

การปรับปรุงการจัดกาารแก้ไขข้อมูลแบบอ้ก



นายสุทธิพงษ์ มีสมสาร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
ภาควิชาวิศวกรรมคอมพิวเตอร์
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
พ.ศ. 2528

ISBN 974-561-411-2

009142

117957916

AN IMPROVEMENT OF DATA COMPRESSION FILE MANAGEMENT

MR.SUTTIPONG MEESOMSARN

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science

Department of Computer Engineering

Graduate School

Chulalongkorn University

1985

หัวข้อวิทยานิพนธ์ การปรับปรุงการจัดการแฟ้มข้อมูลแบบอัตโนมัติ
 โดย นายสุทธิพงษ์ มีสมสาร
 ภาควิชา วิศวกรรมคอมพิวเตอร์
 อาจารย์ที่ปรึกษา รองศาสตราจารย์ สมชาย ทยานยง
 ผู้ช่วยศาสตราจารย์ เมธี ศรีสังวาล



บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยนี้เป็น
 ส่วนหนึ่งของการศึกษาคำถามหลักสูตรปริญญาโทบัณฑิต

Signature of Dean

.....คณบดีบัณฑิตวิทยาลัย
 (รองศาสตราจารย์ ดร.สุประคิษฐ์ นูนนาค)

คณะกรรมการสอบวิทยานิพนธ์

Signature of Chairperson

.....ประธานกรรมการ
 (ผู้ช่วยศาสตราจารย์ สุเมธ วัชรระชัยสุรพล)

Signature of Member

.....กรรมการ
 (รองศาสตราจารย์ สมชาย ทยานยง)

Signature of Member

.....กรรมการ
 (รองศาสตราจารย์ ไกรวิจิต ตันศิริเมธ)

Signature of Member

.....กรรมการ
 (ผู้ช่วยศาสตราจารย์ เมธี ศรีสังวาล)

อธิบดีของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

หัวข้อวิทยานิพนธ์

การปรับปรุงการจัดการแฟ้มข้อมูลแบบอ็ค

ชื่อนิสิต

นายสุทธิพงษ์ มีสมสาร

อาจารย์ที่ปรึกษา

รองศาสตราจารย์ สมชาย ทยานยง

อาจารย์ที่ปรึกษาร่วม

ผู้ช่วยศาสตราจารย์ เมธี ศรีสังวาล

ภาควิชา

วิศวกรรมคอมพิวเตอร์

ปีการศึกษา

2527



บทคัดย่อ

ข้อมูลที่บันทึกอยู่บนจานแม่เหล็กในระบบคอมพิวเตอร์ทั่ว ๆ ไปส่วนมากจะมีความซ้ำซ้อนกัน
กลไกหรือขั้นตอนที่ใช้ในการแปลงรหัสข้อมูล เพื่อลดการซ้ำซ้อนอาจเพิ่มประสิทธิภาพความจุของข้อมูลที่
ถูกบันทึกได้ถึง 2 หรือ 3 เท่า อย่างไรก็ตามถ้าการอ็คข้อมูลเป็นไปอย่างอัตโนมัติจะทำให้ต้นทุนในการ
พัฒนาทางด้านการแสวงหาค้นหาสูงขึ้น กลไกการอ็คข้อมูลที่มองไม่เห็นนี้สามารถใช้ได้กับโครงสร้างของข้อมูล
ทุกชนิด ซึ่งเนื้อหาของข้อมูลส่วนที่ยังไม่ได้ถูกใช้งานหรือสำรองไว้ไม่ควรจะใช้เนื้อที่ในการบันทึกข้อมูล
มากจนเกินไป

การจิกแฟ้มข้อมูลแบบอ็คของเครื่องคาคาพอยต์เป็นเทคนิคในการจิกแฟ้มข้อมูลแบบหนึ่ง
เพื่อลดความซ้ำซ้อนของข้อมูลแบบคาแรคเตอร์ รหัสซัน หรือ ซีรีย โดยทำการอ็คเฉพาะข้อมูลที่เป็นที่ว่าง
เท่านั้น ซึ่งบริษัทคาคาพอยต์ได้เรียกเทคนิคนี้ว่า สเปซ คอมเพรสชัน ดังนั้นเมื่อมีข้อมูลที่เป็นที่ว่างตั้งแต่
3 บิตขึ้นไป จะถูกแทนด้วยตัวอักษรที่ใช้ในการควบคุม ซี เอ็ม ที และตามด้วยไบท์ที่ใช้มีข้อมูลที่
ที่ว่างที่ถูกทำการอ็คแต่ละคู่ของ ซี เอ็ม ที และไบท์ที่ใช้มีที่ว่างสามารถทำการอ็คข้อมูลที่
มากที่สุดถึง 255 บิต

