



ใบรับรองวิทยานิพนธ์
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมไฟฟ้า)

ปริญญา

วิศวกรรมไฟฟ้า

วิศวกรรมไฟฟ้า

สาขา

ภาควิชา

เรื่อง วิธีการจัดกลุ่มใหม่ (วิธีความน่าจะเป็นฟัซซี่ซิมินสองขั้นตอน) สำหรับจัดกลุ่มโพรไฟล์
ภาระของความต้องการการใช้ไฟฟ้า

A Novel Clustering Methods (Two-stage Fuzzy-Possibilistic C-Means) for Electricity-
Demand Load Profile Clustering

นามผู้วิจัย นายกรณ์ย์ ตันทวิวงศ์

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(อาจารย์พิสุทธิ์ รพีศักดิ์, Ph.D.)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(ผู้ช่วยศาสตราจารย์ดุลย์พิเชษฐ์ ฤกษ์ปรีดาพงศ์, Ph.D.)

หัวหน้าภาควิชา

(รองศาสตราจารย์วิชัย สุระพัฒน์, วศ.ม.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญญา ธีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ เดือน พ.ศ.

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วิทยานิพนธ์

เรื่อง

วิธีการจัดกลุ่มใหม่ (วิธีความน่าจะเป็นฟัซซีที่มีนสองขั้นตอน) สำหรับ
จัดกลุ่มโพรไฟล์ภาระของความต้องการการใช้ไฟฟ้า

A Novel Clustering Methods (Two-stage Fuzzy-Possibilistic C-
Means) for Electricity- Demand Load Profile Clustering

โดย

นายกรัณย์ ตันทวีวงศ์

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์
เพื่อความสมบูรณ์แห่งปริญญาวิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมไฟฟ้า)

พ.ศ. 2556

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

กรณี ดันทวิวงศ์ 2556: วิธีการจัดกลุ่มใหม่ (วิธีความน่าจะเป็นฟิชชันซิมินสองขั้นตอน)
สำหรับจัดกลุ่มโพรไฟล์ภาระของความต้องการการใช้ไฟฟ้า ปรินญาวิศวกรรมศาสตร
มหาบัณฑิต (วิศวกรรมไฟฟ้า) สาขาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า อาจารย์ที่
ปรึกษาวิทยานิพนธ์หลัก: อาจารย์พิสุทธิ รัชศักดิ์, Ph.D. 52 หน้า

วิทยานิพนธ์นี้แนะนำเสนอวิธีการจัดกลุ่มใหม่ (วิธีความน่าจะเป็นฟิชชันซิมินสองขั้นตอน)
สำหรับจัดกลุ่มโพรไฟล์ภาระของความต้องการการใช้ไฟฟ้า โดยข้อมูลที่นำมาศึกษานั้นเป็นข้อมูล
โพรไฟล์ภาระของผู้ใช้ไฟฟ้าประเภทกลุ่มอุตสาหกรรมหนึ่งจำนวน 11 บริษัท ข้อมูลย้อนหลัง 1 ปี
ซึ่งผลที่ได้จากการจัดกลุ่มโดยวิธีความน่าจะเป็นฟิชชันซิมินสองขั้นตอนจะถูกนำมาเปรียบเทียบกับ
วิธีต่างๆ ดังนี้ วิธีฟิชชันซิมิน สามารถจัดกลุ่มโพรไฟล์ภาระได้ 3 กลุ่ม วิธีความน่าจะเป็นฟิชชันซิมิน
สามารถจัดกลุ่มโพรไฟล์ภาระได้ 3 กลุ่ม วิธีฟิชชันซิมินสองขั้นตอน สามารถจัดกลุ่มโพรไฟล์ภาระ
ได้ 12 กลุ่ม และวิธีความน่าจะเป็นฟิชชันซิมินสองขั้นตอน สามารถจัดกลุ่มโพรไฟล์ภาระได้ 9 กลุ่ม
ซึ่งผลจากการจัดกลุ่มโพรไฟล์ภาระโดยใช้วิธีที่นำเสนอ คือ วิธีความน่าจะเป็นฟิชชันซิมินสอง
ขั้นตอน จะแสดงถึงลักษณะการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าแต่ละรายได้อย่างกระชับและครอบคลุม

ลายมือชื่อนิติ

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

Karun Tanthawewong 2013: A Novel Clustering Methods (Two-stage Fuzzy-Possibilistic C-Means) for Electricity-Demand Load Profile Clustering. Master of Engineering (Electrical Engineering), Major Field: Electrical Engineering, Department of Electrical Engineering. Thesis Advisor: Mr. Pisut Raphisak, Ph.D. 52 pages.

This thesis proposed a novel clustering methods (two-stage fuzzy-possibilistic C-means) for electricity-demand load profile clustering. The utilizing datasets are an annual load profile of 11 industrial customers. The results are compared with various method i.e. fuzzy C-means, fuzzy-possibilistic C-means and two-stage fuzzy C-means. The results are 3 clusters for fuzzy C-means, 3 clusters for fuzzy-possibilistic C-means, 12 clusters for two-stage fuzzy C-means and 9 clusters for two-stage fuzzy-possibilistic C-means. The result from two-stage fuzzy-possibilistic C-means is cover and compact for whole load profile.

Student's signature

Thesis Advisor's signature

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ อ.พิสุทธิ รัชศักดิ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ผศ.ดร.ดุลย์ พิเศษฐ์ ฤกษ์ปรีดาพงศ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วมที่กรุณาให้คำปรึกษาในการค้นคว้าวิจัย ตลอดจนการตรวจแก้ไขวิทยานิพนธ์จนกระทั่งเสร็จสมบูรณ์ ขอขอบพระคุณ ผศ.ดร.ประมุข อุณหเลขกะ ผู้ทรงคุณวุฒิภายนอก ที่กรุณาช่วยแนะนำและแก้ไขวิทยานิพนธ์เล่มนี้ให้มีความสมบูรณ์มากยิ่งขึ้น และ กราบขอบพระคุณ ผศ.ดร.ศิริโรจน์ ศิริสุขประเสริฐ ประธานการสอบปากเปล่าขั้นสุดท้าย ที่กรุณาให้คำแนะนำและตรวจแก้ไขวิทยานิพนธ์ให้สมบูรณ์ยิ่งขึ้นขอขอบพระคุณคณาจารย์ทุกท่านในภาควิชาวิศวกรรมไฟฟ้า ที่ได้กรุณาอบรมสั่งสอน ในแขนงวิชาต่างๆ

ขอขอบพระคุณคุณพ่อ คุณแม่ พี่ชาย และญาติๆ ที่ให้กำลังใจดีๆ และให้ข้อคิดดีๆ แก่ผู้วิจัยเสมอมา

ขอขอบพระคุณพี่ๆ และน้องๆ นักวิจัย โครงการวิจัยและพัฒนาความชำนาญด้านไฟฟ้ากำลัง มหาวิทยาลัยเกษตรศาสตร์ พี่ๆ เพื่อนๆ ที่เข้าศึกษาในรุ่นเดียวกัน ที่คอยช่วยเหลือและให้กำลังใจในการทำวิทยานิพนธ์เสมอมา

กรณ์ย์ ตันทวิวงศ์
เมษายน 2556

สารบัญ

หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(4)
คำนำ	1
วัตถุประสงค์	2
การตรวจเอกสาร	3
อุปกรณ์และวิธีการ	18
อุปกรณ์	18
วิธีการ	18
ผลและวิจารณ์	28
ผล	28
วิจารณ์	48
สรุปและข้อเสนอแนะ	49
สรุป	49
ข้อเสนอแนะ	49
เอกสารและสิ่งอ้างอิง	50
ประวัติการศึกษา และการทำงาน	52

สารบัญญัตินำ

ตารางที่		หน้า
1	ค่าดัชนี Partition Coefficient เมื่อทำการจัดกลุ่มโดยวิธี FCM	30
2	ค่าดัชนี Classification Entropy เมื่อทำการจัดกลุ่มโดยวิธี FCM	30
3	ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี FCM	31
4	จำนวนสมาชิกของกลุ่มจากการจัดกลุ่มโดยวิธี FCM	32
5	ค่าดัชนี Partition Coefficient เมื่อทำการจัดกลุ่มโดยวิธี FPCM	32
6	ค่าดัชนี Classification Entropy เมื่อทำการจัดกลุ่มโดยวิธี FPCM	33
7	ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี FPCM	34
8	จำนวนสมาชิกของกลุ่มจากการจัดกลุ่มโดยวิธี FPCM	34
9	ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 1	35
10	ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1	36
11	ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2	36
12	ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3	37
13	จำนวนสมาชิกของกลุ่มจากการจัดกลุ่มโดยวิธี Two-stage Fuzzy C-Means	39
14	ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 1	40
15	ดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FuzzyCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1	41
16	ดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2	42

สารบัญตาราง (ต่อ)

ตารางที่	หน้า	
17	ดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3	43
18	ค่า D ที่ใช้ในการทดสอบสมมติฐานของกลุ่มหลักที่ 1	44
19	ค่า D ที่ใช้ในการทดสอบสมมติฐานของกลุ่มหลักที่ 2	45
20	ค่า D ที่ใช้ในการทดสอบสมมติฐานของกลุ่มหลักที่ 3	45
21	จำนวนสมาชิกของกลุ่มจากการจัดกลุ่มโดยวิธี Two-stage FPCM	46
22	ผลการจัดกลุ่มโพรไฟล์ภาวะ	47

