UCSI UNIVERSITY

ENHANCING DESKTOP SEARCH USING VARIABLE SEARCH INDEXING TECHNIQUES

BY

KUAN HON KEAT
1000818171

FACULTY OF MANAGEMENT & INFORMATION TECHNOLOGY
B.Sc. (Hons.) COMPUTING
UCSI UNIVERSITY

2010

Abstract

There are many existing desktop search engines that iteratively improve to increase relevancy of search results while adding features that support towards the searching such as indexing, sorting, preview panel, or even simplicity of high-level layer of the desktop search engine. This project analyses the features used by few popular desktop search engines such as Google Search, Yahoo! Search and so on to identify suitable features for users to use. On the other hand, the project outcome will provide a complete system with better the searching experience by implementing user-defined metadata to distinguish file/folder contents while maintaining a simple user interface.

Chapter 1: Introduction

There are a lot of desktop searches which had developed and evolutionary enhanced with the main objective to obtain better search results from the computer hard drive. In this decade, every computer purchased will come along with their own default desktop search according from with operating system they are produced. Compared to web search tool, desktop search tool differs a lot in terms of features such as scope of search options, the way search results are retrieved, and types of metadata used to be displayed the search results.

A common problem shared by the desktop search tools are the lack of variability in search options, that the users are unable to find the results according to their personal preferences other than those pre-defined by the tools. User may only input for search and browse through the plane of unfiltered results containing irrelevant files and folders included in the results. Other than that, user may also find himself troubled at identifying the contents of the file or folder which he had long forgotten or encounter two or more files or folders with the same name assuming that they had the same file extension. It can be attributed to the unnecessary metadata retrieved and displayed through the search result which is not sufficient to be known by the user for file or folder content identification for accessing the right file or folder.