ABSTRACT

Given that most of the temporal queries issued by users are implicit by nature [1], detecting their underlying temporal intent turns out to be a very interesting problem and a real need to improve the performance of search systems.

In this context, most state-of-the-art methodologies consider any occurrence of temporal expressions in Web snippets and other Web data, as equally relevant to an implicit temporal query. This is obviously not true for most part of the query results.

In this work, we aim to define the temporal intents of implicit temporal queries in order to further improve the Web search process.

For that purpose, we propose a language-independent strategy to associate temporal relevant years to any text query by analyzing its corresponding Web snippets.

The advantage of our approach is that instead of considering all the temporal expressions as equally relevant, as currently common in most of the TIR tasks, we determine which ones are more relevant to the user query.

The contributions of this work can be summarized as follows:

1. We propose a novel approach to properly tag text queries with relevant temporal expressions by relying on a text-based approach and a language-independent methodology.
2. We publicly release a set of queries and ground-truth results to the research community, hence our evaluation results can be compared to future approaches.

CONTRIBUTIONS

The results and discussion

We queried Bing search engine for each of the 42 queries, built aggregated results for the two new data sets and evaluated the performance of the two baseline methods. The results are shown in Table 7, where we computed the precision, recall, F1-measure, and AUC for each method, as well as the mean and median of these measures.

The results show that our method outperforms the baseline methods in terms of precision, recall, F1-measure, and AUC. The mean and median values for our method are higher than those for the baseline methods, indicating that our method is more accurate and robust.

EXPERIMENTAL SETUP

Since no benchmark for (q,d) pairs exists, we built two new data sets:

- **WC_DS**: 80 Web Snippets
- **Class_DS**: Query Snippets

**Class_DS** consists of 42 text queries each.

**WC_DS**

We queried Bing search engine for each of the 42 queries, collecting the top 50 relevant Web results, resulting in 235 distinct (q,d) pairs.

**QLog DS**

We used the QLog search engine to collect results for each query.

We conducted an analysis based on the results for the queries.

**QLog**

We built a confidence interval for the difference of means for paired samples between the number of misclassified dates given by the rule-based method and the BOTE.

**QLog BOTE**: [1.32; 3.20]

**QLog Class**: [1.44; 3.47]

Not surprisingly, results show that query logs are able to return a great number of potential query related years, when compared to Web snippets.

But, interestingly, we found that a large number of these temporally explicit queries consist of misleading temporal relations.

CONCLUSIONS

- In this paper, we proposed a new temporal similarity measure, the GTE, to compute the temporal intents of query dates (q,d) pairs.
- We showed that the combination of the second order similarity measure IS combined with the DICE coefficient shows improved results over all other combinations based on the threshold classification strategy.
- Comparative experiments have also been performed on two different data sets (WC_DS and QLog_DS). Results showed that the Web snippets approach is more effective than the query log based one.
- The results indicate that the introduction of an additional layer of knowledge, may affect the efficiency of a broad set of TIR systems.

REFERENCES