

การค้นคืนสารสนเทศด้วยเทคนิคการกำหนด ความใกล้ชิดของคำค้น

Information Searching with Proximity Operator

ดร.สมาน ลอยฟ้า¹

Dr. Smarn Loipha

บทคัดย่อ

การกำหนดความใกล้ชิดของคำค้น เป็นเทคนิคหนึ่งที่ใช้ในระบบการค้นคืนสารสนเทศ โดยมุ่งค้นหาเอกสารที่ปรากฏคำค้นซึ่งอยู่ใกล้ชิดกัน เทคนิคนี้ช่วยลดจำนวนผลการค้นที่ไม่เกี่ยวข้องและเพิ่มประสิทธิภาพในการค้น เนื่องจากมีฐานข้อมูลเชิงพานิชย์หลายฐานข้อมูลที่สามารถใช้เทคนิคการค้นนี้ ตัวอย่างเช่น Science Direct, IEEE, Emerald Management, Academic Search premier เป็นต้น ซึ่งแต่ละฐานข้อมูลมีความแตกต่างกัน บทความนี้เป็นการนำเสนอแนวคิดเกี่ยวกับเทคนิคการค้นดังกล่าว อาทิ ความหมาย ความสำคัญ ลักษณะ รูปแบบและวิธีการใช้

Abstract

Proximity searching is used in information retrieval systems to locate documents that have keywords occurring near each other. It is one method to reduce the number of search results and improve precision. Many commercial databases allow user to use this technique but their patterns are different, for example,

¹ รองศาสตราจารย์ กลุ่มวิชาการจัดการสารสนเทศและการสื่อสาร คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น
Associate Professor, Information and Communication Management Program,
Faculty of Humanities and Social Sciences, Khon Kaen University

Science Direct, IEEE, Emerald Management, Academic Search premier. This article presents the proximity searching concepts: definition, important, features, format and method of use.

คำค้น : การค้นคืนสารสนเทศ เทคนิคการค้นคืนสารสนเทศ

Keywords : Information Searching, Proximity Operator, Proximity Search

บทนำ

การค้นคืนสารสนเทศในปัจจุบันเป็นการค้นคืนโดยใช้คำค้น ซึ่งเป็นวิธีที่สะดวกสำหรับผู้ใช้งาน (User friendly) และในการค้นคืนสารสนเทศ โดยใช้โปรแกรมค้นหา (Search engines) เป็นเครื่องมือที่มีวัตถุประสงค์ที่สำคัญคือ เพื่อช่วยค้นหาคำตอบที่ถูกต้องตรงกับข้อความสั่งค้น (Query) มากที่สุด จัดลำดับความสำคัญของผลการค้นอย่างถูกต้อง และดำเนินการค้นอย่างรวดเร็ว (Gupta, Ozsoyoglu, and Ozsoyoglu, 2009) อย่างไรก็ตามในระบบการค้นคืนด้วยคำค้น แม้ว่าเอกสารที่ได้รับจากการค้นจะประกอบด้วยคำค้นที่ครบถ้วนและตรงกับคำค้นในข้อความสั่งค้น แต่เนื่องจากระบบไม่ได้คำนึงตำแหน่งของคำค้น ดังนั้น ผลการค้นจึงอาจได้เอกสารที่ไม่เกี่ยวข้อง กล่าวคือ เอกสารที่ค้นได้มีคำค้นปรากฏครบถ้วน แต่เนื้อหาอาจกลายเป็นเรื่องอื่น เพราะคำค้นที่ปรากฏไม่มีความสัมพันธ์กับสิ่งที่ต้องการค้นหา หรือลำดับของคำค้นที่ปรากฏในเอกสารไม่ตรงกับในข้อความสั่งค้น หรือตำแหน่งของคำค้นที่ปรากฏในตัวเอกสารไม่ได้ใกล้เคียงกันเหมือนกับในข้อความสั่งค้น เป็นต้น ซึ่งทำให้ผู้ใช้ไม่ได้รับเอกสารตามที่ต้องการ (Gupta, Ozsoyoglu, and Ozsoyoglu, 2009) นอกจากนี้ ยังทำให้ผู้ใช้ต้องใช้เวลาเพื่อการตรวจสอบเอกสารแต่ละรายการที่ค้นได้ว่ามีความเกี่ยวข้องตรงกับที่ต้องการหรือไม่อีกด้วย

