

บทที่ 6 สรุปผลการวิจัยและข้อเสนอแนะ



6.1 สรุปผลการวิจัย

ผู้วิจัยได้เสนอขั้นตอนวิธีการวิเคราะห์เว็บที่การเข้าใช้เว็บไซต์เพื่อหารูปแบบการเยี่ยมชมเว็บไซต์โดยขั้นตอนวิธีที่นำเสนอสามารถแบ่งออกได้เป็น 3 ส่วน

ส่วนแรกทำหน้าที่สำหรับเก็บรวบรวมข้อมูลการเยี่ยมชมซึ่งสร้างขึ้นโดยแทรกจาวาสคริปต์ลงไปในทุกเว็บเพจในเว็บไซต์ซึ่งสคริปต์ดังกล่าวจะทำหน้าที่เก็บรวบรวมข้อมูลการเยี่ยมชมและกำหนดรหัสประจำตัวผู้ชมลงในเครื่องคอมพิวเตอร์ของผู้ชมในรูปของคุกกี้ซึ่งช่วยให้สามารถติดตามและแยกแยะข้อมูลการเยี่ยมชมในแต่ละเครื่องได้อย่างถูกต้อง และข้อมูลการเยี่ยมชมดังกล่าวจะถูกส่งกลับไปยังเว็บไซต์ทุกครั้งที่มีผู้ชมเบราว์เซอร์เว็บเพจภายในเว็บไซต์

ส่วนที่สองทำหน้าที่สำหรับแปลงข้อมูลที่เก็บรวบรวมมาได้ให้อยู่ในรูปแบบที่เหมาะสมต่อการนำไปประมวลผล ต่อจากนั้นจึงนำข้อมูลที่แปลงรูปแบบเสร็จเรียบร้อยแล้วไปสู่ขั้นตอนการกำหนดทรานแซคชันเพื่อแบ่งข้อมูลการเยี่ยมชมของผู้ชมแต่ละคนออกเป็นทรานแซคชัน ในขั้นตอนการกำหนดทรานแซคชันจะพิจารณาระยะเวลาเยี่ยมชมเป็นส่วนประกอบด้วยโดยจะแบ่งทรานแซคชันของผู้ชมแต่ละคนตามรหัสประจำตัวผู้ชมและระยะเวลาของการเยี่ยมชมในแต่ละครั้งซึ่งจะมีค่าอยู่ในช่วง 2-30 นาที หลังจากนั้นจึงนำข้อมูลการเยี่ยมชมที่เก็บรวบรวมเอาไว้และทรานแซคชันที่ได้ไปวิเคราะห์ทางสถิติเพื่อหาสถิติการเยี่ยมชมในรูปแบบต่างๆต่อไป

ส่วนสุดท้ายทำหน้าที่หากฎความสัมพันธ์และรูปแบบลำดับการเยี่ยมชมจากทรานแซคชันโดยการค้นพบกฎความสัมพันธ์และรูปแบบลำดับที่ว่าเป็นส่วนหนึ่งของเทคนิคในการทำเหมืองข้อมูลซึ่งในงานวิจัยนี้ใช้อัลกอริทึม Apriori ในการหากฎความสัมพันธ์และใช้อัลกอริทึม Apriori All ในการหารูปแบบลำดับ

จากการทดลองพบว่าขั้นตอนวิธีที่นำเสนอสามารถหากฎความสัมพันธ์และรูปแบบลำดับได้จริงและสามารถนำกฎความสัมพันธ์และรูปแบบลำดับดังกล่าวมาใช้เป็นข้อมูลช่วยนักออกแบบเว็บไซต์ในการตัดสินใจปรับโครงสร้างลิงก์ของเว็บไซต์ แต่เนื่องจากการปรับโครงสร้างลิงก์ของเว็บไซต์ตามกฎความสัมพันธ์และรูปแบบลำดับที่ค้นพบเป็นเรื่องที่มีความซับซ้อนมากจึงจำเป็นต้องมีการวิจัยเพิ่มเติม

6.2 ข้อดีและข้อเสียของขั้นตอนวิธี

ขั้นตอนวิธีที่นำเสนอ มีทั้งข้อดีและข้อเสียซึ่งสามารถแจกแจงได้ดังนี้

6.2.1 ข้อดี

- ทำให้สามารถรู้จำนวนผู้ชมที่แท้จริงในแต่ละวัน
- ทำให้สามารถจำแนกข้อมูลการท่องเว็บไซต์ของผู้ชมแต่ละคนได้อย่างถูกต้อง
- ทำให้สามารถค้นพบกฎความสัมพันธ์จากพฤติกรรมกรรมการท่องเว็บไซต์ของผู้ชม
- ทำให้สามารถค้นพบรูปแบบลำดับการท่องเว็บไซต์ของผู้ชม

6.2.2 ข้อเสีย

- การติดตั้งและโยกย้ายระบบมีความยุ่งยากเพราะต้องติดตั้งซอฟต์แวร์ต่างๆที่ใช้ให้ครบถ้วน
- การแทรกสคริปต์ลงไปในเว็บไซต์ทุกเพจในเว็บไซต์มีความยุ่งยาก
- การแทรกสคริปต์ลงไปในเว็บไซต์มีผลทำให้ระยะเวลาที่ใช้ในการโหลดเว็บไซต์เพิ่มขึ้น
- โปรแกรมที่พัฒนาขึ้นใช้เวลาในการประมวลผลนานกว่าเมื่อเปรียบเทียบกับแฮชเกอรัทัวๆไป
- โปรแกรมที่พัฒนาผูกติดกับซอฟต์แวร์ระบบจัดการฐานข้อมูลและภาษาสคริปต์ที่ใช้

6.3 ข้อเสนอแนะ

ผู้วิจัยพบว่าข้อเสนอนี้บางประการที่น่าจะเป็นประโยชน์ และสามารถนำไปใช้ในการพัฒนาขั้นตอนวิธีการค้นพบความรู้จากล็อกไฟล์ เพื่อให้มีประสิทธิภาพในการค้นพบความรู้ที่ดียิ่งขึ้น ดังนี้

- ปรับปรุงวิธีการนับค่าชัพพอร์ตจากทรานแซคชันโดยนำโครงสร้างข้อมูลแบบแฮชทรี (Hash-Tree) มาประยุกต์ใช้
- พัฒนาการวิเคราะห์ข้อมูลทางสถิติให้มีความสมบูรณ์ยิ่งขึ้นโดยนำแฮชเกอรัทัวๆไปมาเป็นต้นแบบ
- ปรับปรุงโปรแกรมให้สามารถใช้ได้กับล็อกไฟล์ที่สร้างจากโปรแกรมเว็บเซิร์ฟเวอร์ประเภทต่างๆ
- พัฒนาการกำหนดรหัสประจำตัวผู้ชมเพื่อให้สามารถใช้ได้กับล็อกไฟล์ที่ไม่มีรหัสประจำตัวของผู้ชม
- พัฒนาระบบจัดเก็บข้อมูลประวัติของผู้ชมแต่ละคนเพื่อนำไปประยุกต์ใช้ในการจำแนกพฤติกรรมกรรมการเยี่ยมชม

- พัฒนาระดับขั้นตอนวิธีการค้นพบกฎการจำแนกโดยทำการจำแนกผู้ชมจากประวัติและพฤติกรรมกรรรมการเยี่ยมชม
- พัฒนาระดับขั้นตอนวิธีปรับโครงสร้างลิงก์ของเว็บไซต์ตามกฎความสัมพันธ์และรูปแบบลำดับที่ค้นพบแบบอัตโนมัติ