Contemporary Engineering Sciences, Vol. 5, 2012, no. 8, 375 - 380

Effects of "Kind of Information" on Networks' Speed Search

Butheyna Zuheir Al-Rawashdeh

Faculty of Computer Science and Information
Ajloun University College
Al- Balqa' Applied University (BAU), Jordan
Butheyna_z@yahoo.com

Nisrein Jamal Abu-Darwish

Faculty of Computer Science and Information
Ajloun University College
Al- Balqa' Applied University (BAU), Jordan
nessreenjamal@yahoo.com

Abstract

This paper presents a study about the effects of the kind of searched information upon networks on the speed of getting such information. The speed is function of kind of information and a mathematical model is constructed here to relate the speed with type of information. The time needed for ordinary text is assumed to be unity, then the time needed for each group of information is measured relative to ordinary text group which called here as weight time. It is found that as the type of information become more complicated i.e. contains pictures, photos, and videos.. the time needed for search and presented such information increases.

Keywords: network, information, search, speed

Introduction

The Internet, and especially World Wide Web, are incredibly popular at homes and offices alike. Because of the absence of centralized control or authority, statistics about the Net lack some degree of certainty. There is no question, however, that the Net is enormous in terms of numbers of users, Web sites, and Web pages. For instance, an estimate of the minimum number of host machines on the Internet is over 16 million, nearly seven million more than a year ago, and over 11 million more than a year before that. Similarly, of the 220 million people in the United States and Canada over the age of 16, 23% (over 50 million) are estimated to use the Internet and 17% (over 37 million) the World Wide Web.[1,2] And there is every reason to believe these numbers will all swell significantly in the next few years, with some analysts suggesting that the Web is doubling in size every 100 to 125 days. Efficient constraint programming search strategies exploit specific aspects or characteristics of a given (instance of a) problem. In Operation Research, relaxation or decomposition strategies exploit the fact that part of the problem can be treated as a classical problem (such as compatible or optimal of problems, shortest path problems, knapsack problems, etc.). This is often called structure in the constraint programming community. A problem is more generally said to be structured if its components (variables and/or constraints) do not all play the same role, or do not have the same importance within the problem. In such a problem, the origin of the complexity relies on the different behavior (or impact) for specific components of the problem. One of the main difficulty in identifying structure in problems is that this structure is not always statically (at the instance level, before solving) present. The interplay between a given instance and the search algorithm itself may define or help to exhibit a hidden structure within the problem. We call it a dynamic structure. It is related to bad initial choices as well as new relationships due to the addition of constraints during the search. This does not make things easy when willing understand the complexity of a problem and use this information to speed up search techniques.[3,4].

There are four different methods for locating information on the Web:
-First, one may go directly to a Web page simply by knowing its location. This is the reason for companies paltering their URLs over their TV, print and radio advertisements.
-Second, the hypertext links emanating from a Web page provide built-in associations to other pages that its author considers to provide related information.
-Third, `narrowcast' services can `push' pages at you that meet your particular user profile. -Fourth, search engines allow users to state the kind of information they hope to find and then furnish information that hopefully relates to that description.

For many years, the field of information retrieval (IR) has devised search algorithms, which take a user's query and furnish him or her with a list of hopefully relevant documents (often ranked according to a `relevance score' calculated by the algorithm). The search engine performance may include the abilities: to form Boolean queries; to specify that a term should (not) or must (not) appear in a Web page; to allow the user to use wildcards and truncation in search statements (to issue queries like compute to stand for computers, computing, compute, etc.) as well as for search algorithms to use automatic stemming (thus equating such items even if the user does not specifically ask that that be done); to search for phrases rather than individual words; to specify the importance of case sensitivity; and to do proximity searching (such as `airplane within five words of Denver'). More advanced capabilities available in some engines include the abilities: to allow the user to write a query in the form of a complete sentence (or short paragraph) which the search engine then parses and exploits; to suggest to the user additional words or phrases to include to refine an initial query (based on his or her relevance judgments of those already presented) or to allow the user to specify certain already retrieved documents to serve as examples of the type he or she would like to see; and to show the user groupings of retrieved documents that reflect how various concepts occur among them. The precise algorithms that search engines use for retrieval are not publicized, but one can infer their approximate workings by reading the Help, Hint or FAQ pages that accompany them as well as by being familiar with the field of IR. In most engines, a Web page typically will be highly ranked if it frequently uses many of the same words (or phrases) found in the query, especially if those are relatively rare words to begin with. The appearance of these items in a page's title, heading, or early in its text tends to raise the relevance scores of the page even further. Web-based information retrieval also differs in some respects from more traditional retrieval. Many researches treated such issue.[3,4,5].

Michael G, 1999, conducted a study to see how effective eight search engines are. Expert searchers sought information on the Web for users who had legitimate needs for information, and these users assessed the relevance of the information retrieved. They calculated traditional information retrieval measures of recall and precision at varying numbers of retrieved documents and used these as the bases for statistical comparisons of retrieval effectiveness among the eight search engines. They also calculated the likelihood that a document retrieved by one search engine was retrieved by other search engines as well.