เนื่องจากการค้นคืนด้วยคำค้นยังมีจุดอ่อนบางประการ จึงทำให้ผลลัพธ์ที่ได้จากการค้นบางส่วนไม่ตรงกับความต้องการของผู้ใช้ และการจัดลำดับความสำคัญของผลการค้นยังไม่เหมาะสม เป็นต้น ดังนั้น การใช้เทคนิคต่างๆ ช่วยในการค้นจึงเป็นสิ่งจำเป็น การกำหนดความใกล้ชิดของคำค้น เป็นเทคนิคหนึ่งที่มีการใช้เพื่อเพิ่มคุณภาพในระบบการค้นด้วยคำค้น (Keyword-oriented retrieval systems) (Schenkel and others, 2007) ซึ่งเทคนิคนี้จะช่วยลดจำนวนผลลัพธ์ที่ไม่เกี่ยวข้องลงและได้ผลลัพธ์ที่ตรงกับความต้องการ

มากขึ้น ปัจจุบันฐานข้อมูลเชิงพาณิชย์จำนวนมากหลายฐานข้อมูลสามารถใช้เทคนิคการกำหนดความใกล้ชิดของคำค้น ดังนั้น เพื่อเป็นการเพิ่มประสิทธิภาพในการค้นคืน บทความนี้จะนำเสนอแนวคิดเกี่ยวกับเทคนิคดังกล่าว

ความหมายของเทคนิคการกำหนดความใกล้ชิดของคำค้น

จากการสำรวจเอกสารที่เกี่ยวข้อง พบว่า มีการให้ความหมายของการค้นคืนแบบกำหนดความใกล้ชิดของคำค้นไว้หลายความหมาย ดังเช่น

การค้นคืนแบบกำหนดความใกล้ชิดของคำค้น (Proximity search) หมายถึง การค้นหาเอกสารโดยที่คำค้นทั้งหมดปรากฏอยู่ในตำแหน่งที่ใกล้กัน (Sadakane and Imai, 2000)

การค้นคืนแบบกำหนดความใกล้ชิดของคำค้น หมายถึง การค้นหาเอกสารโดยใช้คำค้น 2 คำค้นหรือมากกว่า และมีการกำหนดระยะห่างของคำค้น ซึ่งก็คือ การที่มีคำอื่นจำนวนหนึ่งมาแทรกอยู่ระหว่างคำค้น นอกจากนี้ ยังครอบคลุมถึงการกำหนดลำดับของคำค้นด้วย (Proximity search, 2011)

การค้นคืนแบบกำหนดความใกล้ชิดของคำค้น หมายถึง การค้นหาข้อมูลจากฐานข้อมูลโดยใช้ Proximity operator ในข้อความสั่งค้น (Search statement) เพื่อระบุว่าให้ค้นหาเฉพาะระเบียบที่ปรากฏคำค้นตามข้อความสั่งค้น โดยให้มีคำอื่นมาแทรกระหว่างคำค้นได้ตามจำนวนที่กำหนด หรือให้คำค้นปรากฏอยู่ภายในประโยคเดียวกัน หรือในย่อหน้าเดียวกัน (Reitz, 2012)

การค้นคืนแบบกำหนดความใกล้ชิดของคำค้น เป็นเทคนิคการค้นหาด้วยคำค้น 2 คำ โดยคำค้นอาจอยู่ต่อกัน หรืออยู่ใกล้กัน หรืออยู่ในระยะห่างตามที่กำหนด นอกจากนี้ยังรวมถึงการควบคุมลำดับของคำค้น และตำแหน่งที่เฉพาะเจาะจงของคำค้นด้วย (Proximity operators, 2009)

การค้นคืนแบบกำหนดความใกล้ชิดของคำค้น เป็นการค้นหาสารสนเทศที่ใช้คำค้น 2 คำหรือมากกว่า โดยมุ่งการจับคู่ของคำค้นที่ป้อนเข้าไปในลักษณะที่แยกกันและกำหนดระยะห่างของคำค้นทั้งสองอย่างชัดเจน (Shen, Li, & Li, 2008)

