การหาค่าไพวอทของควิกซอร์ตที่เหมาะสมสำหรับข้อมูลแบบจำนวนเต็ม

Suitable Quicksort Pivot for Integer Data

นายศิระ สารากิจ

นายสรัช กิจการนนท์

โครงงานพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา ๒๕๕๖

การหาค่าไพวอทของควิกซอร์ตที่เหมาะสมสำหรับข้อมูลแบบจำนวนเต็ม

Suitable Quicksort Pivot for Integer Data

นายศิระ สารากิจ

นายสรัช กิจการนนท์

โครงงานพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2556

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และให้อ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SUITABLE QUICKSORT PIVOT FOR INTEGER DATA

MR. SIRA SARAKIJ

MR. SARAT KITJAKARNNON

A SPECIAL PROJECT SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR OF SCIENCE

IN COMPUTER SCIENCE

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

ACADEMIC YEAR 2013

เอกสารนี้เป็น ... ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อ ... ของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อโครงงานพิเศษ	การหาค่าไพวอทของควิกซอร์ตที่เหมาะสมสำหรับข้อมูลแบบจำนวนเต็ม Suitable Quicksort Pivot for Integer Data		
ชื่อนักศึกษา	นายศิระ	สารากิจ	53051090
	นายสรัช	กิจการนนท์	53051100
ปริญญา	วิทยาศาสตรบัณฑิต		
สาขาวิชา	วิทยาการคอมพิวเตอร์		
อาจารย์ที่ปรึกษา	ผศ.ดร.ศรัณย์ อินทโกสุม		

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง อนุมัติให้โครงงานพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร วิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ประจำปีการศึกษา 2556

คณะกรรมการสอบ	ลายมือชื่อ
ประธานกรรมการ รศ.ดร.วีระ บุญจริง	
กรรมการ อ.วีระชัย ต้นยะสิทธิ์	
กรรมการและอาจารย์ที่ปรึกษา ผศ.ดร.ศรัณย์ อินทโกสุม	

ลิขสิทธิ์ของคณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อโครงงานพิเศษ	การหาค่าไพวอทของควิกซอร์ตที่เหมาะสมสำหรับข้อมูลแบบจำนวนเต็ม
	Suitable Quicksort Pivot for Integer Data
ชื่อนักศึกษา	นายศิระ สารากิจ 53051090
	นายสรัช กิจการนนท์ 53051100
ปริญญา	วิทยาศาสตรบัณฑิต
สาขาวิชา	วิทยาการคอมพิวเตอร์
ปีการศึกษา	2556
อาจารย์ที่ปรึกษา	ผศ.ดร.ศรัณย์ อินทโกสุม

บทคัดย่อ

ควิกซอร์ตถือได้ว่าเป็นขั้นตอนวิธีการเรียงลำดับที่มีประสิทธิภาพและรวดเร็วโดยอยู่บนพื้นฐานของปัญหาการแบ่งแยกและเอาชนะ โดยได้มีงานวิจัยเสนอแนวทางในการเลือกค่าไพวอทเพื่อใช้ในการพาร์ทิชันข้อมูลซึ่งส่งผลกับประสิทธิภาพของขั้นตอนวิธีควิกซอร์ต อย่างไรก็ตาม ยังไม่มีงานวิจัยใดที่นำเอาเทคนิคการเลือกค่าไพวอทแบบต่าง ๆ มาเปรียบเทียบกับชุดข้อมูลหลายชนิดหลาย ๆ รูปแบบ โดยประโยชน์ของงานวิจัยในลักษณะนี้คือ นักพัฒนาสามารถนำเอางานวิจัยไปประยุกต์ใช้หรือเป็นแนวทางในการหาค่าไพวอทที่เหมาะสมกับข้อมูลของตนเอง จุดประสงค์ของปัญหาพิเศษเล่มนี้คือเพื่อทำการทดลองดังกล่าว กับข้อมูลจำนวนเต็มรูปแบบต่าง ๆ เช่น ข้อมูลสุ่มโดยใช้การแจกแจงยูนิฟอร์ม, ข้อมูลที่เกือบเรียงลำดับ, ข้อมูลที่มีค่าซ้ำกัน เป็นต้น โดยใช้เทคนิคการเลือกค่าไพวอทหลัก ๆ ได้แก่ ค่ามัธยฐาน, ค่าเริ่มต้น, ค่ากลาง, ค่าสุดท้าย และค่าสุ่ม กับข้อมูลจำนวนเต็มที่กล่าวข้างต้น ซึ่งเกณฑ์ในการเปรียบเทียบที่ถูกทดสอบคือ เวลาในการประมวลผลเฉลี่ย, จำนวนครั้งในการเปรียบเทียบเฉลี่ย และจำนวนครั้งในการสลับที่เฉลี่ย ผลลัพธ์ที่ได้จะถูกแสดงและสรุปอยู่ในรูปแบบของกราฟเพื่อให้นักพัฒนาสามารถนำข้อมูลเหล่านี้ไปประกอบการเลือกค่าไพวอทที่เหมาะสมกับข้อมูลของตนเองได้โดยง่าย หรือนักวิจัยอาจนำข้อมูลเหล่านี้ไปพัฒนาขั้นตอนวิธีควิกซอร์ตต่อไป

คำสำคัญ : ขั้นตอนวิธีควิกซอร์ต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Suitable Quicksort Pivot for Integer Data		
Students	Mr. Sira	Sarakij	53051090
	Mr. Sarat	Kitjakarnnon	53051100
Degree	Bachelor of Science		
Major Program	Computer Science		
Academic Year	2013		
Advisor	Asst.Prof.Dr.Sarun	Intakosum	

ABSTRACT

Quicksort is an efficient and fast sorting algorithm that is based on divide and conquer approach. The previous researches showed that the way to select the pivot which is used to partition the data can affect the performance of the algorithm. However, there was not any research that compared the proposed pivot selection techniques on the various kinds of data sets. This kind of research is important since developers can have a guideline in order to choose the suitable pivot that fit well for their data. This special project aims to do such an experiment. The various patterns of integer such as uniform distributed data, nearly sorted data,and a data set that contains a number of repeated data, were prepared. The major pivot selection techniques such as median, first, middle and last values were applied on the prepared data set. The comparison aspects that were tested are, average execution time, comparison, and swapping time.The results are shown and summarized in the form of graphs so that developers can use this information to easily select the proper pivot for their desired data. The researchers may also use this information to improve the quick sort algorithm.

Keywords : Quicksort Algorithm

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

การจัดทำปัญหาพิเศษหัวข้อการหาค่าไพวอทของควิกซอร์ตที่เหมาะสมสำหรับข้อมูลแบบจำนวนเต็ม ฉบับนี้สามารถดำเนินงานจนสำเร็จลุล่วงไปได้ด้วยดี เนื่องจากผู้จัดทำได้รับความช่วยเหลือและสนับสนุนจากบุคคลผู้มีพระคุณหลายท่าน ผู้จัดทำขอกล่าวคำขอบพระคุณบุคคลดังต่อไปนี้

ขอขอบพระคุณ บิดา มารดา และสมาชิกในครอบครัวทุกคน ผู้ให้การอบรม สั่งสอน และคอยสนับสนุนในด้านการศึกษา ช่วยเหลือและให้กำลังใจเสมอมา

ขอขอบพระคุณ ผศ.ดร.ศรัณย์ อินทโกสุม อาจารย์ที่ปรึกษาปัญหาพิเศษนี้ได้เสียสละเวลาให้คำปรึกษา คำแนะนำ ตรวจและแก้ไขปัญหาต่างๆ ด้วยดีตลอดการทำปัญหาพิเศษนี้

ขอขอบพระคุณ รศ.ดร.วีระ บุญจริง อาจารย์ประธานกรรมการการทำหัวข้อปัญหาพิเศษที่ได้ให้คำปรึกษาและคำแนะนำในการแก้ปัญหาต่างๆรวมถึงการตรวจสอบและแก้ไขการเขียนรายงานปัญหาพิเศษ

ขอขอบพระคุณ อ.วีระชัย ตันยะสิทธิ์ อาจารย์กรรมการการทำหัวข้อปัญหาพิเศษที่คอยให้คำแนะนำและคำปรึกษา

ขอขอบพระคุณ อาจารย์สาขาวิทยาการคอมพิวเตอร์ทุกท่าน ที่อบรมสั่งสอน ให้ความรู้ทั้งภาคทฤษฎีและภาคปฏิบัติให้แก่ทางคณะผู้จัดทำตลอดระยะเวลา 4 ปี จนกระทั่งปัญหาพิเศษนี้ได้สำเร็จลุล่วงไปได้ด้วยดี

สุดท้ายนี้ต้องขอบคุณ เพื่อน ๆ ทุกคน ที่ให้คำปรึกษาและให้กำลังใจเสมอมา

ศิระ สารากิจ

สรัช กิจการนนท์

พระจอมเกล้าเจ้าคุณทหาร ...โนโลยี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญภาพ	VII
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญของการทำโครงการพิเศษ	1
1.2 วัตถุประสงค์ของการทำโครงการพิเศษ	2
1.3 ขอบเขตของโครงการพิเศษ	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	2
1.5 ขั้นตอนการดำเนินงาน	3
บทที่ 2 ความรู้พื้นฐานและทฤษฎีที่เกี่ยวข้อง	4
2.1 ขั้นตอนวิธีควิกซอร์ต (Quicksort)	4
2.1.1 ความรู้เบื้องต้นของขั้นตอนวิธีควิกซอร์ต	4
2.1.2 Pseudo code ของขั้นตอนวิธีควิกซอร์ต	11
2.2 การเลือกค่าไพวอท	12
2.2.1 งานวิจัยที่เกี่ยวข้องกับการเลือกค่าไพวอท	12
2.2.2 ประเภทของไพวอทแบบต่างๆ	13
2.2.3 มัธยฐาน (Median)	13
2.2.3.1 Median of Three	14
2.2.3.2 Median of Five	14
2.2.3.3 Median of Seven	14
2.2.3.4 Median of Nine	14
2.3 การสุ่มค่าแบบ Uniform Distribution	16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

	หน้า
บทที่ 3 ออกแบบและขั้นตอนการทดลอง	17
3.1 วัตถุประสงค์การทดลอง	17
3.2 อุปกรณ์ในการดำเนินการทดลอง	17
3.3 ข้อมูลการทดลอง	18
3.3.1 ข้อมูลประเภทที่ 1	18
3.3.1.1 ตัวอย่างข้อมูลประเภทที่ 1	18
3.3.2 ข้อมูลประเภทที่ 2	18
3.3.2.1 ตัวอย่างข้อมูลประเภทที่ 2	18
3.3.3 ข้อมูลประเภทที่ 3	18
3.3.3.1 ตัวอย่างข้อมูลประเภทที่ 3	19
3.3.4 ข้อมูลประเภทที่ 4	19
3.3.4.1 ตัวอย่างข้อมูลประเภทที่ 4	19
3.3.5 ข้อมูลประเภทที่ 5	19
3.3.5.1 ตัวอย่างข้อมูลประเภทที่ 5	20
3.4 การเลือกค่าไพวอทในการทดลอง	20
3.4.1 ค่าเริ่มต้น	20
3.4.2 ค่าสุดท้าย	20
3.4.3 ค่าสุ่ม	20
3.4.4 ค่ากึ่งกลาง	20
3.4.5 ค่ามัธยฐาน	20
3.4.5.1 Median of Three	20
3.4.5.2 Median of Five	20
3.5 มาตรวัดผลการทดลอง	20
บทที่ 4 ผลการทดลอง	21
4.1 ผลการทดลอง	21
4.1.1 ผลการทดลองของข้อมูลประเภทที่ 1	21
4.1.2 ผลการทดลองของข้อมูลประเภทที่ 2	23
4.1.3 ผลการทดลองของข้อมูลประเภทที่ 3	26

สารบัญ(ต่อ)

หน้า

4.1.4 ผลการทดลองของข้อมูลประเภทที่ 4 29

4.1.5 ผลการทดลองของข้อมูลประเภทที่ 5 31

4.2 สรุปผลการทดลอง 34

บทที่ 5 สรุปและข้อเสนอแนะ 35

5.1 สรุป 35

5.2 ข้อเสนอแนะ 36

เอกสารอ้างอิง 37

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

ภาพที่		หน้า
2.1	ตัวอย่างผลลัพธ์ของงานวิจัย Comparative Analysis of Some Pivot Selection Schemes for Quicksort Algorithm	12
2.2	ตำแหน่งข้อมูลของ Median of Five	14
2.3	ตัวแทนข้อมูลของ Median of Five	15
2.4	การสุ่มค่าแบบ Uniform Distribution	16
4.1	จำนวนครั้งในการเปรียบเทียบของข้อมูลประเภท uniform distribution	22
4.2	จำนวนครั้งในการเปรียบเทียบของข้อมูลประเภท uniform distribution	22
4.3	จำนวนครั้งในการสลับที่ของข้อมูลประเภท uniform distribution	23
4.4	เวลาในการประมวลผลของข้อมูลเกือบเรียงลำดับจากน้อยไปมาก	24
4.5	จำนวนครั้งในการเปรียบเทียบของข้อมูลเกือบเรียงลำดับจากน้อยไปมาก	25
4.6	จำนวนครั้งในการสลับที่ของข้อมูลเกือบเรียงลำดับจากน้อยไปมาก	26
4.7	เวลาในการประมวลผลของข้อมูลเกือบเรียงลำดับแบบย้อนกลับจากมากไปน้อย	27
4.8	จำนวนครั้งในการเปรียบเทียบของข้อมูลเกือบเรียงลำดับแบบย้อนกลับจากมากไปน้อย	28
4.9	จำนวนครั้งในการสลับที่ของข้อมูลเกือบเรียงลำดับแบบย้อนกลับจากมากไปน้อย	28
4.10	เวลาในการประมวลผลของข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 5 ตัว	29
4.11	จำนวนครั้งในการเปรียบเทียบของข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 5 ตัว	30
4.12	จำนวนครั้งในการสลับที่ของข้อมูลข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 5 ตัว	31
4.13	เวลาในการประมวลผลของข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 20 ตัว	32
4.14	จำนวนครั้งในการเปรียบเทียบของข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 20 ตัว	33
4.15	จำนวนครั้งในการสลับที่ของข้อมูลข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 20 ตัว	33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของการทำโครงการพิเศษ

