

## บทคัดย่อ

รหัสโครงการ TRG5880089  
ชื่อโครงการ การค้นหาช่วงเวลาที่น่าสนใจของการปรากฏและไม่ปรากฏขึ้นของรูปแบบ  
ชื่อนักวิจัย โกเมศ อัมพวัน ห้องปฏิบัติการนวัตกรรมการประมวลผล คณะวิทยาการสารสนเทศ  
มหาวิทยาลัยบูรพา  
อีเมล [komate@gmail.com](mailto:komate@gmail.com)  
ระยะเวลาโครงการ 2 ปี

ณ ปัจจุบัน การค้นหารูปแบบที่ปรากฏบ่อยและสม่ำเสมอเป็นหัวข้อหนึ่งที่สำคัญทางด้านการทำเหมืองข้อมูล โดยหัวข้อดังกล่าวมุ่งเน้นที่การค้นหาเซตของรูปแบบ (เซตรายการ) ที่น่าสนใจด้วยการพิจารณาเกี่ยวกับความบ่อยและความสม่ำเสมอของปรากฏขึ้นของรูปแบบนั้น ๆ โดยเซตรายการ (รูปแบบ) หนึ่ง ๆ จะเป็นเซตรายการที่น่าสนใจก็ต่อเมื่อเซตรายการนี้ปรากฏขึ้นในฐานข้อมูลบ่อยเกินกว่าค่าขีดแบ่งสนับสนุนและปรากฏอย่างสม่ำเสมอกว่าค่าขีดแบ่งสม่ำเสมอที่ผู้ใช้กำหนด การค้นหารูปแบบดังกล่าวสามารถถูกประยุกต์ใช้ในหลายแง่มุม อาทิเช่น การวิเคราะห์พฤติกรรมการซื้อขายของสินค้า การวิเคราะห์พฤติกรรมของการใช้อินเทอร์เน็ต การวิเคราะห์ความสัมพันธ์และความเกี่ยวเนื่องของยีนในโครโมโซม และอื่น ๆ แต่อย่างไรก็ตาม การค้นหารูปแบบที่ปรากฏบ่อยและสม่ำเสมอยังคงไม่ได้พิจารณาช่วงของการปรากฏและ/หรือไม่ปรากฏขึ้นของรูปแบบที่ซึ่งจะช่วยให้สามารถวิเคราะห์ถึงสาเหตุของการปรากฏและ/หรือไม่ปรากฏขึ้นของรูปแบบนั้น ๆ การค้นหาช่วงของการปรากฏ (ไม่ปรากฏขึ้น) ของรูปแบบสามารถประยุกต์ใช้ได้หลากหลาย อาทิเช่น การประยุกต์ใช้ในธุรกิจค้าปลีกจะสามารถช่วยให้ผู้จัดการห้างร้านสามารถทราบถึงช่วงเวลาที่ยearsสินค้าต่าง ๆ หรือไม่ถูกซื้ออย่างต่อเนื่อง การประยุกต์ใช้ในการแพทย์สามารถช่วยให้แพทย์สามารถทราบถึงช่วงเวลาของการใช้ยา ช่วงเวลาของการเกิดขึ้นของอาการเจ็บป่วยที่เกิดขึ้น (หรือไม่เกิดขึ้น) กับผู้ป่วย และอื่น ๆ

จากข้างต้น งานวิจัยนี้จึงมีแนวคิดที่จะนำเสนอแนวความคิดในการค้นหาช่วงที่น่าสนใจของการปรากฏ (ไม่ปรากฏขึ้น) ของรูปแบบ นิยาม ปัญหา และนำเสนอขั้นตอนวิธีที่มีประสิทธิภาพในการค้นหาช่วงที่น่าสนใจทั้งหมด นอกจากนั้น งานวิจัยนี้ยังได้นำเสนอการทดลองเพื่อทดสอบการค้นหาค้นหาช่วงที่น่าสนใจของการปรากฏ (ไม่ปรากฏขึ้น) ของรูปแบบ และขั้นตอนที่นำเสนอ โดยทำการทดลองกับข้อมูลจำลองและข้อมูลจริงที่เป็นข้อมูลที่เผยแพร่อย่างแพร่หลาย ท้ายสุด ได้นำเสนอการวิเคราะห์ความซับซ้อนในการคำนวณในเชิงเวลาและหน่วยความจำที่ใช้ในการประมวลผล

คำสืบค้น : รูปแบบที่น่าสนใจ ช่วงของการขาดหายไป พฤติกรรมของการขาดหายไป

## Abstract

Project code	TRG5880089
Project title	Discovering interesting periods based on patterns' appearance/ disappearance behavior
Investigator	Komate Amphawan CIL, Faculty of Informatics, Burapha University
E-mail	<a href="mailto:komate@gmail.com">komate@gmail.com</a>
Project Period	2 years

Frequent-regular pattern mining is nowadays an important topic in data mining community. It aims to find out the set of interesting patterns by investigating how often and how regular occurrence of patterns. In this framework, a pattern is interesting if it is more frequently occurs than a user-given frequency threshold and also more regularly occurs than a user-given regularity threshold. This task can be applied in a wide range of applications such as buying behavior analysis, surfing internet behavior analysis, analysis of related genes in chromosomes, etc. However, this approach does not consider periods based on patterns' appearance/disappearance behavior at which periods enable us to know about time interval that patterns or most of items appear or disappear in databases and then we can further analyze the reason of appearance/disappearance to improve/extent/reduce the periods of appearance/ disappearance. Discovering interesting periods can be applied into several applications such as i) in a retail market, the manager of a super-store may want to know that which period that group of products have been sold consecutively/have not been sold for a while, 2) for medical analysis, doctors may want to track the period of taking medicine, period of symptoms that occurs/not occurs on a patient and so on.

Therefore, the objective of this research is to introduce a novel concept of discovering interesting periods based on patterns' appearance/disappearance behavior. An efficient algorithm is also proposed to quickly discover such interesting periods. The performance study on both real and synthetic data will be conducted to show the effective and efficient of the proposed algorithm in terms of runtime and memory consumption to discover the interesting periods based on users' interest.

Keywords : Interesting patterns, Disappearance gap, Absence behavior