การบูรณาการ UBR เพื่อค้นหาเว็บเซอร์วิสเชิงความหมาย

ธีระยุทธ ทองเครือ* และสมจิตร อาจอินทร์*

บทคัดย่อ

เว็บเซอร์วิสที่ถูกใช้งานอย่างแพร่หลายมากขึ้นในปัจจุบันทำให้การค้นหาเว็บเซอร์วิสมีความจำเป็นมากขึ้น การค้นหาเว็บเซอร์วิสในปัจจุบันสามารถค้นหาจาก UBR (Universal Business Registry) ซึ่ง UBR คือ แหล่งเก็บข้อมูลเว็บเซอร์วิส ปัญหาการค้นหาเว็บเซอร์วิสมี 2 ปัญหา ปัญหาแรก UBR มีผู้ให้บริการหลายแห่ง ซึ่งถ้าผู้ใช้ต้องการเลือกเว็บเซอร์วิสจาก UBR เหล่านี้สามารถทำได้ยาก ปัญหาที่สอง คือ การค้นหาเว็บเซอร์วิสจาก UBR ไม่สนใจความหมายของคำ ผู้เขียนได้นำเสนอแนวคิดในการค้นหาเว็บเซอร์วิสโดยการบูรณาการข้อมูลเว็บเซอร์วิสจาก UBR ที่มีอยู่หลายแห่ง การนิยามออนโทโลยีเพื่อเป็นแนวคิดกลางสำหรับการค้นหา และการจับคู่ออนโทโลยีกับข้อมูลที่ถูกบูรณาการบนพื้นฐานของความหมาย วิธีการดังกล่าวทำให้เกิดข้อมูลที่ชัดเจน และมีความหมายต่อมนุษย์มากขึ้น ซึ่งทำให้การค้นหาเว็บเซอร์วิสเชิงความหมายสามารถเป็นไปได้บนมาตรฐานเว็บเซอร์วิสที่มีอยู่ในปัจจุบัน

คำสำคัญ: เว็บเซอร์วิส การบูรณาการ ออนโทโลยี การจับคู่ออนโทโลยี การค้นหาเชิงความหมาย

1. บทนำ

นอกจากเราจะใช้เว็บเพื่อการให้บริการข้อมูลผ่านทางเว็บเพ็จแล้ว เรายังใช้เว็บเพื่อให้บริการแอพพลิเคชั่นได้อีกด้วย ซึ่งแอพพลิเคชั่นดังกล่าว ถูกเรียกว่า “เว็บเซอร์วิส” (Web Service) [1] มาตรฐานเว็บเซอร์วิสปัจจุบัน คือ ภาษา WSDL (Web Services Description Language) และ UDDI (Universal Description, Discovery, and Integration) เว็บเซอร์วิสจะอธิบายคุณสมบัติของตนเองด้วยภาษา WSDL และประกาศการให้บริการบน UBR (Universal Business Registry) ด้วยภาษา UDDI ดังนั้น UBR จึงเป็นศูนย์รวมข้อมูลเว็บเซอร์วิสเพื่อใช้สำหรับการค้นหา ปัจจุบัน UBR มีผู้ให้บริการ 3 ราย คือ IBM, Mocrosoft และ SAP ผู้ใช้สามารถเลือก UBR ใด UBR หนึ่งเพื่อค้นหาเว็บเซอร์วิสที่ตนเองต้องการ โดยส่งคำร้องไปยัง UBR ที่ต้องการค้นหา หลังจากนั้น UBR ทำการค้นหาตามความต้องการของผู้ใช้ และส่งผลลัพธ์กลับมายังผู้ใช้ ปัญหาในการค้นหาเว็บเซอร์วิสมี 2 ปัญหา ปัญหาแรก หากผู้ใช้ต้องการเลือกเว็บเซอร์วิสจาก UBR อื่น เพราะต้องการเปรียบเทียบผลลัพธ์เพื่อเลือกเว็บเซอร์วิสที่ต้องการมากที่สุด ผู้ใช้อาจเกิดความยุ่งยากในการค้นหา เปรียบเทียบ และเลือกผลลัพธ์ซึ่งได้จาก UBR ที่หลากหลาย ปัญหาที่สอง การค้นหาเว็บเซอร์วิสจาก UBR ไม่สนใจความหมายของคำ UBR จะพิจารณาคำที่ผู้ใช้ต้องการกับข้อมูลที่เก็บอยู่เท่านั้น โดยไม่คำนึงถึงความหมายของคำ ทำให้เว็บเซอร์วิสที่ทำงานเหมือนกันแต่มีชื่อต่างกันไม่ถูกเลือกเป็นผลลัพธ์

