

ทศพร ศิรามพุช 2556: การผสมผสานเงื่อนไขในการแบ่งกลุ่มกระแสดัชนีแบบกึ่งมี  
ผู้สอน ปริญญาวิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมคอมพิวเตอร์) สาขาวิศวกรรม  
คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก:  
รองศาสตราจารย์กฤษณะ ไวยมัย, D.U. 57 หน้า

ในปัจจุบันการแบ่งกลุ่มกระแสดัชนีมีการศึกษาวิจัยและใช้งานอย่างแพร่หลาย อย่างไรก็ตาม เทคนิคเหล่านี้ยังขาดการใช้ความช่วยเหลือจาก ความรู้ของผู้เชี่ยวชาญ หรือ ประสบการณ์จากในอดีต ซึ่งสามารถประยุกต์ใช้ใน แอปพลิเคชันจริง ข้อดีของการใช้เทคนิคเหล่านี้จะช่วยให้ประสิทธิภาพและความแม่นยำของกลุ่มข้อมูลผลลัพธ์ดีขึ้น โดยงานวิจัยนี้นำเสนอการแบ่งกลุ่มกระแสดัชนีที่ใช้ความรู้มาช่วยในรูปแบบของ เงื่อนไข (Constraints) ชื่อว่า CE-Stream

เงื่อนไขในระดับข้อมูลถูกนำมาใช้เพื่อช่วยเหลือการแบ่งกลุ่มข้อมูลให้ดียิ่งขึ้น ได้แก่ การใช้เงื่อนไขแบบเชื่อมต่อในการแยกตัวของกลุ่มข้อมูล และ การใช้เงื่อนไขแบบไม่เชื่อมต่อในการส่งมอบข้อมูลสมาชิก รวมไปถึง การออกแบบเงื่อนไขให้รองรับกับ พฤติกรรมของกระแสดัชนีลักษณะต่างๆ อาทิเช่น การพิจารณาการใช้งานของเงื่อนไข (Constraint Activation) การเลือนหายของเงื่อนไข (Constraints Fading) และ การหมดอายุของเงื่อนไข (Constraints Outdating) ยิ่งไปกว่านั้น CE-Stream ยังช่วยลด อัตราการแบ่งแยกกลุ่มข้อมูลที่ไม่จำเป็นให้น้อยลงในระหว่างขั้นตอนการแบ่งกลุ่มข้อมูล และ การส่งมอบสมาชิกข้อมูลที่ผิดอีกด้วย

จากผลการทดลองกับชุดข้อมูลมาตรฐาน Covertype และ KDDcup'99 ซึ่งเป็น ชุดข้อมูลการแบ่งกลุ่มประเภทป่าฝนจากภูมิประเทศ และ ชุดข้อมูลการตรวจจับการบุกรุกเครือข่ายตามลำดับ วัดผลด้วยค่าเอฟเมเชอร์และค่าความบริสุทธิ์ พบว่าผลการทดลองเมื่อเทียบกับเทคนิค E-Stream ดั้งเดิมมีประสิทธิภาพดีขึ้น