จากคำนิยามดังกล่าว สรุปได้ว่า การค้นคืนแบบกำหนดความใกล้ชิดของคำค้น เป็นเทคนิคที่ใช้กับคำค้นมากกว่าหนึ่งคำค้น โดยสามารถกำหนดระยะห่างระหว่างคำค้น

และลำดับของคำค้น ทั้งนี้เพื่อให้คำค้นอยู่ในระยะที่ใกล้ชิดกัน ซึ่งเทคนิคนี้อยู่บนแนวคิดที่ว่า ความใกล้ชิดกันของคำค้นที่ปรากฏในเอกสาร จะแสดงถึงความสัมพันธ์ระหว่างคำค้น กล่าวคือ หากกำหนดคำค้นให้อยู่ในระยะที่ใกล้ชิดกัน ผลลัพธ์ที่ได้จะมีความเกี่ยวข้องกัน สูงกว่าคำค้นที่อยู่ห่างกันหรืออยู่อย่างกระจัดกระจาย

การค้นคืนแบบกำหนดความใกล้ชิดของคำค้น เป็นวิธีการที่ทำให้มั่นใจว่าคำค้นหนึ่งจะปรากฏอยู่ถัดจากอีกคำค้นหนึ่งตามเงื่อนไขที่กำหนด คือ มีคำอื่นมาแทรกได้จำนวนกี่คำ หรือให้คำค้นเหล่านั้นปรากฏอยู่ภายในประโยคเดียวกัน อยู่ในย่อหน้าเดียวกัน หรืออยู่ในเขตข้อมูลเดียวกัน เช่น เขตข้อมูลชื่อเรื่อง หรือ บทคัดย่อ เป็นต้น

ความสำคัญของเทคนิคการกำหนดความใกล้ชิดของคำค้น

โดยทั่วไป การค้นคืนด้วยคำสำคัญ (Keyword Search) มีวัตถุประสงค์ คือ เพื่อค้นหาเอกสารที่ตรงกับคำสำคัญ อย่างไรก็ตาม การค้นคืนที่ให้ประโยชน์อย่างแท้จริง ระบบจะต้องมีองค์ประกอบพื้นฐาน ดังนี้ 1) ผลการค้นจะต้องไม่พลาดเอกสารที่เกี่ยวข้อง (Relevant answers) 2) มีประสิทธิภาพ และ 3) ต้องเสนอผลการค้นตามลำดับที่มีความสัมพันธ์กันสูง ซึ่งก็ยังไม่พบว่า มีระบบใดที่มีคุณลักษณะดังกล่าวที่สมบูรณ์ (Golenberg, Kimelfeld, Sagiv, 2008) ระบบการค้นคืนสารสนเทศ นอกจากจะทำหน้าที่ในการเข้าถึงสารสนเทศในฐานข้อมูลแล้ว ยังจะต้องเสนอผลการค้นที่มีความสัมพันธ์กับคำค้นที่ป้อนลงไปด้วย ในการค้นคืนด้วยเทคนิคตรรกะบูลีน (Boolean Operators) แม้ผลลัพธ์ที่ได้จะมีความสอดคล้องกับคำค้นที่ป้อนลงไป แต่เทคนิคตรรกะบูลีนไม่ได้เน้นการระบุตำแหน่งคำค้น และลำดับของคำค้น ซึ่งทำให้ผลลัพธ์บางส่วนมีเนื้อหาไม่ตรงกับความต้องการของผู้ใช้ ซึ่งเทคนิคการค้นคืนแบบกำหนดความใกล้ชิดของคำค้น จะช่วยแก้จุดอ่อนดังกล่าวได้ และช่วยให้ได้รับสารสนเทศที่ตรงกับความต้องการมากยิ่งขึ้น เนื่องจากเป็นเทคนิคที่เน้นการกำหนดระยะห่างของคำค้น และลำดับของคำค้น ซึ่งจะช่วยให้สามารถค้นหาเอกสารที่มีความสัมพันธ์และตรงกับข้อความสั่งค้น (Gupta, Ozsoyoglu, and Ozsoyoglu, 2009) เทคนิคนี้ยังมีประโยชน์สำหรับการค้นในสภาพที่คำๆ เดียวกันปรากฏอยู่ในเอกสารหลากหลายที่ โดยเฉพาะการค้นเกี่ยวกับเพลง หนังสือ และบทความจากวารสาร เป็นต้น ซึ่งในการค้น หากใช้คำๆ นั้นร่วมกับคำอื่นโดยใช้เทคนิคดังกล่าวจะช่วยให้ได้ผลลัพธ์ที่น้อยลงและมีประสิทธิภาพมากขึ้น นอกจากนี้ ยังมีประโยชน์ในการค้น

