Search Result Summaries Improved by Structure and Multimedia

Brent Wenerstrom and Mehmed Kantardzic
Department of Comp. Eng. and Comp. Sci., University of Louisville, Louisville, Kentucky, U.S.A.

Keywords: Structured, Automatic, Summaries.

Abstract: We previously introduced ReClose which provides summaries with both better content and better visual display for search engine results. We now seek to further improve summaries with the addition of structured text and multimedia, more specifically tables, lists, buttons and images. Currently search engine provided summaries rarely use structured text and images. We show in this paper that structured text and images lead to faster comprehension by search engine users and lead to visually more appealing summaries. 70% of non-expert users made decisions more quickly using summaries preserving document structure and 65% of all users preferred summaries preserving structure to plain text summaries.

1 INTRODUCTION

The inspiration for this study comes from the jumbled mess of text that is shown to a user when summary text is extracted from a table. For example consider Figure 1(a) in which the summary text shows a table of football players invited to the NFL combine. A search engine user may correctly guess that the text comes from a table, or one may guess it is only comma separated values shown in the original document. Compare now the impression given by viewing the same text as a table as the original author of this text meant it to be viewed as in Figure 1(b). In the summary preserving structure there is no mistaking the authors intent of the summary text.

We previously proposed our system ReClose (Wenerstrom and Kantardzic, 2011b) for summary generation. ReClose improved summary generation through a two part summary including query-biased and query-independent portions. This was further enhanced by adding color depth to keywords (Wenerstrom and Kantardzic, 2011a). Together these enhancements to summarization led users to be 20% more accurate when selecting relevant documents using ReClose summaries compared to using using Google summaries. We propose to extend ReClose by incorporating document structure into summaries.

In addition to tables we also explore the usage of lists, images, buttons, text fields and hyperlinks. In the first example (Table 1) the purpose of the sentence "iPhone iPad Android" to be three radio buttons so that one may search based on different devices.

From Figures 1(a) and 1(b) we see that structured

Official Invite — NFL Combine
www.nflcombine.net/players/official-invite-list
jr: player: goes by: school: camp # pos: acho, emmanuel: texas: lb01: ob: adams, adam: joe: arkansas: wo01: wo: adams, michael: mike: ohio st: ol01: ot: alecxih ...

(a) Traditional search engine result observed in Bing.

Official Invite — NFL Combine
www.nflcombine.net/players/official-invite-list
Jr

—Player	—Goes by	—School	—Camp #	—Pos—
—ACHO, EMMANUEL—		—TEXAS	—LB01	—OB —
—ADAMS, ADAM	—JOE	—ARK.	—WO01	—WO—
—ADAMS, MICHAEL	—MIKE	—OHIO ST—	OL01	—OT —
—ALECXIH,	—	—	—	— —

(b) Summary preserving document structure example.

Figure 1: Search engine summary with (b) and without (a) structure.

text and multimedia adds insight not present in current search result summaries. We propose the addition of structured text and multimedia to current search results which will have the following benefits for users:

- Additional semantics added to search results.
- Users will have a better understanding of the summarized web pages.
- Users will be better satisfied with search results.

This work offers two contributions. First it provides an alternative basic unit to automatic summarization in search. Rather than all forms of summarization being based only on text, this work expands that view to include HTML entities such as tables and lists that contain text. Second, this work provides ex-

Wenerstrom B. and Kantardzic M..
Search Result Summaries Improved by Structure and Multimedia.
DOI: 10.5220/0004117901810186
In *Proceedings of the International Conference on Knowledge Discovery and Information Retrieval* (KDIR-2012), pages 181-186
ISBN: 978-989-8565-29-7

Table 1: Comparison of summaries from a traditional search engine versuse summaries with structure.

Buttons and Text Field	Chomp. Download Chomp iPhone Android. iPhone iPad Android	Chomp Download Chomp iPhone Android iPhone iPad Android
Images and Lists	tenerife property sw1. exclusive land for sale; exclusive land for sale; exclusive land for sale; exclusive land for sale; exclusive land for sale; exclusive land for ...	*Tenerife Property SW1*  exclusive land for sale  exclusive land for sale  exclusive land for ...

perimental validation that automatic summarization systems which output summaries with structured text and multimedia in addition to unstructured text outperform unstructured text alone.

2 SUMMARY FROM HTML

Tranditional search engine summaries coming from the likes of Google, Yahoo and Bing generally are made up of plain text. This text often is focused on the usage of the searched keywords within a document.

