Personalised Search Time Prediction using Markov Chains

Vu Tran
University of Duisburg-Essen
Germany

David Maxwell
University of Glasgow
Scotland

Norbert Fuhr
University of Duisburg-Essen
Germany

Leif Azzopardi
University of Strathclyde
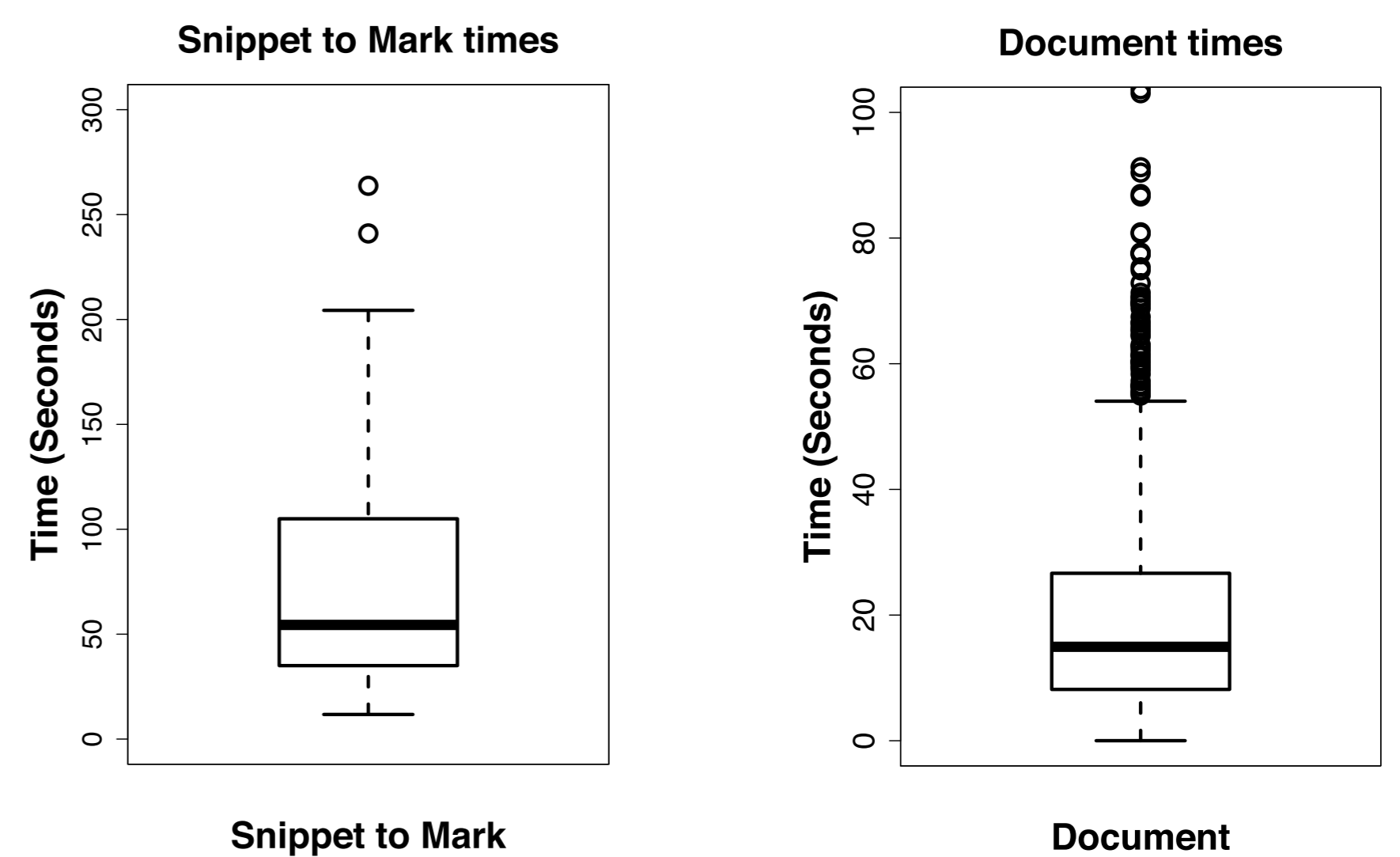*Scotland*

