Expert Systems with Applications 37 (2010) 82–91

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

An adaptive fuzzy information retrieval model to improve response time perceived by e-commerce clients

A.O. Ajayi *, G.A. Aderounmu, H.A. Soriyan

Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria

ARTICLE INFO

Keywords:
Information retrieval
Fuzzy logic
Response time
Fuzzy expert system

ABSTRACT

In this paper, an adaptive fuzzy logic-based information retrieval model is presented to enable users retrieve exact and specific information they sort after. The proposed IR model takes into consideration the limited bandwidth between ISP and its users; and the characteristics (processor speed, memory size, resolution, availability of anti-virus, etc.) of clients' devices in ensuring that a customer has a fruitful and eventful session while conducting business online. The model was designed using unified modelling language and implemented using Borland JBuilder. A performance evaluation of the proposed information retrieval system using two evaluation measures was conducted. The experimental result indicated that the model has an acceptable performance.

1. Introduction

The primary goal of an Information Retrieval (IR) system is to retrieve the relevant documents to a query (Canfora & Cerulo, 2004). A complex algorithm is used to search through the information, retrieve and deliver the results to the user. An IR process begins with the submission of a query, which describes a user's topic of interest and ends with a set of ranked results deemed by the IR's ranking scheme to be the most relevant to the query. Several IR models abound, for instance, the logic-based Boolean model (Waller & Kraft, 1979) is one of the earliest IR systems. This model owes its popularity to its clean formalism and simplicity, but it has the following drawbacks: binary decision criterion without any notion of a grading scale, and the difficulty of translating the query into Boolean expressions (Baeza-Yates & Ribeiro-Neto, 1999; Cleverdon, 1983; Harman, 1992). A number of extended Boolean models which employ extended Boolean operators, have also been developed to provide ranked output (Canfora & Cerulo, 2004). Vector model is another popular retrieval model with the following advantages: a term weighting scheme to improve retrieval performance by sorting the documents according to their degree of similarity to the query, and a partial matching strategy, which approximates the query conditions. However, the vector model lacked clean formalism and simplicity (Rubens, 2006). Similarly, different ways of helping users specify their information needs more effectively have been studied (Belkin et al., 2003; Kelly, Dollu, & Fu, 2005). These approaches are effective in best-match retrieval systems where longer queries generally lead to more relevant search results. However, according to White and Morris (2007), longer queries may actually hurt retrieval performance, leading to a small number of potentially irrelevant results being retrieved. Caching techniques (Brin & Page, 1988; Saraiva et al., 2001) were used to optimize search engines. However, according to Romano, Quaglia, and Ciciani (2004) caching is less effective for web applications. Other approaches, which include probabilistic (Wong & Yao, 1995), fuzzy set (Ogawa, Morita, & Kobayashi, 1991), neural network (Lin, Soergel, & Marchionini, 1991), genetic algorithm (Gordon, 1988), and graph theory (Flake, Lawrence, Giles, & Coetzee, 2002) have been proposed in IR. However, these approaches do not take into consideration traffic and end user system's capability when presenting the search results. Response time is a measure of the period of time between entry of a request by a user and completion of processing for this request. It is one of the major performance issues in the Web. The response time, according to Olshefski and Nieh (2006) is a key indicator of the end user satisfaction in using any web applications. Thus, customers seeking quality online services have choices, they may take their business elsewhere when the system response time exceeds an acceptable threshold.

The response time experienced by web application users, most especially, in developing countries still needs improvement due to the obsolete infrastructure, limited bandwidth, and unreliable computing devices, which are associated to this part of the world. A response time of five minutes (5 min) for instance, to get a query result, is a sharp contrast to the response time specified in Nielsen

* Corresponding author. Address: Department of Computer Science and Engineering, Obafemi Awolowo University, Room 107, Computer Buildings, Ile-Ife, Nigeria. Tel.: +234 8035625962.

E-mail addresses: anuajayi@yahoo.com, anuajayi@oauife.edu.ng (A.O. Ajayi), gaderoun@oauife.edu.ng (G.A. Aderounmu), hasoriyan@yahoo.com, hsoriyan@oauife.edu.ng (H.A. Soriyan).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.05.071