

Literature Review on Genetic Algorithms for Internet Search Techniques

Prof. Rashmi Lad

Assistance Professor,
MAEER's MIT Arts, Commerce and Science College,
Alandi (D), Pune.
E-mail: rylad@mitacsc.ac.in

Prof. Vaishali Amberwadikar

Assistance Professor,
MAEER's MIT Arts, Commerce and Science College,
Alandi (D), Pune.
E-mail: vvamberwadikar@mitacsc.ac.in

Prof. Mayuri Bapat

Assistance Professor,
MAEER's MIT Arts, Commerce and Science College,
Alandi (D), Pune.
E-mail: bapat.mayuri@gmail.com

Abstract

As the number of documents and servers on Internet grows with the enormous speed, it becomes necessary to design efficient algorithms and tools for search and retrieval of documents. This paper describes several existing solutions to the problem and number of search techniques in genetic algorithms for Internet search.

Keywords: Genetic Algorithm, Crossover, Mutation, Internet etc.

Introduction

Among the huge number of documents and servers on Internet, it is hard to quickly locate documents that contain potentially useful information. Therefore, the key factor in software development nowadays should be the design of applications that efficiently locate and retrieve Internet documents that best meet user's requests. The accent is on intelligent content examination and selection of documents that are most similar to those submitted by the user, as input.

One approach to this problem is indexing all accessible Web pages and storing this information into the database. When the application is started, it extracts keywords from the user supplied documents and consults the database to find documents in which given keywords appear with the greatest frequency. This approach, besides the need to maintain a huge database, suffers from the poor performance – it gives numerous documents totally unconnected to the user's topic of interest.

The second approach is to follow links from a number of documents submitted by the user and to find the most similar ones, performing a genetic search on the Internet. Namely, application starts from a set of input documents, and by following their links, it finds documents that are most similar to them. This search and evaluation are performed using genetic algorithms as a heuristic search method. If only links from input documents are followed, it is the Best First Search,