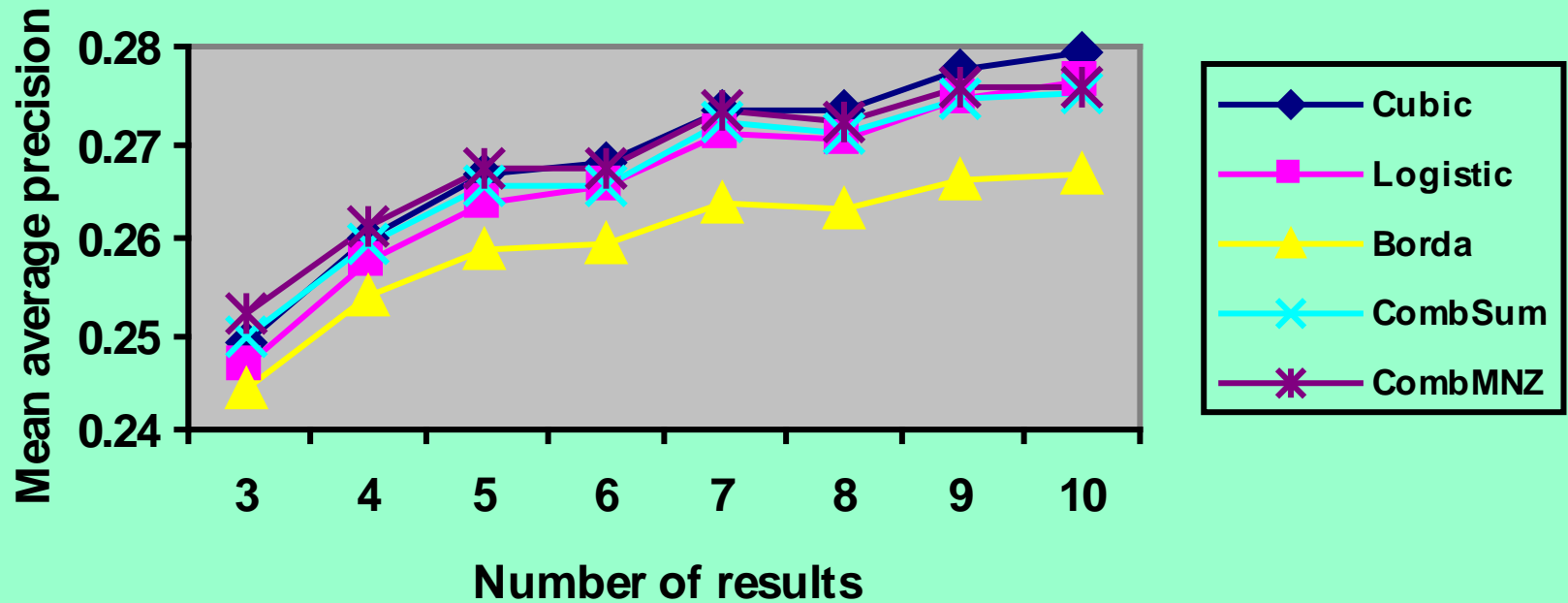
Experimental results for three groups of data (Euclidean distance between actual and estimated curves)

| Group | Cubic | Logistic |
|-----------|--------|----------|
| TREC 9 | 0.3453 | 0.9616 |
| TREC 2001 | 0.2810 | 0.8003 |
| TREC 2004 | 0.1576 | 0.5867 |