Miriam J. 2007, summarized much of what is known from the communication and information literacy fields about the skills that Internet users need to assess the credibility of online information. The article reviews current recommendations for credibility assessment, empirical research on how users determine the credibility of Internet information, and describes several cognitive models of online information evaluation. Based on the literature review and critique of existing models of credibility assessment, recommendations for future online credibility education and practice are provided to assist users in locating reliable information online. The article concluded by offering ideas for research and theory development on this topic in an effort to advance knowledge in the area of credibility assessment of Internet-based information.

Hadrien C, 2005, claimed that, inspired by previous work on impact-based search strategies for constraint programming, that using an explanation-based constraint solver may lead to collect invaluable information on the intimate dynamic and static structure of a problem instance. They define several impact graphs to be used to design generic search guiding techniques and to identify hidden structures of instances. Finally, they discussed how dedicated OR solving strategies (such as

Benders decomposition) could be adapted to constraint programming when specific relationships between variables are exhibited.

Results and discussion

The searched information are divided here into six groups: texts, texts with figures and photos, graphs and photos, videos, PDF group, and most complicated information contains all last types.

The time weight relative to ordinary text (which assumed here to be 1 and all are of nearly same size) needed for each group to be searched in the internet is shown in table 1 below.

Table (1) time weight needed for information groups

Information group	Wight of time
Text	1
Text with figures & photos	1.43
Graphs and photos or pictures	1.25
Videos	2.5
PDF	1.1
Most complicated file	5

Figure 1 shows a comparison between weighted times needed for information groups.

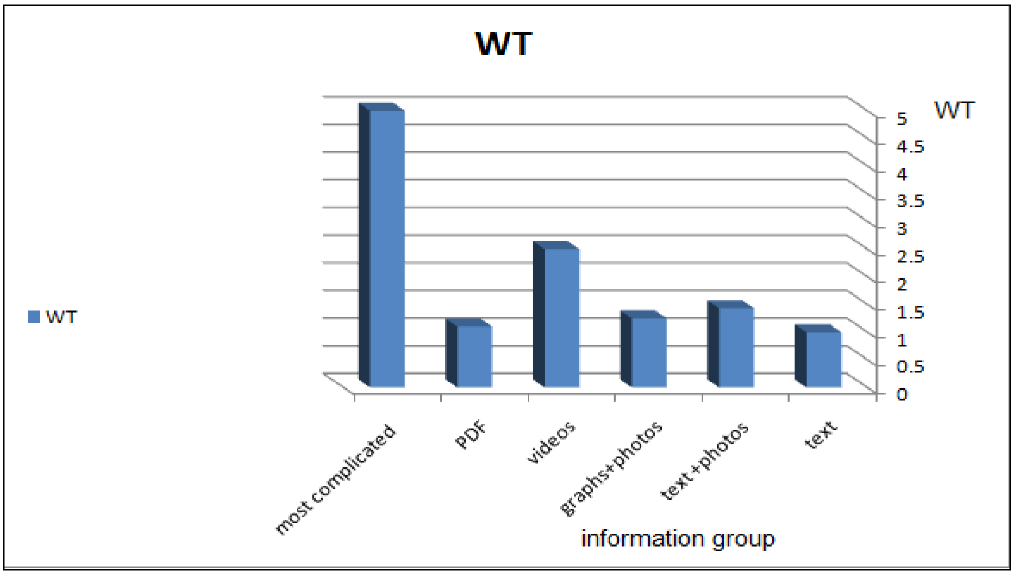

Figure1 weighted time for information groups

Conclusions

The time needed to search and retrieve information depends on its complicated and its contents, as the type of information contains more complicated contents like videos, pictures, photos and other types need more programs or utilities to be down loaded the time increases.

References

[1] C. Hadrien and J. Narendra, 2005, Identifying and exploiting problem structures using explanation-based constraint programming. email: fhcambaza,jussieng@emn.fr _Ecole des Mines de Nantes { LINA CNRS FRE 2729

[2] F. Boussemart, F. Hemery, C. Lecoutre, and L. Sais. Boosting systematic search by weighting constraints. In Proceedings ECAI'04, pages 482{486, 2004.

[3] G. Michael, P. Praveen, 1999, Finding information on the World Wide Web: the retrieval effectiveness of search engines, Information Processing and Management 35 (1999) 141-180.

[4] J. F. Benders. Partitionning procedures for solving mixed-variables programming problems. Numerische Mathematik, 4:238{252, 1962.

[5] J. Miriam, B. Santa, 2007, Making Sense of Credibility on the Web: Models for Evaluating Online Information and Recommendations for Future Research, Journal Of The American Society For Information Science And Technology, 58(13):2078–2091, 2007. 4 rue Alfred Kastler { BP 20722 { F-44307 Nantes Cedex 3, France.

Received: May, 2012