Search Time Predictions

- **Predicting the time to finding the next document**
- Time-based evaluation measures
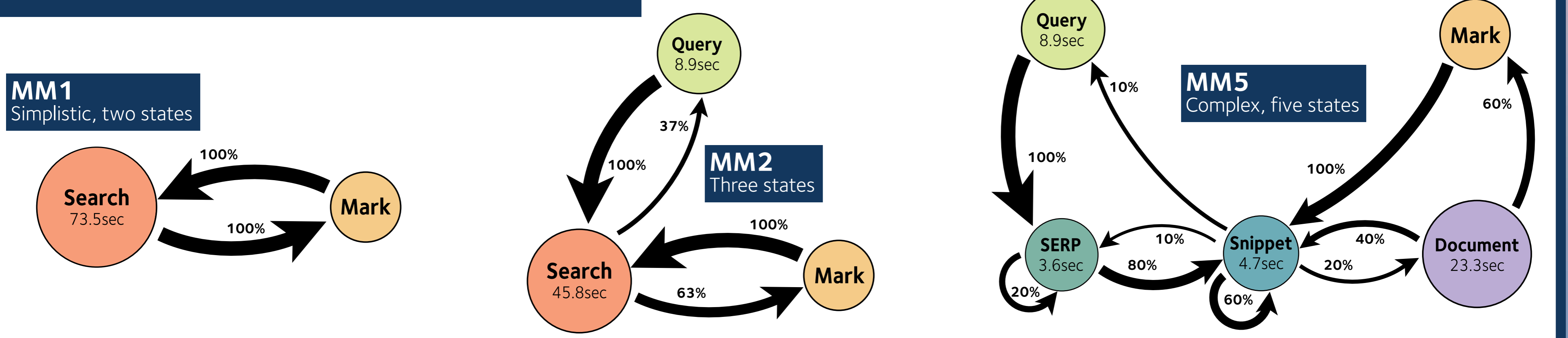- Personalised estimations

Goals

- Provide a **baseline for further research**
- Focus on **situation- and action-specific estimates**
- Improving **search effectiveness by guiding the user**

Data Used

- **User interaction data** was obtained from a **prior user study** (see Maxwell and Azzopardi, IIiX 2014). **72 sessions** were used, where subjects were asked to find as many relevant documents as they could in 20 minutes.

- **Document examination times** varied considerably across the user study participants. We capped these times to **3.5 standard deviations.**

- **Snippet-to-mark times** were also highly variable, and difficult to predict with our models.

Snippet to Mark times / Document times

Models of Search Interaction

MM1
Simplistic, two states

Search 73.5sec — 100% → Mark
Mark — 100% → Search

MM2
Three states

Query 8.9sec
100% / 37%
Search 45.8sec — 100% / 63% → Mark

MM5
Complex, five states

Query 8.9sec — 10%
100%
Mark 60%
SERP 3.6sec — 10% / 80% / 20% — Snippet 4.7sec — 40% / 20% / 60% — Document 23.3sec
100%

Experimental Results

- **Our models are able to make feasible predictions.**
- **User-specific models perform significantly better** than global models.
- Observing users for **only 100 seconds** already offers **good results** for our predictions.
- **More complex models** (e.g. MM5) **offered better performance** than more simplistic models (e.g. MM1).
- **More complex models can be considered to yield even better predictions.**

