

Effectiveness of Temporal Snippets

Omar Alonso
A9.com
Palo Alto, U.S.A
oralonso@gmail.com

Ricardo Baeza-Yates
Yahoo! Research
Barcelona, Spain
rbaeza@acm.org

Michael Gertz
Institute of Computer Science
University of Heidelberg
Heidelberg, Germany
gertz@informatik.uni-
heidelberg.de

ABSTRACT

We introduce the notion of time-centered snippets, called TSnippet, as document surrogates for document retrieval and exploration. We propose an alternative document snippet based on temporal information that can be useful for supporting exploratory search. The idea of using sentences that contain the most frequent chronons (units of time) can be used for constructing document surrogates. We conducted a series of experiments to evaluate this new approach using a crowdsourcing approach. The evaluation against two Web search engines shows that our technique produces good snippets and users like to see time-sensitive information in search results.

Categories and Subject Descriptors: H.3.1 [Information Storage and Retrieval]: Content Analysis and Indexing - Linguistic processing. H.5.2. [Information Interfaces]: User Interfaces.

General Terms: Design, Experimentation.

Keywords: temporal information, snippets, summaries.

1. INTRODUCTION

Temporal information is available in every document either explicitly, e.g., in the form of temporal expressions, or implicitly in the form of metadata. Recognizing such information and exploiting it for document retrieval and presentation purposes are important features that can significantly improve the functionality of search applications. We are primarily interested in utilizing and presenting various time information embedded in documents as relevance cues to highlight the importance of a document given a query. It would be interesting to investigate a system that is aware of time for determining the relevance of documents and that is able to show results in a temporal context. It then can be useful to filter the “trendy stuff” from the rest.

Current snippets available in Web search engines present a couple of lines with highlighted keywords and some context. There are no complete sentences, only bits of text that one must mentally construct as a sentence.

In the early days of the Web, a summary consisted on the first 80 characters of a document. Nowadays modern engines use *query-biased summaries* or variations of *keyword-in-context*. Those summaries or sentence fragments contain query terms that are extracted from the document and are presented in the context in which the query terms appear. In contrast to the query-dependent summary, a well known approach is to provide a *generic summary*, which is a short fragment that represents the key points in a document.

One can argue that a document snippet that leverages temporal information would be an interesting alternative for some document

search and exploration tasks. Intuitively, it makes sense to include time in a snippet. For example, a well known fact about Einstein is that he won the 1921 Nobel Prize in Physics. We define a *temporal snippet* or *time-centered snippet* as a short fragment of text that contains two or three sentences with the most relevant chronons in chronological order. But how can one use all the temporal information that is available in a document to select relevant sentences that can be used to construct such a snippet?

In this paper we study how time information appears in documents and how it can be used to generate short summaries that can be used for relevance cues. We emphasize three themes: i) automatic detection and extraction of temporal information from documents, ii) construction of time-centric snippets as document surrogates, and iii) validation of our approach using experiments and evaluation.

Research on temporal annotations is very recent, and it is well covered in the book by Mani *et al.* [5]. Projects that are close to our work are the user studies on snippet length described by Kaiser in [2] and readability of short summaries by Kanungo in [3]. We use Amazon Mechanical Turk (MTurk, <http://www.mturk.com/>), a crowdsourcing platform for our evaluation study. Very recently, crowdsourcing has emerged as a viable alternative to conduct large scale evaluation of different types of experiments for a wide range of applications like relevance evaluation [1] and user studies [4]. This recent research has shown that MTurk results are reliable and very useful for gathering extra feedback, as we will see in the evaluations.

2. EVALUATING SNIPPETS

We start our research by performing a user study about snippets in Web search. The goal of this experiment is to determine whether users find snippets useful or not. The experiment consists of showing the user a snippet example and asking the following questions.

1. Do you think that current web search snippets are useful in your daily search activities?
2. Assuming you are performing a search. If the snippet shows the information that you were looking for, do you still click on the page/URL?
3. Please provide any examples where snippets didn't show adequate information for some query/task.
4. What other features would you like to see in snippets?

19 workers participated in this particular Human Intelligence Task (HIT). For question 1, 89% answered “yes”, and 11% “no”. For question 2, 68% answered “yes” and 31% “no”. As we can

see, even if the users find the snippet useful they still click on the page to get the right information. Of all the users, 78% provided feedback in question 3 (Table 1) and 84% entered comments for question 4 (Table 2). We manually examined and categorized all comments.

