QUESTION ANSWERING
FOR REVENUE DEPARTMENT
SUPERVISORY SYSTEM

WALLOP ANGKANASOPIT

A RESEARCH PROJECT SUBMITTED
IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(COMPUTER SCIENCE)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2007

COPYRIGHT OF MAHIDOL UNIVERSITY
ระบบค้นหา - ตอบปัญหาในการใช้ระบบสารสนเทศของเจ้าหน้าที่กรมสรรพากร
(THE QUESTION ANSWERING SYSTEM FOR THE REVENUE DEPARTMENT
SUPERVISORY SYSTEM)

วัลลภ อังคณาโสภิต 4737977 SCCS/M

วท.ม. (วิทยาการคอมพิวเตอร์)

คณะกรรมการควบคุมโครงการวิจัย: ศรีสิริ วงศ์สว่าง, Ph.D., ชาญยศ ปลื้มปัติวิริยะเวช, Ph.D

บทคัดย่อ

ในโครงการวิจัยนี้เราได้เสนอตัวแบบระบบค้นหาคำตอบ เรียกว่า “ระบบค้นหาคำตอบสำหรับระบบกำกับดูแลผู้เสียภาษี ของกรมสรรพากร” โครงการวิจัยนี้มีจุดประสงค์เพื่อเพิ่มประสิทธิภาพของระบบกำกับดูแลผู้เสียภาษีเพื่ออยู่ในปัจจุบันบนอินเทอร์เน็ตของกรมสรรพากรผู้ใช้ระบบดังนี้ได้แก่เจ้าหน้าที่ของกรมสรรพากรเพื่อหาคำตอบและรับคำตอบจากระบบทุกวันจะมีคำถามเข้ามาในระบบบางคำถามเมื่อยู่แล้วหรือมีความหมายใกลเคียงกับคำถามเดิมที่มีอยู่แล้วในระบบแต่จะไม่สามารถแจ้งให้เจ้าหน้าที่กรมสรรพากรทราบได้ว่าคำถามลักษณะเดียวกันหรือคล้ายกันมีอยู่แล้วหรือยัง นอกจากนี้การวิเคราะห์คำถามในระบบพบปัญหาที่น่าสนใจหลายปัญหา ปัญหาหลักที่น่าสนใจอย่างหนึ่งเมื่อทำงานกับภาษาไทยคือ โครงสร้างภาษาไทยมีความท้าทายสูง ด้วยเหตุนี้ในโครงการวิจัยนี้ได้นำเอาเทคนิคของการสืบค้นข้อมูลมาใช้ในการออกแบบระบบค้นหาคำถาม ซึ่งตั้งคำถามยังช่วยให้เจ้าหน้าที่ของกรมสรรพากรค้นหาคำถามที่มีอยู่แล้วในระบบ นอกจากนี้ผู้ใช้งานสามารถที่จะถามคำถามได้หลายรูปแบบ เช่น Keyword หรือ Boolean expression หรือใช้ natural language จากผลการทดลองได้ผลเป็นที่น่าพอใจและตัวแบบนี้สามารถเพิ่มประสิทธิภาพของระบบกำกับดูแลผู้เสียภาษีที่ใช้อยู่ในปัจจุบันได้

130 หน้า
THE QUESTION ANSWERING SYSTEM FOR THE REVENUE DEPARTMENT SUPERVISORY SYSTEM.

WALLOP ANGKANASOPIT 4737977 SCCS/M

M.Sc.(COMPUTER SCIENCE)

RESEARCH PROJECT ADVISORS: DAMRAS WONGSAWANG, Ph.D, CHARNYOTE PLUEMPITIWIRIYAVEJ, Ph.D.

ABSTRACT

This research project proposed the question answering (QA) system, called the Question Answering system for the Revenue Department (RD) Supervisory System (QARDSS). The goal is to improve the performance of the current QA system, the E-Supervisory (E-Sup) system, that is a web application running on intranet of RD. The E-Sup system helps the RD staff posing their questions and getting the answers back. Every day so many questions are posed to the board. Some of them are repeated or similar to those previously posed and already answered. However, the E-Sup has no capability to check for these. Furthermore, analysis of the questions reveals several interesting problems such as one of the main concerns when working with Thai natural language is that the structure of Thai language itself is highly ambiguous. For this reason, this research project introduced the information retrieval techniques to be used in the QARDSS system. The QARDSS system helps the RD staff to find repeated or similar questions that already have answers. Beside these, QARDSS also has many features to help users input their questions in many formats such as simple keyword, Boolean expression and natural language. The experimental results show that QARDSS performed satisfactorily and can improve the performance of the current E-Sup system.

KEY WORDS: QUESTION ANSWERING SYSTEM / INFORMATION RETRIEVAL / STRING MATCHING ALLOW ERROR / VECTOR MODEL / KEYWORD EXTRACTION / EDIT DISTANCE

130 pp.