In this paper we will mainly focus on HTML documents, though the ideas extend to all types of web documents. An HTML document is made up of HTML tags and text content. These HTML tags provide text formatting, text structure, hyperlinks, meta data about the document and more. Traditional search engine summaries when extracted from web page content only use the plain text content of the document.

Traditional search engines currently extract text from image descriptions encoded in the HTML tag of images. When testing the usefulness of multimedia and structure we replace text extracted from images, with the original image.

HTML tables may be used to align text or to display tabular data. The classification of HTML tables into these two categories is not herein described. We retain table structure in our experiments for tables found to display tabular data. When table data is extracted without structure meanings can be lost and the true value of the content may become fuzzy as was observed in Figure 1.

Text taken from a list without the list structure may in fact change the meaning of the content. For example on a web page describing CMMS software two bullet points were observed as “work order management” and “preventative maintenance”. When the structure of the list was lost through a search engine’s extraction of content for summary, it was no longer clear that there were two separate entities. On first observation it was difficult to understand what “work order management preventative maintenance” was. Preserving lists in summaries retains the original meaning of web page authors.

When text appears on buttons or hyperlinks it has different meaning compared to plain text. For example a button that reads “Sign Up” shows that the action is immediately available, while in plain text where and how to sign up are not apparent. Hyperlinks show that there is more information available on a linked document. Some web pages provide value by linking to high quality content. When a web page is summarized and text is not shown as a hyperlink the value of these linking documents is lost on the search engine user. We preserve buttons and hyperlinks in summaries.

2.1 HTML Parsing

For the purpose of experimentation we take previously selected text from a web page and injects structure and multimedia elements found in the original HTML web page. We do this by parsing the original HTML web page into a convenient data structure which contains a string representation of the available text content. The chosen summary text is matched against the string, then our data structure injects the HTML that was found between and surrounding the sequence of characters found to match.

Parsing an HTML web page involves using regular expressions as found in Java to loop through each HTML tag and process the text content followed by the HTML tag. Each HTML tag is processed to remove unnecessary formatting. The text content is also preprocessed. The string representation of the text content removes all non-letter, non-number characters and changes all letters to a uniform lower case. Search engines may use a different letter case than

the original web page, add or remove space characters and change punctuation. One reason punctuation is changed by the search engines is to simulate table or list boundaries found in the HTML.

When a substring of text content matches the chosen summary text, then a new summary is produced by going character by character and adding to a buffer the characters referenced by that character information object. If the character is the first character, then preceding non-letter, non-number characters are added first, like a starting quotation mark. Then each middle character adds what characters and tags may lie between each of them, such as space characters and HTML tags that we intend to reproduce. When the last character is reached, than ending context is also added like the ending punctuation. This may be repeated if multiple text spans are found, in which case ellipses (. . .) are added between text spans.

This whole process produces near identical text to the text submitted with the addition of HTML tags that provide text structure and multimedia.

3 RELATED WORKS

The research community has proposed numerous approaches to displaying search results. Some of these approaches have been more radical such as documents appearing in a 2-dimensional particle space (Kelkar et al., 2009) or using word clouds to represent results (Dörk et al., 2009). While other approaches have aimed to make small changes to the summaries such as adding images and/or lists to summaries (Joho and Jose, 2008).

Several studies have focused on the inclusion of images in search results. (Xue et al., 2008) found users had higher precision and recall scores when labeling search results using summaries that include both summary text and images. (Li et al., 2008) showed users to be 30% faster at informational searches when search results included images. (Dziadosz and Chandrasekar, 2002) found users using thumbnails took longer to make relevance decisions but were more accurate.

Other approaches have been proposed at providing context beyond thumbnails and images. (Wenerstrom and Kantardzic, 2011b) proposed a two part text summary, named ReClose, which includes both the traditional query-biased summary and also a web page summary, providing additional context. This ReClose system was then enhanced with color-coding (Wenerstrom and Kantardzic, 2011a). The extent of keyword usage and warnings about topic departure were encoded with words highlighted in different shades. The result was the color-enhanced ReClose summaries leading to users that were 20% more accurate than when using Google summaries.

White et al. (White et al., 2001) add the number of links, name of first non-text object and the size of the document to search results. Dörk et al. (Dörk et al., 2009) proposed visual widgets for representing search results. These widgets included maps, timelines and word clouds.