สารบัญภาพ

ภาพที่	หน้า
1 ตัวอย่าง Load Profile ของผู้ใช้ไฟฟ้ารายหนึ่งวันที่ 11/10/10	3
2 กระบวนการค้นหาความรู้ในฐานข้อมูล	4
3 หลักการพื้นฐานของการจัดกลุ่ม	6
4 การจัดกลุ่มโดยวิธี Fuzzy C-Means	8
5 การจัดกลุ่มโดยวิธี Fuzzy-Possibilistic C-Means	11
6 การจัดกลุ่มโดยวิธี Two-stage Fuzzy C-Means	14
7 ลำดับขั้นตอนวิธีการจัดกลุ่มโพรไฟล์ภาระ	19
8 ตัวอย่างข้อมูล โพรไฟล์ภาระ	20
9 ตัวอย่างข้อมูล โพรไฟล์ภาระที่มีลักษณะผิดปกติ (มีค่าเป็น 0 ตลอด)	21
10 ตัวอย่างข้อมูล โพรไฟล์ภาระที่มีลักษณะผิดปกติ (ไม่มีค่าในบางช่วงเวลา)	21
11 ตัวอย่างข้อมูล โพรไฟล์ภาระที่ไม่ได้ทำการบรรทัดฐานข้อมูล	23
12 ตัวอย่างข้อมูล โพรไฟล์ภาระที่ทำการบรรทัดฐานข้อมูลแล้ว	24
13 การจัดกลุ่มโดยวิธี Two-stage Fuzzy-Possibilistic C-Means	26
14 ข้อมูลโพรไฟล์ภาระในรอบ 1 ปีของผู้ใช้ไฟฟ้า 11 ราย	28
15 ข้อมูลโพรไฟล์ภาระที่ทำการตัดวันหยุดและวันที่มีข้อมูลผิดปกติออกแล้ว	29
16 ข้อมูลโพรไฟล์ภาระที่ทำการบรรทัดฐานแล้ว	29
17 ค่าดัชนี PC และ CE เมื่อทำการจัดกลุ่มโดยวิธี FCM ที่ค่า C แตกต่างกัน	31
18 Typical Load Profile ของการจัดกลุ่มโดยวิธี FCM	32
19 ค่าดัชนี PC และ CE เมื่อทำการจัดกลุ่มโดยวิธี FPCM ที่ค่า C แตกต่างกัน	33
20 Typical Load Profile ของการจัดกลุ่มโดยวิธี FPCM	34
21 Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 1	35

สารบัญญภาพ (ต่อ)

ภาพที่		หน้า
22	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1	36
23	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2	37
24	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3	38
25	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage Fuzzy C-Means	39
26	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 1	41
27	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1	42
28	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2	43
29	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3	44
30	Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM	46

วิธีการจัดกลุ่มใหม่ (วิธีความน่าจะเป็นฟัซซีซิมบิลิสติกสองขั้นตอน) สำหรับจัดกลุ่มโพรไฟล์ ภาระของความต้องการการใช้ไฟฟ้า

A Novel Clustering Methods (Two-stage Fuzzy-Possibilistic C-Means) for Electricity-Demand Load Profile Clustering

คำนำ

ปัจจุบันการใช้พลังงานไฟฟ้ามีแนวโน้มเพิ่มสูงขึ้นเรื่อยๆ ภาคการผลิตไฟฟ้าจึงจำเป็นต้องทำการวางแผนทั้งในส่วนของการผลิต หรือการจัดการพลังงานไฟฟ้าที่ผลิตได้ให้ผู้ใช้ไฟฟ้าสามารถนำไปใช้ได้อย่างมีประสิทธิภาพและคุ้มค่าที่สุด ดังนั้นเพื่อให้เกิดผลประโยชน์ต่อทั้งฝ่ายการผลิตไฟฟ้าและฝ่ายผู้ใช้ไฟฟ้านั้น การไฟฟ้าต้องทราบว่าต้นทุนในการผลิตหรือการส่งจ่ายไฟฟ้านั้นเกิดขึ้นจากอะไรบ้าง สามารถปรับปรุงหรือแก้ไขในส่วนใดได้บ้าง และทางฝ่ายผู้ใช้ไฟฟ้าเองก็ต้องทราบว่าจะสามารถใช้พลังงานไฟฟ้าอย่างไรให้เกิดประสิทธิภาพสูงสุด ซึ่งจะส่งผลต่อรายจ่ายค่าไฟฟ้าที่ลดลงตามไปด้วย

วิทยานิพนธ์นี้จึงได้ทำการศึกษาวิธีการจัดกลุ่มโพรไฟล์ภาระ (Load Profile) ของผู้ใช้ไฟฟ้า เนื่องจากว่าหากเราสามารถจัดกลุ่มได้ว่าผู้ใช้ไฟฟ้ามีลักษณะการใช้ไฟฟ้าอย่างไรบ้าง ทางกรไฟฟ้าสามารถที่จะวางแผนได้ว่าจะทำการผลิตและส่งจ่ายไฟฟ้าอย่างไร ส่วนทางผู้ใช้ไฟฟ้าเองก็ได้ทราบถึงลักษณะการใช้ไฟฟ้าของตนเองว่ามีการใช้ไฟฟ้าอย่างไร สามารถปรับปรุงแก้ไขลักษณะการใช้ไฟฟ้าได้อย่างไรให้เกิดประสิทธิภาพสูงสุด อีกทั้งหากมีการบริหารจัดการการใช้ไฟฟ้าได้ดีแล้ว การไฟฟ้าก็สามารถที่จะชะลอหรือลดการลงทุนในส่วนของการสร้างโรงจักรไฟฟ้าและการสร้างระบบส่งจ่ายไฟฟ้าเพิ่มเพื่อรองรับการใช้ไฟฟ้าที่สูงในบางช่วงเวลา ซึ่งส่งผลโดยตรงทำให้ค่าไฟฟ้ามีราคาถูกเนื่องจากการไฟฟ้าไม่ต้องลงทุนเพิ่มในส่วนดังกล่าว

วิทยานิพนธ์นี้ใช้วิธีการขุดค้นข้อมูล (Data Mining) ในการกลั่นกรองรูปแบบของข้อมูล (Data Pattern) จากข้อมูลดิบ และใช้เทคนิควิธีการจัดกลุ่ม (Clustering Analysis) มาทำการจัดกลุ่มโพรไฟล์ภาระของผู้ใช้ไฟฟ้า หลังจากทำการจัดกลุ่มโพรไฟล์ภาระแล้วทำให้ทราบว่าผู้ใช้ไฟฟ้ามีลักษณะการใช้ไฟฟ้าเป็นอย่างไรบ้าง ซึ่งวิทยานิพนธ์นี้ได้เลือกใช้ข้อมูลการใช้ไฟฟ้าย้อนหลัง 1 ปีของผู้ใช้ไฟฟ้าประเภทโรงงานอุตสาหกรรมจำนวน 11 ราย

วัตถุประสงค์

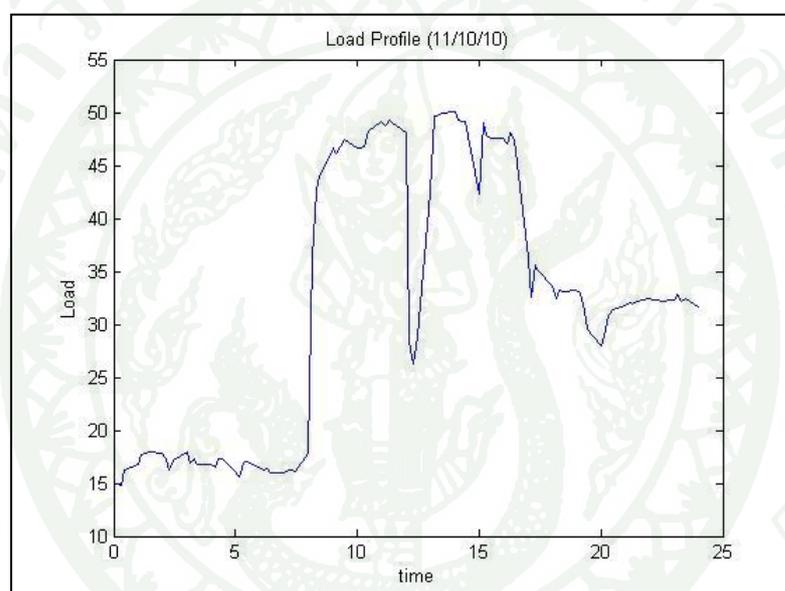
1. ศึกษาวิธีการจัดกลุ่มโพรไฟล์ภาระ (Load Profile Clustering) โดยใช้วิธีการต่างๆ เพื่อให้เหมาะสมกับข้อมูลโพรไฟล์ภาระของผู้ใช้ไฟฟ้าให้มากที่สุด
2. นำเสนอวิธีการจัดกลุ่มใหม่ โดยใช้วิธีความน่าจะเป็นฟัซซีที่มีนสองขั้นตอน (Two-stage Fuzzy-Possibilistic C-Means)



การตรวจเอกสาร

โพรไฟล์ภาระ

ชำนาญ (2548) โพรไฟล์ภาระ (Load Profile) คือ ลักษณะหรือพฤติกรรมการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าแต่ละประเภทที่มีค่าความต้องการไฟฟ้าทุกๆ ช่วงเวลาที่กำหนด อาจเป็น 1 ชั่วโมง 30 นาที หรือ 15 นาที เป็นต้น ดังแสดงในภาพที่ 1



ภาพที่ 1 ตัวอย่าง Load Profile ของผู้ใช้อิฟฟ้ารายหนึ่งวันที่ 11/10/10

การผลิตกำลังไฟฟ้าในระบบไฟฟ้ากำลังนั้น ถูกกำหนดโดยผู้ใช้อิฟฟ้าซึ่งถือว่าเป็นโหลดของระบบ แต่การใช้อิฟฟ้านั้นมิได้คงที่อยู่ทุกขณะ ปริมาณการใช้อิฟฟ้าสูงบ้างต่ำบ้างในช่วงเวลาหนึ่งๆ ซึ่งเป็นปัญหาสำคัญในการจัดระบบผลิต ระบบส่งและระบบจ่ายกำลังไฟฟ้าให้เหมาะสม เพราะสิ่งสำคัญคือจะต้องพยายามรักษาระดับแรงดันในการส่งจ่ายไฟฟ้าไปยังผู้ใช้อิฟฟ้าให้อยู่ในระดับที่ไม่สูงหรือต่ำเกินไปจนเกิดความเสียหายต่ออุปกรณ์เครื่องใช้อิฟฟ้าได้ โดยจะต้องมีการวางแผนและจัดโปรแกรมการเดินเครื่องหรือหยุดเดินเครื่องกำเนิดไฟฟ้าแต่ละเครื่องให้เหมาะสมกับการใช้อิฟฟ้าในแต่ละช่วงเวลา หรือถ้ามองในแง่ของเศรษฐศาสตร์ วิธีการดังกล่าวยังช่วยให้เกิดความประหยัด

