



Summarisation and novelty in Mobile Information Access

Simon Sweeney and **Fabio Crestani**

Dept. Computer and Information Sciences
University of Strathclyde, Glasgow, Scotland

David E. Losada

Depto. de Electronica y Computacion
Universidad de Santiago de Compostela, Spain

Outline

- Demands of Mobile Information Access;
- Effective short summary: include only novel content?
- Related Work;
- Framework to investigate novel summaries;
- Generating novel summaries;
- Experimental collection, measures and design;
- Experimental procedure;
- Results and analysis;
- Conclusions and future work.

Demands of Mobile Information Access

Anytime, anywhere information access:

need to deliver information on request, in a form easily digested on the move.

Design of content important:

upper threshold dictated by what is legible to the human eye, and inherent requirements on device sizes to be portable.

Automatic text summarisation:

- condenses a document, reduces need to refer to full text;
- removes need to manually construct device-friendly content;
- dimensions: perspective, composition, orientation, source, length, ...novelty?

Effective short summary: include only novel content?

Producing a summary containing only novel sentences (assuming sentence extraction summarisation) may result in a loss of context?

Will novel sentences alone provide sufficient basis to determine relevance, or do we need to include additional sentences to provide context?

Full text of a document: (i) relevant sentences, (ii) novel sentences and the (iii) remaining sentences;

Relevant summary may contain redundant information, whereas a novel summary contains only sentences both relevant and novel.

Related Work

Topic Detection and Tracking (TDT):

for the purposes of new event, or first story detection. Concern is *event-based* novelty detection;

TREC Novelty tracks'02-04:

concerned with *topic-based* novelty detection. Here, the focus is novelty detection at the sentence level;

Summarisation and novelty detection:

Maximal Marginal Relevance (MMR) was used to reduce redundancy while maintain query relevance for retrieved document re-ranking and passage selection.

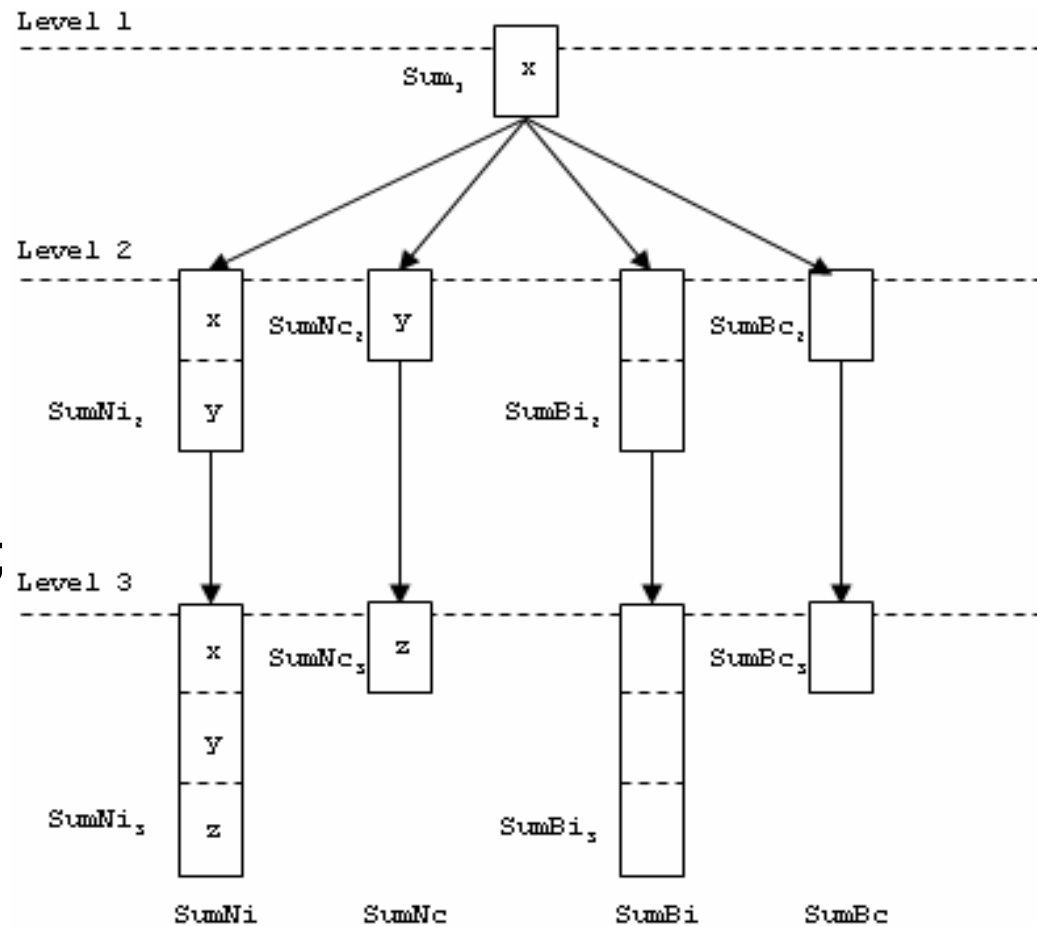
Framework to investigate novel summaries

