Chapter 34
A Fireworks Algorithm for Modern Web Information Retrieval with Visual Results Mining

Hadj Ahmed Bouarara
*Tahar Moulay University of Saida Algeria,
Algeria*

Abdelmalek Amine
*Tahar Moulay University of Saida Algeria,
Algeria*

Reda Mohamed Hamou
*Tahar Moulay University of Saida Algeria,
Algeria*

Amine Rahmani
*Tahar Moulay University of Saida Alegria,
Algeria*

ABSTRACT

The popularization of computers, the number of electronic documents available online /offline and the explosion of electronic communication have deeply rocked the relationship between man and information. Nowadays, we are awash in a rising tide of information where the web has impacted on almost every aspect of our life. Merely, the development of automatic tools for an efficient access to this huge amount of digital information appears as a necessity. This paper deals on the unveiling of a new web information retrieval system using fireworks algorithm (FWA-IR). It is based on a random explosion of fireworks and a set of operators (displacement, mapping, mutation, and selection). Each explosion of firework is a potential solution for the need of user (query). It generates a set of sparks (documents) with two locations (relevant and irrelevant). The authors experiments were performed on the MEDLARS dataset and using the validation measures (recall, precision, f-measure, silence, noise and accuracy) by studying the sensitive parameters of this technique (initial location number, iteration number, mutation probability, fitness function, selection method, text representation, and distance measure), aimed to show the benefit derived from using such approach compared to the results of others methods existed in literature (taboo search, simulated annealing, and naïve method). Finally, a result-mining tool was achieved for the purpose to see the outcome in graphical form (3d cub and cobweb) with more realism using the functionalities of zooming and rotation.

INTRODUCTION

In the last decade, the appearance of internet and the development of electronic communication have made the world as a global village. Since the mid 90's, the Web become a source of information increasingly crucial, which has renewed interest for information retrieval. Meanwhile, the data sources are numerous and the amount of information available in electronic format does not cease to increase especially the unstructured data (free texts written in natural language).

Some years ago, information retrieval system (IRS) was used by only a set of people, but in our modern life a big change has occurred and IRS is often used by billions of peoples over the world. It is defined as a response to the need of users (query) by selecting a subset of documents from a wide range of documents that are closest to this query. The IRS is used in different application domains such as web, forums, business, and social network. Unfortunately, in the current digital society, the overload and the relevance of information have become a major stake where most of IRSs are face to several drawbacks in terms of:

- **The Quality of Relevance and Usefulness:** The access to the relevant information tailored to the needs of a specific user becomes both difficult and necessary.
- **The Parameters Choice:** The majority of information retrieval systems are based on the parameters (similarity measure and data representation technique). A poor choice of these parameters may cause degradation in the quality of results.
- **The Ambiguity of Natural Language:** This point is a major problem because the documents are written by humans.
- **The Multiplicity of Data:** The content of documents occur from different disciplines (biological, multimedia, sports, computer science ...etc.).

- **User Interface:** The majority of current search engines are based on a primary interaction that provides the display of the results as an ordered list. In this case the user views only the 1 or 2 pages and ignores the others but the desired information may be appeared in the 5 or 6 list.

The engineers and decision makers are confronted daily with increasing complex problems (NP-difficult) which affect generally all sectors (the design of mechanical systems, image preprocessing, information retrieval, and clustering ... ... etc.) where the classical techniques are unable to find an effective solution.

Nowadays, the world celebrates the birth of new interesting paradigms known as meta-heuristic methods, which have demonstrated their strength face to the optimization problems by finding the optimum of a function from a finite number of existing solutions. These are interesting techniques that reflect the process of any phenomenon in our daily life on an algorithm treatable by machine.

The content of our work is to shed light on the limits previously detailed by adapting a new meta-heuristic technique called fireworks algorithm (FWA) introduced by Tan in (tan, 2010) to solve the problem of information retrieval using a new visualization tool. This study seeks to the several objectives:

- The development of new version of fireworks algorithm to solve the information retrieval problem.
- Studying the influence of each parameter (iteration number, location number, mutation probability, and distance measure, text representation, and selection method, fitness function) in order to achieve the best performance.
- Comparing the obtained results with the results of other works existed in literature.

Related Content

Towards Private-Public Research Partnerships Combining Rigor and Relevance in DWH/BI Research: The Competence Center Approach

Anne Cleven, Robert Winterand Felix Wortmann (2010). *International Journal of Business Intelligence Research (pp. 60-71).*

www.irma-international.org/article/towards-private-public-research-partnerships/43682

Measuring Effectiveness: A DEA Approach Under Predetermined Targets

Heinz Ahnand Ludmila Neumann (2014). *International Journal of Business Analytics (pp. 16-28).*

www.irma-international.org/article/measuring-effectiveness/107067

Global Optimization of Economic and Social Policy

Sten A. Thore (2014). *Encyclopedia of Business Analytics and Optimization (pp. 1090-1099).*

www.irma-international.org/chapter/global-optimization-of-economic-and-social-policy/107307

Closing the Gap between Data Mining and Business Users of Business Intelligence Systems: A Design Science Approach

Ana Azevedoand Manuel Filipe Santos (2012). *International Journal of Business Intelligence Research (pp. 14-53).*

www.irma-international.org/article/closing-gap-between-data-mining/74733

Data Guided Public Healthcare Decision Making

Ramalingam Shanmugam (2014). *Encyclopedia of Business Analytics and Optimization (pp. 591-601).*

www.irma-international.org/chapter/data-guided-public-healthcare-decision-making/107262