

Enhancing Relevance Estimation of Source Document via Query-Biased Source Document Summary

Mudassar .I. Sayyed¹

Dr. Sarika Sharma²

ABSTRACT

Aim of this paper is to investigate if query-biased source document summary enhances relevance estimation of source document. For this purpose a summarization system has been developed, and a query-biased source document summary is generated automatically for every distal source document connected to current page via hyperlink. User estimates expected gain and cost of hyperlinked information via browsing cues present on and around hyperlink. The system presented in this paper provides additional automatically tailored browsing cues to users in the form of summary so as to assist users in relevance estimation of source document.

To investigate the relevance estimation capabilities of summary a within subject study was conducted where by users performed certain tasks on experimental summarization system and on a baseline system. The performance of user for relevance estimation parameters on both the systems was compared. The results from the evaluation indicate that query-biased source document summary enhances the relevance estimation of source document.

Keywords – Query-biased, summarization, relevance estimation, complementing hyperlink, browsing cues.

I. INTRODUCTION

Browsing is exploring the World Wide Web by following one hyperlink to another. At every encountered hyperlink a tradeoff has to be made based on two questions [1]:

What gain can I expect from a specific information nugget?

What is the likely cost to discover and consume that information?

1. Assistant Professor, Imscd & r, Ahmednagar, S.P. Pune University

Email: sayyedmudassar@gmail.com Ph: 9823100756

2. Professor & Director, JSPM's Eniac Institute of Computer Application, Wagholi, S.P. Pune University

Email: sarika4@gmail.com, Ph: 020-65101860

Both of these questions involve estimates which user can make either from experience or from design cues. F

or estimating from experience familiarity of search task is necessary. If search task is novel user has to estimate via browsing cues. Browsing cues[2] include *title* of hyperlink, *text* around hyperlink, and any other web component used to waft out information scent[3] of source document through hyperlink. Browsing cues are the only means present on current page to estimate relevance of distal source document. There are two problems associated with browsing cues:

- No guidelines are present for framing browsing cues
- Browsing cues are static in nature

In absence of any general guidelines browsing cues are usually distorted and pose challenge in estimation of expected gain and likely cost of consuming information. For user to make informed decisions cues should be designed carefully or else user demonstrate random behaviour; walk over or click to see if desired information is available. Browsing cues are static, consider the following hypothetical example: A searcher is searching for 'relevance' on Wikipedia.org. A webpage titled 'relevance' is displayed on which following sentence is written:

“The concept of relevance is studied in many different fields, including cognitive sciences, logic, and library and information science”.

In this sentence 'library and information science' is a hyperlink with browsing cues; the *title* - *library and information science* and *text* – the sentence in which this hyperlink is embedded. At the hyperlink searcher has to decide if the hyperlinked source page is about 'concept of relevance in library and information science' or it is simply a page about 'library and information science'. There are two

means to remove uncertainty posed by two problems associated with browsing cues discussed above, one – to visit the source document and verify the content as done in current browsing environment, two – to present extended browsing cues which tailors itself according to user's current information need. In this paper we investigate this second option.

We present a summarization system which complements every organic hyperlink with query-biased source document summary (extended browsing cues). We hypothesize that this summary enhances relevance estimation of source document and enables user to accurately estimate expected gain from distal source page and likely cost to consume that information.

1.1 RELATED WORK

We used sentence extraction methods[4]–[7] to summarize source document. Tombros[8] has shown the benefits of query-biased summary. When it comes to using Query-biased source document summary as extended browsing cues to enhance relevance estimation of source document our research closely follows previous work done by White et al[7] and Jones [9]. White et al[7] introduced a summarization system WebDocSum specifically designed for search engines. WebDocSum used a sentence extraction method to produce domain independent source document summary, in our summarization system we have used sentence scoring rules proposed in WebDocSum. Jones [9] developed a *focusing browsing* application using *term cloud* to enhance relevance estimation in browsing environment. Term cloud unlike sentence extraction provides summary preview of source document in form of cloud of words or phrases taken from source document. What makes our system unique is its ability to integrate searching with browsing. When it comes to generation of summary our system uses same metrics as used by White et al[7] in WebDocSum and extends the functionality to browsing as proposed by Jones [9] in *focusing browsing*. Our summarization system uses sentence extraction methods to generate query-biased source document summary for complete search session. We assume a search session starts from search engine and if necessary is extended to goal directed browsing.