Category	User comment
Poor quality	If search terms are too generic, impossible to find information through snippet. Usually have to search through several sites to find the information - A common problem is having a snippet that is out of context. - There are some sites that basically give an introduction as their snippet, which is sometimes annoying. - I do not click on a link when the snippet just contains random keywords, especially if some of them are not at all related to my search. - When snippets are too general or when they don't support the title in any logical way.
Depth, sentence length	Not showing when a famous person was born/died. - When looking for profound information usually it doesn't all show up in the snippet - If it does not fully answer my question, I will click on the page. - If the full sentence didn't complete with the information I needed. I would then click into the site to read the rest of the sentence.
Metadata	If I'm looking for hotels close by the University I don't think I would be able to find any hotels by clicking on the links. - I remember seeing some results in a foreign language, but it does not specify which language it most likely is, so I don't know which language to translate from on my translator extension.

Table 1: Summary of user feedback for question 3.

In summary, users have concerns about sentence length, lack of time-sensitive information (time stamp or temporal data inside the document), and the quality in some cases that is too generic. In the next section, we introduce temporal snippets that address some of those issues.

3. TEMPORAL SNIPPETS

A key idea underlying our approach is to extract temporal information from documents and to make this information explicit for further processing, in particular the construction of temporal snippets. In the following, we first present the notions of time and timelines, followed by a discussion of the concept of temporal expressions and the extraction of such expressions from documents.

3.1 Chronons and Timelines

As our base timeline, we assume a consecutive sequence of day *chronons*, such as “April 15, 2008; April 16, 2008; April 17, 2008” and so on. Contiguous sequences of chronons in such a base timeline can be grouped to obtain more coarse-grained *timelines*, which are then based on weeks, months, years etc. In the following, we assume a set $\mathcal{T} = \{T_y, T_m, T_w, T_d\}$ of timelines constructed based on year, month, week, and day chronons, respectively.

The timelines and chronons associated with timelines described above now provide us with a basis to precisely “anchor” temporal information that is embedded in documents in the form of temporal

Category	User comment
Time information	Time and date, last time the web site was updated. - Publication date or last time search engine crawled the site.
Sentence length	It might be good if they were a line or two longer because sometimes the info in the snippet is cut off and hard to understand. - Maybe scrolling in snippets - Full sentences. - A way to enlarge the text field and read more of the snippet - It might be neat if people could customize the number of lines for each snippet.
Metadata	The main focus of the web site (retailer, information, directory, etc). - I would like to see information about school newspapers, students' activities, etc. So if I want to check out any of these items at the following links, I will know that I can. - I guess I'd like to know if there is some hidden cost involved in the web site like with the surveys or downloading music and paying a monthly fee things like that - Some sort of user rating of the web site with regard to both the web site and the search term used in conjunction.

Table 2: Summary of user feedback for question 4.

expressions. Such expressions are recognized by an entity extraction approach using a time-based linguistic analysis. Expressions can be mapped to temporal entities and terms defined in some temporal ontology. Following the framework proposed by Schilder and Habel [6], we distinguish between explicit, implicit, and relative temporal expressions.

Explicit Temporal Expressions. Explicit temporal expressions describe chronons in some timeline, such as an exact date or year. For example, the text “October 2008” is an explicit temporal expression anchored in the timeline T_m .

Implicit Temporal Expressions. Depending on the capabilities of the entity extraction approach and in particular its underlying time ontology, implicit temporal expression, such as names of holidays or events can be anchored in a timeline as well. For example, the token sequence “Columbus Day 2006” in the text of a document can be mapped to the expression “October 12, 2006”, which is then anchored in the timeline T_d . In general, implicit temporal expressions require that at least a year chronon appears in the context of a named event.

Relative Temporal Expressions. Relative temporal expressions represent temporal entities that can only be anchored in a timeline in reference to another, already anchored temporal expression. That is, their anchoring depends on a chosen point of time reference or narration. For example, the expression “today” alone cannot be anchored in any timeline. However, it can be anchored if the document is known to have a creation date as a reference. Then it is likely that the expression can be mapped to that date. There are many instances of implicit temporal expressions, such as the names of weekdays (e.g., “on Friday”) or months (e.g., “in April”) or references to such points in time like “last Saturday”.

In this section, we describe a technique for constructing a time-centered snippet that leverages temporal information. This snippet can be used as preview that outlines the main events described in a document. Briefly, the temporal snippet aims at presenting the most relevant sentences that contain time information using the extraction technique defined in the previous section.

Assume a document d from a document collection \mathcal{D} . The following temporal snippet algorithm called *TSnippet* determines the top n ($1 \leq n \leq m$) sentences from a list $S_d = [s_1, s_2, \dots, s_m]$ of m sentences in the document. The algorithm consists of two parts: sentence candidate selection and sentence ranking.

Sentence Candidate Selection. The most obvious attempt for generating a temporal snippet is to select the first sentence that contains a temporal expression that is not the document timestamp. Surprisingly, the result for this naïve approach is not that bad for documents that have a good structure. Unfortunately, this simple approach cannot be generalized for all types of documents due to the following aspects:

- Photo captions: a document may include a figure or photo in the first couple of sentences, like `Einstein in 1947`.
- Very long sentences: if the sentence detector fails to identify a document object like a table, it would produce a very long sentence that contains every cell item. This is not semantically correct, and in the event of a chronon in a cell, this incorrect sentence is selected as snippet.
- The first sentence does not contain an explicit temporal expression but only a relative one.