Key decisions:

summary levels restricted to 3, primarily to avoid overburdening users;

title included to assist associating summary levels with the document;

query-biased summaries with setting of 7% of source document length.



Generating novel summaries (1)

Novelty detection:

simple word counting methods (*NewWords*) performed no worse than other more complex methods.

Process:

- start from list of sentences ranked according to relevance;
- establish a list of sentences to form history log of previously seen summary text, and a set of candidate sentences;
- generate a WordsSeen list from history log;
- compute novelty score for the candidate sentence;
- apply weighting to novelty score to emphasize over previous (relevance) scoring matrix;
- rank candidate sentences according to novelty + relevance score.

(cont...)

Generating novel summaries (2)

Process (cont.):

- pick N top scoring sentences to form summary, where:
 - $SumN_i / SumB_i$:
combination of the sentences taken from the history log, and the top N scoring candidate sentences;
 - $SumN_c / SumB_c$:
the top N scoring candidate sentences only.
- finally, sort summary sentences into position order as they occur in the original document.

Sample summary for a typical document

Level 1: Sum₁(Generic)

UN warns that arrears issue will again threaten US vote, future

- UNITED NATIONS (AP) _ Washington's continued failure to pay its bills will again threaten its vote in the General Assembly next year and will lead to a backlash against enacting U.S.-demanded reforms, the United Nations warned. (0)
- Secretary-General Kofi Annan, who has been outspoken in the past week in criticizing the United States, said in a statement that the U.S. Congress and administration had reneged on personal promises to pay its bills this budget season. (1)
- But Congress failed to act on a separate spending bill concerning the dlrs 1.3 million the United Nations says the United States owes in back payments. (5)

Level 2: SumNc₂

UN warns that arrears issue will again threaten US vote, future

- The United States now accounts for two-thirds of the outstanding U.N. arrears. (15)
- The United Nations has managed to keep its operations going by borrowing money from a separate peacekeeping fund once the regular budget runs out, usually in September. (16)
- ``Where we stand today is that a large number of other member states are underwriting the United States' dues in the United Nations by agreeing to permit us to borrow from peacekeeping funds that are really owed to them,' the official said. (19)

Level 3: SumNc₃

UN warns that arrears issue will again threaten US vote, future

- President Bill Clinton has threatened to veto the arrears bill because it contains a provision denying U.S. contributions to international family-planning organizations that advocate abortion rights. (7)
- In a related issue, Congress failed to allot any funding for the U.N. Population Fund _ a decision that will mean ``the unnecessary death and suffering of women who are deprived of the information and means to plan their families,' the agency's executive director, Nafis Sadik said in a statement. (8)
- Annan has suggested asking the General Assembly to decide whether it wants to continue the practice, but the issue hasn't been placed on the assembly's agenda yet. (20)