II QUERY-BIASED SOURCE DOCUMENT SUMMARY

We implemented our summarization system as a Google chrome extension. The summary is generated for every organic hyperlink present on page encountered by user in complete session starting from search engine. The summary is displayed only when user mouse hover's a particular hyperlink for 0.5 seconds. This delay was intentionally programmed so as to check for accidental mouse hover. Figure 1 shows a sample summary generated across a hyperlink with query words 'sachin tendulkar'. We assume search engine as a starting point of search session. The rationale behind our assumption is that 93% of online experiences begin with a search engine[10] and a research[11] showing more than eight in ten American internet users search for information on search engines[11]. Our summarization system excludes paid/ sponsored hyperlinks as 70-80% of users ignore paid ads [10]. Following section describes the summary generation process.

1. The focus was on the clash between **Sachin Tendulkar**, the world's most dominating batsman and Shane Warne, and the world's leading spinner, both at the peak of their careers, clashing in a full-fledged Test series after 7 long years.
2. Sudhir Kumar Chaudhary, a fan of **Tendulkar** from Bihar, has followed almost all the matches played by **Sachin** since 2003 and has sacrificed his personal life to watch **Sachin's** play.
3. In the second Test at Port of Spain, **Sachin Tendulkar** scored 117 in the first innings, his 29th Test century in his 93rd Test match, to equal Sir Donald Bradman's record of 29 Test hundreds.
4. Thus **Sachin Tendulkar** became the youngest player to debut for India in Tests at the age of 16 years and 205 days and also the youngest player to debut for India in ODIs at the age of 16 years and 238 days.

Figure 1: a sample query-biased source document summary with query words highlighted

2.1 Summarization

We adopted sentence extraction methods [4]–[7] to select top four sentence from each source document to compose a summary. *All the sentences from respective source document are extracted and*

initially assigned a score of zero. We use full stop, question mark, and exclamation mark to identify end

of the sentence. We used sentence scoring rules proposed by White et al. [7]

2.2 Title Method

This method assumes that the author of a document reveals the main concept in the title of their document[7]. The assumption is based on the study[8] conducted on title of news articles which tends to reveal the major subject of the underlying article. Author of a document has complete control over the document title. Title is an important part of Search Engine Optimization hence it is a standard practice to include important document relevant keywords in title. For every unique non-stop word from title of a document present in a sentence a positive score is added to a sentence score.

2.3 The Location Method

This method assumes that:

- Sentences located under heading convey significant information about content of document.
- Opening and closing sentences of paragraph are important [5].

For every unique non-stop word present from heading in sentence a positive score is added to a sentence score. This score is less than the title score, and also is in decreasing order of heading font. Scores goes on decreasing from <H1> to <H6>. Every opening and closing sentence of a paragraph is assigned a positive score. The opening sentence of document is assigned greater score than the score assigned to opening sentence of each paragraph.

2.4 The Text Formatting Method

Author of Web document tends to emphasize important words using text formatting tags viz. < b >, < i >, and < u >[7], [8]. In this method every sentence with embedded text formatted words are assigned positive score.

2.5 The Query Biased Method

This method[7][8] takes into consideration the search query used by searcher on search user interface. Every sentence containing the query word is assigned a positive score. The larger the query

words present in the sentence the larger is sentence score.

2.6 Cumulative sentence score

The four methods described above are applied to each sentence of a document in the sequence shown in figure 2.