การเรียงลำดับของข้อมูลถือได้ว่าเป็นกระบวนการที่สำคัญอันหนึ่ง การเรียงลำดับข้อมูลถูกใช้ในงานหลากหลายประเภท อีกทั้งยังเป็นขั้นตอนพื้นฐานของขั้นตอนวิธีต่าง ๆ ซึ่งปัญหาการเรียงลำดับนั้นได้มีผู้เสนอขั้นตอนวิธีต่าง ๆ มากมาย แต่ละขั้นตอนจะมีข้อดีข้อเสียแตกต่างกันไปขึ้นอยู่กับประเภทของงานและข้อมูลที่ใช้เรียงลำดับ

ขั้นตอนการเรียงลำดับควิกซอร์ตถือได้ว่าเป็นขั้นตอนวิธีการเรียงลำดับแบบหนึ่งที่มีประสิทธิภาพ และถูกใช้เป็นขั้นตอนการเรียงลำดับในงานอื่น ๆ ทางคอมพิวเตอร์มากมาย โดยประสิทธิภาพในการเรียงลำดับนั้นขึ้นอยู่กับเทคนิคการเลือกค่าไพวอทที่เหมาะสมสำหรับการแบ่งข้อมูลซึ่งส่งผลกระทบต่อจำนวนครั้งของการเปรียบเทียบในการเรียงลำดับ จำนวนครั้งในการสลับที่จำนวนครั้งการเรียกซ้ำ และเวลาที่ใช้เรียงลำดับข้อมูล โดยได้มีงานวิจัยมากมายที่พยายามนำเสนอวิธีปรับปรุงประสิทธิภาพของควิกซอร์ตอยู่เสมอ ๆ

จากงานวิจัย [3, 4, 5] และหนังสือ [1, 2] ได้เสนอเทคนิคการเลือกค่าไพวอทหลากหลายประเภท แต่ยังไม่มีงานวิจัยใดที่นำขั้นตอนวิธีควิกซอร์ตโดยเลือกค่าไพวอทหลาย ๆ ประเภทมาทดลองกับข้อมูลจำนวนเต็มรูปแบบต่าง ๆ

ปัญหาพิเศษนี้จัดทำขึ้นเพื่อศึกษาและวิจัยการเลือกค่าไพวอทของควิกซอร์ตที่เหมาะสมกับข้อมูลจำนวนเต็มประเภทต่าง ๆ โดยจะใช้เกณฑ์การเลือกค่าไพวอทหลาย ๆ แบบ เช่น ค่าเริ่มต้น ค่าสุดท้าย ค่ากึ่งกลาง ค่าสุ่มจากข้อมูล ค่ามัธยฐาน กับข้อมูลประเภทอาเรย์ของตัวเลข เช่น ข้อมูลสุ่ม ข้อมูลที่มีการเรียงลำดับแล้ว ข้อมูลที่เรียงลำดับแล้วบางส่วน ข้อมูลซ้ำ เป็นต้น โดยผลลัพธ์จะแสดงให้เห็นถึงเวลาเฉลี่ยในการเรียงลำดับในข้อมูลแบบต่าง ๆ จำนวนครั้งในการเปรียบเทียบกับค่าไพวอท และจำนวนครั้งในการสลับที่ โดยแยกตามประเภทข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์ของการทำโครงการพิเศษ

1.2.1 พัฒนางานวิจัยโดยศึกษาการเลือกค่าไพวอทของควิกซอร์ตแบบต่าง ๆ มาทดลองกับข้อมูลหลาย ๆ ประเภท เพื่อวัดประสิทธิภาพและผลลัพธ์ของขั้นตอนวิธีควิกซอร์ต ได้แก่ เวลาในการประมวลผล, จำนวนครั้งในการสลับที่, จำนวนครั้งในการเปรียบเทียบ

1.2.2 เพื่อพิสูจน์ว่าการเลือกค่าไพวอทโดยใช้ค่ามัธยฐานจะสามารถหาค่าไพวอทที่เหมาะสมและช่วยลดจำนวนครั้งในการเปรียบเทียบได้ดีกว่าการเลือกค่าไพวอทแบบอื่น ๆ

1.2.3 นำผลการวิจัยมาหาข้อสรุปว่าการเลือกค่าไพวอทแบบใด จึงจะเหมาะสมกับข้อมูลแต่ละประเภท

1.3 ขอบเขตของโครงการพิเศษ

1.3.1 การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง

1.3.2 ข้อมูลใช้ทดสอบเป็นข้อมูลประเภทอาเรย์ของตัวเลข

1.3.3 เขียนโปรแกรมทดสอบด้วยภาษาจาวา

1.3.4 ขอบเขตของข้อมูลทดสอบประกอบไปด้วย ข้อมูลที่มีการเรียงลำดับถูกต้อง ข้อมูลที่เกือบเรียงลำดับถูกต้อง ข้อมูลที่เกือบเรียงลำดับแบบย้อนกลับ ข้อมูลซ้ำ ข้อมูลสุ่มแบบ uniform distribution

1.3.5 พิสูจน์ผลลัพธ์ที่ได้จากการทดลองว่า ผลลัพธ์ที่ได้จากการเลือกค่าไพวอท มีผลอย่างไรกับข้อมูลทดสอบแบบต่าง ๆ และค่าไพวอทแบบไหนถึงจะเหมาะสมกับข้อมูลแต่ละแบบ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 ได้ข้อสรุปว่าการเลือกค่าไพวอทของควิกซอร์ตแบบใด จึงจะเหมาะสมกับข้อมูลแต่ละประเภทในแง่ของเวลาในการประมวลผล, จำนวนครั้งในการเปรียบเทียบ และจำนวนครั้งในการสลับที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4.2 เป็นแนวทางให้กับผู้ที่สนใจศึกษาการเลือกค่าไพวอทของควิกซอร์ตแบบต่าง ๆ

1.5 ขั้นตอนการดำเนินงาน

1.5.1 ศึกษางานวิจัยที่เกี่ยวข้องกับขั้นตอนวิธีควิกซอร์ตเพื่อทราบถึงปัญหา และข้อจำกัดของขั้นตอนวิธีควิกซอร์ต

1.5.2 ศึกษางานวิจัยที่เกี่ยวข้องกับการเลือกค่าไพวอทของควิกซอร์ตแบบต่าง ๆ

1.5.3 ออกแบบการทดลอง และข้อมูลที่ใช้ในการทดสอบ

1.5.4 พัฒนาโปรแกรมขั้นตอนวิธีควิกซอร์ต

1.5.5 สร้างข้อมูลที่ใช้ในการทดสอบ

1.5.6 ทดสอบโปรแกรม และรวบรวมผลลัพธ์ของโปรแกรม

1.5.7 วิเคราะห์ผลลัพธ์ของโปรแกรม และหาข้อสรุป

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ความรู้พื้นฐานและทฤษฎีที่เกี่ยวข้อง

ความรู้พื้นฐานและทฤษฎีที่เกี่ยวข้องสำหรับการวิจัยเรื่องการหาค่าไพวอทของควิกซอร์ตที่เหมาะสมสำหรับข้อมูลแบบจำนวนเต็มในครั้งนี้ ผู้วิจัยได้ศึกษาค้นคว้ารวบรวมข้อมูลจากวารสาร บทความ ตำรา วิทยานิพนธ์ รวมทั้งเว็บไซต์ โดยนำเสนอตามลำดับดังนี้ (2.1) ขั้นตอนวิธีควิกซอร์ต (2.2) การเลือกค่าไพวอท (2.3) การสุ่มค่าแบบ uniform distribution

2.1 ขั้นตอนวิธีควิกซอร์ต (Quicksort)

ขั้นตอนวิธีควิกซอร์ตเป็นขั้นตอนวิธีที่อยู่บนพื้นฐานของปัญหาที่เรียกว่าแบ่งแยกและเอาชนะ (devide and conquer) โดยหลักการของปัญหานี้คือการแบ่งปัญหาใหญ่ที่แก้ไขได้ยากนั้นออกเป็นปัญหาย่อย ๆ หลาย ๆ ปัญหาที่มีขนาดเล็กกว่า หาคำตอบย่อยของปัญหาเล็ก ๆ แล้วนำคำตอบย่อย ๆ มารวมเป็นคำตอบของปัญหาใหญ่ ซึ่งอาจจะซับซ้อนน้อยกว่า ซึ่งควิกซอร์ตถือได้ว่าใช้เวลาเร็วที่สุดในเชิงปฏิบัติ โดยควิกซอร์ตนั้นใช้เวลาเฉลี่ยในการประมวลผลกรณีที่ดีที่สุด (best-case) เป็น $O(N \log N)$ และในกรณีที่แย่ที่สุด (worst-case) เป็น $O(N^2)$

2.1.1 ความรู้เบื้องต้นของขั้นตอนวิธีควิกซอร์ต

ควิกซอร์ต (Quicksort) เป็นกระบวนการในการเรียงลำดับ โดยเลือกข้อมูลจากกลุ่มข้อมูลขึ้นมาหนึ่งค่าเป็นค่าหลัก หรือค่าไพวอท (pivot) แล้วทำการหาตำแหน่งของข้อมูลภายในกลุ่มข้อมูลชุดนั้น หลังจากนั้นทำการแบ่งข้อมูลออกเป็น 2 ส่วน

1. กลุ่มข้อมูลด้านซ้ายเก็บข้อมูลที่มีค่าน้อยกว่าหรือเท่ากับค่าไพวอท
2. กลุ่มข้อมูลด้านขวาเก็บข้อมูลที่มีค่ามากกว่าค่าไพวอท

โดยมีค่าไพวอทอยู่ตรงกลางระหว่างกลุ่มข้อมูลด้านซ้ายและด้านขวา โดยถ้าหากกลุ่มข้อมูลทางด้านซ้ายมีค่าน้อยกว่าหรือเท่ากับค่าไพวอทและกลุ่มข้อมูลทางด้านขวามีค่ามากกว่าค่าไพวอท แสดงว่าค่าไพวอทอยู่ในตำแหน่งที่ถูกต้องแล้ว หลังจากนั้นนำกลุ่มข้อมูลทั้ง 2 ด้านไปดำเนินการเปรียบเทียบในลักษณะเดียวกันจนกลุ่มข้อมูลทั้งหมดเรียงลำดับจากน้อยไปหามาก โดยมีรายละเอียดขั้นตอนวิธีดังตัวอย่าง 2.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่าง 2.1

ชุดข้อมูลมีทั้งหมด 9 ตัว โดยกำหนดให้ค่าไพวอท (pivot) คือข้อมูลตัวแรกในกลุ่มข้อมูล

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	27	15	22	37	11	59	18	50	42

ขั้นที่ 1 : ทำการแบ่งข้อมูลออกเป็นสองส่วนโดยใช้ค่าไพวอทที่กำหนดคือ 27 ซึ่งเป็นข้อมูลตัวแรกในกลุ่มข้อมูลแทนด้วยตัวหนาและกำหนดตัวชี้เป็น L และ R ซึ่ง L คือ ข้อมูลตัวแรกของกลุ่มข้อมูล และ R คือ ข้อมูลตัวสุดท้ายของกลุ่มข้อมูล

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	**27**	15	22	37	11	59	18	50	42
ตัวชี้	L								R

ขั้นที่ 2 : เปรียบเทียบค่า ณ ตำแหน่ง R ถ้าหากว่าตัวชี้ R มีค่ามากกว่าไพวอท ซึ่งก็คือ 27 ให้เลื่อนตัวชี้ R แต่ถ้าหากตัวชี้ R มีค่าน้อยกว่าตัวชี้ L ให้ทำการสลับค่ากัน ซึ่งจะทำการเปรียบเทียบไปจนกว่าตัวชี้ L เท่ากับ R คือ อยู่ในตำแหน่งเดียวกัน

รอบที่ 1 : เลื่อนตัวชี้ R เนื่องจากค่าที่ตัวชี้ L น้อยกว่าค่าที่ตัวชี้ R

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	**27**	15	22	37	11	59	18	50	42
ตัวชี้	L								**<-R**

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	**27**	15	22	37	11	59	18	50	42
ตัวชี้	L							**R**	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รอบที่ 2 : เลื่อนตัวชี้ R

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	**27**	15	22	37	11	59	18	50	42
ตัวชี้	L							**<-R**	

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	**27**	15	22	37	11	59	18	50	42
ตัวชี้	L						**R**		

รอบที่ 3 : ค่าที่ตัวชี้ R มีค่าน้อยกว่า ค่าที่ตัวชี้ L ให้ทำการสลับตำแหน่งข้อมูล

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	**27**	15	22	37	11	59	**18**	50	42
ตัวชี้	L						R		

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	37	11	59	**27**	50	42
ตัวชี้	L						R		