4 USER EVALUATION

We tested the effectiveness of structured text and multimedia on a number of search engine users. The evaluation process was made up of two major steps. The first step involved users reviewing 30 search results one at a time. Each search result displayed a unique search query, search question and search engine summary. The user had three options to use when answering if they "should" click on a result: yes, no or unsure. Responses and timing information was saved per summary. The second step asked which system a user preferred with the only options being the summaries preserving document structure or traditional summaries.

The search queries used were taken from two main sources. First 150 queries were obtained from the website SearchHippo[1] on Feb. 14 and 22. Second we obtained the top 20 trending queries on Google Hot Trends[2] on Feb. 22, 23, 24 and 28. From these queries, duplicates and most non-English queries were removed. Our final count was 185 queries.

All 185 queries were submitted to a traditional search engine through Bing's Search API[3]. We observed the URLs and summaries for the top 40 search results for each of the 185 queries and downloaded the linked web pages. We were able to download 7,214 web pages. Of these results only 440 summaries currently contained one of the target HTML formating tags. From these 440 we removed those where the formatting would not be noticeable, i.e. small images, blank images, tables used for text placement or small tables. We then randomly chose 30 summaries from each of four groups: forms, tables, lists and images. There were only 10 usable form summaries. Some of the summaries used HTML from multiple groups. The final pool of potential summaries included 94 summaries.

Each user within our evaluation was shown 30 summaries randomly selected from our pool of 94.

[1] http://www.searchhippo.com

[2] http://www.google.com/trends/hottrends

[3] http://www.bing.com/toolbox/bingdeveloper/

15 of these summaries were only made up of text and came directly from Bing. 15 of the summaries were shown to the users with structured text and/or multimedia present from the original web page. The 30 search results were randomly ordered on a single page mixing the ordering of Bing summaries with structured summaries.

We used two pools of people expecting potentially different responses from each group. Our first group was made up of those who effectively think in the abstract. The first pool is made up of 22 students from the CECS department at the University of Louisville. CECS students can be effective at viewing source code and placing that source code in mental models of a system. This form of abstract thinking we do not expect to be common in the general population. The second group, who represent the more general population, we obtained by sending out approximately 70 invites through Facebook to acquaintances and friends of friends. Our invites resulted in 34 participants for our user evaluation.

The user evaluation was posted online. Users completed the evaluation using their own computers outside of a lab setting to better recreate more natural search engine usage among our subjects.

5 EXPERIMENTAL RESULTS

We evaluated the effectiveness of structured text and multimedia within search summaries in three dimensions: user time, user accuracy and user preferences. The values in each of these dimensions have provided evidence for the usage of structured text and multimedia.

5.1 User Decision Time

The more quickly a user can find desired answers or resources the more successful the search engine has been and the more satisfied search engine users will be. The more quickly the decision can be made without losing accuracy, the better quality the results are. From the user evaluation conducted, we have obtained 1650 click decision. Each click decision provides information on whether a user would click or not and how long it took to decide.

From the results we see a broad range of timings for each click decision from 1.3 seconds to 2.8 hours and a median time of 13.3 seconds. The 2.8 hour click decision time is without a doubt an outlier and was removed from our time analysis data. We also removed from our time analysis the other 15 timings with values higher than 78 seconds, which were rare

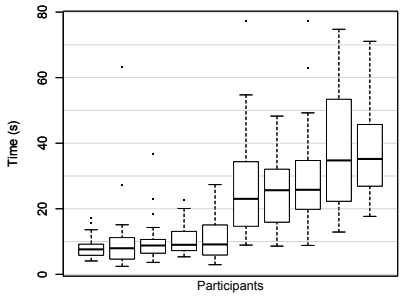


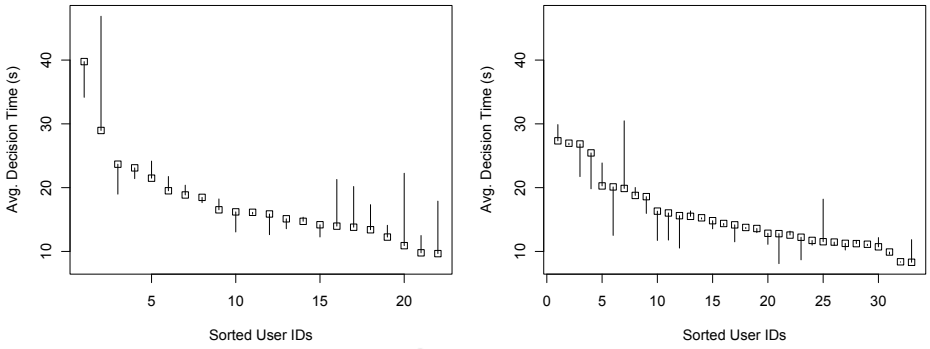
Figure 2: Box plots of the participants with the five lowest median timings and the five highest median timings.