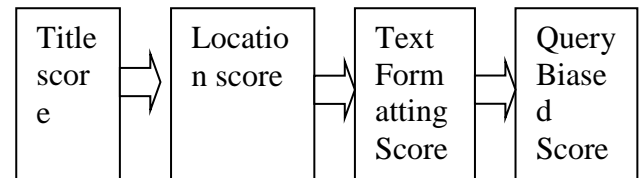


Figure 2: Cumulative sentence scoring methodology

All the methods apply individual scores to the sentence and the final sentence score is calculated summing up all the scores.

2.7 Top Ranked sentences as source document summary

Based on the cumulative sentence score, top four scored sentences from each hyperlinked document are selected as a sentence summary. In case of tie in sentence scores the sentence which appeared first in document is selected. Like this we generate a summary for every source document connected to current page via hyperlink and display particular summary only when user mouse hover's hyperlink to that page, all other summaries are kept hidden.

III EXPERIMENTAL METHODOLOGY

We used a within subject repeated measure design to compare our experimental system with baseline system. All the subjects used both the systems. Each subject was asked to select a work task from task list specifically designed to correlate with simulated work task situation as proposed by Borlund [12]. Simulated work tasks are designed so as to reflect real life information seeking situations. For each system subject had to select a new task; inter system task repetition was not allowed. In this section we detail the experimental methodology.

3.1 Systems used

We compared experimental system enhanced with source document summary with baseline system without such enhancement.

3.1.1 Baseline system

Any available web browser is eligible to be a baseline system. For this study we used Google chrome as a baseline system. When using a baseline system, at every hyperlink encountered subject had to judge the relevance of connected page (source document) using browsing cues surrounding the hyperlink. On SERP these cues are the *title*, *url*, and *abstract* (short text snippet with query words embedded) of source page. In browsing environment these cues are the anchor text of hyperlink and usually the sentence in which this hyperlink is embedded. When user is foraging for information he has to decide if clicking particular hyperlink will give expected information or not. In baseline system the relevance assessment is solely based on browsing cues provided by search engine or author of webpage.

3.1.2 Experimental system

In experimental system every organic hyperlink is complemented with query-biased source document summary. User has to mouse hover on a particular hyperlink for 0.5 seconds to make source document summary visible. This summary is clickable – user can either click hyperlink or summary to visit the source document. This system is implemented as a Google chrome extension. Section 3 detailed how this summary is generated.

3.2 Equipment and Place of Experiment

The experiment was conducted using two desktop computers one running a baseline system while another installed with experimental system. Experiment was conducted in a lab partitioned by a half glass partition. Subjects performed experiment in one half and we watched the experiment from other half through glass. Subjects were instructed to raise their hand if they completed the task or if they wanted to talk to us.

3.3 Document Domain

World Wide Web is used as document domain. We assume search engine as a starting point of search. We excluded sponsored links and advertisements from document domain. Both sponsored link and advertisement are clearly marked by Google search engine using symbols as shown in figure 3 and figure 4 respectively.

Sponsored ⓘ

Figure 3: Symbol used for sponsored Link

Ads

Figure 4: Symbol used for sponsored Link

A hypothetical scenario: subject enters search query in Google search user interface, Google returns the results as a ranked list on search engine result page. First title on this page is sponsored and second is advertisement. Subject is not allowed to visit these titles. Subject finds fourth title relevant to his current information need if subject is using a baseline she has to click the title to visit the source document, else if subject is using an experimental system she can view summary for relevance assessment and then visit source document if necessary.

3.4 Subjects

Forty subjects (16 Female, 24 Male) were selected at random from a list of 73 volunteers. Our volunteers were all master's students from a management institute. The basic requirement to volunteer for experiment was working knowledge of English. Mean age of subjects was 22.05 years [min 21, max 25]. Every subject was comfortable using search engine and Google search engine in particular.