***** เมื่อเกิดการสลับค่าขึ้น จะต้องเปลี่ยนตัวชี้ในการเลื่อนด้วย ในกรณีนี้ เริ่มต้น เราทำการเลื่อนตัวชี้ R มาตลอด เมื่อเกิดการสลับค่าเกิดขึ้น หลังจากนั้นถ้าหากต้องเลื่อนตำแหน่ง จะเปลี่ยนไปเลื่อนตัวชี้ L แทน โดยจะทำการเลื่อนก็ต่อเมื่อค่าที่ตัวชี้ L น้อยกว่าค่าที่ตัวชี้ R ถ้าหากเกิดการสลับอีกครั้งก็จะเปลี่ยนไปเลื่อนตัวชี้ R อีกครั้ง สลับกันไปเรื่อย ๆ**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รอบที่ 4 : เลื่อนตัวชี้ L เนื่องจากค่าที่ตัวชี้ L น้อยกว่าค่าที่ตัวชี้ R ตามเงื่อนไข ***

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	37	11	59	**27**	50	42
ตัวชี้	L->						R		

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	37	11	59	**27**	50	42
ตัวชี้		L					R		

รอบที่ 5 : เลื่อนตัวชี้ L

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	37	11	59	**27**	50	42
ตัวชี้		L->					R		

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	37	11	59	**27**	50	42
ตัวชี้			L				R		

รอบที่ 6 : เลื่อนตัวชี้ L

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	37	11	59	**27**	50	42
ตัวชี้			L->				R		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	37	11	59	**27**	50	42
ตัวชี้				L			R		

รอบที่ 7 : ค่าที่ตัวชี้ R มีค่าน้อยกว่า ค่าที่ตัวชี้ L ให้ทำการสลับตำแหน่งข้อมูล

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	37	11	59	**27**	50	42
ตัวชี้				L			R		

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	**27**	11	59	37	50	42
ตัวชี้				L			R		

รอบที่ 8 : เลื่อนตัวชี้ R ตามเงื่อนไข ***

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	**27**	11	59	37	50	42
ตัวชี้				L			<-R		

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	**27**	11	59	37	50	42
ตัวชี้				L		R			

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รอบที่ 9 : เลื่อนตัวชี้ R

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	**27**	11	59	37	50	42
ตัวชี้				L		<-R			

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	**27**	11	59	37	50	42
ตัวชี้				L	R				

รอบที่ 10 : ค่าที่ตัวชี้ R มีค่าน้อยกว่า ค่าที่ตัวชี้ L ให้ทำการสลับตำแหน่งข้อมูล

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	**27**	11	59	37	50	42
ตัวชี้				L	R				

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	11	**27**	59	37	50	42
ตัวชี้				L	R				

รอบที่ 11 : เลื่อนตัวชี้ L ตามเงื่อนไข ***

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	11	**27**	59	37	50	42
ตัวชี้				L->	R				

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	11	27	59	37	50	42
ตัวชี้					L,R				

ขั้นตอนที่ 3 : เมื่อค่าที่ตัวชี้ L เท่ากับค่าที่ตัวชี้ R จะทำการแบ่งกลุ่มข้อมูลออกเป็น 2 กลุ่ม คือ กลุ่มข้อมูลที่ทุกค่ามีค่าน้อยกว่าไพวอท และกลุ่มข้อมูลที่ทุกค่ามีค่ามากกว่าตัวไพวอท

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	18	15	22	11	27	59	37	50	42
ตัวชี้					L,R				

กลุ่มที่ 1 : กลุ่มข้อมูลที่มีค่าน้อยกว่าไพวอท

ลำดับที่	1	2	3	4
ข้อมูล	18	15	22	11
ตัวชี้				

กลุ่มที่ 2 : กลุ่มข้อมูลที่มีค่ามากกว่าไพวอท

ลำดับที่	1	2	3	4
ข้อมูล	59	37	50	42
ตัวชี้				

หลังจากนั้นให้ทั้งกลุ่มที่ 1 และกลุ่มที่ 2 กลับไปทำขั้นตอนที่ 1 ใหม่จนกว่าข้อมูลจะเรียงกันตามตาราง

ลำดับที่	1	2	3	4	5	6	7	8	9
ข้อมูล	11	15	18	22	27	37	42	50	59

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 Pseudo code ของขั้นตอนวิธีควิกซอร์ต