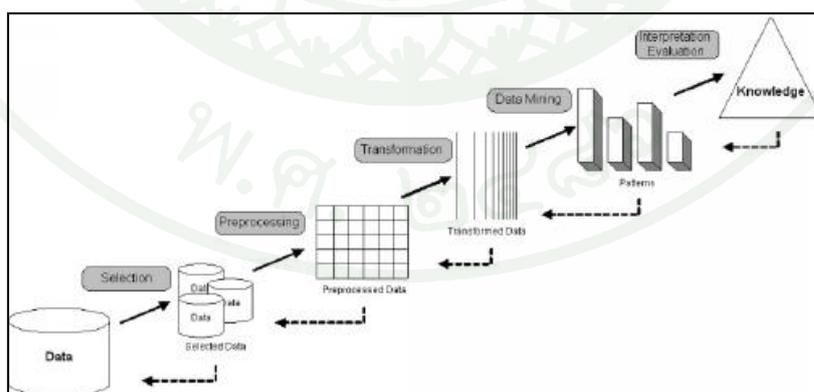
เช่น ในช่วงเวลาที่มีการใช้ไฟฟ้าน้อยๆ ก็ไม่จำเป็นต้องเดินเครื่องกำเนิดไฟฟ้าพร้อมกันทุกเครื่อง นอกจากนี้ถ้ามีการศึกษา Load Profile อย่างต่อเนื่อง ในลักษณะเก็บข้อมูลทางสถิติ ยังช่วยให้สามารถพยากรณ์การขยายตัวของการใช้ไฟฟ้าในอนาคตได้ เพื่อจะได้มีการเตรียมการสร้างแหล่งผลิต และจัดหาพลังงานดิบทดแทนหรือสำรองไว้ล่วงหน้าเพื่ออนาคตด้วย ดังนั้น การที่จะทราบพฤติกรรมหรือลักษณะการใช้ไฟฟ้าได้นั้น ต้องศึกษาจาก Load Profile

ในวิทยานิพนธ์นี้ใช้ข้อมูลซึ่งบันทึกมาจากมิเตอร์อิเล็กทรอนิกส์ซึ่งติดตั้งอยู่กับผู้ใช้ไฟฟ้า ข้อมูลที่เก็บจากมิเตอร์อิเล็กทรอนิกส์นั้นเป็นข้อมูลปริมาณการใช้ไฟฟ้าในช่วงเวลาทุกๆ 15 นาที เพราะฉะนั้น ในหนึ่งวัน จะมีข้อมูล Load Profile ทั้งหมด 96 ค่า

การขุดค้นข้อมูล

Kumar *et al.* (2005) การขุดค้นข้อมูล (Data Mining) เป็นกระบวนการของการกลั่นกรองทางสารสนเทศ (Information) ที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่ เพื่อทำนายแนวโน้มและพฤติกรรมโดยอาศัยข้อมูลในอดีต

การขุดค้นข้อมูลเป็นส่วนหนึ่งในขั้นตอนของกระบวนการค้นหาความรู้ในฐานข้อมูล (Knowledge Discovery) ซึ่งประกอบไปด้วยขั้นตอนดังแสดงในภาพที่ 2



ภาพที่ 2 กระบวนการค้นหาความรู้ในฐานข้อมูล

ที่มา : Kumar *et al.* (2005)

1. การเลือกข้อมูล (Selection)

การเลือกข้อมูลหรือแบ่งข้อมูลตามเงื่อนไขที่กำหนดจากฐานข้อมูลที่น่าสนใจ เช่น เลือกลูกค้าตามเขตที่อยู่ของลูกค้า หรือ เลือกข้อมูลตามชนิดสินค้าที่ลูกค้ามีการสั่งซื้อ เป็นต้น

2. การเตรียมข้อมูล (Preprocessing)

ขั้นตอนการเตรียมข้อมูล หรือเรียกว่าการทำความสะอาดข้อมูล (Data Cleaning) โดยการแยกข้อมูลที่ไม่เกี่ยวข้อง ข้อมูลที่มีความซับซ้อนข้อมูลที่ทำการบันทึกที่ไม่ถูกต้อง หรือข้อมูลที่ไม่สอดคล้องกันออกไป และทำการรวบรวมข้อมูลให้พร้อมที่จะดำเนินการในขั้นตอนการเปลี่ยนรูปข้อมูลต่อไป

3. การเปลี่ยนรูปข้อมูล (Transformation)

ขั้นตอนการเปลี่ยนรูปข้อมูลหรือรวบรวมข้อมูลให้อยู่ในรูปแบบที่เหมาะสมสำหรับการขุดค้นข้อมูล เช่น ในรูปของผลรวม หรือผลสรุปจากข้อมูลที่ได้มาทั้งหมด

4. การขุดค้นข้อมูล (Data Mining)

ขั้นตอนนี้เป็นขั้นตอนที่สำคัญ เพราะเป็นขั้นตอนที่ทำการค้นกรองรูปแบบของข้อมูล (Data Pattern) จากข้อมูลดิบที่มี ในขั้นตอนนี้มีการนำเทคนิคต่างๆ เพื่อเข้ามาช่วยในการดึงรูปแบบที่ซ่อนอยู่ในข้อมูลมาใช้งาน

5. การแปลผล และการประเมิน (Interpretation Evaluation)

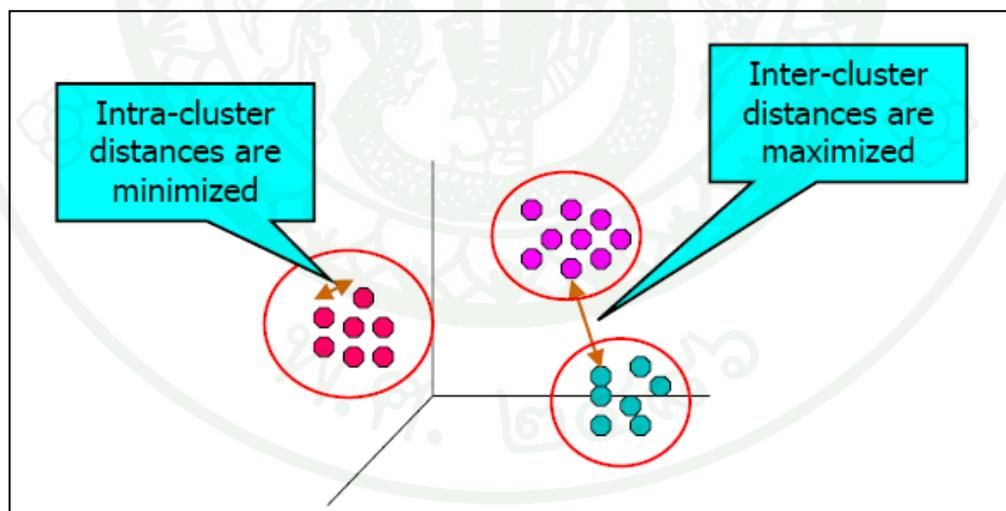
ทำการแปลงรูปข้อมูลที่ได้จากขั้นตอนการขุดค้นข้อมูลให้อยู่ในรูปของความรู้ หรือการสรุปผลที่ได้จากขั้นตอนการขุดค้นข้อมูล เพื่อนำเสนอต่อองค์กรหรือผู้ใช้งาน จนสามารถนำไปใช้ให้เกิดประโยชน์ได้

การจัดกลุ่ม

โดยทั่วไปแล้ว จุดมุ่งหมายของการจัดกลุ่ม (Clustering) นั้นเพื่อที่จะแบ่งข้อมูลออกเป็นกลุ่มๆ ซึ่งการจัดกลุ่มข้อมูลนั้น ถูกนำไปใช้ประโยชน์ในหลายๆด้าน เช่น Pattern Recognition, Image Processing, Fraud Detection และ Marketing เป็นต้น

Lo *et al.* (2005) วิธีการในการจัดกลุ่มนั้น โดยทั่วไปจะมี 2 วิธีหลักๆ คือ Crisp Partitioning และ Soft Partitioning ซึ่ง Crisp Partitioning นั้น ข้อมูลแต่ละข้อมูลจะถูกจัดให้ขึ้นอยู่กับกลุ่มใดกลุ่มหนึ่งเพียงกลุ่มเดียวเท่านั้น โดยจะมีเส้นแบ่งระหว่างกลุ่มที่แน่นอน ในขณะที่ข้อมูลส่วนใหญ่ นั้นมีความคลุมเครืออยู่ในตัว ทำให้ไม่สามารถที่จะจัดให้อยู่ในกลุ่มใดกลุ่มหนึ่งได้อย่างชัดเจน เพราะฉะนั้น Soft Partitioning จึงเป็นวิธีที่ถูกนำมาใช้ในการแก้ปัญหานี้

Kumar *et al.* (2005) การจัดกลุ่มนั้นจะมีหลักการพื้นฐานอยู่ 2 ข้อ คือ ระยะห่างระหว่างสมาชิกในกลุ่มเดียวกัน (Intra-cluster) ต้องมีค่าน้อย และระยะห่างระหว่างสมาชิกของคนละกลุ่ม (Inter-cluster) ต้องมีค่ามาก ดังแสดงในภาพที่ 3



ภาพที่ 3 หลักการพื้นฐานของการจัดกลุ่ม

ที่มา : Kumar *et al.* (2005)

วิธีการจัดกลุ่มนั้นมีหลากหลายวิธี ซึ่งแต่ละวิธีก็จะมีข้อดี-ข้อเสีย แตกต่างกันไป และแต่ละวิธีก็จะเหมาะสมกับข้อมูลที่แตกต่างกันออกไป ตัวอย่างของวิธีต่างๆ เช่น Fuzzy C-Means, Fuzzy-Possibilistic C-Means และ Two-stage Fuzzy C-Means เป็นต้น

1. Data Pre-Processing

Gerbec *et al.* (2002) ข้อมูลดิบที่เราต้องการที่จะนำมาทำการจัดกลุ่มนั้น จะต้องผ่านกระบวนการ Data Pre-Processing ก่อน เพื่อทำการกำจัดข้อมูลที่ไม่สมบูรณ์ออกไป เนื่องจากข้อมูลที่ไม่สมบูรณ์เหล่านี้จะส่งผลให้เกิดการจัดกลุ่มที่ผิดพลาดได้ หลังจากนั้นต้องทำการบรรทัดฐานข้อมูล (Data Normalization) เพื่อปรับให้ข้อมูลทั้งหมดอยู่ในช่วงเดียวกัน ซึ่งการบรรทัดฐานข้อมูลนั้น มี 2 วิธี ดังสมการที่ 1-2