We therefore define the following criteria for selecting candidate sentences from a given document $d \in \mathcal{D}$:

- The sentence must contain at least one temporal expression.
- All temporal expressions in the sentence must be explicit expressions.
- The sentence length is greater than *sentence-min* and less than *sentence-max*. The parameters *sentence-min* and *sentence-max* should be adjusted according to a particular document collection.

What is now needed is a ranking function that, given a number of candidate sentences with temporal expressions, produces a sorted list of relevant sentences.

Sentence Ranking. How can one define a meaningful ranking function for a set of candidate sentences from a document d ? One approach would be to use the sentence length and a low ratio of position/length as parameters. Among other solutions, one is to select the sentence that contains the most frequently occurring chronon in d and which is also the first sentence that contains such chronon d .

A much better approach is to identify the features that are interesting and to define a ranking function that combines these features. In our approach, we select the following features:

- p : position of the temporal expression within the sentence (ratio of position/sentence length).
- s : sentence number.
- sl : sentence length in tokens.
- co : chronon appearance order (e.g., first, second, etc.).
- cf : chronon frequency in document.
- cfs : chronon frequency in sentence.

The ranking function is a linear feature combination where each parameter is tunable:

$$rank(d, q) := \alpha p + \beta s + \gamma sl + \delta co + \epsilon cf + \epsilon cfs + \cos(d, q)$$

Making it query dependent is a matter of adding a similarity measure to the overall ranking function so that it does take into account the presence of terms in the final result. For this, we add the cosine similarity feature to the linear combination.

Based on the components described above, the *TSnippet* algorithm thus can be summarized as follows:

1. Segment the document d into sentences s_1, \dots, s_m , and treat each sentence as a separate document.
2. Select the candidate sentences according to sentence selection criteria.
3. Apply the ranking function to selected sentences.
4. Select the top- n sentences sorted by chronon as temporal snippet.

The next obvious step is to evaluate the quality of our new snippet against Web search engines.

4. EVALUATING TSNIPPETS

The goal of this experiment is to evaluate the quality of temporal snippets. In this experiment, we selected 10 random informational queries from Wikipedia featured articles as follows: DNA, European Commission, FairTax, French Texas, German Occupation Luxembourg World War I, Hey Jude, Hurricane Juan, City of Manchester Stadium, Super Nintendo Entertainment System, Pluto. A HIT in MTurk, as presented in Figure 1, consisted in presenting 3 snippets (*TSnippet*, Google, and Yahoo) and ask workers (5 workers per HIT) the following questions:

1. Based on the snippets presented in the previous section, do you think a snippet that contains time/temporal information is useful in general for Web search? Please answer according to the content of the snippet (text between Title and URL).
2. Based on the snippets presented before, which search engine covers adequate temporal information about the [query] page on Wikipedia?
3. If the snippet shows the information that you were looking for (facts, etc.), will you be inclined to click on the page/URL?
4. Please provide any feedback about document snippets.

For question 1, 90% of the workers answered “yes” (Yes, it is useful. I can see some facts/events that help me understand how relevant the document is) and 10% answered “no” (No, it is not useful. I don’t see how time information can be useful to assess how relevant the document is).

For question 2, workers selected *TSnippet* 80% of the time and Yahoo 20%. Of the 80% were *TSnippet* was selected, 38% was a single choice, 50% was tied with Yahoo and 12% tied with Google.

Question 3 was unanimously answered in a positive way (99% yes), which implies that no matter how good the snippet is, people still click on the page most of the time to satisfy their information needs. A summary of the feedback for question 4 is in Table 3.

We also conducted a follow-up questionnaire about scenarios for *TSnippets* on MTurk. Workers considered a good alternative for ordinary snippets in case of time-centric search and exploration such