ขั้นตอนวิธีควิกซอร์ตแบบพื้นฐาน จะมี pseudo code ดังนี้

```
quicksort(a[], low, high) {
        if (high > low) {
                pivot = partition(a, low, high);
                quicksort(a, low, pivot-1);
                quicksort(a, pivot+1, high);
        }
}
partition(a[], low, high) {
        pivot = findPivot();
        left = low;
        right = high;
        int pivot_item = a[pivot];
        while (left < right) {
                while ((a[left] <= pivot_item) && (left < high)) left++;
                while ((a[right] > pivot_item) && (right > low)) right--;
                if (left < right) swap(a, left, right);
        }
        a[pivot] = a[right];
        a[right] = pivot_item;
        return right;
}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 การเลือกค่าไพวอท

2.2.1 งานวิจัยที่เกี่ยวข้องกับการเลือกค่าไพวอท

การเลือกค่าไพวอทที่ดีนั้นส่งผลให้ควิกซอร์ตทำงานได้เร็วขึ้น ซึ่งหนังสือส่วนใหญ่แนะนำให้ใช้ค่าแรกของกลุ่มข้อมูลเป็นค่าไพวอทจึงอาจทำให้ประสิทธิภาพของควิกซอร์ตลดลงในกรณีที่ข้อมูลมีการเรียงลำดับอยู่แล้ว จึงได้มีการนำเสนอขั้นตอนวิธีในการหาค่าไพวอทแบบต่าง ๆ

เพื่อปรับปรุงไม่ให้เกิดกรณีที่แย่ที่สุดของควิกซอร์ต จึงได้มีการนำเสนองานวิจัยเพื่อปรับปรุงขั้นตอนวิธีการเลือกค่าไพวอทแบบต่าง ๆ เพื่อหลีกเลี่ยงค่าต่ำสุดหรือสูงสุดของกลุ่มข้อมูลนั้น ๆ โดยมีขั้นตอนในการเลือกเช่น การเลือกค่าไพวอทจากการสุ่มตำแหน่งข้อมูล [3] , การเลือกค่าไพวอทจากตำแหน่งกลางของข้อมูล [4] , การเลือกค่าไพวอทจากตัวแทนข้อมูล (median-of-five) [5]

จากงานวิจัย *Comparative Analysis of Some Pivot Selection Schemes for Quicksort Algorithm* [5] ได้มีการศึกษาการเลือกค่าไพวอทวิธีต่าง ๆ จากค่ามัธยฐาน (Median-of-three, Median-of-five, Median-of-seven, Median-of-nine) กับข้อมูลสุ่มประเภทจำนวนเต็มด้วยขนาดต่าง ๆ เพื่อหาเวลาเฉลี่ยของวิธีที่น้อยที่สุด

Table 1: Average execution time (sec) of the six schemes

Array size	Median 50	Median 51	Median 70	Median 71	Median 90	Median 91
100K	0.16	0.21	0.31	0.32	0.32	0.51
150K	0.25	0.34	0.47	0.47	0.49	0.77
200K	0.34	0.47	0.65	0.66	0.66	1.06
250K	0.44	0.60	0.84	0.84	0.86	1.35
300K	0.53	0.72	1.04	1.05	1.05	1.66
350K	0.58	0.80	1.16	1.18	1.18	1.83

ภาพที่ 2.1 ตัวอย่างผลลัพธ์ของงานวิจัย *Comparative Analysis of Some Pivot Selection Schemes for Quicksort Algorithm*

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยข้อสรุปจากงานวิจัย [5] กล่าวว่าการเลือกใช้ค่ามัธยฐานชนิด Median-of-five จะลดเวลาในการประมวลผลเฉลี่ยลง 23 - 30% และถ้าหากข้อมูลมีขนาดใหญ่ก็จะมีโอกาสเลือกค่าไพวอทที่เหมาะสมและมีประสิทธิภาพมากที่สุด

2.2.2 ประเภทของไพวอทแบบต่าง ๆ

จากหลายงานวิจัยในหัวข้อ 2.2.1 พบว่ามีเทคนิคการเลือกค่าไพวอทหลายวิธีที่ช่วยให้การทำขั้นตอนวิธีควิกซอร์ต (Quicksort) มีประสิทธิภาพมากที่สุดและหลีกเลี่ยงผลลัพธ์ในกรณีที่เลวร้ายที่สุด (กรณีที่ค่าไพวอทเป็นค่าที่น้อยที่สุด หรือมากที่สุดของกลุ่มข้อมูล) โดยมีวิธีการหาค่าไพวอทดังนี้

2.2.2.1 **ค่าเริ่มต้น** คือ ค่าที่ได้จากการนำค่าที่เป็นข้อมูลตัวแรกของกลุ่มข้อมูลมาเป็นค่าไพวอท

2.2.2.2 **ค่าสุดท้าย** คือ ค่าที่ได้จากการนำค่าที่เป็นข้อมูลตัวสุดท้ายของกลุ่มมาเป็นค่าไพวอท

2.2.2.3 **ค่าสุ่ม** คือ ค่าที่ได้จากการสุ่มตำแหน่งของข้อมูลของกลุ่มมาเป็นค่าไพวอท

2.2.2.4 **ค่ากึ่งกลาง** คือ ค่าที่ได้จากการนำค่ากลางของกลุ่มข้อมูลมาเป็นค่าไพวอท

2.2.2.5 **ค่ามัธยฐาน** คือ ค่าที่ได้จากการแจกแจงกลุ่มข้อมูล โดยนำเอาค่าที่อยู่ตำแหน่งตรงกลางของกลุ่มข้อมูลมาเป็นค่าไพวอท

โดยค่ามัธยฐานนั้นมีหลายประเภท รายละเอียดจะอยู่ในหัวข้อ 2.2.3

2.2.3 มัธยฐาน (Median)

การเลือกไพวอทที่ดีนั้นจะส่งผลให้ความเร็วในการทำงานของควิกซอร์ตนั้นดีขึ้น โดยจะมีผลกับ 3 ปัจจัยที่สำคัญของอัลกอริทึมก็คือ จำนวนครั้งในการเปรียบเทียบ จำนวนครั้งในการสลับที่ จำนวนครั้งในการเรียกซ้ำ โดยถ้าเราสามารถลดการเรียกซ้ำได้ ก็จะลดการเปรียบเทียบและการสลับที่ได้ โดยจากหนังสือหลายเล่มนั้นได้เสนอวิธีการเลือกค่าไพวอทโดยเลือกจากค่าเริ่มต้นเป็นหลักซึ่งอาจจะได้ผลไม่ดีกับข้อมูลบางประเภทเช่น ข้อมูลที่เรียงลำดับแล้ว โดยได้มีงานวิจัย [5] เสนอวิธีการเลือกค่าไพวอทโดยใช้ค่ามัธยฐานเช่น median-of-three, median-of-five, median-of-seven, median-of-nine โดยแต่ละวิธีมีรายละเอียดและขั้นตอนวิธีดังนี้

2.2.3.1 Median-of-three คือ การเลือกค่ามัธยฐานจากข้อมูล 3 ตัว โดยมาจากค่าเริ่มต้น, ค่ากลาง และค่าสุดท้ายของกลุ่มข้อมูล แล้วนำมาเรียงลำดับหาค่ากลางเป็นค่าไพวอท

2.2.3.2 Median-of-five คือ การเลือกค่ามัธยฐานจากข้อมูล 5 ตัว โดยมาจากค่าเริ่มต้น, ค่ากลาง, ค่าสุดท้าย และอีก 2 ค่าได้จาก (ค่าเริ่มต้น + ค่าสุดท้าย) / 4 และ 3 * ((ค่าเริ่มต้น + ค่าสุดท้าย) / 4) ของกลุ่มข้อมูล แล้วนำมาเรียงลำดับหาค่ากลางเป็นค่าไพวอท โดยอัลกอริทึมนี้ถูกใช้โดย Breast et al (2000) และ Cerin (2002)

2.2.3.3 Median-of-seven คือ การเลือกค่ามัธยฐานจากข้อมูล 7 ตัว โดยมาจากค่าเริ่มต้น, ค่ากลาง, ค่าสุดท้าย, (ค่าเริ่มต้น + ค่าสุดท้าย) / 4, 3 * ((ค่าเริ่มต้น + ค่าสุดท้าย) / 4) และอีก 2 ค่าได้จากการเลือกแบบสุ่มแล้วนำมาเรียงลำดับหาค่ากลางเป็นค่าไพวอท

2.2.3.4 Median-of-nine คือ การเลือกค่ามัธยฐานจากข้อมูล 9 ตัว โดยมาจากค่าเริ่มต้น, ค่ากลาง, ค่าสุดท้าย, (ค่าเริ่มต้น + ค่าสุดท้าย) / 8, (ค่าเริ่มต้น + ค่าสุดท้าย) / 4, 3* ((ค่าเริ่มต้น + ค่าสุดท้าย) / 8), 5*((ค่าเริ่มต้น + ค่าสุดท้าย) / 8), 7*((ค่าเริ่มต้น + ค่าสุดท้าย) / 8)

โดยมีตัวอย่างของการหาค่าไพวอทโดยใช้ค่ามัธยฐานดังต่อไปนี้

ตัวอย่าง 2.2

การเลือกค่าไพวอทจากค่ามัธยฐาน (Median-of-five) จากกลุ่มข้อมูลที่มีข้อมูล 9 ตัว

ลำดับที่	0	1	2	3	4	5	6	7	8
ข้อมูล	27	15	22	37	11	59	18	50	42

ค่าเริ่มต้น	(ค่าเริ่มต้น + ค่าสุดท้าย) / 4	ค่ากลาง	3*((ค่าเริ่มต้น + ค่าสุดท้าย) / 4)	ค่าสุดท้าย

ภาพที่ 2.2 ตำแหน่งข้อมูลของ Median of five

กำหนดให้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

v[0] = ค่าเริ่มต้น, v[1] = ค่าสุดท้าย, v[2] = (ค่าเริ่มต้น + ค่าสุดท้าย) / 2, v[3] = (ค่าเริ่มต้น + ค่าสุดท้าย) / 4, v[4] = 3*((ค่าเริ่มต้น + ค่าสุดท้าย) / 4)

จะได้

v[0] = 0, v[1] = 8, v[2] = (0 + 8) / 2 = 4, v[3] = (0 + 8) / 4 = 2, v[4] = 3*((0 + 8) / 4) = 6

เมื่อแทนตำแหน่งจะได้ค่าเป็น

v[0] = 27, v[1] = 42, v[2] = 11, v[3] = 22, v[4] = 18

v[0]	v[3]	V[2]	V[4]	V[1]

ภาพที่ 2.3 ตัวแทนข้อมูลของ Median of five

หาค่ามัธยฐานของ Median-of-five โดย

1. เปรียบเทียบค่าระหว่าง v[0] และ v[1] โดยถ้า v[0] มากกว่า v[1] ให้สลับที่ระหว่าง v[0] และ v[1]

2. เปรียบเทียบค่าระหว่าง v[1] และ v[2] โดยถ้า v[1] มากกว่า v[2] ให้สลับที่ระหว่าง v[1] และ v[2]

3. เปรียบเทียบค่าระหว่าง v[2] และ v[3] โดยถ้า v[2] มากกว่า v[3] ให้สลับที่ระหว่าง v[2] และ v[3]

