An Initial Investigation into Fixed and Adaptive Stopping Strategies

David Maxwell and Leif Azzopardi
University of Glasgow, Scotland

Kalervo Järvelin and Heikki Keskustalo
University of Tampere, Finland

Abstract

Most models, measures and simulations often **assume that a searcher will stop at a predetermined place** in a ranked list of results. **Real-world searchers however adapt** their interactions with a ranked list based upon their interpretation of it. This poster presents our initial simulated analysis of stopping rules, **exploring a fixed depth and two adaptive strategies** in the context of ad-hoc topic retrieval.

Stopping Rules and Strategies

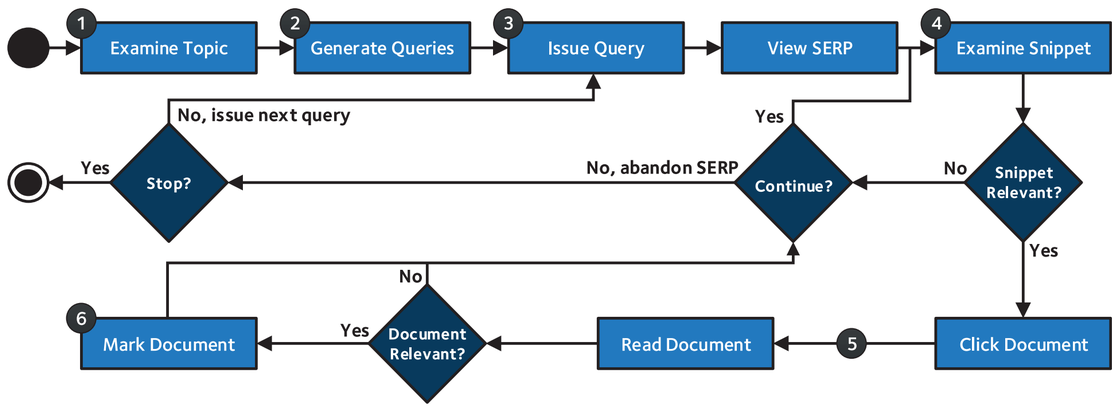
Knowing when to stop is considered a fundamental aspect of human behaviour. Studies have shown that we stop searching when when we feel that what we have found is *"good enough"*. **Researchers have however devised a series of stopping rules**, two of which *(frustration and disgust rules)* we operationalised for this study.

SS1: A **fixed depth** stopping strategy

SS2: Stop examining the results list after observing **a total** of *x* non-relevant snippets

SS3: Stop examining the results list after observing *x* non-relevant snippets **in a row**