วัตถุประสงค์ในการทำวิทยานิพนธ์นี้ เพื่อศึกษาการทำงานของเครื่องคาคาพอยต์ในการ
อ็คข้อมูลที่อยู่ที่ว่าง และปรับปรุงวิธีการอ็คข้อมูลเสียใหม่โดยทำให้สามารถอ็คข้อมูลทุกบิตที่ซ้ำ ๆ กันได้

รวมทั้งข้อมูลที่เป็นที่ว่างก๊วย ขั้นตอนการทำงานต่าง ๆ ได้ถูกนำมาอธิบายในรูปของผังงาน เพื่อช่วยให้
เข้าใจได้ดียิ่งขึ้น โปรแกรมที่ใช้ในการทำวิทยานิพนธ์นี้เขียนขึ้นด้วยภาษาเอสเซมบลี และภาษาคาคาบัค
ของเครื่องคาคาพอยท์ และคอมไพเลอร์เอสเซมบลี ของเครื่อง ไอ บี เอ็ม

Thesis Title	An Improvement of Data Compression File Management
Name	Mr. Suttipong Meesomsarn
Thesis Advisors	Associated Professor Somchai Thayarnyong Assistance Professor Matee Srisangwan
Department	Computer Engineering
Academic Year	1984

ABSTRACT

Data stored on disks in commercial computer systems generally contains significant redundancy. A mechanism or procedure which recodes the data to lessen the redundancy could possibly double or triple the effective data densities in stored data. Moreover, if compression is automatic, it can also aid in the rise of software development costs. A transparent compression mechanism could permit the use of "sloppy" data structures, in that empty space or sparse encoding of data would not greatly expand the use of storage; however, that requires a good compression procedure.

A Datapoint's Compression file technique is developed to lessen the redundancy of character repetition or character series, but only blank is considered, and this technique is called "Space Compression". Any series of three or more blanks within the data is replaced by an CMP (Space Compress) character followed by a blank - count byte, the blank -

count byte indicates the number of blanks removed (compressed). Up to 255 blanks can be removed by one CMP character and one blank - count byte. If more than 255 bytes are removed, two or more CMP and blank - count byte pairs are inserted.

The objective of this thesis is to study the Datapoint's space compression technique and improve this technique to compress not only blank but also any characters. The algorithms of compression technique are described by flowcharts and listing of disassembled source programs to more understand. Programs used in this thesis are written in Datapoint's Assembly and Databus languages and IBM'S DOS/VS or DOS/VSE Assembly language.



กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สำเร็จได้ด้วยความช่วยเหลือจากรองศาสตราจารย์ สมชาย ทยานมยง ผู้อำนวยการสถาบันบริการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย และผู้ช่วยศาสตราจารย์ เมธี ศรีสังวาล อาจารย์ผู้ควบคุมการทำวิทยานิพนธ์ ที่ได้กรุณาแนะนำหัวข้อวิทยานิพนธ์แนะแนวทางการทำวิจัย การเขียนและการตรวจสอบแก้ไขจนสำเร็จเป็นเล่ม ผู้เขียนขอกราบขอบพระคุณท่านอาจารย์เป็นอย่างสูง

นอกจากนี้ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ และคุณพูลสุข มีสมสาร ที่ได้กรุณาจัดพิมพ์วิทยานิพนธ์จนเสร็จ เป็นรูปเล่มที่สมบูรณ์