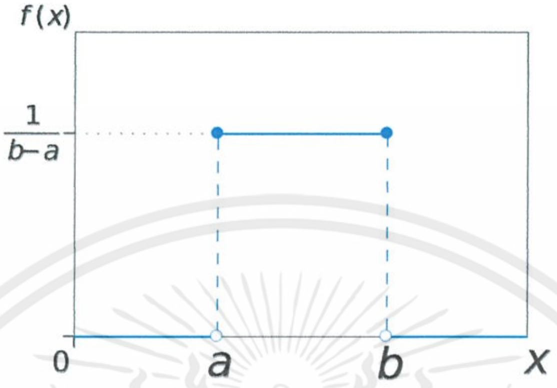
4. เปรียบเทียบค่าระหว่าง v[3] และ v[4] โดยถ้า v[3] มากกว่า v[4] ให้สลับที่ระหว่าง v[3] และ v[4]

โดยผลลัพธ์ ค่าตัวแทนทั้ง 5 จะเรียงตำแหน่งกันจากน้อยไปหามาก ทำให้ค่า v[2] เป็นค่าที่อยู่กึ่งกลางของกลุ่มข้อมูลมากที่สุด โดยขั้นตอน median-of-five จะเลือกค่า v[2] เป็นค่าไพวอทของกลุ่มข้อมูลนั้น ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 การสุ่มค่าแบบ uniform distribution

การสุ่มค่าแบบ uniform distribution คือ การสุ่มทุกค่าของจำนวนเต็มที่อยู่ในช่วง a (minimum) ถึง b (maximum) โดยที่ a < b ด้วยความน่าจะเป็นเท่าๆ กัน



ภาพที่ 2.4 การสุ่มค่าแบบ uniform distribution

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

ออกแบบและขั้นตอนการทดลอง

การวิจัยเรื่องการเลือกค่าไพวอทของควิกซอร์ตที่เหมาะสมสำหรับข้อมูลแบบจำนวนเต็ม เป็นการวิจัยเชิงทดลอง การทดลองนี้เป็นการทดลองโดยการเลือกค่าไพวอทแบบต่าง ๆ ของขั้นตอนวิธีควิกซอร์ต โดยจะทดลองกับข้อมูลประเภทจำนวนเต็ม (Integer) โดยผู้วิจัยได้ทำการออกแบบการทดลองและดำเนินตามขั้นตอนดังนี้ (3.1) วัตถุประสงค์การทดลอง (3.2) อุปกรณ์ในการดำเนินการทดลอง (3.3) ข้อมูลทดลอง (3.4) การเลือกค่าไพวอทในการทดลอง (3.5) มาตรวัดผลการทดลอง

3.1 วัตถุประสงค์การทดลอง

เพื่อเป็นการเปรียบเทียบประสิทธิภาพของการเลือกค่าไพวอทที่เหมาะสมกับข้อมูลประเภทต่าง ๆ ซึ่งเป็นการพัฒนาต่อยอดจากงานวิจัย *Comparative Analysis of Some Pivot Selection Schemes for Quicksort Algorithm* [5] ซึ่งได้ทำการทดลองกับข้อมูลแบบสุ่มประเภทจำนวนเต็มเพียงอย่างเดียว และผลลัพธ์แสดงถึงเวลาการประมวลผลเฉลี่ยเท่านั้น

วัตถุประสงค์ของงานวิจัยนี้คือต้องการต่อยอดจากงานวิจัยดังกล่าว โดยปรับปรุงข้อมูลทดสอบให้มีความหลากหลายเช่น ข้อมูลสุ่มแบบ uniform distribution, ข้อมูลแบบเกือบเรียงลำดับจากน้อยไปมาก, ข้อมูลสุ่มโดยมีโอกาสซ้ำกัน เป็นต้น และมีมาตรวัดผลเพิ่มเติมจากเดิมที่ใช้แค่เวลาเฉลี่ยในการประมวลผลเป็นตัววัด ผู้วิจัยได้เพิ่มมาตรวัดอื่น ๆ ได้แก่ จำนวนครั้งในการเปรียบเทียบ และจำนวนครั้งในการสลับที่

3.2 อุปกรณ์ในการดำเนินการทดลอง

เขียนโปรแกรมควิกซอร์ตด้วยภาษาจาวา ทดลองบนเครื่อง Intel Core i7-2620M 2.7 GHz และแรม 8 GB บนระบบปฏิบัติการ Windows 7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 ข้อมูลทดลอง

การทดลองนี้จะใช้ข้อมูลจำนวนเต็ม 5 ประเภท ดังนี้

3.3.1 ข้อมูลประเภทที่ 1 เป็นข้อมูลที่ได้จากการสุ่มแบบ uniform distribution ในช่วง 1 - 1000 มีขนาดของข้อมูลเป็น 1000, 5000 และ 10000 ตัว โดยแต่ละขนาดประกอบไปด้วย 10 ชุดข้อมูลทดสอบที่แตกต่างกัน

จุดประสงค์ในการทดลองของข้อมูลประเภทที่ 1 คือ เพื่อทดสอบหาค่าไพวอทที่เหมาะสมกับข้อมูลที่มีการกระจายอยู่ในช่วง 1 - 1000 โดยเป็นการกระจายที่ไม่ได้มีการกำหนดรูปแบบ และมีขนาดของข้อมูลที่แตกต่างกัน

3.3.1.1 ตัวอย่างข้อมูลประเภทที่ 1

941	616	526	230	350	391	133	483	708	784	803	201
513	109	120	39	179	712	405	790	615	886	615	350

3.3.2 ข้อมูลประเภทที่ 2 เป็นข้อมูลที่เกือบเรียงลำดับจากน้อยไปมาก โดยที่มีข้อมูลบางส่วนที่ยังไม่ได้เรียงลำดับ ขนาดข้อมูลทั้งหมดมี 10000 ตัว โดยมีเปอร์เซ็นต์ของจำนวนข้อมูลที่ยังไม่ได้เรียงลำดับเป็น 0% (เรียงลำดับแล้วทั้งหมด), 1%, 5%, 70% และ 90% ของจำนวนข้อมูลทั้งหมด โดยแต่ละเปอร์เซ็นต์ประกอบไปด้วย 10 ชุดข้อมูลทดสอบที่แตกต่างกัน

จุดประสงค์ในการทดลองของข้อมูลประเภทที่ 2 คือ เพื่อให้การหาค่าไพวอทแต่ละแบบนั้นมีโอกาสที่จะเกิดกรณีที่แย่ที่สุด (worst-case) เพราะข้อมูลส่วนใหญ่นั้นเกือบเรียงลำดับแล้ว และเพื่อพิสูจน์ว่าการใช้ค่ามัธยฐานหรือค่ากลางนั้นจะสามารถหลีกเลี่ยงกรณีที่แย่ที่สุด ได้อย่างมีประสิทธิภาพจริงหรือไม่

3.3.2.1 ตัวอย่างข้อมูลประเภทที่ 2

1	2	3	4	5	50	102	99	9	10	11	12
13	14	1050	504	17	908	19	78	2054	22	23	24

3.3.3 ข้อมูลประเภทที่ 3 เป็นข้อมูลที่เกือบเรียงลำดับจากมากไปน้อย (reverse) โดยมีข้อมูลบางส่วนที่เรียงลำดับจากน้อยไปมากแล้ว (เรียงลำดับถูกต้อง) ขนาดข้อมูลทั้งหมดมี 10000 ตัว โดยมี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เปอร์เซ็นต์ของจำนวนข้อมูลที่เรียงลำดับจากน้อยไปมากแล้วเป็น 0% (ข้อมูลเรียงลำดับจากมากไปน้อยทั้งหมด), 1%, 5%, 70% และ 90% ของจำนวนข้อมูลทั้งหมด โดยแต่ละเปอร์เซ็นต์ประกอบไปด้วย 10 ชุดข้อมูลทดสอบที่แตกต่างกัน

จุดประสงค์ในการทดลองของข้อมูลประเภทที่ 3 คือ เพื่อให้การหาค่าไพวอทแต่ละแบบนั้นมีโอกาสที่จะเกิดกรณีที่แย่ที่สุด (worst-case) แต่จะแตกต่างจากข้อมูลประเภทที่ 2 เพราะมีข้อมูลบางส่วนที่เรียงลำดับถูกต้องแล้ว (เรียงจากน้อยไปมาก) ดังนั้นการใช้ค่ามัธยฐานหรือค่ากลางนั้นยังสามารถหลีกเลี่ยงกรณีที่แย่ที่สุดได้อย่างมีประสิทธิภาพอยู่หรือไม่ เพราะการเลือกค่าไพวอทแบบอื่น ๆ มีโอกาสที่จะเลือกค่าที่เป็นกรณีที่แย่ที่สุดมากกว่าข้อมูลประเภทที่ 2

3.3.3.1 ตัวอย่างข้อมูลประเภทที่ 3

10000	9999	9998	9997	5	6	7	8	9	9991	9990	9989
9988	9987	15	16	17	18	9982	9981	9980	9979	23	24

3.3.4 ข้อมูลประเภทที่ 4 เป็นข้อมูลแบบสุ่ม และมีข้อมูลซ้ำ โดยข้อมูลที่ซ้ำนั้น จะซ้ำ 5 ตัว ขนาดข้อมูลทั้งหมดมี 10000 ตัว โดยขนาดของจำนวนข้อมูลที่ซ้ำมี 0, 2000, 5000 และ 7000 ตัว โดยแต่ละขนาดประกอบไปด้วย 10 ชุดข้อมูลทดสอบที่แตกต่างกัน

จุดประสงค์ในการทดลองของข้อมูลประเภทที่ 4 คือ เพิ่มโอกาสที่การเลือกค่าไพวอทแต่ละประเภทจะเลือกค่าซ้ำเดิมหลายๆ รอบ เพราะข้อมูลที่ใช้ทดสอบมีข้อมูลซ้ำกันอยู่

3.3.4.1 ตัวอย่างข้อมูลประเภทที่ 4

1491	771	2264	14	2157	1698	2293	771	6143	14	1838	2922
771	2220	771	1000	2863	771	14	2139	14	2187	64	14

3.3.5 ข้อมูลประเภทที่ 5 เป็นข้อมูลแบบสุ่ม และมีข้อมูลซ้ำ โดยข้อมูลที่ซ้ำนั้น จะซ้ำ 20 ตัว ขนาดข้อมูลทั้งหมดมี 10000 ตัว โดยขนาดจำนวนข้อมูลซ้ำมี 0, 2000, 5000 และ 7000 ตัว โดยแต่ละขนาดประกอบไปด้วย 10 ชุดข้อมูลทดสอบที่แตกต่างกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จุดประสงค์ในการทดลองของข้อมูลประเภทที่ 5 คือ เพิ่มโอกาสที่การเลือกค่าไพวอทแต่ละประเภทจะเลือกค่าซ้ำเดิมหลายๆ รอบ โดยข้อมูลชุดที่ 5 นี้จะมีข้อมูลที่ซ้ำมากกว่าข้อมูลชุดที่ 4 ทำให้โอกาสที่จะเลือกค่าไพวอทซ้ำเดิมมากกว่า 1 รอบนั้นมีมากขึ้น

3.3.5.1 ตัวอย่างข้อมูลประเภทที่ 5

60	1016	60	90	60	60	60	9023	60	60	4037	9006
60	60	1049	60	60	3551	60	2622	60	60	8023	4007
8043	60	3591	60	60	60	7670	1017	277	60	60	9014
2953	2019	95	127	4011	70	1695	7653	3585	8025	3597	100

3.4 การเลือกค่าไพวอทในการทดลอง

ในแต่ละประเภทของข้อมูล จะทดลองโดยใช้ค่าไพวอทที่แตกต่างกันออกไปดังนี้

3.4.1 ค่าเริ่มต้น คือ ค่าที่ได้จากการนำค่าที่เป็นข้อมูลตัวแรกของกลุ่มข้อมูลมาเป็นค่าไพวอท

3.4.2 ค่าสุดท้าย คือ ค่าที่ได้จากการนำค่าที่เป็นข้อมูลตัวสุดท้ายของกลุ่มมาเป็นค่าไพวอท

3.4.3 ค่าสุ่ม คือ ค่าที่ได้จากการสุ่มตำแหน่งของข้อมูลของกลุ่มมาเป็นค่าไพวอท

3.4.4 ค่ากึ่งกลาง คือ ค่าที่ได้จากการนำค่ากลางของกลุ่มข้อมูลมาเป็นค่าไพวอท