$$x_1 = \frac{(x_{li} - x_{min})}{x_{max} - x_{min}} \quad (1)$$

$$x_1 = \frac{x_{li}}{x_{max}} \quad (2)$$

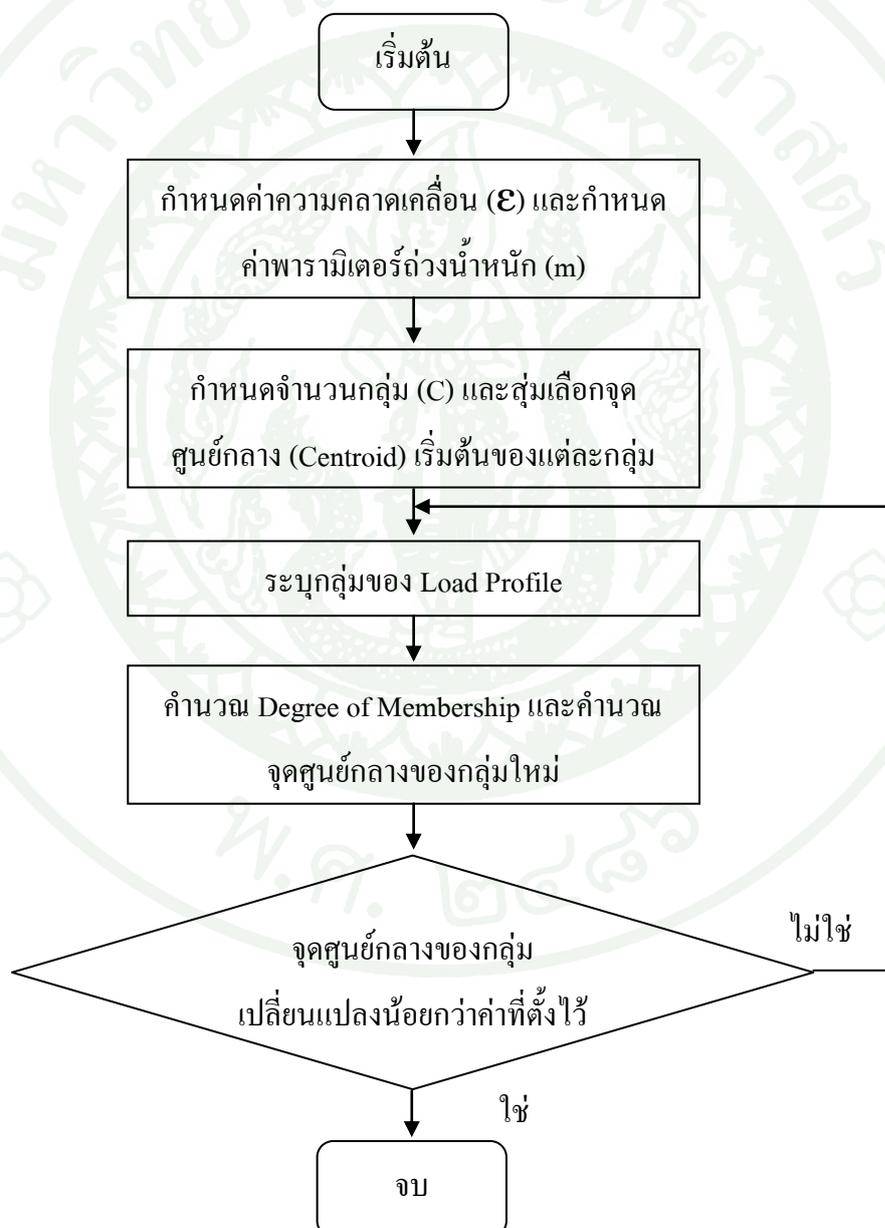
เมื่อ x_1 คือ ข้อมูลที่ทำการบรรทัดฐานแล้ว
 x_{li} คือ ข้อมูลตัวที่ i

จากสมการที่ 1 เป็นการบรรทัดฐานข้อมูลโดยข้อมูลที่ทำการบรรทัดฐานแล้วนั้น จะมีค่าสูงสุดคือ 1 และมีค่าต่ำสุดคือ 0 ส่วนสมการที่ 2 เป็นการบรรทัดฐานข้อมูลโดยข้อมูลที่ทำการบรรทัดฐานแล้วนั้น จะมีค่าสูงสุดคือ 1 แต่ค่าต่ำสุดจะไม่ใช่ 0

ในวิทยานิพนธ์นี้จะเลือกใช้การบรรทัดฐานตามสมการที่ 2 เนื่องจากจะทำให้มองเห็นภาพของ base load ในโพรไฟล์ภาระได้อย่างชัดเจน

2. Fuzzy C-Means

Bezdek *et al.* (1978) วิธีนี้จะมีพารามิเตอร์ Degree of Membership หรือ Membership Grade เป็นตัวบ่งบอกว่า Load Profile แต่ละตัวนั้นมีค่าความเป็นสมาชิกของแต่ละกลุ่มเท่าใด หากมี Degree of Membership กับกลุ่มใดมากที่สุด ก็จะมีค่าความเป็นสมาชิกกับกลุ่มนั้นมากที่สุด ซึ่งการจัดกลุ่มโดยวิธี Fuzzy C-Means แสดงดังภาพที่ 4



ภาพที่ 4 การจัดกลุ่ม โดยวิธี Fuzzy C-Means

2.1 ให้ N เป็นจำนวนข้อมูลที่นำมาจัดกลุ่ม

2.2 ทำการกำหนดจำนวนกลุ่ม (C) ทำการกำหนดค่าเริ่มต้นของจุดศูนย์กลางแต่ละกลุ่ม (V_i) ทำการกำหนดค่าความคลาดเคลื่อน (Epsilon value : ϵ) เพื่อเป็นตัวกำหนดให้หยุดคำนวณต่อ และทำการกำหนดค่าพารามิเตอร์ถ่วงน้ำหนัก (Weighting Parameter : m)

2.3 คำนวณค่าระดับความเป็นสมาชิก (Degree of membership : u_{ik}) สำหรับแต่ละจุดข้อมูลของโพรไฟล์ภาระ ดังสมการที่ 3

$$u_{ik} = \left[\sum_{j=1}^C \left(\frac{D_{ik}}{D_{jk}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (3)$$

เมื่อ D_{ik} คือ ระยะทางระหว่างข้อมูลตัวที่ k กับกลุ่มที่ i

D_{jk} คือ ระยะทางระหว่างข้อมูลตัวที่ k กับกลุ่มที่ j

2.4 ระยะทางระหว่างจุดข้อมูลกับจุดศูนย์กลางแต่ละกลุ่ม ใช้ Euclidean distance ในการหา ดังสมการที่ 4

$$D_{ik}^2 = \|x_k - V_i\|^2 \quad (4)$$

2.5 ทำการปรับค่าจุดศูนย์กลางของกลุ่ม ดังสมการที่ 5

$$V_i = \frac{\sum_{k=1}^N u_{ik}^m x_k}{\sum_{k=1}^N u_{ik}^m} \quad (5)$$

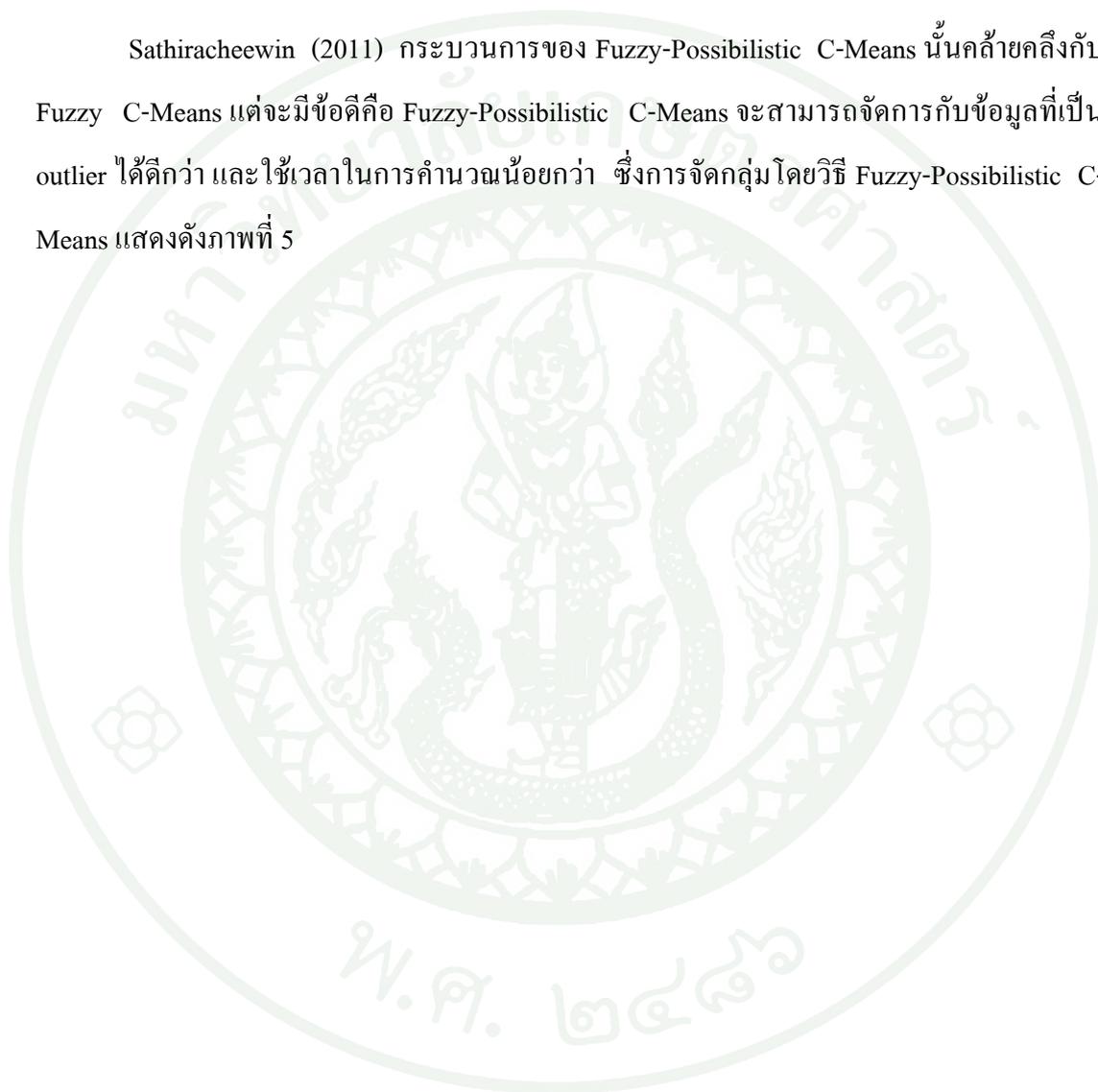
2.6 ทำการ Minimize Objective function ดังสมการที่ 6

