

SEARCH ENGINE GENERATOR PROGRAM FOR A WEB SITE

SOTARAT THAMMABOOSADEE

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(TECHNOLOGY OF INFORMATION SYSTEM MANAGEMENT)
FACULTY OF GRADUATED STUDIES
MAHIDOL UNIVERSITY
2005**

**ISBN 974-04-5808-4
COPYRIGHT MAHIDOL UNIVERSITY**

โปรแกรมสร้างอุปกรณ์สืบค้นข้อมูลในเว็บไซต์ (SEARCH ENGINE GENERATOR FOR A WEBSITE)

โยทศรััตต ธรรมบุษดี 4637214 EGTI/M

วท.ม. (เทคโนโลยีการจัดการระบบสารสนเทศ)

คณะกรรมการควบคุมวิทยานิพนธ์ : พรรณศิริ อธิคมรังษฤษฎ์, M.S.(COMPUTER SCIENCE), ฌนัท พูลสวัสดิ์, M.Sc. (INFORMATICS), รังสิพรรณ มฤคทัต Ph.D. (COMPUTER SCIENCE)

บทคัดย่อ

โปรแกรมประยุกต์ที่ใช้สร้างเครื่องมือสืบค้นข้อมูลในเว็บไซต์ มีเป็นประโยชน์แก่คนทำเว็บไซต์หรือผู้ใช้อินเทอร์เน็ตที่ต้องการสร้างเครื่องมือสืบค้นข้อมูลที่มีปัจจัยการค้นหาและรูปลักษณะภายนอกเป็นของตนเอง

การทำงานของโปรแกรมประกอบด้วยการรวบรวมเว็บเพจโดยตัวรวบรวมเว็บเพจ (Crawler) และ การทำดัชนีของเว็บเพจโดยตัวสร้างดัชนี (Indexer) ผลลัพธ์ที่ได้เป็นฐานข้อมูลในรูปแบบของ MySQL® และเว็บเพจที่ใช้ในการค้นหาข้อมูลซึ่งอยู่ในรูปแบบของ PHP ผู้ใช้งานสามารถตั้งค่าคุณสมบัติของตัวรวบรวมเว็บเพจและตัวสร้างดัชนี รวมถึงกำหนดลักษณะภายนอกของการสืบค้นข้อมูลได้

ความแม่นยำในการค้นหานั้นขึ้นอยู่กับค่าลำดับ 4 ประเภท ได้แก่ ค่าลำดับคำ, ค่าลำดับเว็บเพจ, ค่าลำดับเว็บเพจถ่วงน้ำหนัก และ ค่าลำดับผลตอบแทนจากผู้ใช้ งานวิจัยนี้ทดลองหาค่าน้ำหนักระหว่างค่าลำดับทั้ง 4 นี้และนำมาใช้ร่วมกันเพื่อให้ได้ความแม่นยำมากที่สุดในการค้นหา

ผลจากงานวิจัยคือโปรแกรมประยุกต์ที่สามารถสร้างเครื่องมือสืบค้นข้อมูลในเว็บไซต์ซึ่งสามารถกำหนดปัจจัยการค้นหาและรูปลักษณะภายนอก จากการประเมินโดยผู้ใช้อินเทอร์เน็ตทั่วไปผลการประเมินอยู่ในระดับที่น่าพอใจ และจากการเปรียบเทียบผลการค้นหากับ Google ซึ่งเป็นเครื่องมือสืบค้นข้อมูลในอินเทอร์เน็ตที่มีอยู่เดิม ได้ผลการเปรียบเทียบในระดับที่น่าพอใจ

121 หน้า ISBN: 974-04-5808-4

SEARCH ENGINE GENERATOR PROGRAM FOR A WEBSITE

SOTARAT THAMMABOOSADEE 4637214 EGTI/M

M.Sc. (TECHNOLOGY OF INFORMATION SYSTEM MANAGEMENT)

THESIS ADVISORS: PHANSIRI ATHIKOMRUNGSARIT, M.S. (COMPUTER SCIENCE), KANAT POOLSAWASD, M.Sc. (INFORMATICS), RANGSIPAN MARUKATAT Ph.D. (COMPUTER SCIENCE)

ABSTRACT

This research creates and develops an English search engine generator application for a website. This application is useful for a website developer or internet user who wants to create a search engine for their website. The appearance and search factors of this search engine can be customized.

This application collects retrieved web pages from a website by the crawler function, then web pages are indexed by the indexer function. The database is a MySQL[®] database. PHP files, which are used to implement the search engine, are created. Users can customize the crawler, the indexer and the appearance of the search engine.

A precision search depends on four ranks: WordRank, PageRank, Weighted PageRank and Users' Feedback Ranking Score. In this research, testing was conducted to find the optimal weights of these four ranks.

This research successfully develops a search engine generator for a website, which has adjustable search factors and appearance. The evaluation of the application by internet users resulted in a satisfactory system assessment. The search results comparisons with Google, the existing popular search engine, were at a satisfactory level.

KEY WORDS: SEARCH ENGINE / INFORMATION RETRIEVAL /
WEB PAGE RANKING / WEBSITE INDEXING

121 P. ISBN: 974-04-5808-4