ข้อมูลชื่อบุคคลอีกด้วย เนื่องจาก ชื่อต้นและชื่อสกุลของบุคคลโดยปกติควรจะปรากฏอยู่ใกล้กัน (Gupta, Ozsoyoglu, & Ozsoyoglu, 2009)

ในส่วนที่เกี่ยวกับการจัดลำดับผลการค้น โดยปกติผู้ใช้ส่วนใหญ่จะสนใจเฉพาะผลการค้นที่น่าเสนอในลำดับต้นๆ เท่านั้น ดังนั้น การนำเสนอผลการค้นโดยมีการจัดลำดับตามความเกี่ยวข้องของเนื้อหาจึงมีความสำคัญต่อผู้ใช้ การกำหนดตำแหน่งและระยะห่างของคำค้นถือว่าเป็นองค์ประกอบหนึ่งที่สำคัญต่อการจัดลำดับผลการค้น เนื่องจากระยะห่างของคำค้นจะแสดงถึงความสัมพันธ์ของคำค้นที่ปรากฏในเอกสาร (Goldman, Shivakumar, Venkatasubramanian, & Garcia-Monlina, 1998) หากเอกสารใดที่ปรากฏคำค้นที่ป้อนลงไปโดยคำค้นทั้งสองยังอยู่ใกล้ชิดกัน โอกาสที่จะเป็นตัวแทนแนวคิดของข้อความสังค้น ยังมีสูงมากตามไปด้วย (Bai, et al. 2008) ดังนั้น เทคนิคนี้จึงมีความสำคัญต่อกระบวนการจัดอันดับ (Ranking process) ผลการค้นด้วย (Kimelfeld and Sagiv, 2006)

ลักษณะของเทคนิคการกำหนดความใกล้ชิดของคำค้น

เทคนิคนี้เป็นการค้นหาเอกสารที่คำค้นหนึ่งปรากฏอยู่ใกล้กับอีกคำค้นหนึ่ง ซึ่งโดยทั่วไปจะอยู่ในเอกสารนั้นๆ เท่านั้น ดังนั้น เทคนิคนี้จึงเป็นการพิจารณาถึงระยะห่างของคำค้นที่อยู่ภายในเอกสารนั้นๆ เท่านั้น (Goldman, Shivakumar, Venkatasubramanian, and Garcia-Monlina, 1998) ปกติเทคนิคนี้จะใช้ใน 2 ลักษณะ คือ 1) ระยะห่างของคำค้นโดยไม่คำนึงถึงลำดับของคำค้น และ 2) ระยะห่างของคำค้นโดยคำนึงถึงลำดับของคำค้น ส่วนโอเปอเรเตอร์ที่ใช้ในการกำหนดระยะห่างและตำแหน่งของคำค้น ได้แก่ NEAR, WITHIN และ PRE เป็นต้น

สำหรับระยะห่างของคำค้นมีได้มากน้อยเพียงใดนั้น ขึ้นอยู่กับแต่ละฐานข้อมูลอย่างไรก็ตาม สามารถจำแนกระยะห่างของคำค้นได้เป็น 3 ลักษณะ คือ (Science Direct, 2012)

- 1) ระยะห่างของคำค้นซึ่งอยู่ภายในวลีเดียวกัน (the same phrase) เช่น W/3, W/4 และ W/5 เป็นต้น
- 2) ระยะห่างของคำค้นซึ่งอยู่ภายในประโยคเดียวกัน (the same sentence) เช่น W/10 และ W/5 เป็นต้น

3) ระยะห่างของคำค้นซึ่งอยู่ภายในย่อหน้าเดียวกัน (the same paragraph) เช่น W/30, W/40 และ W/50 เป็นต้น

