

ปัญหาการจับคู่ของรูปแบบที่คล้ายกันของลำดับข้อมูลปรากฏอยู่ในงานประยุกต์มากมาย ประเด็นที่สำคัญในการคำนวณการจับคู่เพื่อให้ได้คู่ของลำดับข้อมูลที่ต้องการประเด็นหนึ่ง คือการย่อ-ขยายและการเลื่อนของลำดับข้อมูล แม้ว่าจะมีการนำเสนอวิธีการที่สามารถแก้ไขปัญหานี้ได้อย่างมีประสิทธิภาพและมีประสิทธิผลแล้ว อย่างไรก็ตามในบางงานประยุกต์ ลำดับข้อมูลอาจถูกเลื่อนเท่านั้น แม้ว่าในปัญหาของงานประยุกต์ดังกล่าวจะสามารถนำงานที่ถูกลำเสนอให้แก้ปัญหทั้งการย่อ-ขยายและการเลื่อนมาใช้ได้ แต่วิธีการเหล่านั้นก็อาจจะมีความซับซ้อนเชิงคำนวณที่สูงเกินกว่าความจำเป็นเนื่องจากลำดับข้อมูลจะถูกเลื่อนเท่านั้น

งานวิจัยนี้ได้มีการนำเสนอขั้นตอนวิธีที่มีประสิทธิภาพ ในการจับคู่ที่คล้ายกันสำหรับลำดับที่มีการเลื่อนเท่านั้น โดยขั้นตอนวิธีที่พัฒนาขึ้น จะเริ่มจากการแปลงทั้งลำดับข้อมูลที่ต้องการค้นหาและลำดับข้อมูลในฐานข้อมูลให้อยู่ในรูปแบบจุด จากนั้นการจับคู่ระหว่างลำดับข้อมูลที่ต้องการค้นหากับลำดับข้อมูลใด ๆ ในฐานข้อมูล ก็สามารถทำได้โดยการเปรียบเทียบจุดต่อจุด ทั้งนี้ได้เพิ่มประสิทธิภาพในการค้นหาดังกล่าวด้วยการนำโครงสร้างแบบอาร์ทริมาประยุกต์ใช้ จากนั้นทดสอบประสิทธิภาพทางด้านความเร็วในการทำงานของขั้นตอนวิธีที่พัฒนาขึ้น โดยเปรียบเทียบความเร็วที่ได้จากขั้นตอนวิธีที่พัฒนาขึ้น กับขั้นตอนวิธีการจับคู่ที่คล้ายกันสำหรับลำดับที่มีการย่อ-ขยายและเลื่อน

จากผลการทดลองเมื่อทำการแปรผันค่าสี่ค่าดังต่อไปนี้ ความคลาดเคลื่อน ความยาวของลำดับ จำนวนลำดับในฐานข้อมูล และจำนวนลำดับข้อคำถาม พบว่าขั้นตอนวิธีการจับคู่ที่คล้ายกันสำหรับลำดับที่มีการเลื่อน ใช้เวลาในการประมวลผลน้อยกว่าขั้นตอนวิธีการจับคู่ที่คล้ายกันสำหรับลำดับที่มีการย่อ-ขยายและการเลื่อนในทุกกรณี โดยระยะห่างของเวลาในการประมวลผลคิดโดยเฉลี่ยเป็น 3.5% ของจำนวนลำดับข้อมูลผลลัพธ์ที่ได้จากการจับคู่ ดังนั้นจึงสามารถสรุปได้ว่าขั้นตอนวิธีการจับคู่ที่คล้ายกันสำหรับลำดับที่มีการเลื่อน มีประสิทธิภาพทางด้านความเร็วในการประมวลผลโดยรวมสูงกว่า ขั้นตอนวิธีการจับคู่ที่คล้ายกันสำหรับลำดับที่มีการย่อ-ขยายและการเลื่อน

Similar pattern matching problem on the data sequences exists in many applications. One of the important issues for computing the matching effectively is scaling and shifting of the given sequences. Though, there are many effective and efficient solutions for resolving such the issue. However, there are some applications which only the shifting issue needs to be addressed. Although, the shifted-only pattern matching can be solved by applying the scaling-and-shifting matching solutions directly. However, the time complexity could be excessive since the scaling computation is not necessary. In this thesis, we propose an efficient algorithm to match similar shifted sequences. The proposed transforms both the query sequences and the data sequences in the given data base into point geometrically. Subsequently, the sequence matching can be processed simply by the point-to-point comparing-computation.

From experimental we have found that the proposed algorithm is more efficient than the existing scaling-and-shifting algorithm in all cases, i.e., when the error, sequences length, the number of data sequences and query sequences are varied. Approximately, the proposed algorithm can perform faster than the scaling-and-shifting 3.5 %. So we can conclude that the proposed shifted-only similar pattern matching algorithm has more efficient than the existing work.