ปัญหาความหลากหลายของ UBR แก้ปัญหาโดยการบูรณาการข้อมูลจาก UBR ที่มีอยู่หลากหลายให้เป็นข้อมูลกลุ่มเดียวกัน มีความถูกต้อง ไม่ซ้ำซ้อน และเป็นข้อมูลที่ได้ในช่วงรันไทม์ ส่วนปัญหาการค้นหาเว็บเซอร์วิสที่ไม่สนใจความหมายของคำ มีสาเหตุจากขาดความชัดเจนในรูปแบบข้อมูลและขั้นตอนการประมวลผลเพื่อการค้นหา [2] การแก้ปัญหา คือ บรรยายขอบเขตเว็บเซอร์วิสที่สามารถค้นหาได้ด้วยออนโทโลยี (Ontology) และการจับคู่ออนโทโลยี (Ontology Mapping) ออนโทโลยีที่บรรยายขอบเขตเว็บเซอร์วิส เรียกว่า “โดเมนออนโทโลยี” (Domain Ontology) ซึ่งบรรยายแนวคิดทั้งหมดของเว็บเซอร์วิสที่สามารถค้นหาได้ เช่น แนวคิดเว็บเซอร์วิสการจองโรงแรม การสอบถามราคาห้อง แนวคิดเหล่านี้เป็นแนวคิดกลางสำหรับผู้ใช้ที่ต้องการค้นหาเว็บเซอร์วิส และยังใช้ในการจัดกลุ่มเว็บเซอร์วิสที่บูรณาการจากแต่ละ UBR อีกด้วย การจับกลุ่มข้อมูลเหล่านี้ เรียกว่าการจับคู่ออนโทโลยี ซึ่งเป็นการพิจารณาความคล้ายคลึงกันระหว่างทั้งสองแนวคิด ซึ่งเมื่อจับคู่ได้แล้วจะทำให้ทราบว่ามีแนวคิดใดบ้างที่เป็นแนวคิดเดียวกัน ทำให้ผู้ใช้สามารถได้รับผลลัพธ์ของการค้นหาที่มีความหมาย มากกว่าการค้นหาด้วยคำเพียงอย่างเดียว เช่น ผู้ใช้ค้นหาเว็บเซอร์วิส HotelBooking ผลลัพธ์จะได้เว็บเซอร์วิส BookingRoom และ ReserveRoom

วัตถุประสงค์ของงานวิจัยนี้แก้ปัญหาสองส่วน ส่วนแรก คือ แก้ปัญหาความหลากหลายของ UBR ส่วนที่สองแก้ปัญหาการค้นหาเว็บเซอร์วิสโดยสนใจความหมายมากกว่าการเปรียบเทียบคำเพียงอย่างเดียว เนื้อหาในบทความนี้กล่าวถึงแนวคิดในการแก้ปัญหาดังกล่าว ในหัวข้อต่อไปจะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ซึ่งอธิบายรายละเอียดภาษา

* ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

UDDI และ WSDL, การจับคู่ออนโทโลยี และงานวิจัยที่เกี่ยวข้อง ในส่วนที่ 3 อธิบายส่วนประกอบที่เกี่ยวข้องกับการค้นหาเว็บเซอร์วิส และส่วนสุดท้ายเป็นบทสรุปของบทความนี้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ภาษา WSDL และ UDDI

ผู้ให้บริการเว็บเซอร์วิสจะใช้ภาษา WSDL เพื่อสร้างเอกสารบรรยายว่าตนมีฟังก์ชันใดบ้างที่สามารถเรียกใช้ได้ พร้อมกับอธิบายพารามิเตอร์และผลลัพธ์จากการทำงานของแต่ละฟังก์ชันนั้น [3] ส่วนขั้นตอนการประกาศรายละเอียดเว็บเซอร์วิสนั้นจะใช้ภาษา UDDI [4] ในการบรรยาย เช่น ชื่อธุรกิจ ที่อยู่ เว็บไซต์ เซอร์วิสที่ให้บริการ เป็นต้น นอกจากนี้ผู้ให้บริการยังสามารถกำหนดกลุ่มและประเภท (Taxonomy) ของเว็บเซอร์วิสได้ มาตรฐานที่ใช้ในการกำหนดประเภทในปัจจุบันที่ใช้อย่างกว้างขวาง คือ NAICS (North American Industry Classification System) และ UNSPSC (The Universal Standard Products and Services Classification) [5] ผู้ให้บริการสามารถกำหนดประเภทโดยใช้มาตรฐานใดก็ได้ หลังจากผู้ให้บริการเว็บเซอร์วิสกำหนดรายละเอียดดังกล่าวเสร็จสิ้น จะส่งข้อมูลไปลงทะเบียนยัง UBR การค้นหาเว็บเซอร์วิสจาก UBR แบ่งเป็น 2 วิธี คือ ค้นหาด้วยคำ (Keywords) และค้นหาด้วยประเภทเว็บเซอร์วิสด้วยมาตรฐาน NAICS หรือ UNSPSC ซึ่งวิธีการหลังนี้เป็นวิธีการที่ได้ผลลัพธ์แม่นยำมากกว่าวิธีแรก