3.5 Work tasks

Subjects were asked to select any two work tasks from a list of 17 work tasks specially designed on the lines of simulated work task situation. These 17 work tasks were all of similar difficulty level as assessed by two independent raters. Cohen's κ was run to determine if there was agreement between two rater's judgments, there was high agreement between the two rater's judgments, $\kappa = .880, p < .0005$. These 17 work tasks were from different categories as holiday planning, career guidance, health, social etc. so as to entertain users with different interest group. One of the work tasks is shown in figure 5.

Figure 5: A simulated work task

You are planning a week long family vacation this year. Your budget for the vacation is Rs 40000. You decide to visit north India. You want to gather information about the places which you can visit, accessibility and accommodation facilities available at those places. You are also concerned about the budget.

Task: Bearing in mind the constraints, your task is to plan a week long family vacation.

3.6 Experimental procedure

The experiment lasted for about 40 minutes, with 10 minutes being allocated to each task and 10 minutes for training session on experimental system. A work task common for all subjects and from out of the

task list was used as a training work task. Subjects were called one at a time. *Subjects had to book mark the webpage they considered relevant to their current information need in a search session.* These web pages were not evaluated by us for their actual relevance. These are the user perceptions of relevance. For each subject following procedure was followed:

- i. An introductory orientation session.
- ii. First system allocation.
- iii. A 10 minutes training session if experimental system.
- iv. Work task selection by subject.
- v. A 10 minutes search session on first system.
- vi. Second system allocation.
- vii. A 10 minutes training session if experimental system.
- viii. Work task selection by subject.
- ix. A 10 minutes search session on second system.
- x. Questionnaire.

3.7 User opinion and interaction.

Two methods were used during the experiment to collect data on user opinion and their interaction with systems. These methods were:

- i. Questionnaire
- ii. Screen recording
BB flashback expresses 5 player; a freely available screen recorder at flashbackrecorder.com was used. Consent of each subject was taken before recording user interaction with system.

3.8 Comparative analysis

The independent variable in this experiment is system type and hence focus of our evaluation was relevance assessment between systems. We compared dependent variable on experimental system with same dependent variable on baseline system for each subject.

3.9 Hypotheses

The purpose of this experiment is to investigate if Query-biased source document summary enhances relevance estimation of source document. We not only asked subjects about their perception of relevance estimation but also measured their performance. Following are the experimental hypotheses of our research:

Hypothesis 1:

Subject takes less time to find relevant information in experimental system as compared with baseline system.

Hypothesis 2:

Subject traverse less number of nodes to find relevant information in experimental system as compared with baseline system.

Hypothesis 3:

Query-biased source document summary is more helpful in relevance estimation of source document when compared with title and abstract.

3.10 Experimental results and analysis

In this section the results of user experiment are described. We focus on results that relate to three experimental hypothesis proposed. A total of 40 subjects participated in this experiment. Results are presented for inter system (Baseline versus Experimental) comparisons. The results presented in this section are based on questionnaire responses and user interaction recorded using screen recorder software. Questionnaire used a 5-points Likert scale to measure the perception of subjects regarding helpfulness of component. With lowest score interpreted as 'did not help'. For instance responses for question 'How much did the following components helped you in assessing relevance of hyperlinked page?' are scored on a scale of 1-5 for responses *None* to *A Lot*. We used paired t-test to analyze likert responses ie ordinal data as in our research mean is a meaningful data summary[13]. No assumptions are made about the normality of data; data is exclusively tested and found to be normally distributed hence parametric statistical tests are used to test the statistical significance. The analysis presented involves within-subject (one subject- two systems) comparisons.

3.10.1 Subject's performance on relevance assessment

Subject's interaction with system was recorded using screen recording software. The interaction was replayed and data was extracted manually for both systems on following parameters; *total number of hyperlinks visited, total number of pages bookmarked, and total session time*. Like this for 40 subjects 120 paired data points (baseline system - 40 subjects x 3 parameters, experimental system - 40 subjects x 3 parameters) were extracted. These parameters are described as:

- i. Total number of hyperlinks visited – the number of hyperlinks subject clicked in his entire search session including SERP.