(16/1650), seem to be extreme values for the decisions made and are more than 3 standard deviations from the mean.

We see large variations in the timings of each individual. Consider Figure 2 where we show box plots of the users with the five smallest medians and the five largest median timings. Here we see that there are individuals where the majority of their timings fall below 20 seconds. While there are others that carefully consider their options and have a majority of their timings over 20 seconds. This shows that it is important to consider the variation in time due to individual speed. This is countered by having all users recorded making decisions for both types of summaries.

When we separate out the two groups of experienced (computer science group) and inexperienced (Facebook group) abstract thinkers, we see very different timings. In both groups we compared the timings when using structured summaries to text only summaries. We did this by comparing averages. To avoid the long right tail heavily impacting averages, we first took the logarithm.

First we compare decision times in the computer science group. This is shown in Figure 3(a). The square in this chart represents the average decision time when users are using plain text summaries. A line is drawn from each square to the average decision time for structured summaries. If the line goes up, then plain text summaries were faster for that user. In the computer science group 9 users were faster using structured summaries, while 13 were faster using plain text summaries. From Figure 3(a) we can see that for most users there was not a big difference between the two. However, there is a noticeable difference for the slowest two and the fastest 7. The fastest group or the 7 with the lowest average decision time, every one of them was faster with plain text summaries. For the power users who make blink decisions using abstract thinking, they were much faster using plain text summaries. Except for a couple extreme cases little other observed differences were present.



(a) Computer science students.

(b) Facebook group.

Figure 3: The average decision time of users using plain-text summaries are drawn as boxes. A line is drawn from each box to the average decision time for summaries using structured text.

A paired t-test was performed to check for a statistical difference, and with a p-value of 0.11, none was found.

We see a very different effect among the Facebook pool, see Figure 3(b). A large majority of the Facebook pool were faster using structured summaries or 23 of 33. Many of the differences were very small, but there are more noticeable lines pointing down in Figure 3(b). First we compared the difference in values using a paired t-test and there was not a significant difference at the alpha=0.05 range with a p-value of 0.096. There is a significant difference (p-value=0.041) using a Wilcoxon signed rank test pairing the averages of users. Visually, it appears that slower users were most positively impacted using structured summaries among the inexperienced abstract thinkers of the Facebook pool.

5.2 User Accuracy

When a user clicks a link that leads to an irrelevant web page, that user has wasted time and this wasted time may negatively reflect on a search engine's results. When judging the impact of structured summaries on decision making we needed to judge the accuracy of the click selections. We had a judge go through and create first the context of each query such as for the query *katherine jenkins* the query question was "Who is Katherine Jenkins (singer)?" Decisions were made based mainly on how well the web page answered the question posed in the search question. In the pool of summaries, nearly 50% are relevant and 50% are not relevant, meaning that guessing all relevant would result in a low accuracy.

Accuracy was little affected by the introduction of structure and multimedia in the summary. Accuracies by summary type and by group are shown in Table 2. In this table we see that the more common search engine user, those inexperienced in abstract thinking were more accurate using structured summaries, while the abstract thinkers were less accurate. The inexperienced abstract thinkers were significantly more accurate than was the experienced group using Fisher's exact test with a p-value of 0.038. Perhaps the experienced abstract thinkers were overly confident in their mental models, and took time to ignore the introduced structure, leading to lower accuracies.

Table 2: Comparison of accuracies between summary types and evaluation groups.

Pool	Structured	Plain Text
Experienced	65.3%	70.0%
Inexperienced	**72.8%**	71.7%

5.3 User Preferences

Overall 67% (37/55) prefer the summaries preserving document structure to plain text summaries. The per group preferences are shown in Table 3. Table 3 shows that 17 out of 22 (77.3%) experienced abstract thinkers preferred the structured summaries. At the same time 20 of 33 (60.6%) of the inexperienced abstract thinkers preferred the structured summaries. This is a much higher, but not significant, percentage of experienced abstract users preferring structured summaries. Perhaps, experienced abstract thinkers see the value in adding structure, though their habits do not yet reflect it. Part of the inexperienced pool does not like change and would prefer the summaries to remain familiar, though they comprehend them more quickly.