Mean Error vs. Observation Time: MM1, MM2 and MM5

Global MM1 / Global MM2 / Global MM5 / User MM1 / User MM2 / User MM5
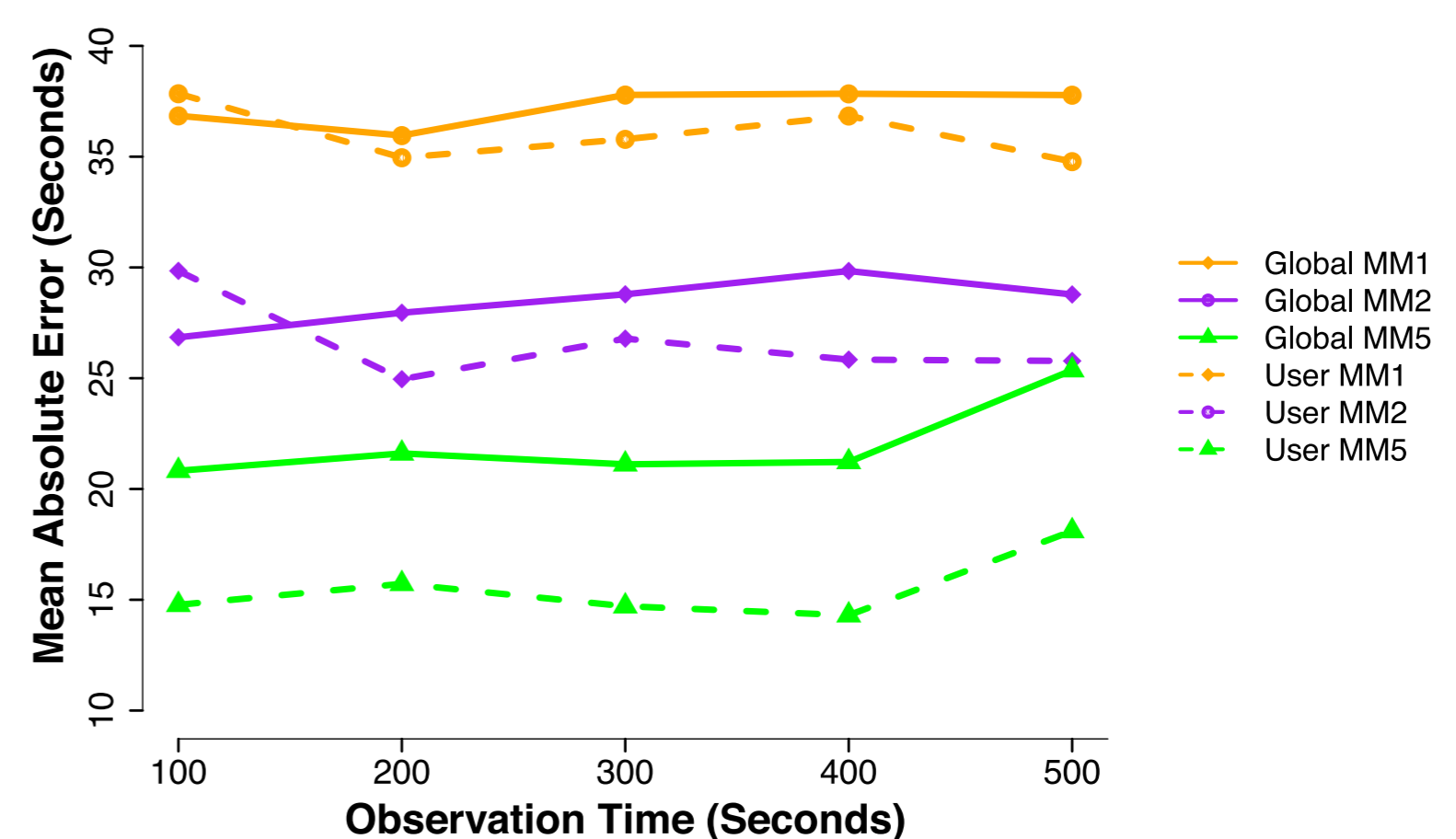
Global models were trained over 66 entire subsamples of session data, and were used as our baselines - remaining samples were used for testing. Personalised models were constructed from cutoff data from each individual subject, producing one personalised model per user.

UNIVERSITÄT DUISBURG ESSEN — *Offen* im Denken

University of Glasgow

University of Strathclyde Glasgow