- ii. Total number of pages bookmarked – all pages which were bookmarked by subject as relevant to his information need.
- iii. Total session time – If subject completed the task successfully the task completion time, else if subject quit or time exhausted session time is considered 10 minutes.

Our data analysis needed two variables which were derived from the three basic variables listed above, table 1 lists the derived variables used for testing hypothesis.

Table 1. Derived variables used for testing hypothesis

Derived variable used in analysis	Variable obtained by:
Time required to find relevant information	$\frac{\text{total session time}}{\text{total number of pages bookmarked}}$
Nodes traversed to find relevant information	$\frac{\text{total number of Hyperlinks visited}}{\text{total number of pages bookmarked}}$

3.10.1.1 Time required finding relevant information

The time required finding relevant information as described in table 1 is the total session time divided by the total number of pages bookmarked. It is the mean time needed by a subject to find one relevant page on a particular system. Table 2 shows the mean time for all subjects on particular system, standard deviation and significance level when comparing means using paired t-test of significance. Read; time as - mean time in minutes to find relevant information and node as - mean number of nodes traversed to find one relevant page.

A paired-samples t test was used to determine whether there was a statistically significant mean change in time required by subject for finding a relevant page when using experimental system compared to a baseline system. Three outliers were detected that were more than 1.5 box-lengths from the edge of the box in a boxplot. Inspection of their values did not reveal them to be extreme and were kept in the analysis. The assumption of normality was not violated, as assessed by Shapiro-Wilk's test ($p = .309$). Subjects took less time for finding relevant page when using experimental system ($M = 4.20, SD = 1.73$ minutes) as opposed to baseline system ($M = 4.57, SD = 1.58$ minutes), a mean

decrease of -0.37 minutes, 95% $CI [-1.14, 0.40]$, $t(39) = -0.970, p > .005(.338)$, $d = -0.15$.

The mean change in time was not statistically significantly different from zero and, therefore, we reject our hypothesis that subject takes less time to find relevant information in experimental system as compared with baseline system.

Table 2

Parameter	Baseline system		Experimental system		Sig
	Mean	Sd	Mean	Sd	
Time	4.57	1.58	4.20	1.73	.338
Node	2.53	0.32	1.53	0.17	.000

3.10.1.2 Nodes traversed to find relevant information

Nodes traversed to find relevant information as described in table 1 is total number of hyperlinks visited in a session divided by total number of pages bookmarked. A paired-samples t test was used to determine whether there was a statistically significant mean difference between the average numbers of nodes traversed by a subject for finding a relevant page when using experimental system compared to baseline system. No outliers were detected that were more than 1.5 box-lengths from the edge of the box in a boxplot. The assumption of normality was not violated, as assessed by Shapiro-Wilk's test ($p = .181$). Subjects traversed less number of nodes for finding relevant page when using experimental system ($M = 1.53, SD = 0.17$ nodes) as opposed to baseline system ($M = 2.53, SD = 0.32$ nodes), a statistical significant mean decrease of -0.99 nodes, 95% $CI [-1.11, -0.87]$, $t(39) = -16.488, p < .005, d = -2.60$.

The mean difference was statistically significantly different from zero and, therefore, we accept our hypothesis that subject traversed less number of nodes to find relevant information in experimental system when compared with baseline system.