3.4.5 ค่ามัธยฐาน คือ ค่าที่ได้จากการแจกแจงกลุ่มข้อมูลโดยนำเอาค่าที่อยู่ตำแหน่งตรงกลางของกลุ่มข้อมูลมาเป็นค่าไพวอท โดยจะใช้ประเภทของค่ามัธยฐานดังนี้

3.4.5.1 median-of-three

3.4.5.2 median-of-five

3.5 มาตรวัดผลการทดลอง

ในการทดลอง ผู้วิจัยจะใช้เวลาในการประมวลผลเฉลี่ย, จำนวนครั้งในการเปรียบเทียบเฉลี่ยและจำนวนครั้งในการสลับที่เฉลี่ย ที่เกิดจากการทดสอบโปรแกรมกับข้อมูลทั้ง 5 ประเภทในหัวข้อ 3.3 เป็นมาตรวัดผลการทดลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

จากบทที่แล้ว ผู้วิจัยได้กล่าวถึงขั้นตอนการทดลอง ประเภทข้อมูล และมาตรวัดการทดลอง ในบทนี้จะเป็นการรายงานผลการทดลองของข้อมูลทั้ง 5 ประเภท กับค่าไพวอทชนิดต่างๆ

4.1 ผลการทดลอง

ในการทดลองจะใช้ข้อมูลทดสอบรูปแบบละ 10 ชุดข้อมูล ตามประเภทข้อมูลแต่ละประเภทที่ได้กล่าวในบทที่ 3 เพื่อหาค่าเฉลี่ยของเวลาในการประมวลผลโดยใช้คำสั่ง stopwatch.start() และ stopwatch.stop() ของจาวา, ค่าเฉลี่ยของจำนวนครั้งในการเปรียบเทียบ, ค่าเฉลี่ยของจำนวนครั้งในการสลับที่ โดยผลการทดลองมีดังนี้

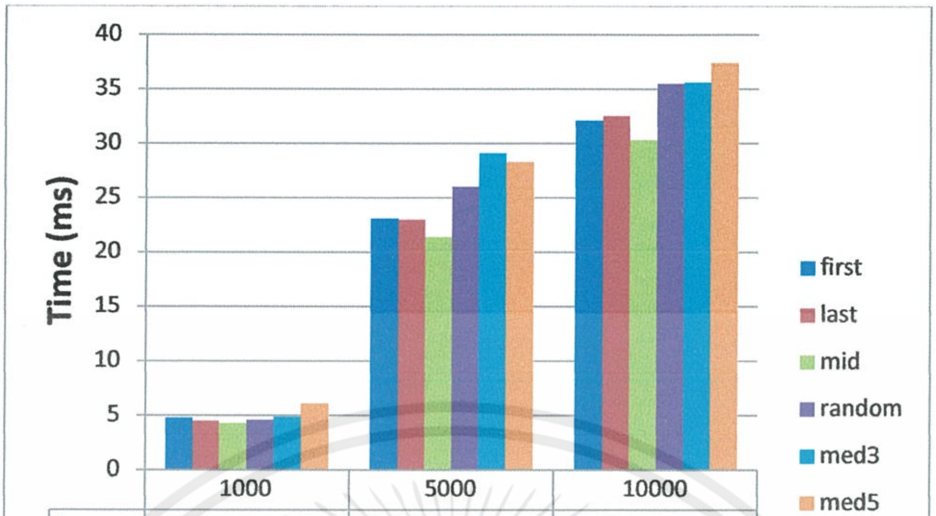
4.1.1 ผลการทดลองของข้อมูลประเภทที่ 1

ข้อมูลประเภทที่ 1 เป็นข้อมูลที่ได้จากการสุ่มแบบ uniform distribution ในช่วง 1 - 1000 มีขนาดของข้อมูลเป็น 1000, 5000 และ 10000 ตัว

จากผลลัพธ์ของข้อมูลประเภทที่ 1 ในภาพที่ 4.1 แสดงให้เห็นว่าการเลือกค่าไพวอทจากค่ากึ่งกลาง (middle) จะได้ค่าเวลาเฉลี่ยเร็วกว่าไพวอทชนิดอื่นๆ เพราะได้ค่ากึ่งกลางของข้อมูลที่เหมาะสมกับข้อมูลประเภท uniform distribution และเมื่อเทียบกับการเลือกค่าไพวอทชนิด median-of-three หรือ median-of-five ที่ต้องใช้เวลาในการหาค่ากลางที่เรียงลำดับแล้วจากการเลือกตัวอย่างข้อมูลในแต่ละพาร์ทิชัน

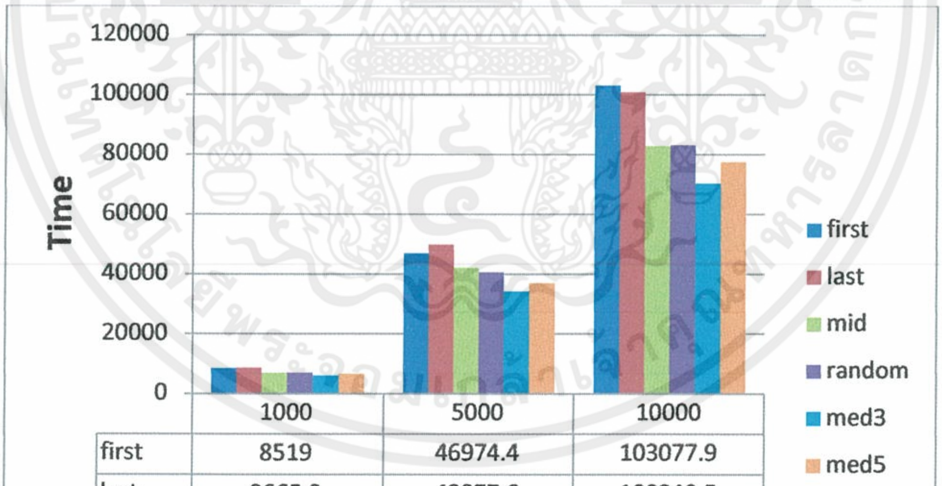
จากผลลัพธ์ภาพที่ 4.2 และ 4.3 การเลือกค่าไพวอทจาก median-of-three และ median-of-five จะใช้จำนวนครั้งในการเปรียบเทียบน้อยกว่าไพวอทชนิดอื่นๆ เพราะการเลือกใช้ median จะทำการหาค่ากึ่งกลางที่เหมาะสมของแต่ละพาร์ทิชันนั้นๆ แล้วทำให้มีประสิทธิภาพมากกว่ากว่าการหาค่าไพวอทแบบอื่นๆ และจำนวนครั้งในการสลับที่นั้นมีค่าฉลี่ยใกล้ๆ เพราะข้อมูลอยู่ในรูปแบบกระจายจำเป็นต้องมีการสลับค่าเกือบทุกตำแหน่ง

เอกสารนี้เป็นเพียงส่วนหนึ่งเท่านั้น ใช้สำหรับงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



	1000	5000	10000
first	4.8	23.1	32.1
last	4.5	23	32.5
mid	4.3	21.4	30.3
random	4.6	26	35.5
med3	4.9	29.1	35.6
med5	6.1	28.3	37.4

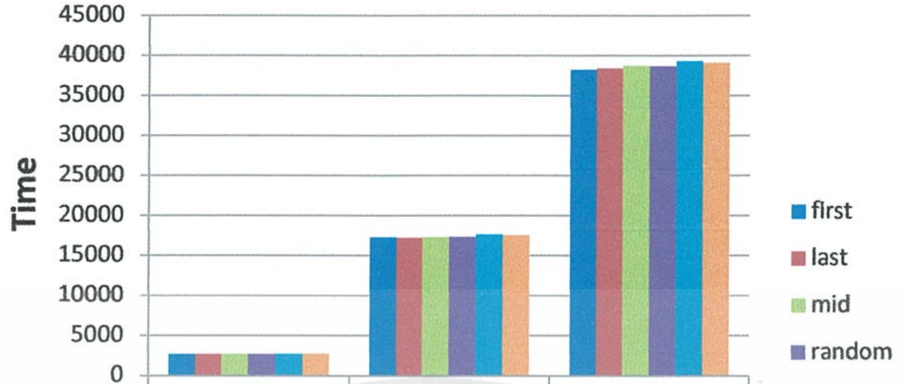
ภาพที่ 4.1 เวลาในการประมวลผลของข้อมูลประเภท uniform distribution



	1000	5000	10000
first	8519	46974.4	103077.9
last	8665.8	49877.6	100840.5
mid	6927.5	42106.8	82795.5
random	7049.5	40537.8	83044.3
med3	6144.9	34266.9	70329.5
med5	6599.3	37040.5	77462.3

ภาพที่ 4.2 จำนวนครั้งในการเปรียบเทียบของข้อมูลประเภท uniform distribution

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



	1000	5000	10000
first	2708.6	17278	38229.1
last	2707.2	17201.4	38399.5
mid	2695.1	17253.8	38721
random	2712.5	17347.9	38664.9
med3	2760.2	17663.2	39319.4
med5	2760.5	17559.6	39047.9

ภาพที่ 4.3 จำนวนครั้งในการสลับที่ของข้อมูลประเภท uniform distribution

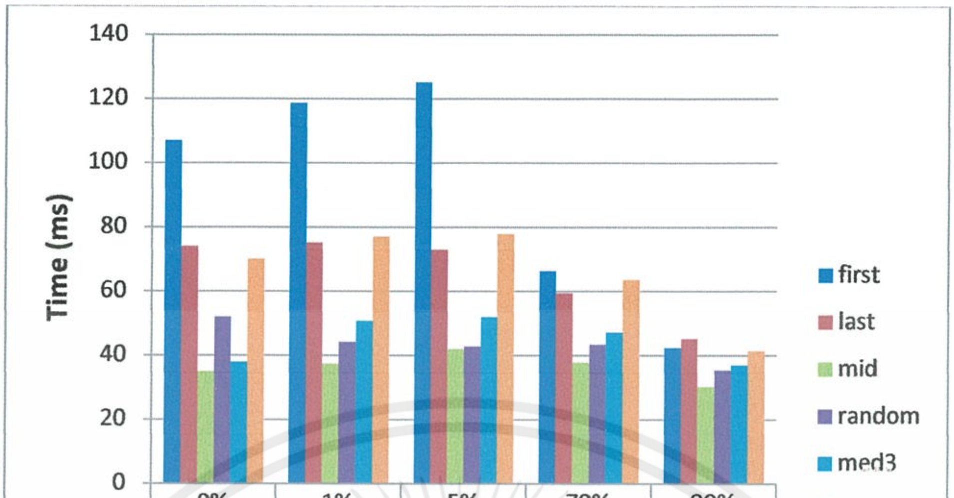
4.1.2 ผลการทดลองของข้อมูลประเภทที่ 2

ข้อมูลประเภทที่ 2 เป็นข้อมูลที่เกือบเรียงลำดับจากน้อยไปมาก โดยที่มีข้อมูลบางส่วนที่ยังไม่ได้เรียงลำดับ ขนาดข้อมูลทั้งหมดมี 10000 ตัว โดยมีเปอร์เซ็นต์ของจำนวนข้อมูลที่ยังไม่ได้เรียงลำดับเป็น 0% (เรียงลำดับแล้วทั้งหมด), 1%, 5%, 70% และ 90% ของจำนวนข้อมูลทั้งหมด

จากผลลัพธ์ในภาพที่ 4.4 แสดงให้เห็นว่าแนวโน้มผลลัพธ์ของข้อมูลเกือบเรียงลำดับจากน้อยไปมากแล้วในเปอร์เซ็นต์ของข้อมูลที่ยังไม่ได้เรียงลำดับ 1% และ 5% กับข้อมูลที่มีการเรียงลำดับแล้วทั้งหมดนั้นไม่มีความแตกต่างกันมากนัก โดยที่การเลือกค่าไพวอทจากค่ากึ่งกลางใช้เวลาในการประมวลผลน้อยที่สุด เพราะมีโอกาสที่จะได้ค่ากึ่งกลางของค่าที่เรียงลำดับแล้วซึ่งเป็นกรณีที่ดีที่สุด (best-case) ตรงกันข้ามกับการเลือกไพวอทโดยใช้ค่าเริ่มต้นซึ่งอาจจะทำให้ได้ค่าที่น้อยที่สุดในแต่ละพาร์ทิชันซึ่งปีนกรณีที่แย่ที่สุด (worst-case)

ในขณะที่ข้อมูลที่มีเปอร์เซ็นต์ยังไม่ได้เรียงลำดับ 70% และ 90% การเลือกค่าไพวอทเริ่มต้นใช้เวลาในการประมวลผลลดลง เนื่องจากโอกาสที่จะได้ค่าน้อยที่สุดในแต่ละพาร์ทิชันลดน้อยลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



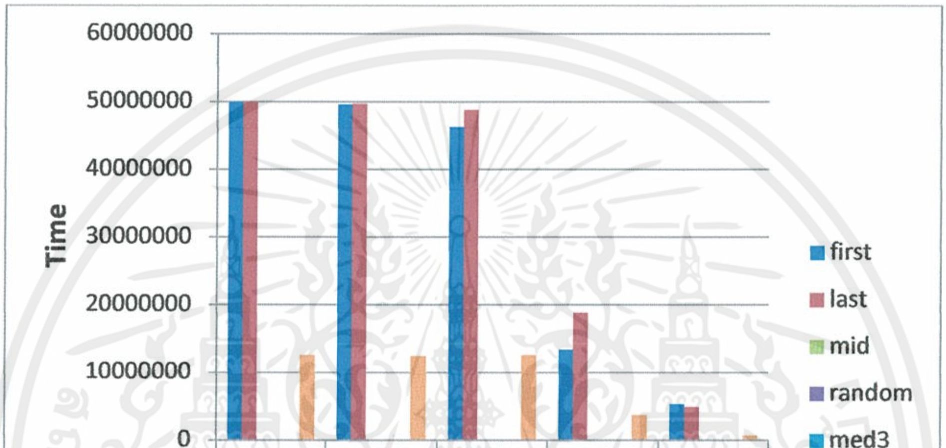
	0%	1%	5%	70%	90%
first	107	118.8	125.2	66.4	42.4
last	74	75.2	73	59.4	45.2
mid	35	37.4	42	37.8	30.2
random	52	44.2	42.8	43.4	35.4
med3	38	50.8	52	47.2	37
med5	70	77	77.8	63.6	41.4

ภาพที่ 4.4 เวลาในการประมวลผลของข้อมูลเกือบเรียงลำดับจากน้อยไปมาก

สำหรับผลลัพธ์จำนวนครั้งในการเปรียบเทียบของข้อมูลแบบเกือบเรียงลำดับจากน้อยไปมาก ดังภาพที่ 4.5 ค่าไพวอทเริ่มต้น และค่าสุดท้ายนั้นมีจำนวนในการเปรียบเทียบมากกว่าไพวอทชนิดอื่นๆ อย่างเห็นได้ชัด ในเปอร์เซ็นต์ของข้อมูลที่ยังไม่ได้เรียงลำดับ 0%, 1% และ 5% เพราะข้อมูลมีการเรียงลำดับแล้วเป็นส่วนมาก เมื่อเลือกค่าไพวอทที่เป็นกรณีแย่ที่สุดมา ทำให้เกิดการเปรียบเทียบจำนวนมากในแต่ละพาร์ทิชัน ในขณะที่การเลือกค่าไพวอทจากค่ากึ่งกลาง, ค่าสุ่ม และ median-of-three จะมีโอกาสได้ค่ากึ่งกลางของข้อมูลที่เรียงลำดับแล้ว ทำให้เกิดการเปรียบเทียบค่าน้อย ในขณะที่เปอร์เซ็นต์ของข้อมูลที่ยังไม่ได้เรียงลำดับ 70% และ 90% นั้น ค่าไพวอทเริ่มต้น และค่าสุดท้าย มีจำนวนครั้งในการเปรียบเทียบลดลงอย่างเห็นได้ชัด เนื่องจากข้อมูลเริ่มมีการกระจาย ทำให้โอกาสเลือกค่าที่เป็นกรณีที่แย่ที่สุดลดลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

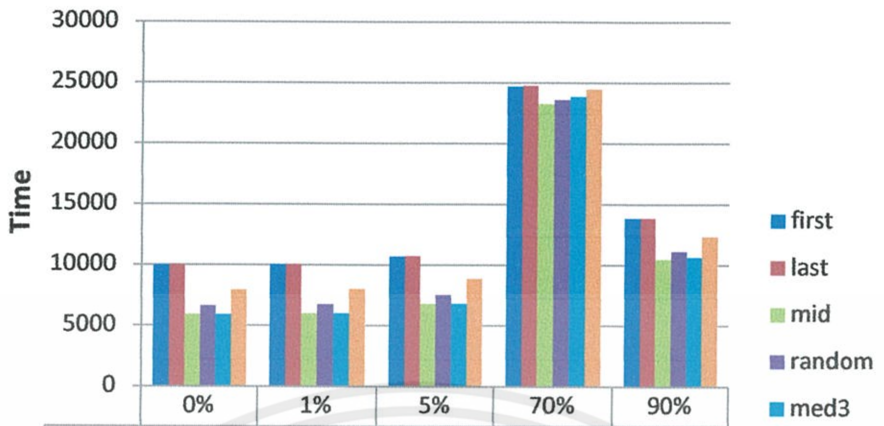
ผลลัพธ์จำนวนครั้งในการสลับที่ของข้อมูลแบบเกือบเรียงลำดับจากน้อยไปมากมีแนวโน้มไปในทิศทางเดียวกับจำนวนครั้งในการเปรียบเทียบคือค่าไพวอทกึ่งกลางและ median-of-three จะใช้จำนวนครั้งในการสลับที่น้อยเนื่องจากในแต่ละพาร์ทิชันของข้อมูลนั้นจะอยู่ในรูปแบบเกือบเรียงลำดับ ทำให้ค่ากลางที่ได้นั้นอยู่ในรูปแบบกรณีที่ดีที่สุดคือ ไม่จำเป็นต้องมีการสลับค่า ดังภาพที่ 4.6



	0%	1%	5%	70%	90%
first	49995000	49567958	46281932	13372495	5339208
last	49995000	49650917	48781250	18858735	4938081
mid	113631	113618	113368.6	109830.2	109026
random	146778	152618.8	152690.2	123881.8	118450
med3	113631	113561	112906.2	96795.8	90203.6
med5	12559309	12444765	12558594	3722283	674455.8

ภาพที่ 4.5 จำนวนครั้งในการเปรียบเทียบของข้อมูลเกือบเรียงลำดับจากน้อยไปมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



	0%	1%	5%	70%	90%
first	9999	10054.8	10707.8	24700.8	13865.6
last	9999	10061.2	10712.4	24770	13885.65
mid	5904	5998.2	6778.2	23238.2	10479.65
random	6619	6746.4	7546	23598	11127.35
med3	5904	6003.8	6814.6	23876.6	10649.75