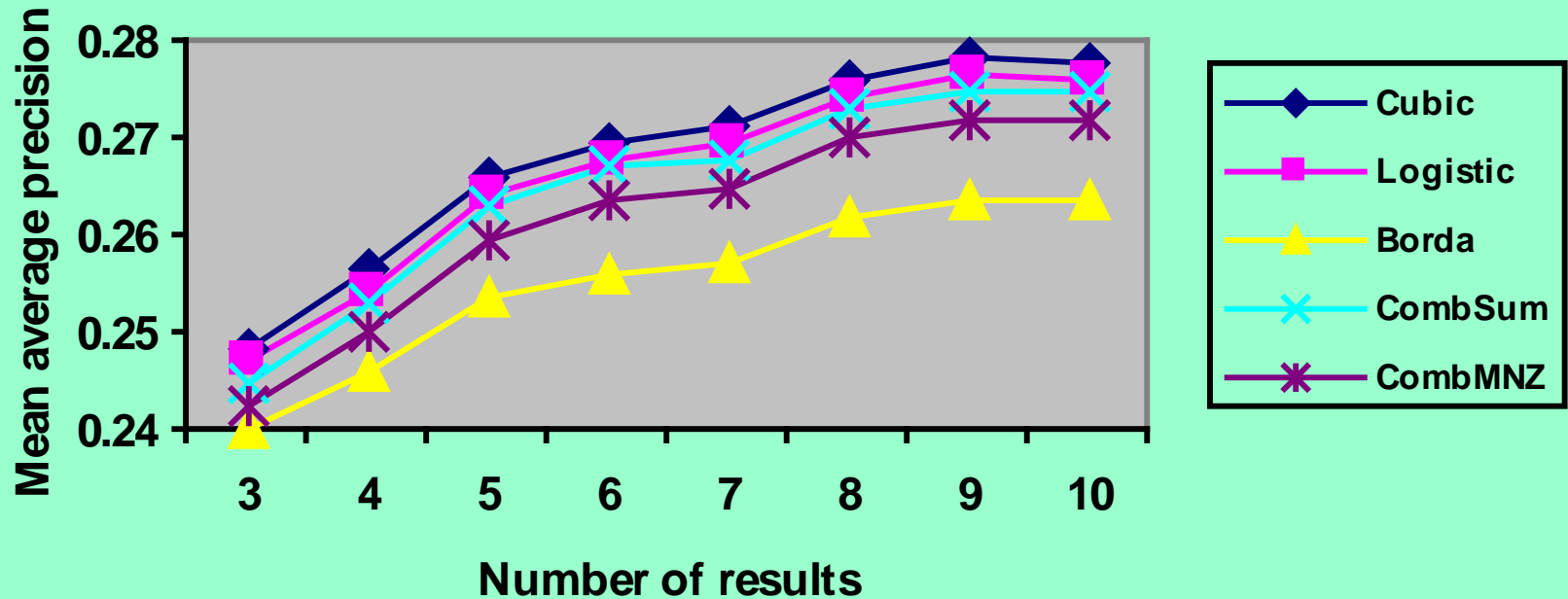
A data fusion experiment

- Three groups of results were used;
- Borda fusion, cubic model, and logistic model using only rank information; Then CombSum was used for fusion;
- CombSum and CombMNZ used score information;
- Mean average precision (MAP) and RP (recall level precision) were used for performance evaluation

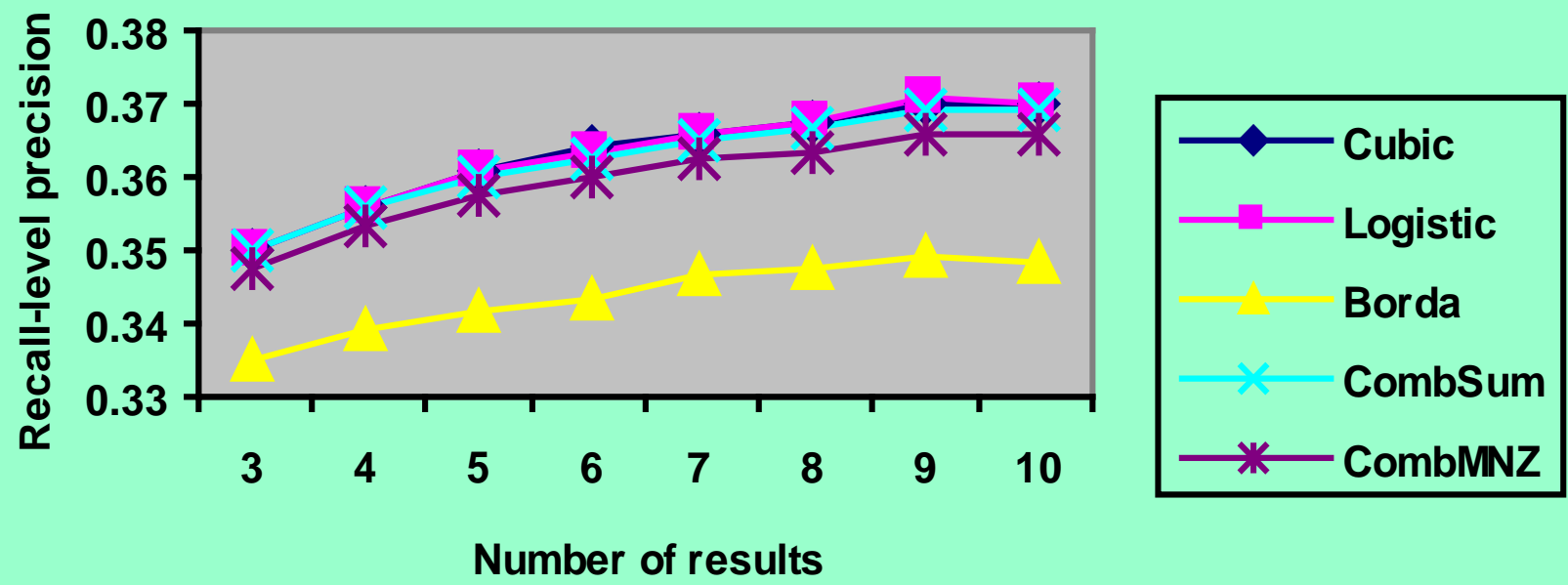
Experimental results (TREC 9, MAP)



Experimental results (TREC 2001, MAP)



Experimental results (TREC 2001, RP)



Conclusions

- The cubic model is more accurate than the logistic model for rank-relevance probability estimation in information retrieval results;
- Both models are effective for data fusion;
- Both of them are better than Borda fusion;
- The cubic model is slightly better than CombSum and CombMNZ;
- The logistic model is as good as CombSum and CombMNZ.

A scenic landscape of misty, jagged mountains. The peaks are dark and rocky, with some sparse greenery. A thick layer of white mist or fog fills the valleys and surrounds the lower parts of the mountains, creating a sense of depth and atmosphere. The sky is a pale, hazy blue. The overall tone is serene and majestic.

Thank you!