	หน้า
มหัศจรรย์ภาษาไทย.....	ง
มหัศจรรย์ภาษาอังกฤษ.....	ฉ
กิจกรรมประกาศ.....	ช
รายการรูปประกอบ.....	ฉ
บทที่	
1. บทนำ.....	1
1.1 ความเป็นมาของปัญหา.....	1
1.2 วัตถุประสงค์และขอบเขตการวิจัย.....	3
1.3 ขั้นตอนและวิธีดำเนินการวิจัย.....	4
1.4 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย.....	4
2. ส่วนประกอบและลักษณะทั่วไปของเครื่องคานาคอชท์.....	5
2.1 ส่วนประกอบและลักษณะทางทฤษฎี.....	5
2.2 ระบบโครงสร้าง.....	15
2.3 ลักษณะการบันทึกข้อมูลของบนจานแม่เหล็ก.....	68
2.4 ลักษณะแฟ้มข้อมูลแบบอื่น.....	71
3. โปรแกรมที่เกี่ยวข้องกับแฟ้มข้อมูลแบบอื่น.....	73
3.1 โปรแกรมที่ใช้ในการสร้างแฟ้มข้อมูลแบบอื่น.....	73
3.2 โปรแกรมที่ใช้ในการแสดงเนื้อหาของแฟ้มข้อมูลแบบอื่น.....	94
3.3 โปรแกรมที่ใช้ในการจัดรูปแบบแฟ้มข้อมูลแบบอื่น.....	104
3.4 โปรแกรมที่ใช้ในการถ่ายเทข้อมูลระหว่างเครื่องจานแม่เหล็กและเครื่องเทปแม่เหล็ก.....	112
4. การสร้างโปรแกรมที่เกี่ยวข้องกับแฟ้มข้อมูลแบบอื่นที่พัฒนาขึ้นใหม่.....	151
4.1 ขั้นตอนการสร้างโปรแกรมที่เกี่ยวข้องกับแฟ้มข้อมูลแบบอื่น.....	151
4.2 การสร้างโปรแกรมที่เกี่ยวข้องกับแฟ้มข้อมูลแบบอื่นที่พัฒนาขึ้นใหม่.....	162

	หน้า
4.3 การศึกษากังโปรแกรมที่สร้างขึ้น.....	234
4.4 วิธีการเรียกใช้โปรแกรมที่สร้างขึ้น.....	235
5. สรุปผลการวิจัยและเสนอแนะ.....	236
5.1 การทดสอบโปรแกรมที่สร้างขึ้น.....	236
5.2 สรุปผลการวิจัย.....	236
5.3 ข้อเสนอแนะ.....	236
เอกสารอ้างอิง.....	237
ภาคผนวก.....	236
ประวัติ.....	323

รายการรูปประกอบ

รูปที่	หน้า
2.1	6
2.2	8
2.3	9
2.4	11
2.5	12
2.6	14
2.7	17
2.8	19
2.9	15
2.10	20
2.11	21
2.12	22
2.13	23
2.14	25
2.15	28
2.16	30
2.17	32
2.18	33
2.19	35
2.20	40
2.21	44
2.22	46
2.23	50

	หน้า
2.24 แสดงส่วนของแคสเซต ไครเวอ์ รุทึน.....	52
2.25 แสดงขั้นตอนการทำงานของคอมมานด์ อินเทอร์พรีเตอร์.....	58
2.26 แสดงส่วนของคอมมานด์ อินเทอร์พรีเตอร์.....	59
3.1 แสดงรายละเอียดของแฟ้มข้อมูลที่ใช้เป็น Configuration file.....	77
3.2 แสดงขั้นตอนการทำงานของโปรแกรม EDIT/CMD.....	78
3.3 แสดงขั้นตอนการทำงานของโปรแกรม LIST/CMD.....	96
3.4 แสดงขั้นตอนการทำงานของโปรแกรม REFORMAT/CMD.....	107
3.5 แสดง MESSAGES จากโปรแกรม TAPE/CMD.....	113
3.6 แสดงขั้นตอนการทำงานของโปรแกรม TAPE/CMD.....	114
4.1 แสดงลักษณะการพิมพ์แบบ OCTAL และ CHARACTER.....	153
4.2 แสดงลักษณะการแบ่งโปรแกรม.....	155
4.3 แสดงตัวอย่างลักษณะของแฟ้มข้อมูลเครื่องหมาย.....	156
4.4 แสดงตัวอย่างโปรแกรมที่ได้จากการ DISASSEMBLY.....	158
4.5 แสดงขั้นตอนการทำงานส่วน Expansion ของโปรแกรม EDIT/CMD....	163
4.6 แสดงคำสั่งการทำงานส่วน Expansion ของโปรแกรม EDIT/CMD....	164
4.7 แสดงขั้นตอนการทำงานส่วน Compression ของโปรแกรม EDIT/CMD....	166
4.8 แสดงคำสั่งการทำงานส่วน Compression ของโปรแกรม EDIT/CMD....	168
4.9 แสดงขั้นตอนการทำงานส่วน Expansion ของโปรแกรม XEDIT/CMD... 171	171
4.10 แสดงคำสั่งการทำงานส่วน Expansion ของโปรแกรม XEDIT/CMD... 172	172
4.11 แสดงขั้นตอนการทำงานส่วน Compression ของโปรแกรม XEDIT/CMD... 175	175
4.12 แสดงคำสั่งการทำงานส่วน Compression ของโปรแกรม XEDIT/CMD... 177	177
4.13 แสดงขั้นตอนการทำงานส่วน Expansion ของโปรแกรม LIST/CMD.... 182	182
4.14 แสดงคำสั่งการทำงานส่วน Expansion ของโปรแกรม LIST/CMD.... 184	184
4.15 แสดงขั้นตอนการทำงานส่วน Expansion ของโปรแกรม XLIST/CMD... 187	187