2.2 การจับคู่ออนโทโลยี

ออนโทโลยี คือ การกำหนดขอบเขตของข้อมูล โดยให้คำนิยามหรือคำจำกัดความ ออนโทโลยีเป็นพื้นฐานในการใช้แนวความคิดร่วมกัน และการประกอบแนวคิดด้วยการใช้ความสัมพันธ์และคุณสมบัติต่าง ๆ [6]

การจับคู่ออนโทโลยี หรือการจับคู่แนวคิด คือ การสร้างความสัมพันธ์ระหว่าง 2 ออนโทโลยีโดยพิจารณาความหมาย เพื่อแก้ปัญหาการสื่อสารระหว่างออนโทโลยีและการใช้ข้อมูลร่วมกัน [7] วิธีการทำงาน คือ นำออนโทโลยีที่ต้องการใช้ข้อมูลร่วมกัน มาทำการประมวลผลจับคู่ เช่น ออนโทโลยีแรกนิยามแนวคิดการจองโรงแรมด้วยคำว่า HotelBooking ขณะที่ออนโทโลยีหนึ่งใช้คำว่า reserveRoom การจับคู่พบว่าความหมายของ booking และ reserve คล้ายคลึงกัน ดังนั้นแนวคิดทั้งสองจะถูกจับคู่ให้เป็นแนวคิดเดียวกัน ซึ่งหากผู้ใช้อ้างถึงแนวคิด HotelBooking ก็หมายถึงว่าอ้างถึงแนวคิด ReserveRoom ด้วย

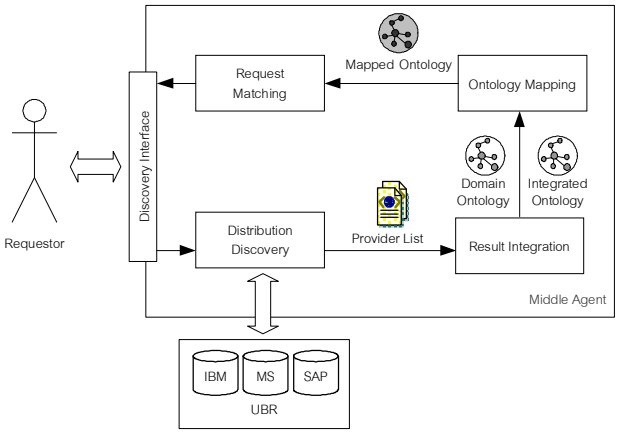
ปัจจุบันมีวิธีการจับคู่ออนโทโลยีถูกนำเสนอหลากหลายวิธี เช่น SMART [8], PROMPT [9], FCA-Merge [10], IF-Map [11] วิธีการเหล่านี้มีขั้นตอนในการจับคู่แตกต่างกันไป ผู้เขียนได้เลือกวิธีของ ONION (ONtology compositION sytesm) [12] ซึ่งเป็นการจับคู่แนวคิดที่มีความคล้ายคลึงกันบนพื้นฐานของความหมาย ขั้นตอนของ ONION จะพิจารณาคู่ของแนวคิดโดยค้นหาความหมายของแนวคิดทั้งสองจากพจนานุกรม เช่น WordNet [13] และคำนวณค่าความเหมือนจากแนวคิดแต่ละคู่ ซึ่งหากค่าเป็น 0 หมายถึงทั้งสองแนวคิดไม่มีความคล้ายคลึงกัน แต่หากค่าตั้งแต่ 0 ขึ้นไป หมายถึงว่าแนวคิดทั้งคู่มีความคล้ายคลึงกัน