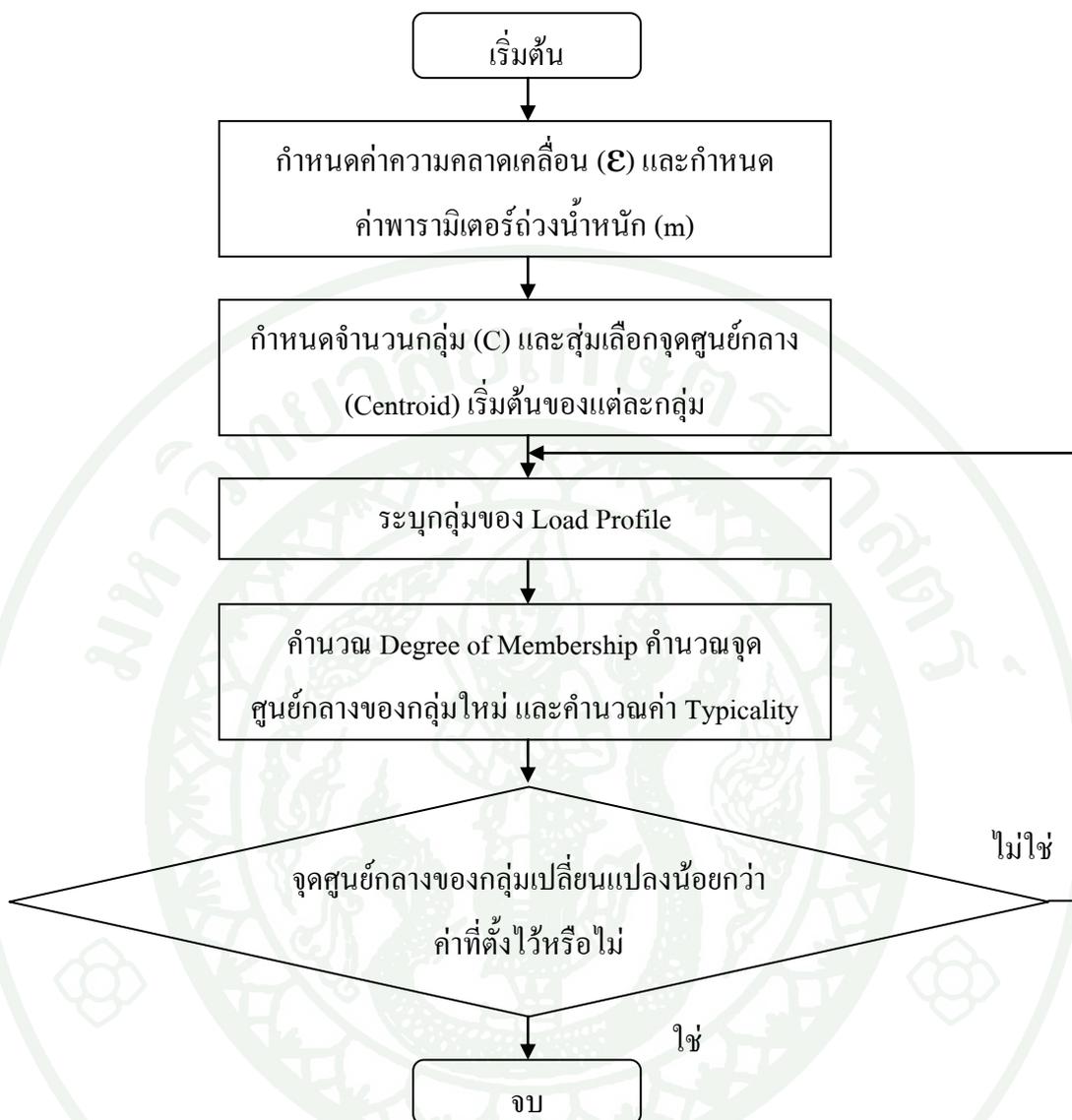
$$\min_{(U,V)} \{ J_m(U,V) = \sum_{i=1}^C \sum_{k=1}^N u_{ik}^m D_{ik}^2 \} \quad (6)$$

ข้อด้อยของ Fuzzy C-Means คือใช้เวลาดำเนินการค่อนข้างมาก และมีความอ่อนไหวต่อข้อมูลที่เป็น noise หรือ outliers ซึ่งจะส่งผลต่อค่า membership

3. Fuzzy-Possibilistic C-Means

Sathiracheewin (2011) กระทบวนการของ Fuzzy-Possibilistic C-Means นั้นคล้ายคลึงกับ Fuzzy C-Means แต่จะมีข้อดีคือ Fuzzy-Possibilistic C-Means จะสามารถจัดการกับข้อมูลที่เป็น outlier ได้ดีกว่า และใช้เวลาในการคำนวณน้อยกว่า ซึ่งการจัดกลุ่มโดยวิธี Fuzzy-Possibilistic C-Means แสดงดังภาพที่ 5





ภาพที่ 5 การจัดกลุ่ม โดยวิธี Fuzzy-Possibilistic C-Means

3.1 ทำการกำหนดจำนวนกลุ่ม (C) ทำการกำหนดค่าเริ่มต้นของจุดศูนย์กลางแต่ละกลุ่ม (V_i) ทำการกำหนดค่าความคลาดเคลื่อน (Epsilon value : ϵ) เพื่อเป็นตัวกำหนดให้หยุดคำนวณต่อ และทำการกำหนดค่าพารามิเตอร์ถ่วงน้ำหนัก (Weighting Parameter)

3.2 คำนวณค่าระดับความเป็นสมาชิก (Degree of membership : u_{ik}) สำหรับแต่ละจุดข้อมูลของโพรไฟล์ภาระ ดังสมการที่ 7

$$u_{ik} = \left[\sum_{j=1}^C \left(\frac{D_{ik}}{D_{jk}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (7)$$

เมื่อ D_{ik} คือ ระยะทางระหว่างข้อมูลตัวที่ k กับกลุ่มที่ i

D_{jk} คือ ระยะทางระหว่างข้อมูลตัวที่ k กับกลุ่มที่ j

3.3 ระยะทางระหว่างชุดข้อมูลกับจุดศูนย์กลางแต่ละกลุ่ม ใช้ Euclidean distance ในการหา ดังสมการที่ 8

$$D_{ik}^2 = \|x_k - V_i\|^2 \quad (8)$$

3.4 ค่า Typicality (t_{ik}) ซึ่ง $\sum_{k=1}^n t_{ik} = 1$ หาได้จากสมการที่ 9

$$t_{ik} = \left[\sum_{j=1}^n \left(\frac{D_{ik}}{D_{ij}} \right)^{\frac{2}{\eta-1}} \right]^{-1} \quad (9)$$

เมื่อ η คือ ค่า Weighting Exponent

3.5 คำนวณจุดศูนย์กลางของกลุ่มที่ i ดังสมการที่ 10

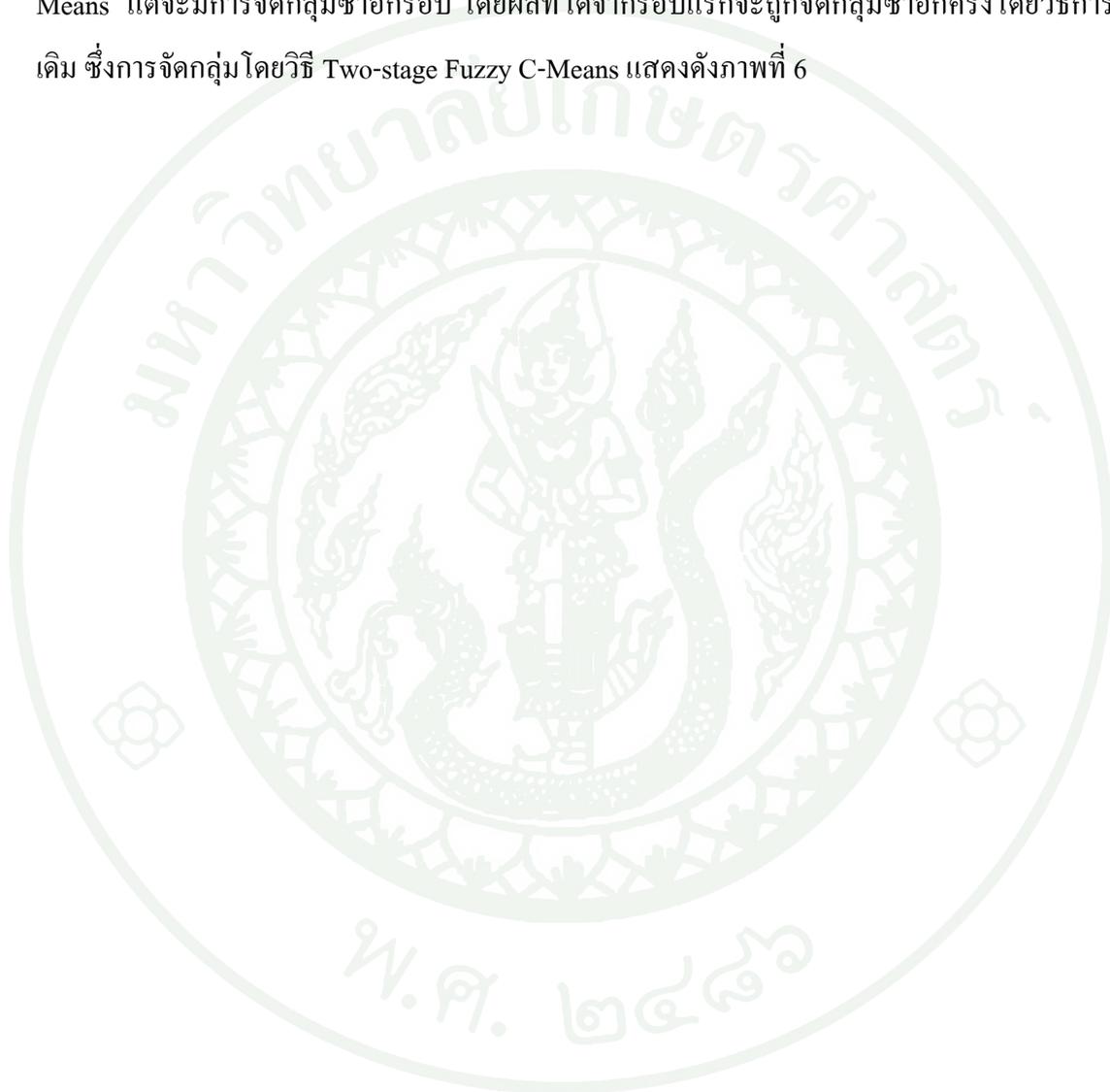
$$V_i = \frac{\sum_{k=1}^n (u_{ik}^m + t_{ik}^m) X_k}{\sum_{k=1}^n (u_{ik}^m + t_{ik}^m)} \quad (10)$$