Level 2: SumBc₂

UN warns that arrears issue will again threaten US vote, future

- A senior U.N. official warned that the United States will have to pay some of its arrears next year if it wants to retain its vote in the General Assembly. (9)
- U.N. bills for 1999 alone won't cover the minimum payment the United States will have to make to keep the vote, meaning it will have to pay off some of the arrears, the official said. (10)
- The United States now accounts for two-thirds of the outstanding U.N. arrears. (15)

Level 3: SumBc₃

UN warns that arrears issue will again threaten US vote, future

- The United States is expected to pay the United Nations just enough this year to reach the minimum it needs to preserve its vote in the 185-member General Assembly. (3)
- Congress is expected to vote on a budget bill Tuesday and Wednesday that would authorize a dlrs 197 million payment that would bring the United States just over the threshold. (4)
- Lawmakers have made payment conditional on sweeping U.N. reforms that the United Nations says have largely been implemented. (6)

Experimental collection, measures & design

- Documents and queries taken from AQUAINT collection (TREC Novelty track) consisting of newswire stories from NYT and APW;
- 20 participants, 5 randomly selected queries, and 10 top-ranking documents per query, 3 summaries per document;
- Total of 150 document summaries (75 with novelty *SumN* and 75 without novelty *SumB*).
- Precision (P), Recall (R), Decision Correctness (DC) and Time spent;
- Experiment was conducted in such a way that each user group experienced using the different system settings.

Experimental procedure

- Users were presented with a retrieved document list in response to a simulated query (TREC topic) and tasked with identifying correctly relevant and non-relevant documents.
- Users were informed that their performance scores would be penalised if they made mistakes, to not biased quick decisions.
- The information presented for each document was the automatically generated summaries.

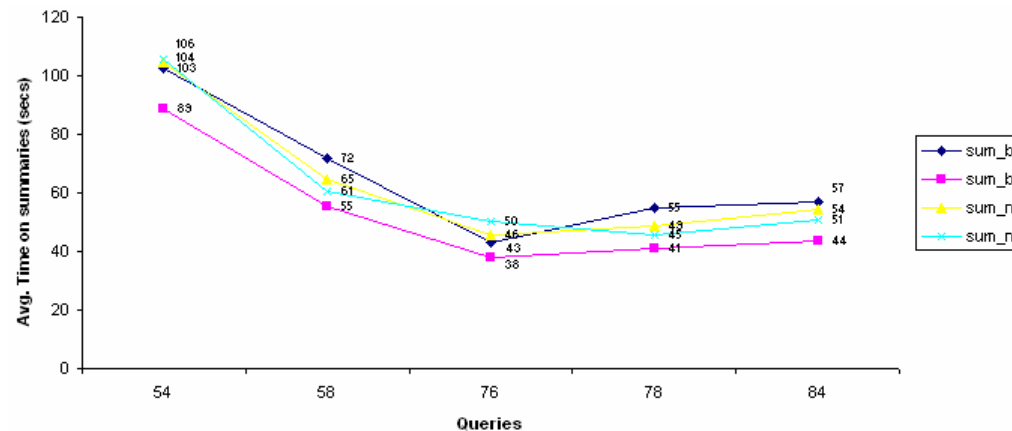
Results: Overall

- Little difference in performance (DC, P, R and in time spent viewing content) between the different summaries.