2.3 งานวิจัยที่เกี่ยวข้อง

ปัจจุบันงานวิจัยที่เกี่ยวกับการค้นหาเว็บเซอร์วิสเชิงความหมายได้รับความสนใจจากนักวิจัยอย่างกว้างขวาง งานวิจัย [14] นำเสนอแนวคิดในการค้นหาเว็บเซอร์วิสด้วยการจับคู่ระหว่างเว็บเซอร์วิสที่ประกาศเอาไว้กับคำร้องผู้ใช้ โดยเตรียมขั้นตอนวิธีในการจับคู่ input และ output ของเว็บเซอร์วิสที่ผู้ใช้ต้องการกับ input และ output ของเว็บเซอร์วิสที่ผู้ให้บริการประกาศเอาไว้ วิธีการนี้มีข้อดี คือ ผู้ใช้ไม่ต้องระบุเงื่อนไขในการค้นหา ผู้ใช้กำหนดเพียงสิ่งที่ต้องการจากเว็บเซอร์วิสเท่านั้น แต่ข้อเสียคือ ผู้ใช้ไม่สามารถค้นหาตามการทำงานของเว็บเซอร์วิสได้ ผู้เขียนแก้ปัญหาวิธีการนี้โดยการเลือกเว็บเซอร์วิสจากชื่อฟังก์ชันที่ผู้ใช้ต้องการ หลังจากนั้นจึงพิจารณาที่ input และ output ของเว็บเซอร์วิส เช่น ฟังก์ชัน HotelBooking และ HotelCancelBooking มี output คือ ResevationInfo เหมือนกัน แต่ทั้งสองมีการทำงานที่แตกต่างกันโดยสิ้นเชิง แนวคิดจากงานวิจัย [14] จะทำให้ได้ผลลัพธ์ไม่ถูกต้อง เนื่องจากพิจารณาเพียง output อย่างเดียว แต่แนวทางของผู้เขียนจะพิจารณาชื่อฟังก์ชันก่อน และพิจารณา input และ output ในลำดับถัดไป ทำให้ได้ผลลัพธ์ที่ถูกต้องมากกว่า

งานวิจัย [2] นำเสนอแนวทางในการค้นหาเว็บเซอร์วิสโดยมีการสร้างออนโทโลยีบน UBR และทำการจับคู่ออนโทโลยีที่สร้างขึ้นกับข้อมูลที่อยู่ในเอกสาร WSDL ซึ่งได้แก่ ชื่อฟังก์ชัน, input และ output นอกจากนี้ยังมีการเพิ่มคำอธิบาย precondition และ effect ของฟังก์ชัน เพื่อให้ผู้ใช้สามารถค้นหาด้วยเงื่อนไขทั้งสองได้ ข้อดีของงานวิจัย [2] คือ ผู้ใช้สามารถใช้เงื่อนไขค้นหาตาม precondition และ effect ของฟังก์ชันได้

และไม่ต้องค้นหาผ่านตัวแทนกลาง แต่ข้อเสีย คือ ผู้ให้บริการเว็บเซอร์วิส ต้องแก้ไขเอกสาร WSDL เพื่อจับคู่ชื่อฟังก์ชัน input และ output กับออนโทโลยีที่สร้างขึ้น ซึ่งเป็นขั้นตอนที่ยุ่งยาก นอกจากนี้ผลลัพธ์จากการค้นหาจะได้รับจาก UBR เพียงแห่งเดียว ไม่มีการรวบรวมข้อมูลจาก UBR อื่นๆ ผู้เขียนได้เสนอวิธีการเพิ่มเติม คือ เสนอวิธีการจับคู่แนวคิดระหว่างออนโทโลยี และข้อมูลในเอกสาร WSDL ในช่วงรันไทม์ ผู้ให้บริการไม่ต้องจับคู่แนวคิดด้วยตนเอง นอกจากนี้ยังนำเสนอวิธีการบูรณาการข้อมูลจาก UBR ที่มีอยู่หลากหลายให้รวมเป็นข้อมูลกลุ่มเดียว เพื่อให้การค้นหาได้ผลลัพธ์เว็บเซอร์วิสที่เป็นทางเลือกมากยิ่งขึ้น



ภาพที่ 1 แบบจำลองการค้นหาเว็บเซอร์วิสเชิงความหมาย

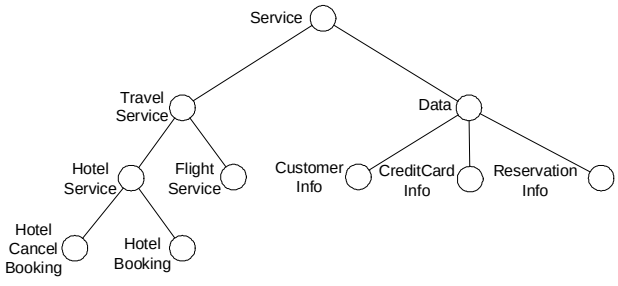
3. แบบจำลองการค้นหาเว็บเซอร์วิสเชิงความหมาย