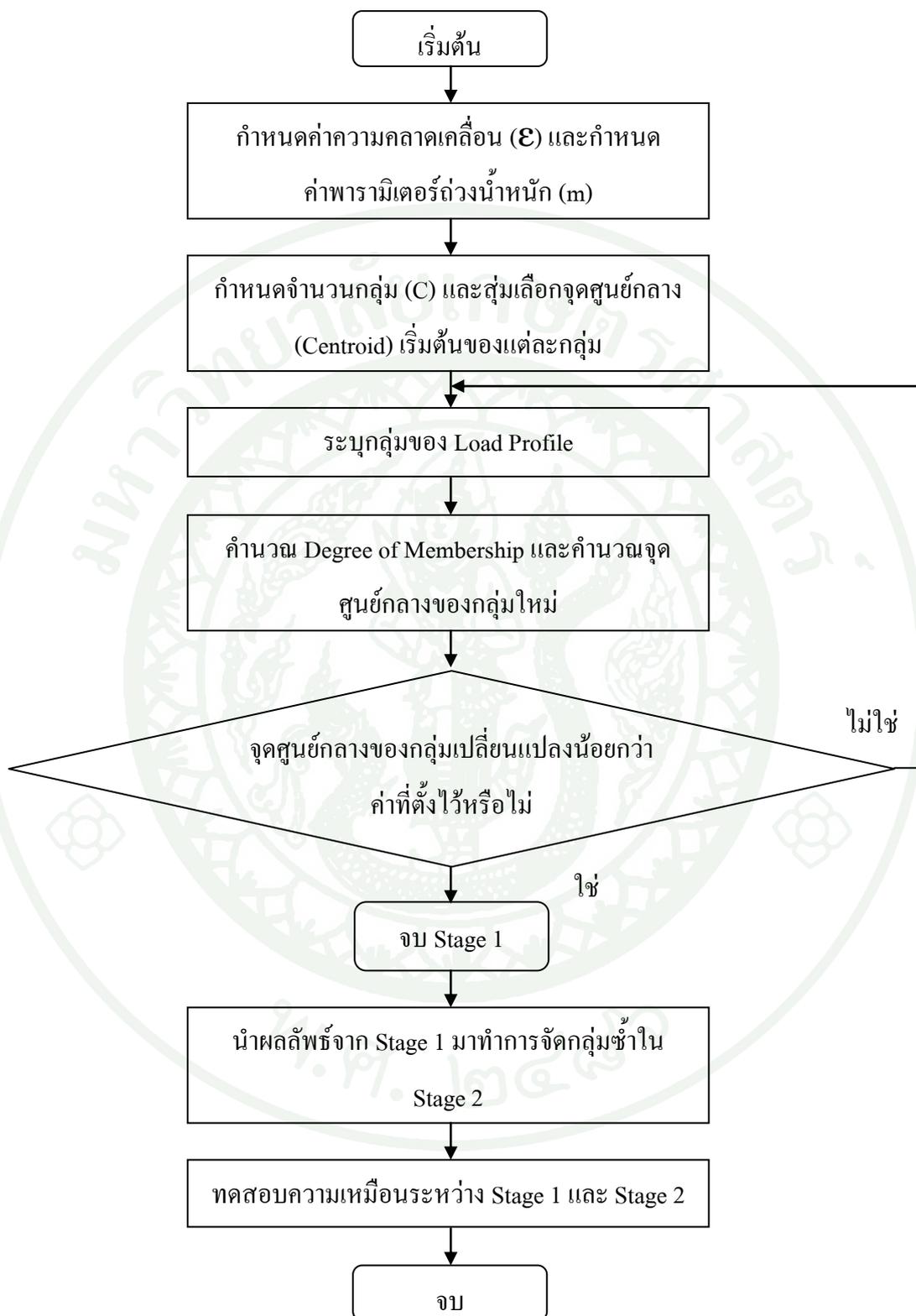
3.6 ทำการ Minimize Objective function ดังสมการที่ 11

$$\min_{(U,T,V)} \{ J_m(U,T,V) = \sum_{i=1}^C \sum_{k=1}^N (u_{ik}^m + t_{ik}^n) D_{ik}^2 \} \quad (11)$$

4. Two-stage Fuzzy C-Means

Lo *et al.* (2005) การจัดกลุ่มโดยใช้วิธี Two-stage Fuzzy C-Means เป็นการจัดกลุ่มซ้ำ (Re-clustering) เพื่อให้ได้รูปแบบที่แท้จริงของข้อมูล ซึ่งกระบวนการทำนั้นเหมือนกับวิธี Fuzzy C-Means แต่จะมีการจัดกลุ่มซ้ำอีกรอบ โดยผลที่ได้จากรอบแรกจะถูกจัดกลุ่มซ้ำอีกครั้งโดยวิธีการเดิม ซึ่งการจัดกลุ่มโดยวิธี Two-stage Fuzzy C-Means แสดงดังภาพที่ 6





ภาพที่ 6 การจัดกลุ่มโดยวิธี Two-stage Fuzzy C-Means

5. Cluster Validation

Hyun and Kyung (1996) ปัญหาสำคัญในการจัดกลุ่มคือ การเลือกจำนวนกลุ่มที่เหมาะสมกับชุดข้อมูลมากที่สุด มีการเสนอดัชนีที่ใช้ในการเลือกจำนวนกลุ่มที่เหมาะสมที่สุดมากมาย แต่ในวิทยานิพนธ์นี้จะใช้ดัชนีที่เป็นที่นิยมกันอย่างกว้างขวาง 3 วิธี คือ Partition Coefficient (PC), Classification Entropy (CE) และ Xie and Beni (XB) ดังสมการที่ 12-14 (Xie and Beni, 1991)

$$PC = \frac{1}{N} \sum_{i=1}^C \sum_{k=1}^N (u_{ik})^m \quad (12)$$

จำนวนกลุ่มที่เหมาะสมที่สุด คือ ทำให้ PC มีค่ามากที่สุด

$$CE = -\frac{1}{N} \sum_{i=1}^C \sum_{k=1}^N u_{ik} \log u_{ik} \quad (13)$$

จำนวนกลุ่มที่เหมาะสมที่สุด คือ ทำให้ CE มีค่าน้อยที่สุด

$$XB = \sum_{i=1}^C \frac{\sum_{k=1}^N u_{ik}^2 \|x_k - v_i\|^2}{N \min_{i \neq j} \|v_k - v_i\|^2} \quad (14)$$

จำนวนกลุ่มที่เหมาะสมที่สุด คือ ทำให้ XB มีค่าน้อยที่สุด

Martin et al. (2008) สำหรับการจัดกลุ่มโดยใช้วิธี Fuzzy-Possibilistic C-Means จะใช้สูตรการคำนวณดัชนี XB ดังสมการที่ 15

$$XB = \sum_{i=1}^C \frac{\sum_{k=1}^N (u_{ik}^2 + t_{ik}^2) \|x_k - v_i\|^2}{N \min_{i \neq j} \|v_k - v_i\|^2} \quad (15)$$

จำนวนกลุ่มที่เหมาะสมที่สุด คือ ทำให้ XB มีค่าน้อยที่สุด

Typical Load Profile (TLP)

เมื่อได้จำนวนกลุ่มที่เหมาะสมสำหรับข้อมูล Load Profile แล้ว จะต้องทำการสร้าง Load Profile ตัวอย่างหรือ TLP เพื่อเป็นตัวแทนของกลุ่มและสามารถนำไปใช้ในการวิเคราะห์เพื่อนำไปใช้ประโยชน์ได้ การคำนวณ TLP ดังสมการที่ 16

$$x_{TLP} = \frac{\sum_{i=1}^k x_i}{k} \quad (16)$$

เมื่อ x_{TLP} คือ Load Profile ตัวอย่างหรือ TLP

x_i คือ Load Profile ในกลุ่ม

k คือ จำนวน Load Profile ในกลุ่ม

Goodness of Fit (GoF)

นิกา (2533) เป็นการทดสอบสมมติฐาน (Hypothesis) โดยในวิทยานิพนธ์นี้เลือกใช้วิธี Kolmogorov-Smirnov Two-Sample Test ซึ่งเป็นการทดสอบกลุ่มตัวอย่าง 2 กลุ่ม เป็นกระบวนการทดสอบว่าสมมติฐานที่ตั้งไว้นั้นเป็นจริงหรือไม่ โดยทำการทดสอบระยะห่าง (Distance Test) ระหว่างกลุ่มตัวอย่าง 2 กลุ่ม ซึ่งมีสมมติฐานดังนี้

1. H_0 คือ ประชากรกลุ่มที่ 1 และกลุ่มที่ 2 มีลักษณะเหมือนกัน
2. H_1 คือ ประชากรกลุ่มที่ 1 และกลุ่มที่ 2 มีลักษณะต่างกัน

โดยคำนวณค่า D จากสมการที่ 17

$$D = \text{maximum}|S_1(X) - S_2(X)| \quad (17)$$

เมื่อ $S_1(X)$ เป็นฟังก์ชันความถี่สะสมสัมพัทธ์ของกลุ่มตัวอย่างที่ 1

$S_2(X)$ เป็นฟังก์ชันความถี่สะสมสัมพัทธ์ของกลุ่มตัวอย่างที่ 2

ชินกร (2552) กรณีที่กลุ่มตัวอย่างมีขนาดใหญ่ ($n_1 > 40, n_2 > 40$ โดยที่ n_1, n_2 ไม่จำเป็นต้องเท่ากัน) และกำหนดระดับนัยสำคัญ = 0.001 จะปฏิเสธสมมติฐาน H_0 เมื่อค่า D ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าวิกฤติของ D ซึ่งคำนวณจากสมการที่ 18

$$D_{critical} = 1.95 \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \quad (18)$$