med5	7951	8029.2	8855.8	24489	12331.25

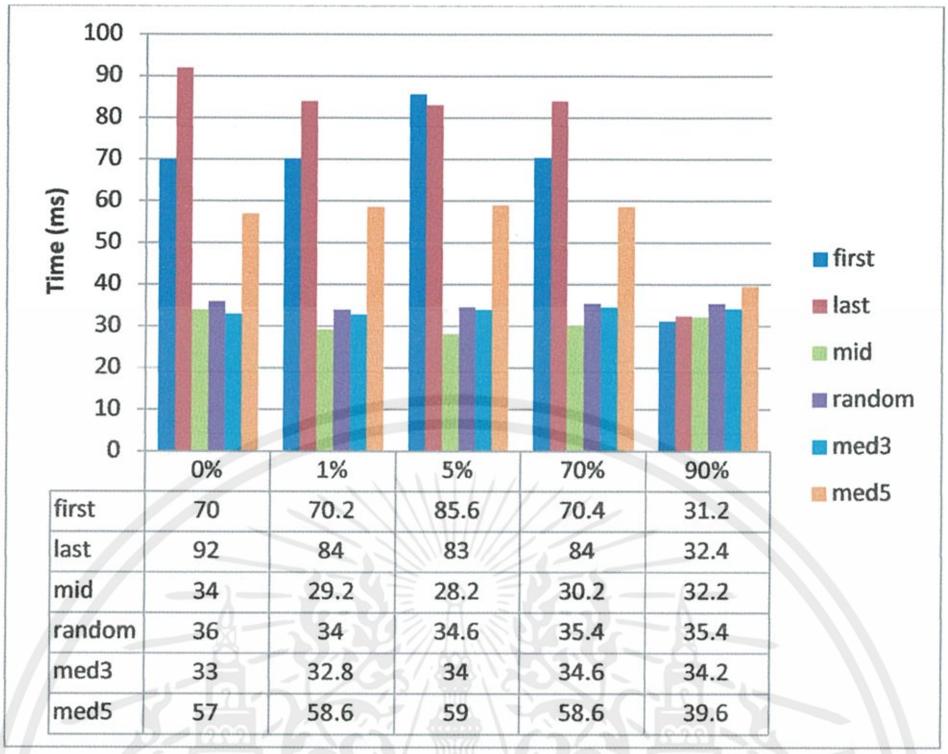
ภาพที่ 4.6 จำนวนครั้งในการสลับที่ของข้อมูลเกือบเรียงลำดับจากน้อยไปมาก

4.1.3 ผลการทดลองของข้อมูลประเภทที่ 3

ข้อมูลประเภทที่ 3 เป็นข้อมูลที่เกือบเรียงลำดับจากมากไปน้อย (reverse) โดยมีข้อมูลบางส่วนที่เรียงลำดับจากน้อยไปมากแล้ว (เรียงลำดับถูกต้อง) ขนาดข้อมูลทั้งหมดมี 10000 ตัว โดยมีเปอร์เซ็นต์ของจำนวนข้อมูลที่เรียงลำดับจากน้อยไปมากแล้วเป็น 0% (ข้อมูลเรียงลำดับจากมากไปน้อยทั้งหมด), 1%, 5%, 70% และ 90% ของจำนวนข้อมูลทั้งหมด

จากผลลัพธ์ภาพที่ 4.7 แสดงให้เห็นว่าชนิดไพวอทที่ใช้เวลาในการประมวลผลกับข้อมูลแบบเกือบเรียงย้อนกลับในภาพรวมน้อยที่สุดคือ ค่ากลาง เพราะว่ามีโอกาสที่จะได้ค่าที่เป็นกึ่งกลางของข้อมูลที่อยู่ในรูปแบบเรียงลำดับแล้วมากที่สุด ทำให้การเปรียบเทียบค่า และการสลับที่ทำได้อย่างมีประสิทธิภาพ ซึ่งจากผลลัพธ์สังเกตได้ว่าเวลาในการประมวลผลของค่าไพวอทกึ่งกลางจะคงที่ไม่ว่าข้อมูลจะเรียงลำดับแบบย้อนกลับมากน้อยแค่ไหน และชนิดไพวอทที่ใช้เวลาในการประมวลผลมากที่สุดในภาพรวมคือ ค่าเริ่มและค่าสุดท้าย มีโอกาสสูงที่จะเลือกค่าที่เป็นกรณีที่แย่ที่สุดของแต่ละพาร์ทิชั่น ในข้อมูลเกือบเรียงลำดับแบบย้อนกลับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

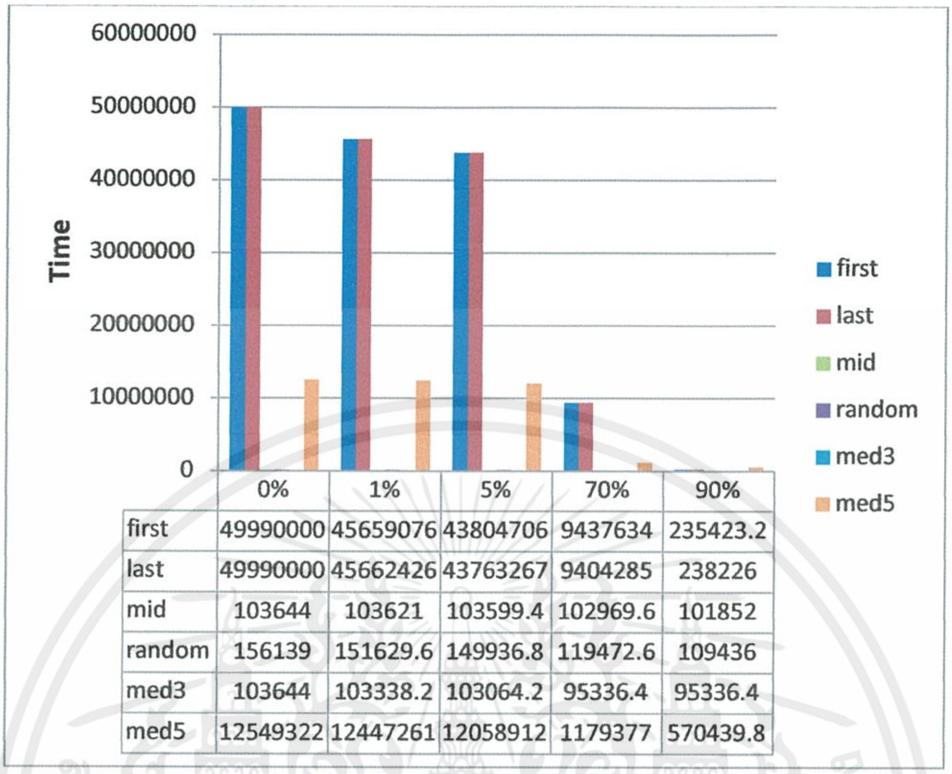


	0%	1%	5%	70%	90%
first	70	70.2	85.6	70.4	31.2
last	92	84	83	84	32.4
mid	34	29.2	28.2	30.2	32.2
random	36	34	34.6	35.4	35.4
med3	33	32.8	34	34.6	34.2
med5	57	58.6	59	58.6	39.6

ภาพที่ 4.7 เวลาในการประมวลผลของข้อมูลเกือบเรียงลำดับแบบย้อนกลับจากมากไปน้อย

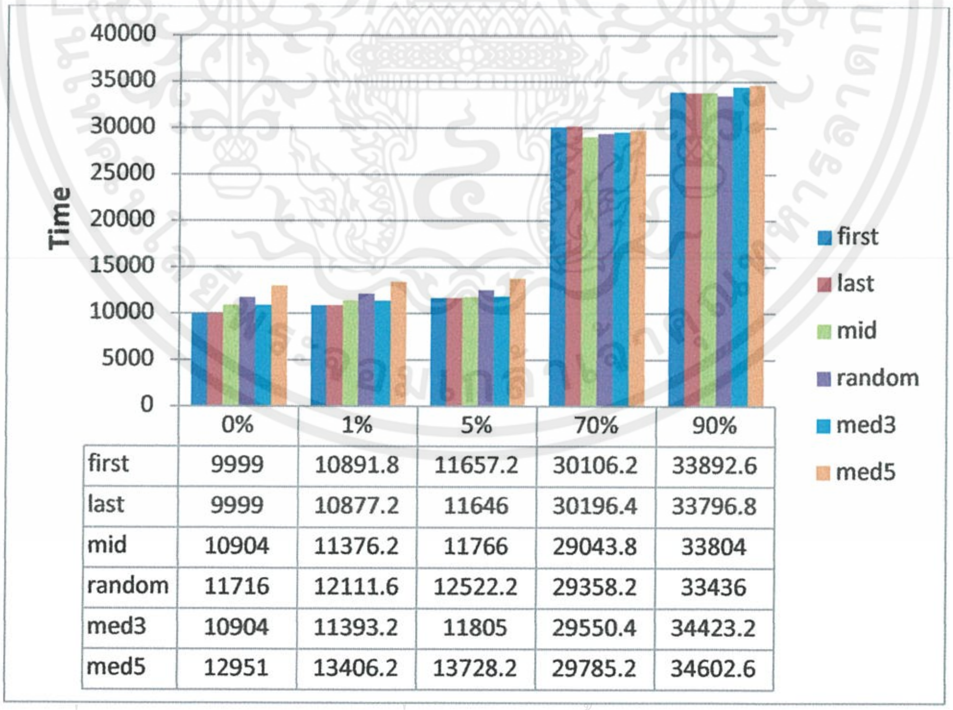
จากผลลัพธ์ภาพที่ 4.8 สามารถสรุปได้ว่าการเลือกค่ากลาง หรือ median-of-three เป็นไพวอทนั้นจะทำจำนวนครั้งในการเปรียบเทียบนั้นลดลง เนื่องจากการเลือกทั้งสองแบบนั้นจะหลีกเลี่ยงกรณีที่แย่ที่สุดในแต่ละพาร์ทิชันได้ดี เพราะข้อมูลอยู่ในรูปที่เรียงลำดับแล้ว โดยจากผลลัพธ์ median-of-three จะมีประสิทธิภาพดีกับข้อมูลที่ไม่ได้เรียงแล้วทั้งหมด และ ผลลัพธ์จำนวนครั้งในการสลับที่ในภาพที่ 4.9 จะเห็นได้ว่าจะมีค่าเฉลี่ยที่ใกล้กันเพราะข้อมูลเรียงลำดับแบบย้อนกลับ ต้องมีการสลับที่เกือบทุกตำแหน่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



	0%	1%	5%	70%	90%
first	49990000	45659076	43804706	9437634	235423.2
last	49990000	45662426	43763267	9404285	238226
mid	103644	103621	103599.4	102969.6	101852
random	156139	151629.6	149936.8	119472.6	109436
med3	103644	103338.2	103064.2	95336.4	95336.4
med5	12549322	12447261	12058912	1179377	570439.8

ภาพที่ 4.8 จำนวนครั้งในการเปรียบเทียบของข้อมูลเกือบเรียงลำดับแบบย้อนกลับจากมากไปน้อย



	0%	1%	5%	70%	90%
first	9999	10891.8	11657.2	30106.2	33892.6
last	9999	10877.2	11646	30196.4	33796.8
mid	10904	11376.2	11766	29043.8	33804
random	11716	12111.6	12522.2	29358.2	33436
med3	10904	11393.2	11805	29550.4	34423.2
med5	12951	13406.2	13728.2	29785.2	34602.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

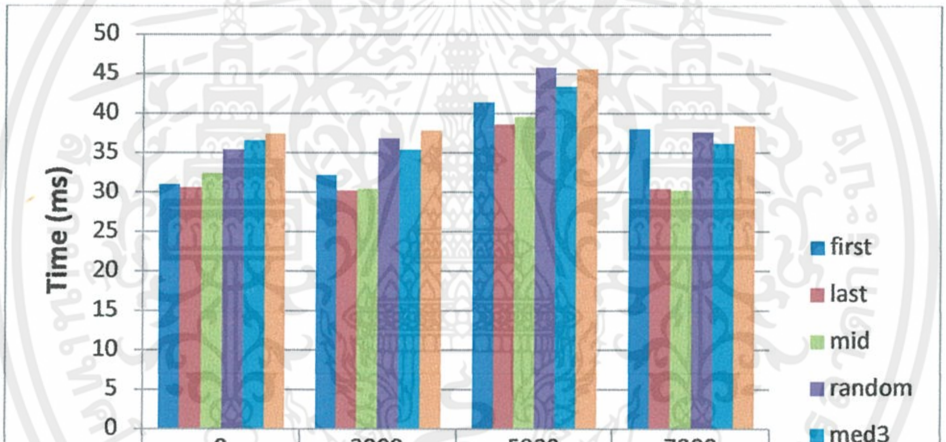
ภาพที่ 4.9 จำนวนครั้งในการสลับที่ของข้อมูลเกือบเรียงลำดับแบบย้อนกลับจากมากไปน้อย

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.4 ผลการทดลองของข้อมูลประเภทที่ 4

ข้อมูลประเภทที่ 4 เป็น เป็นข้อมูลแบบสุ่ม และมีข้อมูลซ้ำ โดยข้อมูลที่ซ้ำนั้น จะซ้ำ 5 ตัว ขนาดข้อมูลทั้งหมดมี 10000 ตัว โดยขนาดของจำนวนข้อมูลที่ซ้ำมี 0, 2000, 5000 และ 7000 ตัว

จากผลลัพธ์ภาพที่ 4.10 แสดงให้เห็นว่าไพวอทชนิด ค่าสุดท้ายและค่ากลาง ใช้เวลาในการประมวลผลน้อยที่สุดกับข้อมูลแบบสุ่มที่มีโอกาสซ้ำ เพราะมีโอกาสที่จะได้ค่าที่เหมาะสม และใช้เวลาในการหาค่าไพวอทน้อย เมื่อเทียบการใช้ค่าไพวอทจากการสุ่ม หรือ median เพราะมีโอกาสที่จะเลือกค่าที่ซ้ำกันได้ และต้องใช้เวลาในการคำนวณหาค่าไพวอท โดยจากผลลัพธ์ เมื่ออัตราส่วนที่ข้อมูลจะซ้ำกันมากขึ้น เวลาที่ใช้ในการประมวลผลของค่าสุดท้าย หรือค่ากลาง กับค่าสุ่ม หรือ median จะแตกต่างอย่างชัดเจน



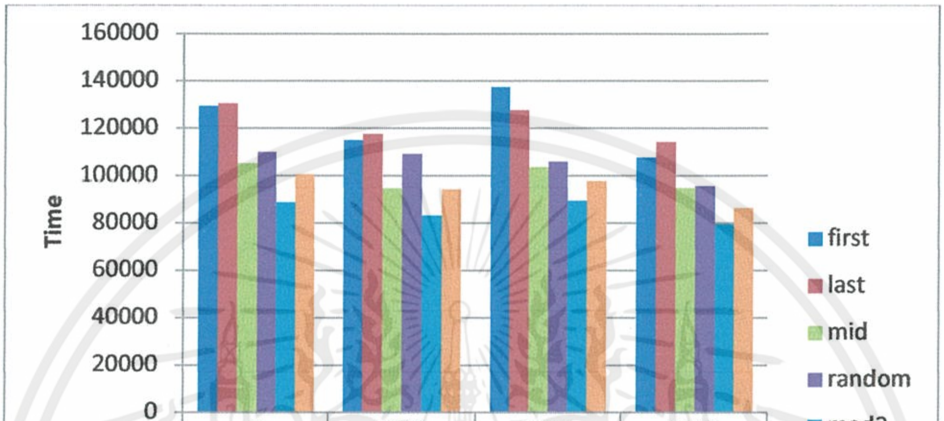
	0	2000	5000	7000
first	31	32.2	41.4	38
last	30.6	30.2	38.6	30.4
mid	32.4	30.4	39.6	30.2
random	35.4	36.8	45.8	37.6
med3	36.6	35.4	43.4	36.2
med5	37.4	37.8	45.6	38.4

ภาพที่ 4.10 เวลาในการประมวลผลของข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 5 ตัว

สำหรับผลลัพธ์จำนวนครั้งในการเปรียบเทียบของข้อมูลประเภทที่ 4 จากผลลัพธ์ภาพที่ 4.11 แสดงให้เห็นว่าการใช้ค่า median เป็นค่าไพวอทนั้นจะใช้จำนวนครั้งในการสลับที่น้อยกว่าชนิดไพวอทอื่นๆ และมีความคงที่ไม่ว่าข้อมูลจะมีอัตราส่วนในการซ้ำมาก หรือน้อย เพราะค่า median จะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

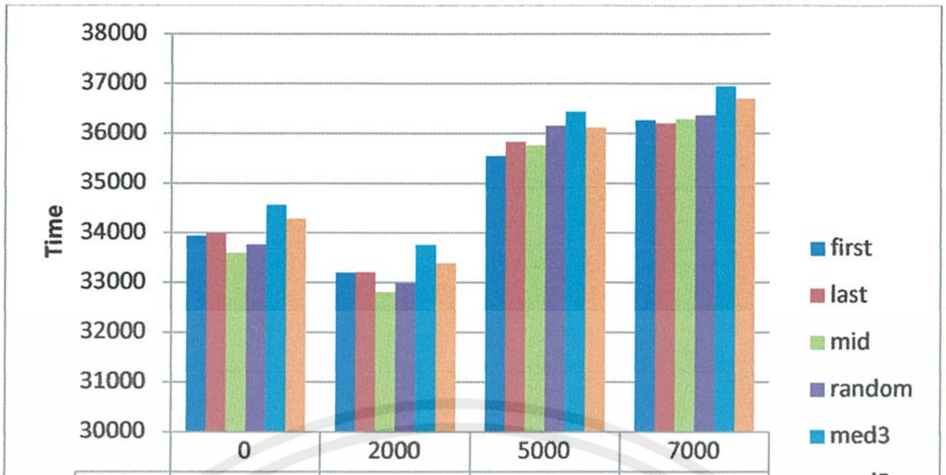
เลือกค่ากลางที่เหมาะสมของแต่ละพาร์ทิชัน โดยนำเอากลุ่มตัวอย่างมาประมวลผล ทำให้สามารถหลีกเลี่ยงค่ากลางที่เป็นตัวซ้ำ หรือค่ากลางที่เป็นกรณีที่แย่ที่สุด และผลลัพธ์จำนวนครั้งในการสลับที่ในภาพที่ 4.12 มีจำนวนค่าเฉลี่ยที่ใกล้เคียงกัน เพราะข้อมูลเป็นการการกระจายแบบสุ่ม และมีเลขซ้ำ ทำให้ต้องมีการสลับที่เกือบทุกตำแหน่งของข้อมูล



	0	2000	5000	7000
first	129459	115072.8	137592.2	107744
last	130532.6	117637.2	127761	114223.6
mid	105284.2	94612.8	103643	94751
random	109968	109228.6	105877.8	95614
med3	88800.4	83325.6	89536	79569.8
med5	100482.4	94269.2	97765.4	86292

ภาพที่ 4.11 จำนวนครั้งในการเปรียบเทียบของข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 5 ตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



	0	2000	5000	7000
first	33941	33200.8	35547	36271.4
last	33987.6	33208.6	35830.6	36203.2
mid	33594.4	32807.8	35764	36286.6
random	33765.2	32986.4	36157	36361.2
med3	34566.2	33755.6	36438	36951.6
med5	34280.4	33388	36112.4	36703.4

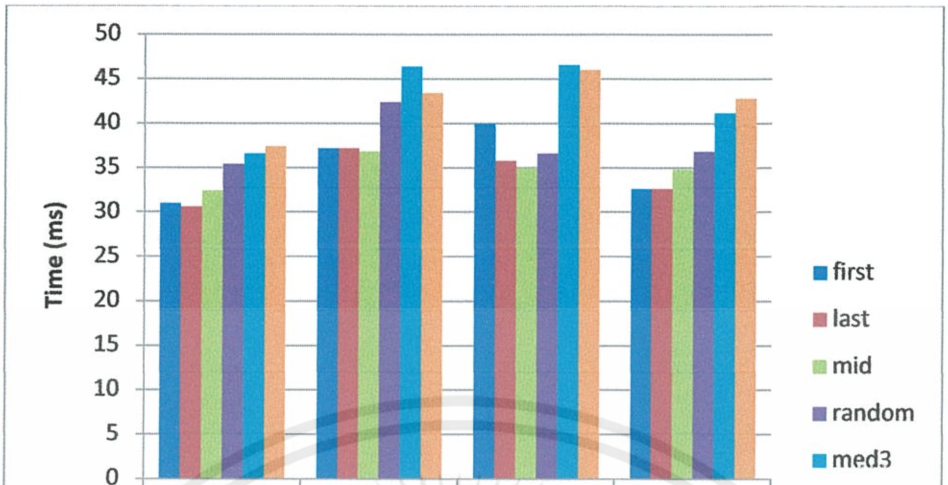
ภาพที่ 4.12 จำนวนครั้งในการสลับที่ของข้อมูลข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 5 ตัว

4.1.5 ผลการทดลองของข้อมูลประเภทที่ 5

ข้อมูลประเภทที่ 5 เป็นข้อมูลแบบสุ่ม และมีข้อมูลซ้ำ โดยข้อมูลที่ซ้ำนั้น จะซ้ำ 20 ตัว ขนาดข้อมูลทั้งหมดมี 10000 ตัว โดยขนาดจำนวนข้อมูลซ้ำมี 0, 2000, 5000 และ 7000 ตัว

จากผลลัพธ์ในภาพที่ 4.13 แสดงให้เห็นว่าเมื่อปรับข้อมูลให้ตัวเลขแต่ละตัวมีจำนวนเลขซ้ำมากขึ้น จากชุดการทดลองที่ 4 ที่ให้แต่ละตัวซ้ำ 5 ตัว โดยชุดการทดลองที่ 5 ที่ให้แต่ละตัวซ้ำ 20 ตัว ทำให้การหา ไพวอทโดยใช้ค่าเริ่มต้น และค่าสุดท้ายใช้เวลาในการประมวลผลน้อยที่สุด เพราะการหาค่าไพวอทชนิดอื่นๆ นั้นมีโอกาสที่จะเลือกข้อมูลที่ซ้ำ และต้องใช้เวลาในการประมวลผลนาน ตัวอย่างเช่น median จะใช้เวลาในการประมวลผลนานกับประเภทข้อมูลที่มีเลขซ้ำ โดยที่ยิ่งข้อมูลมีโอกาสซ้ำมาก เวลาในการประมวลผลก็จะมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