แบบจำลองการค้นหาเว็บเซอร์วิสเชิงความหมายเป็นตัวกลางระหว่างผู้ใช้และ UBR จากภาพที่ 1 ประกอบด้วย ผู้ใช้ (Requestor), UBR ที่เกี่ยวข้อง และตัวแทนกลาง (Middle Agent) เมื่อผู้ใช้ต้องการค้นหาเว็บเซอร์วิส ผู้ใช้จะส่งคำร้องผ่านอินเตอร์เฟสของตัวแทนกลาง (Discovery Interface) ไปยัง Distribution Discovery ซึ่งทำหน้าที่ติดต่อกับ UBR ที่เกี่ยวข้อง เพื่อนำข้อมูลจากแต่ละ UBR มาทำการบูรณาการบนส่วน Result Integration ซึ่ง Result Integration มีหน้าที่สร้าง Integrated Ontology ซึ่งเป็นผลลัพธ์จากการบูรณาการข้อมูลจาก UBR หลังจากนั้น Ontology Mapping จะนำเอา Integrated Ontology มาจับคู่กับโดเมนออนโทโลยี (Domain Ontology) การทำงานในส่วนนี้ทำให้ได้ออนโทโลยีที่เชื่อมโยงระหว่างโดเมนออนโทโลยีกับ Integrated Ontology (Mapped Ontology) เพื่อใช้ในการค้นหาในส่วน Request Matching ผลจากการจับคู่ออนโทโลยี Request Matching จะสามารถสืบค้นบนโดเมนออนโทโลยี แทนการสืบค้นบน Integrated Ontology ซึ่งจะได้ผลลัพธ์ของแนวคิดที่มีความคล้ายคลึงกัน

3.1 โดเมนออนโทโลยี

โดเมนออนโทโลยี คือ ออนโทโลยีที่ใช้บรรยายแนวคิดของเว็บเซอร์วิสที่ใช้สำหรับการค้นหา จากภาพที่ 2 แสดงตัวอย่างโดเมนออนโทโลยี ซึ่งลักษณะเป็นโครงสร้างลำดับชั้น สัญลักษณ์วงกลมแสดงแนวคิดหรือโหนด ส่วนเส้นเชื่อมระหว่างโหนดแทนความสัมพันธ์แบบคลาสย่อย โหนดที่อยู่ในระดับล่างจะมีความจำเพาะเจาะจงมากกว่าโหนดที่อยู่ในระดับบนขึ้นไป

ผู้เขียนแบ่งการบรรยายแนวคิดของโดเมนออนโทโลยีเป็น 2 ส่วน ส่วนแรก บรรยายประเภทของเซอร์วิสที่สามารถค้นหาได้ ส่วนที่สอง อธิบายชนิดข้อมูลที่สามารถเป็นไปได้ในส่วน input และ output ของฟังก์ชันบนเว็บเซอร์วิส จากภาพที่ 2 โหนด HotelService แทนแนวคิดเซอร์วิสประเภทโรงแรม ซึ่งเป็นส่วนหนึ่งของแนวคิดเซอร์วิสเกี่ยวกับการท่องเที่ยว (TravelService) และยังมีแนวคิดที่เฉพาะเจาะจงลงไป เช่น เซอร์วิสการจองโรงแรม (HotelBooking) และเซอร์วิสยกเลิกการจองโรงแรม (HotelCancelBooking) การบรรยายประเภทของเซอร์วิส จะระบุกลุ่มของเซอร์วิสตามมาตรฐาน UNSPSC และ NAICS ให้กับแต่ละแนวคิด เพื่อใช้ในส่วน Distribution Discovery เช่น แนวคิด HotelService มีรหัส UNSPSC คือ 90.11.15.01 (Hotels) และรหัส NAICS คือ 72111 (Hotels and Motels)



ภาพที่ 2 โดเมนออนโทโลยี

3.2 Distribution Discovery

หลังจากที่ผู้ใช้ส่งคำร้องเพื่อขอค้นหาเว็บเซอร์วิสผ่านทางอินเตอร์เฟสแล้ว จะถูกส่งมายัง Distribution Discovery การทำงานในส่วนนี้จะนำคำร้องไปเลือกแนวคิดที่ผู้ใช้ต้องการ

จากโดเมนออนโทโลยี และสร้างคำร้องเพื่อค้นหาข้อมูลจาก UBR เช่น ผู้ใช้เลือกเว็บเซอร์วิส HotelBooking ซึ่งมีคุณสมบัติกลุ่มเซอร์วิส UNSPSC คือ Hotels และ NAICS คือ Hotels and Motels คำร้องที่สร้างขึ้นแสดงดังภาพที่ 3(ก) ในส่วนอิลีเมนต์ keyedReference บอกกลุ่มของเซอร์วิสที่ต้องการ ส่วนอิลีเมนต์ findQuanlifier ซึ่งกำหนดเป็น orAllKeys คือ เลือกเซอร์วิสที่เป็นมาตรฐานกลุ่มใดกลุ่มหนึ่งจากอิลีเมนต์ categoryBag หลังจากนั้นคำร้องเหล่านี้จะถูกส่งไปยังแต่ละ UBR ซึ่งรายการ UBR ที่ใช้ถูกกำหนดไว้ใน UBRList.xml ดังภาพที่ 3(ข) อิลีเมนต์ Location คือ ตำแหน่งของ UBR เมื่อได้รับผลลัพธ์จาก UBR ครบถ้วนแล้ว ข้อมูลผู้ให้บริการจะถูกส่งไปยัง Result Integration เพื่อทำการบูรณาการต่อไป