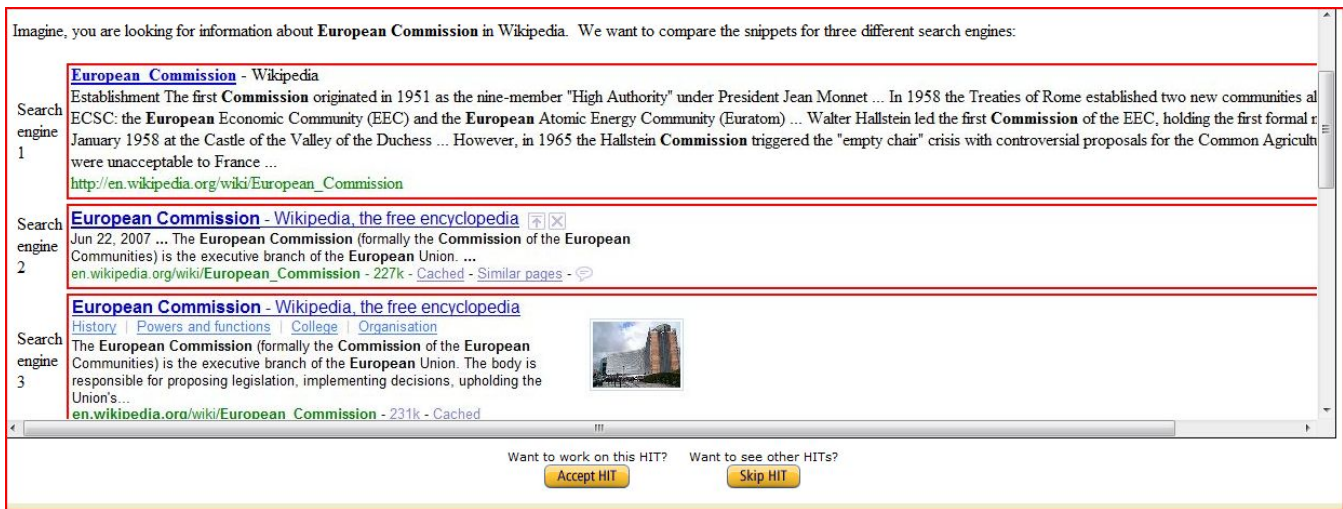


Figure 1: Example of a HIT in MTurk for the query [European Commission]. The user has three snippets to evaluate.

task that include, “historical research”, “doing research” or “detailed search”.

It also important to note the effect on evaluation due to the change in snippet. We ran the same experiment nearly a year ago, and Yahoo had the worst numbers. The fact that they have improved their snippet shows on the evaluation.

Category	User comment
Time information	I tend to wonder how 'current' information is if I see that a document, for e.g., was created in 2002 and we're now in 2009, but I might click anyway if it looks relevant to my research. - The best snippets should distinguish between the date of the article and the dates in the article. In the example above, the date of 2009 is irrelevant, but it could be important if you are looking for the most recent article. - Search engine 1 had more specific information about the Commission, which I would expect to see after I clicked on the URL; however, Search engine 2, having a date in the beginning, let's me know if the information that I'm about to see is recent, which is important for an accurate search.
Presentation	Very interesting to see how the information is presented differently. - I like that you can choose history etc only if you want to. - I find snippets very useful in general, they sometimes provide the information I'm looking for right on the results page, saving me an additional page load.

Table 3: Summary of user feedback for TSnippet.

5. CONCLUSIONS AND OUTLOOK

We have presented a technique for constructing temporal snippets as document surrogates. By studying how temporal expressions appear progressively in a document, we designed a temporal snippet algorithm called TSnippet that provides useful summaries, as demonstrated by our evaluation. Temporal snippets are an alternative to present time referencing summaries and should be viewed

as complementary and not as a replacement to current snippets. Ideally, we would like a search engine to be aware of the temporal information and present search results in a temporal context.

Our study demonstrates that TSnippet can be an interesting alternative for constructing document previews that highlight temporal information. We showed that our technique works and users like it when it was compared against existing Web search engines. More research needs to be done regarding the experiment design for MTurk. We have shown one approach and many other UI factors can be incorporated.

Time information (temporal and/or creation date) is very important according to MTurk workers. In some cases, creation date is key. In others, easy facts can be presented on the search results page without requiring users to click on the URL. This shows that time information is an indication for users to make decisions about how good a snippet is.

6. REFERENCES

- [1] O. Alonso, D. E. Rose, and B. Stewart: Crowdsourcing for Relevance Evaluation. *SIGIR Forum* (42):2, 12–18, 2008.
- [2] M. Kaisser, M. Hearst, and J. Lowe: Improving Search Results Quality by Customizing Summary Lengths. In *Proc. of Association for Computational Linguistics (ACL-08): HLT*, 701–709, 2008.
- [3] T. Kanungo and D. Orr: Predicting the readability of short Web summaries. In *Proc. 2nd ACM Int. Conf. on Web Search and Data Mining (WSDM)*, 202–211, 2009.
- [4] A. Kittur, E. H. Chi, and B. Suh: Crowdsourcing user studies with Mechanical Turk. In *Proc. 26th SIGCHI Conference on Human factors in computing systems*, 453–456, 2008.
- [5] I. Mani, J. Pustejovsky, and R. Gaizauskas. *The Language of Time*. Oxford University Press, 2005.
- [6] F. Schilder and C. Habel: From Temporal Expressions to Temporal Information: Semantic Tagging of News Messages. In *ACL'01 Workshop on Temporal and Spatial Information Processing*, 2001.