รูปแบบการกำหนดความใกล้ชิดของคำค้น

ในการเขียนข้อความสั่งค้นโดยใช้เทคนิคการกำหนดความใกล้ชิดของคำค้น จะนำเสนอใน 2 ลักษณะ คือ

1. การระบุระยะห่างของคำค้นโดยไม่คำนึงถึงลำดับของคำค้น ปกติจะใช้โอเปอร์เรเตอร์ NEAR หรือใช้สัญลักษณ์ N ซึ่งความหมายของ NEAR หมายถึง ให้ระบบค้นหาเอกสารที่ประกอบด้วยคำค้นตามที่ปรากฏในข้อความสั่งค้น โดยไม่คำนึงถึงลำดับก่อนหลังของคำค้น ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นได้ตามที่ระบุ

สำหรับรูปแบบการเขียนข้อความสั่งค้น จะเขียนโอเปอร์เรเตอร์ NEAR คั่นระหว่างคำค้นทั้งสอง โดยหลังคำว่า NEAR ให้ใช้เครื่องหมาย / และตามด้วยตัวเลขที่ระบุระยะห่างระหว่างคำค้นทั้งสอง ตัวอย่างข้อความสั่งค้น เช่น dogs NEAR/3 cats หรือ dogs N/3 cats ซึ่งจากข้อความสั่งค้น หมายความว่า ให้ระบบค้นหาเอกสารที่มีคำค้นทั้ง dogs และ cats โดยระหว่างคำค้นทั้งสองคำมีคำอื่นมาแทรกได้ไม่เกิน 3 คำ และไม่คำนึงถึงลำดับของคำค้น คือ คำค้นใดจะมาก่อนหรือตามหลังก็ได้

2. การระบุระยะห่างของคำค้นโดยคำนึงถึงลำดับของคำค้น ปกติจะใช้โอเปอร์เรเตอร์ WITHIN หรือใช้สัญลักษณ์ W ซึ่งความหมายของ WITHIN หมายถึง ให้ระบบค้นหาเอกสารที่ประกอบด้วยคำค้นตามที่ปรากฏในข้อความสั่งค้น โดยคำนึงถึงลำดับก่อนหลังของคำค้น ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นได้ตามที่ระบุ

สำหรับรูปแบบการเขียนข้อความสั่งค้น จะเขียนโอเปอร์เรเตอร์ WITHIN คั่นระหว่างคำค้นทั้งสอง โดยหลังสัญลักษณ์ W ให้ใช้เครื่องหมาย / และตามด้วยตัวเลขที่ระบุระยะห่างระหว่างคำค้นทั้งสอง ตัวอย่างข้อความสั่งค้น เช่น dogs W/3 cats ซึ่งจากข้อความสั่งค้น หมายความว่า ให้ระบบค้นหาเอกสารที่มีคำค้นทั้ง dogs และ cats โดยที่คำค้นทั้งสองคำมีคำอื่นมาแทรกได้ไม่เกิน 3 คำ และคำนึงถึงลำดับของคำค้น คือ คำค้นแรกต้องมาก่อน

ฐานข้อมูลเชิงพาณิชย์ที่ใช้เทคนิคการกำหนดความใกล้ชิดของคำค้น

ฐานข้อมูลเชิงพาณิชย์หลายฐานข้อมูลที่สามารถใช้เทคนิคการกำหนดความใกล้ชิดของคำค้น ตัวอย่างเช่น Academic Search Premier, Education Research Complete, ABI Inform Complete, Biological Sciences, CINAHL plus with full-text, ProQuest Agriculture Journal, Science Direct, IEEE, Emerald, และ The Cochrane Library เป็นต้น สำหรับโอเปอเรเตอร์ที่ใช้ในแต่ละฐานข้อมูลมีทั้งเหมือนและแตกต่างกัน ซึ่งได้แก่ NEAR (N), Within (W), PRE (P) และ NEXT เป็นต้น ซึ่งมีรายละเอียดดังตารางที่ 1 (EBSCO, 2012; ProQuest, 2012; ScienceDirect, 2012; IEEE, 2012; Emerald, 2012; The Cochrane Library, 2012)