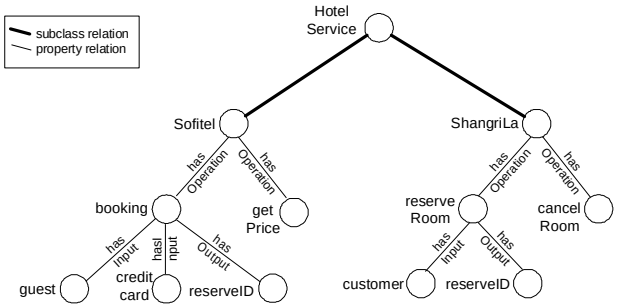
```
<find_business>
   <findQualifier>orAllKeys</ findQualifier>
   <categoryBag>
      <keyedReference tModelKey="uuid:db77455.."
                  keyValue="90.11.15.01 "
                  keyName=" Hotels"/>
      <keyedReference tModelKey="uuid:cd15325.."
                  keyValue="72111"
                  keyName="Hotels and Motels"/>
   </categoryBag>
</find_business>                                    (ก)
```

```
<Operators>
  <Operator>
     <Name>IBM</Name>
     <Location>http://uddi.ibm.com/beta/inquiryapi</
Location>
  </Operator>
  <Operator>
     <Name>Microsoft</Name>
     <Location>http://uddi.microsoft.com/inquire</
Location>
  </Operator>
  …
</Operators>                                        (ข)
```

ภาพที่ 3 (ก) คำร้องค้นหาเว็บเซอร์วิสจาก UBR
(ข) UBRList.xml

3.3 Result Integration

ผลลัพธ์จากแต่ละ UBR ที่ถูกส่งมายัง Result Integration จะถูกบูรณาการเป็น Integrated Ontology ขั้นตอนการบูรณาการเริ่มจากการค้นหาผู้ให้บริการที่ซ้ำซ้อนกัน จากภาพที่ 4 แสดงตัวอย่างของ Integrated Ontology หลังจากนั้นจะนำข้อมูลจากเอกสาร WSDL ได้แก่ ชื่อฟังก์ชัน, input และ output จากภาพที่ 4 โหนด Sofitel เมื่อนำข้อมูลจากเอกสาร WSDL จะเชื่อมด้วยความสัมพันธ์ hasOperation กับโหนด booking และ getPrice ซึ่งทั้งสองเป็นเซอร์วิสของ Sofitel นอกจากนี้จะเห็นว่าโหนด booking ยังมีความสัมพันธ์แบบ hasInput และ hasOutput กับโหนดอื่นๆ ซึ่งความสัมพันธ์เหล่านี้หมายถึง input และ output ของเซอร์วิส โดยโหนด booking มี input คือ guest และ creditcard และ output คือ reservationID Result Integration จะนำข้อมูลจากเอกสาร WSDL มาสร้างเป็นความสัมพันธ์บน Integrated Ontology จนครบทุกโหนด ซึ่ง Integrated Ontology ที่สมบูรณ์จะนำไปใช้จับคู่ออนโทโลยีในส่วนต่อไป



ภาพที่ 4 Integrated Ontology

3.4 Ontology Mapping

Integrated Ontology ที่ได้จากส่วน Result Integration เป็นเสมือนตัวแทนข้อมูลจาก UBR ทั้งหมด ขั้นตอนของ Ontology Mapping คือการใช้วิธีของการจับคู่ออนโทโลยีตามวิธี ONION ซึ่งกล่าวในหัวข้อที่ 2.2 จากภาพที่ 5 ด้านซ้ายของภาพ คือ โดเมนออนโทโลยี ด้านขวาของภาพ คือ Integrated Ontology ขั้นตอนของ ONION จะพิจารณาหาค่าความเหมือนของแต่ละคู่แนวคิด เช่น HotelBooking กับ booking หากแนวคิดคู่ใดมีค่าความเหมือนมากกว่า 0 จะทำการจับคู่แนวคิดนั้น จากภาพที่ 5 เส้นประที่ลากจากโหนด HotelBooking ไปยังโหนด booking และ reserveRoom หมายความว่า booking และ reserveRoom ถูกจับคู่กับแนวคิด HotelBooking ค่า 1.0 ที่อยู่บนเส้นแสดงค่าความเหมือนของแนวคิดทั้งสอง ขั้นตอนการจับคู่จะทำจนกว่าจะครบทุกโหนดในโดเมนออนโทโลยี และส่งไปยัง Request Matching เพื่อการค้นหาผลลัพธ์ต่อไป