อุปกรณ์และวิธีการ

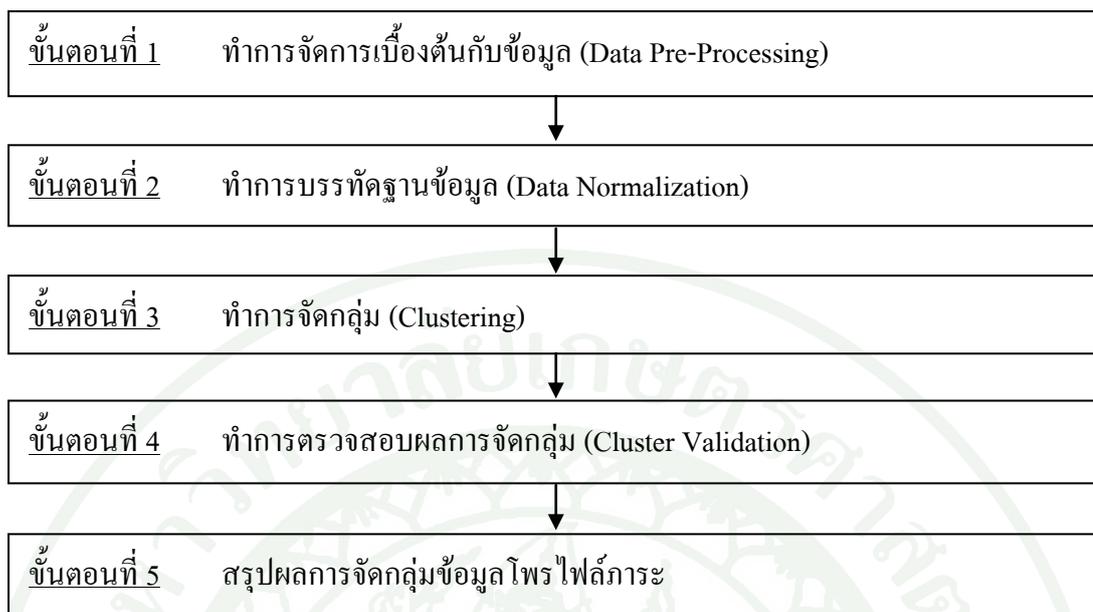
อุปกรณ์

วัสดุอุปกรณ์ที่จำเป็นสำหรับทำวิทยานิพนธ์นี้ ประกอบด้วย

1. เครื่องคอมพิวเตอร์โน้ตบุ๊ก ระบบปฏิบัติการ Microsoft Windows XP Service Pack 3
2. โปรแกรม Microsoft Office 2007 (Excel, Word)
3. โปรแกรม MATLAB R2010b

วิธีการ

วิธีการจัดกลุ่มใหม่ (วิธีความน่าจะเป็นฟัซซี่ซิมินสองขั้นตอน) สำหรับจัดกลุ่มโพรไฟล์ภาระของความต้องการการใช้ไฟฟ้า ถูกสรุปไว้ดังแสดงในภาพที่ 7 ขั้นตอนที่ 1 ทำการจัดการเบื้องต้นกับข้อมูล (Data Pre-Processing) ทำการคัดกรองข้อมูลที่ไม่สมบูรณ์ หรือข้อมูลที่มีปัญหาออกไป เพื่อไม่ให้ส่งผลกระทบต่อการจัดกลุ่ม ขั้นตอนที่ 2 ทำการบรรทัดฐานข้อมูล (Data Normalization) ก่อนที่จะเข้าสู่กระบวนการจัดกลุ่ม เพื่อให้ข้อมูลทั้งหมดอยู่ในช่วงเดียวกัน ขั้นตอนที่ 3 ทำการจัดกลุ่ม (Clustering) โดยใช้วิธีต่างๆ ได้แก่ Fuzzy C-Means, Fuzzy-Possibilistic C-Means, Two-stage Fuzzy C-Means และ Two-stage Fuzzy-Possibilistic C-Means ขั้นตอนที่ 4 ทำการตรวจสอบผลการจัดกลุ่ม (Cluster Validation) โดยใช้ดัชนีต่างๆ ได้แก่ Partition Coefficient (PC), Classification Entropy (CE) และ Xie and Beni (XB) ขั้นตอนที่ 5 สรุปผลการจัดกลุ่มข้อมูลโพรไฟล์ภาระ



ภาพที่ 7 ลำดับขั้นตอนวิธีการจัดกลุ่มโพรไฟล์ภาระ

1. การจัดการเบื้องต้นกับข้อมูล (Data Pre-Processing)

ข้อมูลโพรไฟล์ภาระที่นำมาใช้ในวิทยานิพนธ์นี้ ถูกบันทึกโดยใช้เครื่องวัดแบบดิจิทัล (Digital Meter) ซึ่งเก็บค่าโพรไฟล์ภาระทุกๆ 15 นาที เพราะฉะนั้นใน 1 วัน จะมีค่าโพรไฟล์ภาระ 96 ค่า ตัวอย่างของข้อมูลโพรไฟล์ภาระ แสดงดังภาพที่ 8

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	dd/mm/yy	0:15	0:30	0:45	1:00	1:15	1:30	1:45	2:00	2:15	2:30	2:45	3:00	3:15	3:30
2	4/1/2011	5.148	5.124	5.172	5.16	5.148	5.196	5.16	5.172	5.184	5.16	5.16	5.184	5.196	5.16
3	5/1/2011	10.98	11.304	12.18	10.332	10.284	10.2	10.08	10.164	10.176	10.152	10.944	12.024	10.548	10.92
4	6/1/2011	13.032	11.376	12.024	12.6	10.86	12.096	12.192	10.896	12.612	11.904	11.136	12.96	11.628	11.616
5	7/1/2011	11.784	11.532	10.452	10.332	10.452	10.356	10.596	12.312	10.86	10.212	13.236	12.42	12.24	13.272
6	10/1/2011	12.24	12.156	13.596	12.072	11.352	12.06	11.232	11.748	12.456	13.44	12.564	11.784	13.488	12.456
7	11/1/2011	13.308	12.876	11.94	11.892	12.06	12.408	11.22	12.876	12.672	11.568	12.084	12.612	11.58	12.636
8	12/1/2011	13.632	12.156	12.96	13.356	10.644	11.544	12.432	10.872	10.848	10.944	11.34	12.324	13.608	13.08
9	13/1/2011	12.24	12.576	11.196	13.068	12.42	11.496	12.936	12.54	10.344	11.712	11.148	11.496	12.96	11.604
10	14/1/2011	13.128	12.504	13.56	12.588	12.228	12.96	12.144	11.676	12.528	11.208	11.328	12.612	10.884	11.916
11	17/1/2011	14.916	13.524	13.092	12.552	12.252	12.204	12.144	13.716	14.22	13.488	14.916	14.76	12.756	14.628
12	18/1/2011	12.24	11.664	12	14.592	13.788	13.008	14.424	13.908	13.056	14.388	13.968	13.188	14.556	14.172
13	19/1/2011	15.276	14.448	14.016	15.012	14.484	13.416	14.964	14.184	13.956	15	13.416	12.42	14.136	11.64
14	20/1/2011	14.532	15.156	14.112	14.184	15	13.968	13.632	13.62	13.668	13.728	13.776	13.908	14.688	14.844
15	21/1/2011	15.444	15.624	14.568	13.56	13.392	12.564	12.516	14.088	13.476	12.528	14.34	14.916	14.124	14.892
16	24/1/2011	4.656	4.572	4.584	4.524	4.512	4.5	4.488	4.5	4.524	4.584	4.584	4.596	4.548	4.62
17	25/1/2011	14.148	14.076	13.44	12.984	12.96	12.708	13.68	14.016	14.04	14.316	14.412	14.364	13.596	13.38
18	26/1/2011	12.012	11.616	11.64	13.524	13.944	13.92	13.92	13.656	11.652	11.652	11.688	11.64	11.88	13.032
19	27/1/2011	11.916	13.044	13.26	11.184	10.44	12.588	13.044	12.984	13.056	12.06	12.084	12.084	12.576	13.188
20	28/1/2011	15.288	15.54	14.508	14.292	11.808	12.588	14.484	14.964	14.856	15.6	15.372	13.884	14.964	14.448
21	31/1/2011	14.1	14.04	13.812	12.324	11.1	13.38	13.896	13.476	13.128	11.496	12.156	12.756	11.628	13.86
22	1/2/2011	17.028	16.176	15.444	14.844	14.868	16.416	16.68	16.092	16.044	17.004	15.468	13.284	16.128	16.872
23	2/2/2011	15.912	15.876	16.944	16.704	13.836	13.68	16.248	16.26	15.756	16.764	16.74	15.636	16.704	16.86
24	3/2/2011	16.284	15.636	14.952	16.128	13.8	13.956	16.92	16.56	15.816	16.788	16.956	15.78	15.444	14.556
25	4/2/2011	17.592	17.076	16.656	17.004	17.496	16.08	15.48	14.916	13.704	16.56	17.34	15.912	14.388	14.892

ภาพที่ 8 ตัวอย่างข้อมูลโพรไฟล์ภาระ

จากภาพที่ 8 ซึ่งเป็นตัวอย่างของข้อมูลโพรไฟล์ภาระ โดยข้อมูลในคอลัมน์ A จะแสดงถึงวันที่ของโพรไฟล์ภาระ ซึ่งจะแสดงวันที่อยู่ในรูปของวัน/เดือน/ปี (dd/mm/yy) ส่วนข้อมูลตั้งแต่คอลัมน์ B เป็นต้นไป เป็นข้อมูลโพรไฟล์ภาระราย 15 นาที (Quarter-hourly) ซึ่งมีหน่วยเป็นกิโลวัตต์ (kW)

จากข้อมูลโพรไฟล์ภาระ ดังกล่าวแล้วจะมีข้อมูลโพรไฟล์ภาระบางวันที่มีลักษณะแตกต่างออกไปจากวันอื่นๆอย่างสิ้นเชิง หรือมีลักษณะผิดปกติจากข้อมูลอื่นๆ ดังแสดงในภาพที่ 9 และ 10 ซึ่งในที่นี้เราจะถือว่าเป็นข้อมูลที่ไม่สมบูรณ์ เราจึงต้องทำการค้นหาและกำจัดข้อมูลเหล่านี้ออกไปก่อนเพื่อไม่ให้ส่งผลกระทบต่อการจัดกลุ่มโพรไฟล์ภาระในขั้นตอนต่อไป