Group	Type	DC	P	R	Time (secs)
1 & 4	$SumB_i$	0.764	0.822	0.845	66
2 & 3	$SumB_c$	0.768	0.850	0.798	53
2 & 3	$SumN_i$	0.776	0.809	0.852	64
1 & 4	$SumN_c$	0.760	0.803	0.752	63

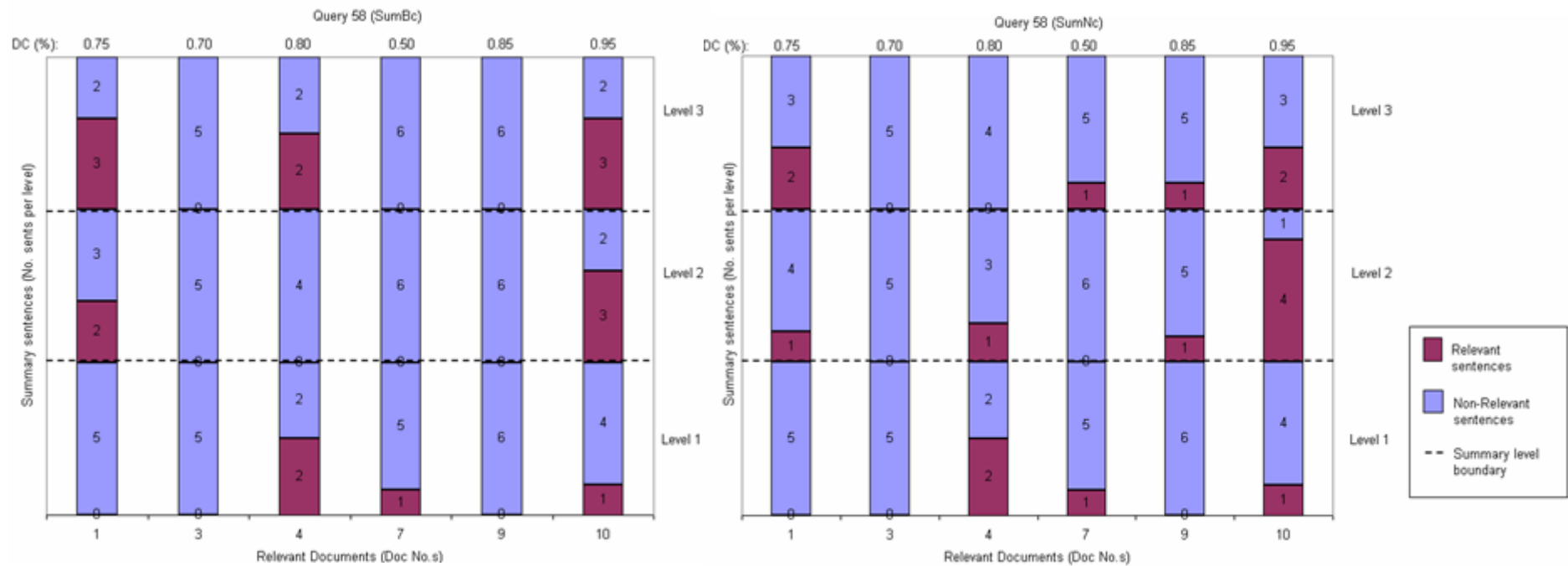
Results: Query level

- Pattern of improvement over the first query, performance levelling out for intermediate queries and a drop for the final query;
- Comparing queries in terms of the average time spent, the first query takes the greatest amount of time with a decrease in time spent on all other queries;
- Despite spending less time, users perform no worse in making relevance decisions for the later queries.



Results: Relevant sentences in summaries

- The percentage of relevant sentences in summaries does impact on users' accuracy in DC;



Conclusions

- Findings suggest that there is little difference in performance (DC, P, R and time) between the experimental summaries;
- Performance of the baseline summaries, for the task we set our users, questions the benefits of using novel summaries;
- However, this may be due in part to the high levels of performance achieved with the baseline summaries, and the novelty detection algorithm adopted;
- For mobile information access, given issues of bandwidth and screen size, short summaries that include only novel information, are most suited.

Future work

- Investigate performance of a more refined approach to novelty detection beyond a simple count of new words.
- Study the effects of permitting users to make decisions at any level, to investigate summary level preference, and impact decision accuracy.
- Investigate other dimensions of summarisation, extending the notion of ``Show me more'' to include personalisation.