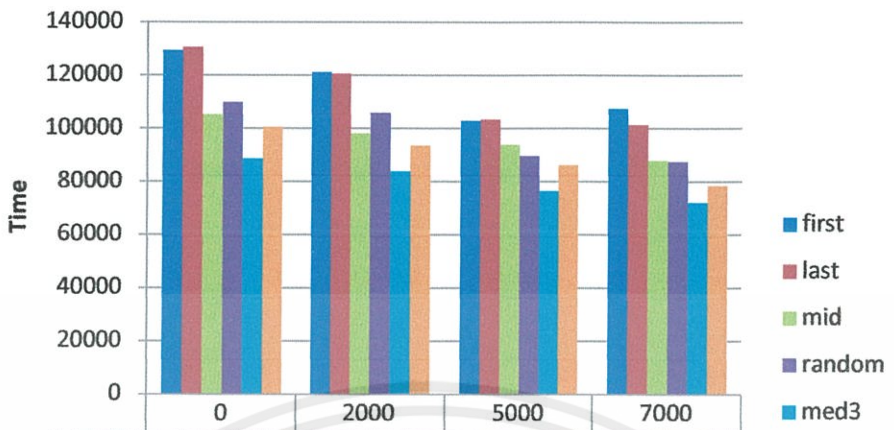


	0	2000	5000	7000
first	31	37.2	40	32.6
last	30.6	37.2	35.8	32.6
mid	32.4	36.8	35	34.8
random	35.4	42.4	36.6	36.8
med3	36.6	46.4	46.6	41.2
med5	37.4	43.4	46	42.8

ภาพที่ 4.13 เวลาในการประมวลผลของข้อมูลสุ่ม โดยมีข้อมูลซ้ำหมายเลขละ 20 ตัว

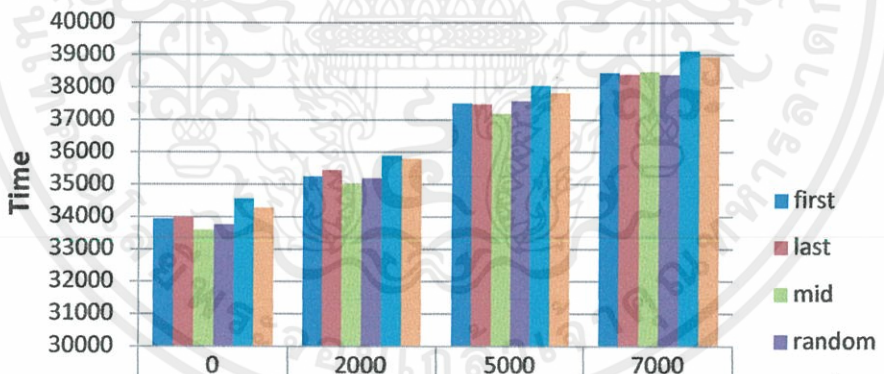
ผลลัพธ์จำนวนครั้งในการเปรียบเทียบจากภาพที่ 4.14 แสดงให้เห็นว่า การใช้ค่า median จะใช้จำนวนครั้งในการเปรียบเทียบน้อยกว่าประเภทอื่น เพราะค่า median จะเลือกค่ากลางที่เหมาะสมของแต่ละ พาร์ทิชัน โดยนำเอากลุ่มตัวอย่างมาประมวลผล ทำให้สามารถหลีกเลี่ยงค่ากลางที่เป็นตัวซ้ำ หรือค่ากลางที่เป็นกรณีที่แย่ที่สุด ซึ่งจากผลลัพธ์ในภาพที่ 4.11 และ 4.14 ชี้ให้เห็นว่าการเลือกค่าไพวอท โดยใช้ค่า median กับข้อมูลที่มีโอกาสซ้ำมากๆ จะยิ่งลดเวลาในการเปรียบเทียบค่าไพวอทกับข้อมูล และผลลัพธ์จำนวนครั้งในการสลับที่ในภาพที่ 4.15 มีจำนวนค่าเฉลี่ยที่ใกล้เคียงกัน เพราะข้อมูลเป็นการการกระจายแบบสุ่ม และมีเลขซ้ำ ทำให้ต้องมีการสลับที่เกือบทุกตำแหน่งของข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



	0	2000	5000	7000
first	129459	121218	102990.4	107522.8
last	130532.6	120652.4	103402	101420.4
mid	105284.2	98166.6	93892.2	88024.4
random	109968	105970	89711.6	87500.6
med3	88800.4	84000.4	76528.4	72198.6
med5	100482.4	93587.6	86247	78391.4

ภาพที่ 4.14 จำนวนครั้งในการเปรียบเทียบของข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 20 ตัว



	0	2000	5000	7000
first	33941	35252.2	37507	38454.8
last	33987.6	35443.4	37477	38405.2
mid	33594.4	35019	37176.6	38481.4
random	33765.2	35195.8	37579.2	38395.8
med3	34566.2	35900.4	38058.6	39124.2
med5	34280.4	35788	37823.4	38948

ภาพที่ 4.15 จำนวนครั้งในการสลับที่ของข้อมูลข้อมูลสุ่มโดยมีข้อมูลซ้ำหมายเลขละ 20 ตัว

เอกสารนี้เป็น ... ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 สรุปผลการทดลอง

จากผลการทดลองที่ได้กล่าวในข้างต้นสามารถสรุปว่า ถ้าใช้เวลาในการประมวลเป็นเกณฑ์วัดประสิทธิภาพ การเลือกค่าไพวอทโดยใช้ค่ากึ่งกลางจะให้เวลาในการประมวลผลเฉลี่ยกับทุกประเภทข้อมูลทดลองน้อยที่สุด เพราะว่าการเลือกค่าไพวอทโดยใช้ค่ากึ่งกลางจะหลีกเลี่ยงค่าที่อาจทำให้เกิดกรณีที่แย่ที่สุด และใช้เวลาในการหาค่าไพวอทน้อย เมื่อเทียบกับการหาค่าไพวอทโดยใช้ค่ามัธยฐาน

แต่ถ้าหากใช้จำนวนครั้งในการเปรียบเทียบเป็นเกณฑ์วัดประสิทธิภาพ การเลือกค่าไพวอทโดยใช้ค่า มัธยฐานจะใช้จำนวนครั้งในการเปรียบเฉลี่ยกับทุกประเภทข้อมูลน้อยที่สุด เพราะการใช้ค่ามัธยฐานจะช่วยให้การหาข้อมูลที่เป็นกึ่งกลางของกลุ่มข้อมูลทำได้อย่างมีประสิทธิภาพ เนื่องจากการใช้ค่ามัธยฐานเป็นวิธีในการเลือกค่าไพวอทจะนำเอาข้อมูลบางส่วนของกลุ่มข้อมูลมาประมวลผลก่อน ทำให้ค่าไพวอทที่ได้เป็นค่าที่ใกล้เคียงกับค่ากลางมากที่สุด และจำนวนครั้งในการสลับที่เฉลี่ยของไพวอทแต่ละชนิดมีค่าใกล้เคียงกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปและข้อเสนอแนะ

งานวิจัยนี้ได้ทำการทดลองหาค่าไพวอทของควิกซอร์ตที่เหมาะสมกับข้อมูลจำนวนเต็มประเภทต่างๆ เช่น ข้อมูลที่มีการเรียงลำดับ ข้อมูลที่เกือบเรียงลำดับ ข้อมูลที่เกือบเรียงลำดับแบบย้อนกลับ ข้อมูลซ้ำ ข้อมูลสุ่ม กับไพวอทชนิดต่างๆ เช่น ค่าเริ่มต้น, ค่าสุดท้าย, ค่ากึ่งกลาง, ค่าสุ่ม และค่ามัธยฐาน เพื่อหาค่าเฉลี่ยของเวลาในการประมวลผล, จำนวนครั้งในการเปรียบเทียบ และจำนวนครั้งในการสลับที่

5.1 สรุป

จากงานวิจัย [3, 4, 5] และหนังสือ [1, 2] ได้เสนอเทคนิคการเลือกค่าไพวอทหลากหลายประเภทเช่น ค่าเริ่มต้น, ค่าสุดท้าย, ค่ากึ่งกลาง, ค่าสุ่ม และค่ามัธยฐาน ผู้วิจัยได้ทำการทดลองขั้นตอนวิธีควิกซอร์ตโดยใช้เทคนิคการเลือกค่าไพวอทในข้างต้นกับข้อมูลจำนวนเต็มประเภทต่าง ๆ จากผลการทดลอง สามารถสรุปได้ดังนี้

5.1.1 จำนวนครั้งในการสลับที่ : ถ้าประเภทข้อมูลอยู่ในรูปเกือบเรียงลำดับ หรือเรียงลำดับอยู่แล้ว การหาค่าไพวอทโดยใช้ค่ากลาง หรือค่ามัธยฐาน จะใช้จำนวนครั้งในการสลับที่น้อยกว่าชนิดอื่นๆ แต่ถ้าข้อมูลอยู่ในรูปกระจาย หรือสุ่มแล้ว จำนวนครั้งเฉลี่ยในการสลับที่ของค่าไพวอทแต่ละชนิดจะไม่แตกต่างกันมาก

5.1.2 จำนวนครั้งในการเปรียบเทียบ : จำนวนครั้งในการเปรียบเทียบโดยใช้ค่ามัธยฐานเป็นไพวอทนั้นจะได้จำนวนครั้งเฉลี่ยที่น้อยกว่าการเลือกค่าไพวอทชนิดอื่นๆ ในทุกประเภทข้อมูลเพราะว่าการใช้ค่ามัธยฐานจะทำการหาค่ากลางที่มีประสิทธิภาพโดยนำเอากลุ่มตัวอย่างของข้อมูลในแต่ละพาร์ทิชันมาประมวลผลแล้วจึงเลือกค่ากลางเป็นค่าไพวอท ทำให้ค่าไพวอทที่เป็นค่ากลางที่ได้จากการมองภาพรวมของชุดข้อมูลแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.1.3 เวลาในการประมวลผล : เวลาในการประมวลผลโดยใช้ค่าไพวอทเป็นค่ากึ่งกลางจะใช้เวลาในการประมวลผลน้อยกว่าการเลือกค่าไพวอทชนิดอื่นๆ กับข้อมูลที่มีการสุ่ม หรือมีรูปแบบการเรียงลำดับแล้ว เพราะการเลือกค่ากึ่งกลางมีโอกาสหลีกเลี่ยงการได้ค่าไพวอทที่เป็นกรณีที่แย่ที่สุดในข้อมูลที่มีการเรียงลำดับ และไม่เสียเวลาในการประมวลผลเพื่อหาค่าไพวอทในข้อมูลแบบสุ่ม แต่สำหรับข้อมูลที่มีการสุ่มแล้วให้มีหมายเลขซ้ำกันได้นั้นการใช้ค่าไพวอทชนิดค่าเริ่มต้น หรือค่าสุดท้ายนั้นจะใช้เวลาในการประมวลผลน้อยกว่าค่ากึ่งกลาง หรือค่าอื่นๆ เพราะในการเลือกค่าไพวอทกึ่งกลางกับข้อมูลประเภทซ้ำกันได้นั้น มีโอกาสเลือกค่าที่ซ้ำกันได้มากกว่าค่าเริ่มต้นหรือค่าสุดท้าย ทำให้ต้องเสียเวลาในการประมวลผลเพื่อหาค่าไพวอทใหม่ โดยขึ้นอยู่กับอัตราการซ้ำของข้อมูล

5.2 ข้อเสนอแนะ

ในการวิจัยครั้งนี้ได้ทำการทดลองกับข้อมูลที่เป็นประเภทจำนวนเต็มขนาดไม่เกิน 10000 ตัวเพียงอย่างเดียว ผู้สนใจสามารถประยุกต์นำเอางานวิจัยไปใช้กับข้อมูลที่เป็นประเภทสายอักขระ (String) หรือข้อมูลจำนวนเต็มที่มีขนาดใหญ่ขึ้น เพื่อศึกษาและประเมินประสิทธิภาพการเลือกค่าไพวอทของควิกซอร์ตที่เหมาะสมกับข้อมูลประเภทสายอักขระ หรือข้อมูลจำนวนเต็มขนาดใหญ่ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] Cormen T., Leiserson C., Rivest R. and Stein C. Introduction to Algorithms. 2nd ed. The MIT Press, 2001.

[2] Levitin A. Introduction to The Design & Analysis of Algorithms. 3rd ed. Pearson Education, Inc., 2012

[3] Hoare, C. A. R. (1962). "Algorithm 65: Find". Comm. ACM 4 (7): 321–322.

[4] R.S. Scowen, "Algorithm 271: Quickersort," Comm. ACM 8, 11, pp 669-670, Nov. 1965.

[5] Mohammed A.; and Othman M., 2007. Comparative Analysis of Some Pivot Selection Schemes for Quicksort Algorithm. Information Technology Journal 6 (3): 424 - 427.

[6] David R. Martin. **"Quick Sort"** [Online]. Available: http://www.sorting-algorithms.com/quick-sort

[7] LaFata J. **"Sorting Algorithm Shootout"** [Online]. Available: http://blog.quibb.org/2009/10/sorting-algorithm-shootout/

[8] Urbaniak G. **"Research Randomizer"** [Online]. Available :http://www.randomizer.org/

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้