3.5 Request Matching

การทำงานในส่วน Request Matching เป็นการจับคู่ระหว่างคำร้องที่ผู้ใช้ส่งมากับโดเมนออนโทโลยีที่ผ่านการ

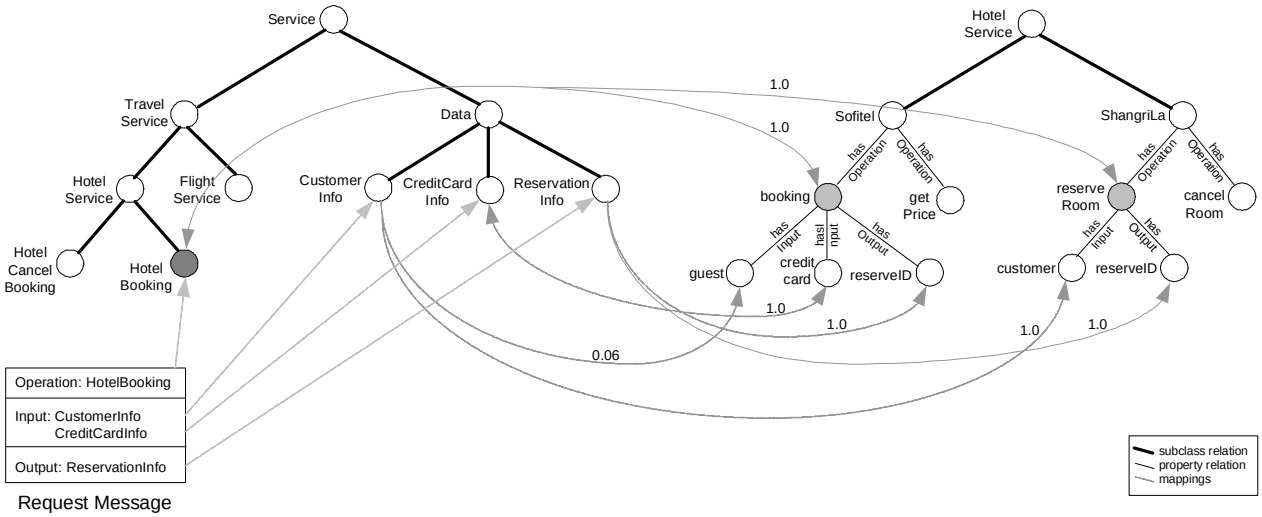
จับคู่จาก Ontology Mapping คำร้องของผู้ใช้จะแปลงให้เป็น Query เพื่อใช้ในการค้นหาแนวคิดที่ต้องการจากโดเมนออนโทโลยี จากภาพที่ 5 ส่วนล่างด้านซ้ายมือของภาพ คือ Request Message ซึ่งอธิบายได้ว่า ผู้ใช้ต้องการเซอร์วิสการจองโรงแรม (HotelBooking) ที่มี input คือ CustomerInfo และ CreditCardInfo และ output คือ ReservationInfo ขั้นตอนแรกจะทำการ Query เซอร์วิสที่ถูกจับคู่กับแนวคิดที่ผู้ใช้ต้องการ ซึ่งในที่นี้ คือ booking และ reserveRoom ขั้นตอนที่สอง ทำการ Query หา input และ output ที่ถูกจับคู่กับแนวคิดเดียวกันกับที่ผู้ใช้ต้องการ จากภาพที่ 5 เซอร์วิส booking มี input คือ guest ซึ่งถูกจับคู่กับ CustomerInfo ขั้นตอนที่สาม คือ การหาค่าความเหมือน โดยค่าความเหมือนของแต่ละเซอร์วิสนั้น คือ ผลรวมของค่าความเหมือนของเซอร์วิส, input และ output จากภาพที่ 5 ค่าความเหมือนจะแสดงบนเส้นประ ซึ่งผลรวมของค่าความเหมือนของเซอร์วิส booking คือ 1+1+1+0.06 = 3.06 และเซอร์วิส reserveRoom คือ 1+1+1 = 3 ดังนั้นเซอร์วิส booking จะตรงกับความต้องการของผู้ใช้มากกว่าเซอร์วิส reserveRoom