We received 38 comments, though it was optional, providing a lot of feedback about users' thoughts and reactions. Tables and images received the most attention from users. Some users really liked table formatted data such as in the comment, "Tabulated data is ALWAYS easier to read in **table** format…" One reason for liking or disliking tables is the speed at which

Table 3: Preferences of experienced and inexperienced evaluation groups. Bolded values are significantly better than 50% using a binomial test and alpha of 0.05.

Pool	Structured	Plain Text
Experienced	**77.3%**	22.7%
Inexperienced	60.6%	39.4%

tables are comprehended. One participant wrote, "Because the information is organized in a **table** or has illustrations, you can determine more quickly how relevant the info is." However, one participant felt that tables added additional mental work as was mentioned before, "**tables** require me to decipher another layout. It's slower than just reading the information without extra images, lists, etc."

Images received a similar reception. Some users really liked the images, "Also, the embedded **images** were useful in determining the content of the web page." Another user felt that images were a big distraction, "...however I noticed that the instant I saw an **image** my eye jumped immediately to the url with the **image**, bypassing the one I was viewing. I needed words to be sure it was what I was looking for."

Overall users preferred summaries when possible to include structure and multimedia. As one user put it, the summaries preserving document structure were "More intuitive and easier to read and understand." There will always be those that are familiar with the current approach and would prefer no change take place, but these users queried in our evaluation overall would benefit from the addition of structured text and multimedia to search engine summaries.

6 CONCLUSIONS

In this study we introduced an extension to our current search engine summary generation system, ReClose (Wenerstrom and Kantardzic, 2011b). This extension introduces the use of structured text and multimedia in the form of HTML tables, lists, text fields, buttons, hyperlinks and images. We compared these additions experimentally to traditional summaries obtained from Bing.

The results of our users evaluation suggest that users will comprehend search results more quickly, lose no accuracy and prefer the structured summaries to plain text. From our users evaluation a significant number (23 of 33 or 69.7%) of inexperienced abstract thinkers were faster using structured summaries. While the experienced pool was faster (13/22 or 59.1%) using traditional summaries. No significant difference existed between the accuracies of user choices between traditional summaries and structured, multimedia summaries. Overall a significant number (37 of 55 or 67.3%) of users preferred structured, multimedia summaries to the traditional, plain-text summaries. We now plan to explore the use of HTML entities as the basic unit of summary generation rather than text entities like sentences.

REFERENCES

Dörk, M., Williamson, C., and Carpendale, S. (2009). Towards visual web search: Interactive query formulation and search result visualization. In *Workshop on Web Search Result Summarization and Presentation*.

Dziadosz, S. and Chandrasekar, R. (2002). Do thumbnail previews help users make better relevance decisions about web search results? In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '02, pages 365–366, New York, NY, USA. ACM.

Joho, H. and Jose, J. M. (2008). Effectiveness of additional representations for the search result presentation on the web. *Information Processing & Management*, 44(1):226 – 241. Evaluation of Interactive Information Retrieval Systems.

Kelkar, S., John, A., and Seligmann, D. (2009). Visualizing search results as web conversations. In *Workshop on Web Search Result Summarization and Presentation*.

Li, Z., Shi, S., and Zhang, L. (2008). Improving relevance judgment of web search results with image excerpts. In *Proceeding of the 17th international conference on World Wide Web*, WWW '08, pages 21–30, New York, NY, USA. ACM.

Wenerstrom, B. and Kantardzic, M. (2011a). Enhanced visualization for web-based summaries. In *Very Large Data Search (VLDS) 2011 First International Workshop on Searching and Integrating New Web Data Sources*, pages 38–43.

Wenerstrom, B. and Kantardzic, M. (2011b). ReClose: Web page summarization combining summary techniques. *International Journal of Web Information Systems*, 7:333–359.

White, R., Jose, J. M., and Ruthven, I. (2001). Query-biased web page summarisation: a task-oriented evaluation. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '01, pages 412–413, New York, NY, USA. ACM.

Xue, X.-B., Zhou, Z.-H., and Zhang, Z. M. (2008). Improving web search using image snippets. *ACM Transactions on Internet Technology*, 8:21:1–21:28.