ตารางที่ 1 Proximity Search Operators ที่ใช้ในฐานข้อมูลเชิงพาณิชย์บางฐานข้อมูล

ฐานข้อมูล	โอเปอเรเตอร์	ตัวอย่างการเขียนข้อความสั่งค้น (Query)	ความหมาย
Academic Search Premier, Education Research, CINAHL plus with full-text (EBSCO)	NEAR (N)	basketball N5 players	ให้ระบบค้นหาเอกสารที่ประกอบด้วยคำค้นทั้ง 2 คำ (basketball และ players) โดยไม่คำนึงถึงลำดับก่อนหลังของคำค้น ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นทั้งสองคำได้ไม่เกิน 5 คำ
	Within (W)	oil W3 clean-up	ให้ระบบค้นหาเอกสารที่ประกอบด้วยคำค้นทั้ง 2 คำ (oil และ clean-up) โดยคำนึงถึงลำดับก่อนหลังของคำค้นป้อนลงไป คือให้ปรากฏ คำค้นคำแรกก่อน และตามด้วยคำคนที่สอง ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นทั้งสองคำได้ไม่เกิน 3 คำ

ABI Inform Complete, Biological Sciences, ProQuest Agriculture Journal (ProQuest)	NEAR (N)	nursing NEAR/3 women or nursing N/3 women	ให้ระบบค้นหาเอกสารที่ประกอบด้วย คำค้นทั้ง 2 คำ (nursing และ women) โดยไม่คำนึงถึงลำดับก่อนหลังของคำค้น ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นทั้ง สองคำได้ไม่เกิน 3 คำ
	PRE (P)	nursing PRE/3 women or nursing P/3 women	ให้ระบบค้นหาเอกสารที่ประกอบด้วย คำค้นทั้ง 2 คำ (nursing และ women) โดยคำนึงถึงลำดับก่อนหลังของคำค้นที่ ป้อนลงไป คือ ให้ปรากฏคำค้นคำแรก ก่อน และตามด้วยคำค้นที่สอง ทั้งนี้ให้มี คำอื่นมาแทรกระหว่างคำค้นทั้งสองคำ ได้ไม่เกิน 3 คำ
Science Direct	WITHIN (W)	pain W/10 morphine	ให้ระบบค้นหาเอกสารที่ประกอบด้วย คำค้นทั้ง 2 คำ (pain และ morphine) โดยไม่คำนึงถึงลำดับก่อนหลังของคำค้น ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นได้ ไม่เกิน 10 คำ
	PRE	pain PRE/3 morphine	ให้ระบบค้นหาเอกสารที่ประกอบด้วย คำค้นทั้ง 2 คำ (pain และ morphine) โดยคำนึงถึงลำดับก่อนหลังของคำค้น คือ คำค้นแรกต้องมาก่อน ทั้งนี้ให้มีคำอื่น มาแทรกระหว่างคำค้นได้ไม่เกิน 3 คำ
IEEE	NEAR	implantable NEAR/3 cardiac	ให้ระบบค้นหาเอกสารที่ประกอบด้วย คำค้นทั้ง 2 คำ (implantable และ cardiac) โดยไม่คำนึงถึงลำดับของคำค้น ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นได้ ไม่เกิน 3 คำ
	ONEAR	implantable ONEAR/3 cardiac	ให้ระบบค้นหาเอกสารที่ประกอบด้วย คำค้นทั้ง 2 คำ (implantable และ cardiac) โดยคำนึงถึงลำดับของคำค้น คือ คำค้นแรกต้องมาก่อน ทั้งนี้ให้มีคำอื่น มาแทรกระหว่างคำค้นได้ไม่เกิน 3 คำ