3.10.2 User perception of relevance estimation

Users were asked to rate the helpfulness of components on a scale of 1 to 5, with 1 being *none* and 5 being *a lot* helpful. We used paired t-test to analyze likert responses; ordinal data as in our research mean is a meaningful data summary[13]. Users had to rate three components; title, abstract, and summary. For testing

hypothesis we required comparisons between title - summary, and abstract – summary. We tested these pairs using two paired sample t-tests. The results are:

Title - Summary

A paired-samples t-test was used to determine whether there was a statistically significant mean difference between the helpfulness of components when title was compared to summary. No outliers were detected that were more than 1.5 box-lengths from the edge of the box in a boxplot. The assumption of normality was not violated, as assessed by Shapiro-Wilk's test ($p = .110$). Subjects considered summary to be more helpful ($M = 3.10$, $SD = 1.15$) as opposed to the title ($M = 1.23$, $SD = 0.423$), a statistically significant mean increase of 1.88, 95% CI [1.512, 2.238], $t(39) = 10.43$, $p < .005$, $d = 1.65$.

Abstract - summary

A paired-samples t-test was used to determine whether there was a statistically significant mean difference between the helpfulness of components when abstract was compared to summary. No outliers were detected that were more than 1.5 box-lengths from the edge of the box in a boxplot. The assumption of normality was not violated, as assessed by Shapiro-Wilk's test ($p = .104$). Subjects considered summary to be more helpful ($M = 3.10$, $SD = 1.15$) as opposed to the abstract ($M = 2.18$, $SD = 1.20$), a statistically significant mean increase of 0.93, 95% CI [0.453, 1.397], $t(39) = 3.968$, $p < .005$, $d = 0.62$.

Based on above two test results we safely accept our hypothesis that Query-biased source document summary is more helpful in relevance estimation of source document when compared with title and abstract.

IV . DISCUSSIONS

In previous section we presented the experimental results obtained, in this section we present in depth analysis of these results. First we discuss the number of nodes traversed by subjects finding relevant information then the time and lastly the user's perception of relevance.

4.1 Nodes traversed to find relevant information

We hoped to improve the relevance estimation of source document by complementing every organic hyperlink with source document summary. In experimental system subjects were equipped with query-biased source document summary in case if they needed extended

browsing cues. We found summary to be useful in relevance estimation as subjects traversed significantly less $p < .005$ number of nodes. For baseline system average source documents visited for hitting one book mark was 2.53 whereas the same is 1.53 in experimental system; it took approximately 25 node visits in baseline system to book mark 10 pages whereas in experimental system it took only 15. The traversal in experimental system was more directed and confident than in baseline system. We attribute this confidence to enhanced relevance estimation via summary.

4.2 Time required finding relevant information

We hoped to reduce the average search session time but found that there was no significant mean difference $p > .005$ (0.338) between the time required by subject for finding a relevant page when using experimental system and time required by subject for finding a relevant page when using baseline system. This was very well justified by the fact that in experimental system even though subjects traversed less number of nodes ie they visited less number of source documents they spent more time reading the summary. We confirmed these results by randomly selecting five subjects and replaying their interaction with experimental system and found that subjects viewed average 4 (3.8 to be exact) summaries before visiting the source document. These results show - in experimental system subjects spent more time assessing relevance before clicking the hyperlink, rather than after clicking the hyperlink.

4.3 User perception of web components for relevance estimation

We hoped and found that summary helped subjects in assessing the relevance of source document. The extended browsing cues, automatically tailored to user's information need presented in form of summary helped subjects assess the relevance of source document. In our analysis we considered lowest score 1 as being not helpful and 5 being A lot helpful average summary score was 3.10 , average abstract score was 2.18, and average title score was 1.23. User's perceived summary being most helpful and title as least helpful when assessing relevance estimation of source document. Our findings validates the length hypothesis[14] which states that the quality of representation is directly proportional to its length.

V CONCLUSIONS

The main aim of this research paper was to investigate if Query-biased source document summary enhances relevance estimation of source document. We found that summary did enhance the relevance estimation of source document. User took their time in assessing the relevance of source document and made informed decisions about node traversal. User's traversal was goal oriented rather than random. Our system helped user's in estimating expected gain from distal information nugget and estimating the expected cost of acquiring that information. Our results added new dimension to earlier research that Query-biased summary are not only helpful in searching[7] or only browsing[9] but for complete search session starting from search engine and usually ended with browsing. We also validated length hypothesis[14] which states that the quality of representation is directly proportional to its length. We did this by asking users to rate helpfulness of web components title, abstract, and summary. These components are arranged in increasing order of their length with title of hyperlink being shortest in length and summary (extended browsing cues) being more descriptive and longest.