4. บทสรุป

ปัญหาการค้นหาเว็บเซอร์วิสในปัจจุบันมี 2 ส่วน ส่วนแรก คือ ปัญหาความหลากหลายของ UBR ส่วนที่สอง คือ ปัญหาการค้นหาโดยไม่สนใจความหมายของคำ ปัญหาความหลากหลายของ UBR สร้างความยุ่งยากต่อผู้ใช้ในการเปรียบเทียบและเลือกเว็บเซอร์วิสที่ต้องการ ผู้เขียนได้นำเสนอแนวคิดในการบูรณาการข้อมูลจาก UBR ที่มีอยู่หลากหลาย โดยการประมวลผลข้อมูลเพื่อจัดโครงสร้างข้อมูลใหม่ในช่วงรันไทม์ เพื่อให้โครงสร้างมีความชัดเจน และง่ายต่อการค้นหาเว็บเซอร์วิส ส่วนปัญหาการค้นหาที่ไม่พิจารณาความหมายของคำนั้น ผู้เขียนได้นำการจับคู่ออนโทโลยีมาใช้ เพื่อจับคู่แนวคิดที่มีความหมายคล้ายคลึงกัน วิธีการแก้ปัญหาทั้งสองส่วนมีผลเกี่ยวเนื่องกัน การบูรณาการทำให้ได้ข้อมูลที่ถูกรวบรวมเป็นหนึ่งเดียว การจับคู่แนวคิดทำให้เชื่อมโยงข้อมูลในระดับแนวคิดกับข้อมูลจริง การค้นหาเว็บเซอร์วิสจึงสามารถเลือกแนวคิดที่เกี่ยวข้องได้โดยไม่จำเป็นต้องมีรูปแบบคำที่เหมือนกัน ผู้ใช้จะได้รับผลลัพธ์เว็บเซอร์วิสที่เกี่ยวข้องมากขึ้น และไม่ยุ่งยากในการค้นหาเว็บซอร์วิสจาก UBR ที่มีอยู่หลายแหล่ง บทความนี้นำเสนอเพียงแนวคิดในการค้นหาเว็บเซอร์วิส ซึ่งเป็นส่วนหนึ่งของงานวิจัย การดำเนินการต่อไป คือสร้างแอปพลิเคชั่นในการทำงานกับ UBR ตามแนวคิดที่นำเสนอ พร้อมกับเพิ่มความสามารถในการค้นหา ด้วยตรรกะ ผู้ใช้สามารถระบุเงื่อนไขในการค้นหาได้มากขึ้น ซึ่งเป็นการส่งเสริมผลลัพธ์ของการค้นหาที่มีความหมายมากยิ่งขึ้น

5. เอกสารอ้างอิง

[1] Leo J. Obrst, Michael C. Daconta, Kevin T. Smith, “Understanding Ontologies”, *The Semantic Web*, 2003.



ภาพที่ 5 การจับคู่ระหว่างโดเมนออนโทโลยีและ Integrated Ontology

[2] Kunal Verma, Kaarthik Sivashanmugam, Amit Sheth, John Miller, “Adding Semantics to Web Services Standards”, *International Conference on Web Services 2003*, 2003.

[3] W3C, “Web Services Description Language (WSDL)”, http://www.w3.org/TR/wsdl, 2001.

[4] UDDI.org, “Universal Description, Discovery and Integration of Web Services”, http://www.uddi.org, 2001.

[5] Takahiro Kawamura, Massimo Paolucci, Terry R. Payne, Katia Sycara, “Importing the Semantic Web in UDDI “, *Forthcoming in Proceedings of Web Services, E-business and Semantic Web Workshop*, 2002.

[6] James Hendler, “Agents and the Semantic Web”, *IEEE Intelligent Systems*, vol.16, no.2, pp. 30-37, 2001.

[7] Macro Schorlemmer, Yannis Kalfoglou, “Ontology mapping: the state of the art”, *The Knowledge Engineering Review*, vol.18, no.1, pp. 1-31, 2003.

[8] N.F. Noy, Musen M., “SMART: automated support for ontology merging and alignment”, *Proceedings of the 12th Workshop on Knowledge Acquisition, Modelling and Management (KAW99)*, 1999.

[9] N.F. Noy, Musen M., “PROMPT: algorithm and tool for automated ontology merging and alignment”, *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI00)*, pp. 2000.

[10] Alexander Maedche, Gerd Stumme, “FCA-MERGE: Bottom-Up Merging of Ontologies”, *Proc. 17th Intl. Conf. on Artificial Intelligence (IJCAI ’01)*, 2001.

[11] Macro Schorlemmer, Yannis Kalfoglou, “IF-Map: An Ontology-Mapping Method based on Information-Flow Theory”, *Journal of Data Semantics*, vol.1, no.1, pp. 98-127, 2003.

[12] Gio Wiederhold, Prasenjit Mitra, “Resolving Terminological Heterogeneity in Ontologies”, *Proceedings of Workshop on Ontologies and Semantic Interoperability at the 15th European Conference on Artificial Intelligence (ECAI)*, 2002.

[13] Prince-ton University, Cognitive Science Laboratory, “Wordnet - a lexical database for english”, http://wordnet.princeton.edu/, 1998.

[14] Takahiro Kawamura, Massimo Paolucci, Terry R. Payne, Katia Sycara, “Semantic Matching of Web Services Capabilities”, *Proceedings of the 1st International Semantic Web Conference (ISWC 2002)*, 2002.