Emerald Management	~ (tilde symbol)	"information literacy library"~5	ให้ระบบค้นหาเอกสารที่ประกอบด้วยคำค้นทั้ง 2 คำ (information literacy และ library) โดยไม่คำนึงถึงลำดับก่อนหลังของคำค้น ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นได้ไม่เกิน 5 คำ
The Cochrane Library	NEAR	breast cancer NEAR/5 women	ให้ระบบค้นหาเอกสารที่ประกอบด้วยคำค้นทั้ง 2 คำ (breast cancer และ women) โดยไม่คำนึงถึงลำดับก่อนหลังของคำค้น ทั้งนี้ให้มีคำอื่นมาแทรกระหว่างคำค้นทั้งสองคำได้ ไม่เกิน 5 คำ
	NEXT	cholera NEXT treat*	ให้ระบบค้นหาเอกสารที่ประกอบด้วยคำค้นทั้ง 2 คำ (cholera และ treat) โดยคำนึงถึงลำดับก่อนหลังของคำค้นที่ป้อนลงไป คือ ให้ปรากฏคำค้นคำแรกก่อน และตามด้วยคำค้นที่สอง

นอกจากนี้ โปรแกรมค้นหาข้อมูลบนอินเทอร์เน็ตบางโปรแกรมยังสามารถค้นคืนด้วยเทคนิคนี้เช่นกัน อาทิ Google ซึ่งโอเพอร์เรเตอร์ที่ Google ใช้คือ AROUND โดยมีรูปแบบการใช้ดังนี้ "search term A AROUND (n) search term B" โดย n หมายถึง จำนวนคำอื่นที่มาแทรกระหว่างคำค้นทั้งสอง ดังตัวอย่างเช่น global warming AROUND(3) tsunami เป็นต้น

ข้อจำกัดและปัญหาในการใช้เทคนิคการกำหนดความใกล้ชิดของคำค้น

เทคนิคการกำหนดความใกล้ชิดของคำค้นยังมีข้อจำกัดและปัญหาบางประการ อาทิ มีฐานข้อมูลเชิงพาณิชย์ส่วนหนึ่งเท่านั้นที่สามารถใช้เทคนิคการค้นนี้ บางฐานข้อมูลยอมให้ใช้เฉพาะกับคำค้นเพียง 2 คำค้นเท่านั้น ผลลัพธ์ของการค้นด้วยเทคนิคนี้อาจมีน้อย หรือไม่มีเลย หากข้อความสั้นเกินไป คือมีหลายคำค้น ถ้าข้อความสั้นประกอบด้วยคำค้นเพียง 2 คำค้น ผลการค้น โดยการใช้ NEAR operator กับไม่ใช้เทคนิคนี้ จะไม่แตกต่างกัน แต่เทคนิคนี้มีความสำคัญ หากข้อความสั้นประกอบด้วย คำค้น 3 คำค้น หรือมากกว่า โดยมีทั้งคำค้นหลักและคำค้นรอง และผู้ใช้ต้องการกำหนดว่าคำค้นรองใด

ควรจะปรากฏอยู่ในตำแหน่งใดหรือในระยะห่างจากคำค้นหลักเท่าใด ส่วนข้อจำกัดอื่นๆ ได้แก่ ยังขาดความเป็นมาตรฐานเดียวกันของแต่ละฐานข้อมูล ได้แก่ การกำหนดโอเปอเรเตอร์ซึ่งยังมีความแตกต่างกัน ความหมายของโอเปอเรเตอร์ที่อาจมีความแตกต่างกัน และการกำหนดรูปแบบของการเขียนข้อความสั่งค้นที่มีความแตกต่างกัน เป็นต้น ซึ่งทำให้เป็นภาระแก่ผู้ใช้ในการเรียนรู้การใช้เทคนิคดังกล่าวของแต่ละฐานข้อมูล

บทสรุป

เทคนิคการกำหนดความใกล้ชิดของคำค้น เป็นเทคนิคที่ผู้ใช้สามารถกำหนดระยะห่างระหว่างคำค้นได้ตามที่ต้องการ และถือว่าเป็นเทคนิคที่จะช่วยให้ผู้ใช้มีความพึงพอใจต่อผลของการค้น ด้วยการลดผลการค้นที่ไม่เกี่ยวข้องลง ได้ผลการค้นที่เฉพาะเจาะจงและตรงกับความต้องการมากขึ้น โดยอาศัยการกำหนดตำแหน่งของคำค้นให้ใกล้ชิดกัน อย่างไรก็ตาม เทคนิคนี้จะใช้ได้เฉพาะบางฐานข้อมูลเท่านั้น และอาจสร้างความยุ่งยากให้กับผู้ใช้ เนื่องจาก ขาดความเป็นมาตรฐานเดียวกันในแต่ละฐานข้อมูลทั้งโอเปอเรเตอร์ที่ใช้และรูปแบบการเขียนข้อความสั่งค้น