VI FUTURE WORKS

Our findings established that Query-biased source document summary enhances relevance estimation of source document, but our experimental system failed in reducing the session time. Instead of visiting nodes randomly users spent more time reading summary and deciding which node to visit. We intend to improve on this parameter. We wish to reduce session time by suggesting user a hyperlink which may best satisfy her information need.

REFERENCES

- [1] J. Nielsen, "Information Foraging: Why Google Makes People Leave Your Site Faster," 2003. [Online]. Available: <https://www.nngroup.com/articles/information-scent/>. [Accessed: 03-Jan-2015].
- [2] C. Olston and E. D. H. Chi, "ScentTrails: Integrating browsing and searching on the Web," *ACM Trans. Comput. Interact. TOCHI*, vol. 10, no. 3, pp. 177–197, 2003.
- [3] S. K. Card, P. Pirolli, M. Van Der Wege, J. B. Morrison, R. W. Reeder, P. K. Schraedley, and J. Boshart, "Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability," *Proc. SIGCHI Conf. Hum. factors Comput. Syst. CHI 01*, vol. 3, no. 3, pp. 498–505, 2001.
- [4] G. Salton, A. Singhal, M. Mitra, and C. Buckley, "Automatic Text Structuring and Summarization," *Inf. Process. Manag.*, vol. 33, no. 2, pp. 193–207, Mar. 1997.
- [5] R. Brandow, K. Mitze, and L. F. Rau, "Automatic Condensation of Electronic Publications by Sentence Selection," *Inf. Process. Manag.*, vol. 31, no. 5, pp. 675–685, Sep. 1995.
- [6] R. W. White, J. M. Jose, and I. Ruthven, "Using top-ranking sentences to facilitate effective information access," *J. Am. Soc. Inf. Sci. Technol.*, vol. 56, no. 10, pp. 1113–1125, 2005.
- [7] R. W. White, J. M. Jose, and I. Ruthven, "A task-oriented study on the influencing effects of query-biased summarisation in web searching," *Inf. Process. Manag.*, vol. 39, no. 5, pp. 707–733, 2003.
- [8] A. Tombros and M. Sanderson, "Advantages of query biased summaries in information retrieval," *Proc. 1998 21st Annu. Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, pp. 2–10, 1998.
- [9] G. J. F. Jones and Q. Li, "Focused browsing: Providing topical feedback for link selection in hypertext browsing," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4956 LNCS, pp. 700–704, 2008.
- [10] E. H. Siu, "24 Eye-Popping SEO Statistics - Search Engine Journal," 2012. [Online]. Available: <https://www.searchenginejournal.com/24-eye-popping-seo-statistics/42665/>. [Accessed: 14-Jan-2015].
- [11] S. research center) Fox, "Search Engines | Pew Research Center," 2002. [Online]. Available: <http://www.pewinternet.org/2002/07/03/search-engines/>. [Accessed: 23-Dec-2014].
- [12] P. Borlund, "Experimental components for the evaluation of interactive information retrieval systems," *J. Doc.*, vol. 56, no. 1, pp. 71–90, 2000.
- [13] BCSS, "BCSS: Myths and Nonsense about the t-test," 2015. [Online]. Available: <http://bcss1.blogspot.in/2015/02/myths-and-nonsense-about-t-test.html>. [Accessed: 04-Feb-2015].
- [14] R. S. Marcus, P. Kugel, and A. R. Benenfeld, "Catalog information and text as indicators of relevance," *J. Am. Soc. Inf. Sci.*, vol. 29, no. 1, pp. 15–30, 1978.