The Simulator and Method



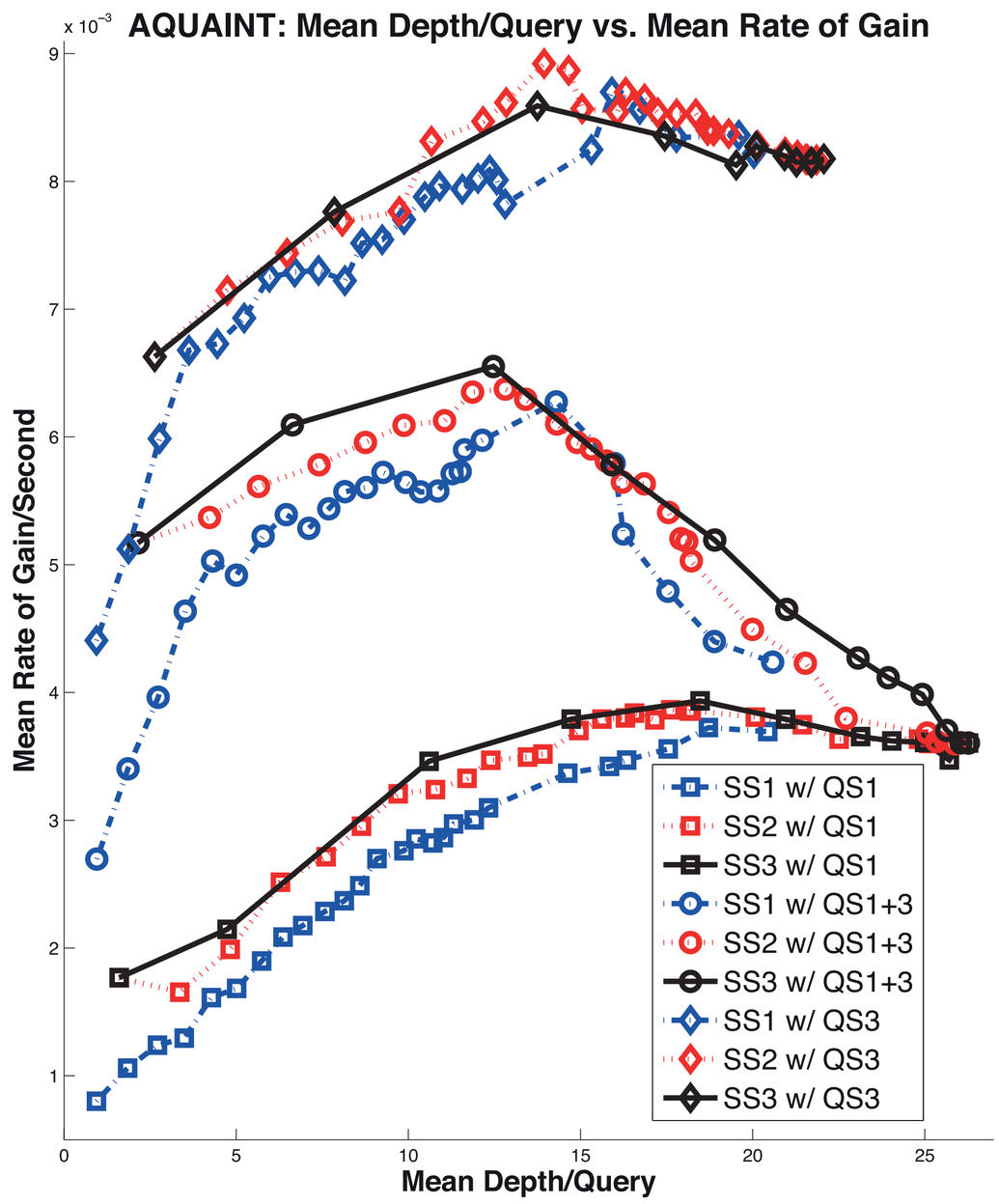
We developed a simulator based upon the user models defined by Baskaya et al. (2013) and Thomas et al. (2014). Simulated searchers **examined 50 topics** over the **AQUAINT** and **WT2g** collections. The goal was to **find as many relevant documents in 20 simulated minutes.** We also used **three querying strategies from** Keskustalo et al. (2009).

QS1: Generation of single-term queries (poor performance).

QS3: Generation of tri-term queries (good performance).

QS1+3: A blended strategy of **QS1** and **QS3**, to examine the robustness of the three stopping strategies.

Results



SS1 achieved its highest level of CG with **high thresholds** (25-50), **much deeper than what is typically used in simulations** (10 snippets/documents per query)

SS2 also achieves its best performance over a **wide range of thresholds**

SS3 appeared to be the **most consistent** of the three stopping strategies, with reasonable thresholds ranging from 3-5 (stop after 3-5 non-relevant snippets in a row)

We deduce that **SS3 appears the most robust.**

Similar findings were found over the WT2g collection

Future Work

Explore a **greater variety of stopping strategies**

Compare stopping strategies to **actual stopping behaviour**

Examine the **relationship between the quality of the results list and interaction probabilities** (i.e. how does performance affect interaction?)

Acknowledgements:
This work is funded by the MUMIA Cost Action, ref no. ECOST-STSM-IC1002-080914-049840.

UNIVERSITY OF TAMPERE

University of Glasgow
VIA VERITAS VITA