บรรณานุกรม

- Balmin, Andrey, et al. (2003). A system for keyword proximity search on XML databases. Proceeding of the 29th VLDB Conference. Berlin.
- Bai, Jing, et al. (2008). Investigation of partial query proximity in web search. WWW 2008 (April 21-25):1183-1184.
- The Cochrane Library. (2012). Search manual. Retrieved March, 23, 2012 from <http://www.thecochranelibrary.com/view/0/SearchManual.html>
- EBSCO. (2012). Proximity search. Retrieved March, 23, 2012 from http://support.Ebsco.com/help/index.php?help_id=55
- Emerald. (2012). Detailed “advanced search” functionality. Retrieved March, 23, 2012 from http://www.emeraldinsight.com/help/user/advanced_functionality.htm

- Feuer, Alan; Savev, Stefan, & Aslam, Javed A. (2009). Implementing and evaluating phrasal query suggestions for proximity search. *Information Systems* 34: 711-723
- Goldman, Roy; Shivakumar, Narayanan; Venkatasubramanian, Suresh; & Garcia-Monlina, Hector. (1998). Proximity search in database. *Proceeding of the 24th VLDB Conference*. New York.
- Golenberg, Konstantin; Kimelfeld, Benny; and Sagiv, Yehoshua. (2008). Keyword proximity search in complex data graphs. *SIGMOD'08* (June 9-12): 927-930.
- Gupta, Chirag; Ozsoyoglu, Gultekin & Ozsoyoglu, Z. Meral. (2009). Efficient k-Word proximity Search. *ISCIS*. 14-16: 123-128.
- IEEE. (2012). Summary of Search Operators. Retrieved March, 23, 2012 from http://ieeexplore.ieee.org/Xplorehelp/Help_searchoperators.html
- Kimelfeld, Benny and Sagiv, Yehoshua. (2006). Finding and approximating top-k answers in keyword proximity search. *PODS'06* (June 26-28):173-175.
- Proximity operators. (2009). Retrieved December 26,2011 from <http://library.alliant.edu/screens/proximity.pdf>
- Proximity search. (2011). Online Dictionary for Library and Information Science. Retrieved December 26,2011 from http://www.abc-clio.com/ODLIS_p.aspx
- Proximity Search. (2011). Retrieved December 26,2011 from [http://search.legis.state.ia.us/NXT/gateway.dll?f=templates\\$fn+help_search-proximity](http://search.legis.state.ia.us/NXT/gateway.dll?f=templates$fn+help_search-proximity)
- Proximity Search. (2011). Wikipedia. Retrieved December 26,2011 from [http://en.wikipedia.org/wiki/Proximity_search_\(text\)](http://en.wikipedia.org/wiki/Proximity_search_(text))
- ProQuest. (2012). การสืบค้นด้วยบรรทัดคำสั่ง. Retrieved March, 23, 2012 from http://search.proquest.com/help/academic.th-TH/Command_Search.html
- Reitz, Joan M. (2012). Proximity. Online dictionary for library and information science. Retrieved February, 20, 2012 from http://www.abc-clio.com/ODLIS/odlis_p.aspx

- Sadakane, Kunihiro & Imai, Hiroshi. (2000). Text retrieval by using k-word proximity Search. International Symposium on Database Applications in Non-Traditional Environments, 1999 (DANTE '99) Proceedings. Retrieved March, 23, 2012 from <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=844958>
- Schenkel, et al. (2007). Efficient text proximity search. SPIRE'07 Proceedings of the 14th international conference on String processing and information retrieval. Heidelberg :Springer-Verlag Berlin, 287-299.
- ScienceDirect. (2012). Connectors and proximity operators-expert search. Retrieved March, 23, 2012 from http://help.sciencedirect.com/flare/Content/stextpt_connectors.htm
- Shen, Haiying;Li, Ze, and Li, Ting. (2008). An investigation on multi-token list basedproximity search in multi-dimensional massive database. Third 2008 International Conference on Convergence and Hybrid Information Technology.