รูปที่

หน้า

4.16 แสดงคำสั่งการทำงานส่วน Expansion ของโปรแกรม XLIST/CMD.... 189

4.17 แสดงขั้นตอนการทำงานส่วน Expansion ของโปรแกรม REFORMAT/CMD. 194

4.18 แสดงคำสั่งการทำงานส่วน Expansion ของโปรแกรม REFORMAT/CMD. 195

4.19 แสดงขั้นตอนการทำงานส่วน Compression ของโปรแกรม REFORMAT/CMD. 196

4.20 แสดงคำสั่งการทำงานส่วน Compression ของโปรแกรม REFORMAT/CMD. 199

4.21 แสดงขั้นตอนการทำงานส่วน Expansion ของโปรแกรม XREFO/CMD... 201

4.22 แสดงคำสั่งการทำงานส่วน Expansion ของโปรแกรม XREFO/CMD... 202

4.23 แสดงขั้นตอนการทำงานส่วน Expansion ของโปรแกรม TAPE/CMD.... 205

4.24 แสดงคำสั่งการทำงานส่วน Expansion ของโปรแกรม TAPE/CMD.... 207

4.25 แสดงขั้นตอนการทำงานส่วน Compression ของโปรแกรม TAPE/CMD.... 208

4.26 แสดงคำสั่งการทำงานส่วน Compression ของโปรแกรม TAPE/CMD.... 210

4.27 แสดงคำสั่งการทำงานเพื่อตรวจสอบ Compression option..... 216

4.28 แสดงขั้นตอนการทำงานส่วน Expansion ของโปรแกรม XTAPE/CMD... 219

4.29 แสดงคำสั่งการทำงานส่วน Expansion ของโปรแกรม XTAPE/CMD... 222

4.30 แสดงขั้นตอนการทำงานส่วน Compression ของโปรแกรม XTAPE/CMD... 224

4.31 แสดงคำสั่งการทำงานส่วน Compression ของโปรแกรม XTAPE/CMD... 226

4.32 แสดงขั้นตอนการติดตั้งโปรแกรม XEDIT/CMD..... 234

4.33 แสดงขั้นตอนการติดตั้งโปรแกรม XLIST/CMD..... 234

4.34 แสดงขั้นตอนการติดตั้งโปรแกรม XREFO/CMD..... 235

4.35 แสดงขั้นตอนการติดตั้งโปรแกรม XTAPE/CMD..... 235

ก.1 แสดงรายละเอียดของแฟ้มข้อมูล DOS/EPT..... 239

ข.1 แสดงรายละเอียดของโปรแกรม INITCORE..... 244

ข.2 แสดงรายละเอียดของโปรแกรม DUMP CORE..... 245

ข.3 แสดงรายละเอียดของโปรแกรม EDITMARK..... 247

รูปที่

หน้า

ค.1 แสดงรายละเอียดของโปรแกรม DUMP/CONV..... 249

ค.2 แสดงรายละเอียดของโปรแกรม PASM5500..... 256

ค.3 แสดงรายละเอียดของโปรแกรม DASM5500..... 263

ง.1 แสดงข้อมูลที่ใช้ในการทดสอบโปรแกรม..... 278

ง.2 แสดงรายละเอียดของแฟ้มข้อมูลที่ไต่จากโปรแกรม EDIT/CMD..... 289

ง.3 แสดงรายละเอียดของแฟ้มข้อมูลที่ไต่จากโปรแกรม XEDIT/CMD..... 294

ง.4 แสดงรายละเอียดของแฟ้มข้อมูลที่ไต่จากโปรแกรม REFORMAT/CMD..... 298

ง.5 แสดงรายละเอียดของแฟ้มข้อมูลที่ไต่จากโปรแกรม XREFO/CMD..... 303

ง.6 แสดงรายละเอียดของแฟ้มข้อมูลที่ไต่จากโปรแกรม TAPE/CMD..... 308

ง.7 แสดงรายละเอียดของแฟ้มข้อมูลที่ไต่จากโปรแกรม TAPE/CMD
 ในกรณี Compressed..... 313

ง.8 แสดงรายละเอียดของแฟ้มข้อมูลที่ไต่จากโปรแกรม TAPE/CMD
 ในกรณี Non-compressed..... 318