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	dd/mm/yy	0:15	0:30	0:45	1:00	1:15	1:30	1:45	2:00	2:15	2:30	2:45	3:00	3:15	3:30
2	2/6/2011	16.116	15.996	16.152	16.812	16.14	15.624	16.344	16.116	15.96	13.872	13.32	14.544	14.148	15.816
3	3/6/2011	15.936	16.044	16.524	16.584	15.888	16.14	16.656	16.164	15.996	13.992	15.204	16.308	15.912	15.768
4	6/6/2011	15.948	18.696	20.016	19.62	17.172	16.236	15.876	16.152	17.928	16.224	15.84	16.236	16.056	16.332
5	7/6/2011	18.696	18.444	19.584	17.124	15.768	16.98	16.464	18.492	18.66	18.672	18.432	18.456	17.808	15.576
6	8/6/2011	18.468	19.692	18.684	18.252	18.972	18.732	17.94	16.764	15.3	16.248	16.98	19.044	18.336	17.904
7	9/6/2011	17.112	16.104	18.348	17.844	17.436	16.62	16.956	17.592	17.04	16.668	17.484	14.46	14.688	16.092
8	10/6/2011	15.468	17.256	18.18	18.192	17.784	17.952	17.808	17.136	15.732	16.668	17.916	18.228	15.756	15.528
9	13/6/2011	6.108	6.228	6.24	6.132	6.3	6.192	6	6.312	6.276	6.276	6.3	6.264	6.264	6.252
10	14/6/2011	15.084	14.868	15.492	15.36	14.544	13.008	13.44	12.276	13.212	14.856	13.956	11.868	14.004	14.772
11	15/6/2011	15.048	15.396	14.592	15.084	15.36	14.448	14.52	15.312	15.312	13.98	12.024	12.108	15.408	14.652
12	16/6/2011	15.444	14.448	12.108	11.856	13.188	13.62	14.244	13.488	12.348	11.964	12	13.728	15.456	14.304
13	17/6/2011	14.568	14.388	14.292	12.168	11.772	13.44	15.12	14.844	21.516	15.804	16.044	16.428	15.888	15.528
14	20/6/2011	7.896	7.908	7.884	7.956	8.076	7.824	7.764	8.04	7.992	8.028	7.896	7.872	8.04	7.92
15	21/6/2011	15.12	12.624	14.388	15.012	15.492	15.18	14.568	13.452	12.888	14.22	15.312	13.272	12.396	14.808
16	22/6/2011	15.396	15.096	15.744	15.336	15.204	15.48	14.58	15.636	15.192	14.88	14.952	12.816	13.644	15.804
17	23/6/2011	15.564	15.96	15.312	15.468	15.756	14.616	15.192	13.584	12.744	15.156	15.396	14.88	15.6	15
18	24/6/2011	15.516	14.112	12.576	14.916	15.468	14.892	15.96	15.3	15.468	16.068	15.024	15.072	16.128	15.048
19	27/6/2011	19.212	19.572	18.204	16.44	16.056	15.18	14.904	17.424	16.812	16.98	17.364	16.656	17.592	15.72
20	28/6/2011	15.396	17.604	18.696	16.116	16.104	15.9	17.652	18.756	18.432	18.096	18.492	16.212	16.284	15.528
21	29/6/2011	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	30/6/2011	18.972	16.692	14.292	14.808	18.096	18.108	17.988	17.304	18.276	14.028	13.86	18.732	17.784	16.62
23	4/7/2011	5.376	5.436	5.328	5.436	5.436	5.124	5.352	5.484	5.352	5.28	5.364	5.208	5.244	5.364
24	5/7/2011	13.644	12.888	13.032	12.336	10.188	11.7	13.464	12.708	12.648	12.912	13.728	12.696	11.52	11.316
25	6/7/2011	11.904	13.26	12.588	10.584	11.508	13.008	12.312	10.44	10.428	11.424	12.36	13.02	13.44	14.004

ภาพที่ 9 ตัวอย่างข้อมูลโพรไฟล์ภาระที่มีลักษณะผิดปกติ (มีค่าเป็น 0 ตลอด)

จากภาพที่ 9 ซึ่งเป็นตัวอย่างข้อมูลโพรไฟล์ภาระที่มีลักษณะผิดปกติ สังเกตข้อมูลของวันที่ 29/06/11 ข้อมูลโพรไฟล์ภาระจะมีค่าเป็น 0 ตลอด ซึ่งตามความเป็นจริงแล้ว แม้ว่าวันนั้นจะไม่มี การผลิตเลย แต่ค่าโพรไฟล์ภาระก็จะต้องมีค่าเล็กน้อย จะไม่เป็น 0 ส่วนสาเหตุที่ทำให้ข้อมูลโพรไฟล์ภาระมีค่าเป็น 0 ตลอดนั้น อาจเป็นความผิดพลาดที่เกิดจากเครื่องวัดหรืออาจเป็นความผิดพลาด ในขั้นตอนการถ่ายโอนข้อมูลจากเครื่องวัดไปยังเครื่องคอมพิวเตอร์

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	dd/mm/yy	0:15	0:30	0:45	1:00	1:15	1:30	1:45	2:00	2:15	2:30	2:45	3:00	3:15	3:30
2	16/9/2011	79.74	81.36	89.46	79.38	79.74	86.04	76.32	80.64	77.22	76.68	76.32	77.94	77.94	75.78
3	19/9/2011	35.1	32.4	32.4	39.42	32.58	32.4	31.86	32.4	32.22	32.22	32.4	31.86	32.4	32.22
4	20/9/2011	83.52	84.6	83.16	84.96	85.32	87.12	87.3	86.22	86.76	81.18	63.36	62.1	62.1	61.92
5	21/9/2011	77.4	83.34	74.88	74.16	83.34	75.78	77.94	83.52	75.42	78.84	83.7	75.42	77.94	83.16
6	22/9/2011	76.32	84.78	80.1	74.16	84.24	80.82	77.22	86.04	79.92	78.3	84.24	75.78	74.34	82.44
7	23/9/2011	76.14	83.34	75.96	74.52	82.8	74.52	73.26	79.56	76.5	74.52	79.02	76.68	73.08	76.86
8	26/9/2011	18.36	17.28	18.18	17.46	17.82	17.82	17.64	18.18	17.46	18.18	17.64	18.18	17.46	18.18
9	27/9/2011	79.74	79.56	78.66	76.68	78.12	78.12	78.66	80.1	81.9	81.36	80.82	81	80.28	81.18
10	28/9/2011														
11	29/9/2011	85.32	84.24	84.6	85.32	80.82	78.12	81.36	78.84	80.28	82.8	74.16	78.84	80.1	73.44
12	30/9/2011	82.62	82.44	81.36	72	73.08	65.7	61.56	62.1	71.46	83.34	82.98	81.9	79.74	81.54
13	3/10/2011	21.24	21.78	21.42	21.78	21.42	20.7	20.16	20.34	20.34	20.34	20.16	19.98	19.26	19.26
14	4/10/2011	70.38	69.84	68.76	68.58	67.5	67.5	68.58	67.68	69.12	70.02	69.84	71.82	69.66	68.4
15	5/10/2011	72.54	72.36	77.94	71.82	69.66	69.3	70.2	69.66	78.48	72.36	72.72	70.92	70.02	70.2
16	6/10/2011	76.68	76.5	74.7	71.46	68.76	68.58	67.86	67.86	70.2	71.82	71.46	71.64	73.62	74.34
17	7/10/2011	74.52	73.26	78.12	76.32	70.02	74.16	75.42	71.1	75.6	81.18	73.26	72.54	80.28	73.44
18	10/10/2011	22.5	22.68	22.86	23.04	22.68	22.68	22.68	22.68	22.32	22.14	21.6	20.7	20.34	20.52
19	11/10/2011	75.42	75.24	75.78	76.14	77.04	77.04	77.58	81.36	79.56	77.04	77.04	78.84	79.02	74.88
20	12/10/2011	77.22	72.36	74.16	77.76	71.64	75.78	78.3	72	75.96	78.12	72.9	74.52	78.66	72
21	13/10/2011	73.44	81.36	75.6	73.26	80.82	73.8	73.8	80.28	71.28	72.72	79.92	71.82	73.98	79.02
22	14/10/2011	91.44	80.46	91.26	86.4	78.3	90.18	91.08	79.92	83.88	86.76	79.2	83.16	88.38	84.06
23	17/10/2011	3.24	3.24	3.42	3.6	3.24	3.24	3.24	3.42	3.42	3.42	3.42	3.24	3.24	3.24
24	18/10/2011	3.24	3.24	3.6	3.24	3.24	3.24	3.24	3.42	3.42	3.6	3.24	3.42	3.24	3.42
25	19/10/2011	3.24	3.42	3.42	3.6	3.42	3.42	3.42	3.42	3.42	3.6	3.24	3.42	3.24	3.24

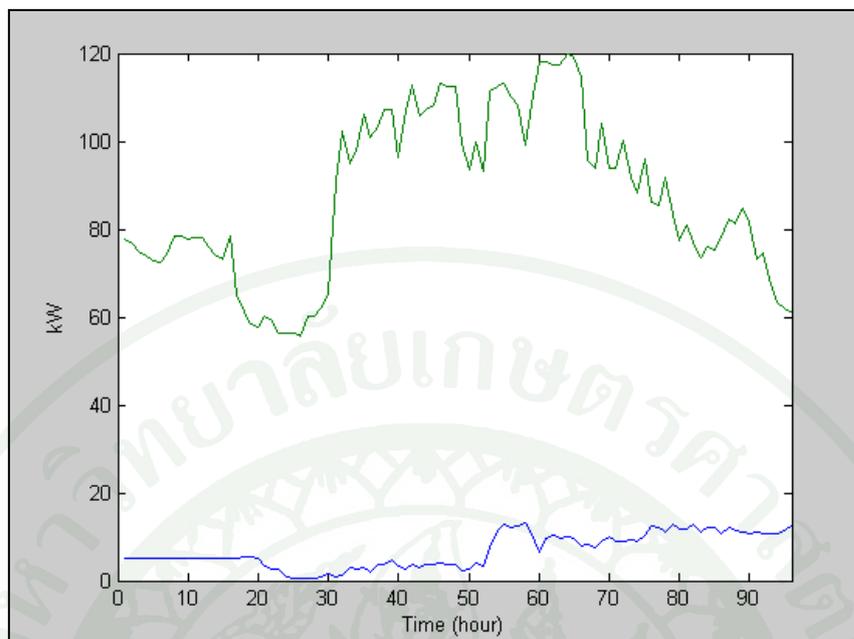
ภาพที่ 10 ตัวอย่างข้อมูลโพรไฟล์ภาระที่มีลักษณะผิดปกติ (ไม่มีค่าในบางช่วงเวลา)

จากภาพที่ 10 ซึ่งเป็นตัวอย่างข้อมูลโพรไฟล์ภาระที่มีลักษณะผิดปกติ สังเกตข้อมูลของวันที่ 28/09/11 ข้อมูลโพรไฟล์ภาระในบางช่วงเวลาจะไม่มีค่า ซึ่งความเป็นจริงแล้วไม่ว่าจะมีการใช้ไฟฟ้ามากหรือน้อย เครื่องวัดก็จะต้องมีการบันทึกค่าตลอดทุกช่วงเวลา สาเหตุที่ทำให้ข้อมูลโพรไฟล์ภาระในบางช่วงเวลาไม่มีค่านั้น อาจเป็นความผิดพลาดที่เกิดจากเครื่องวัด อาจเป็นความผิดพลาดในขั้นตอนการถ่ายโอนข้อมูลจากเครื่องวัดไปยังเครื่องคอมพิวเตอร์ หรืออาจเกิดเหตุการณ์ไฟฟ้าดับในช่วงเวลานั้นๆพอดี

เมื่อทำการกำจัดข้อมูลซึ่งมีความผิดปกติออกไปได้หมดแล้ว ข้อมูลที่พร้อมที่จะถูกดำเนินการในขั้นต่อไป

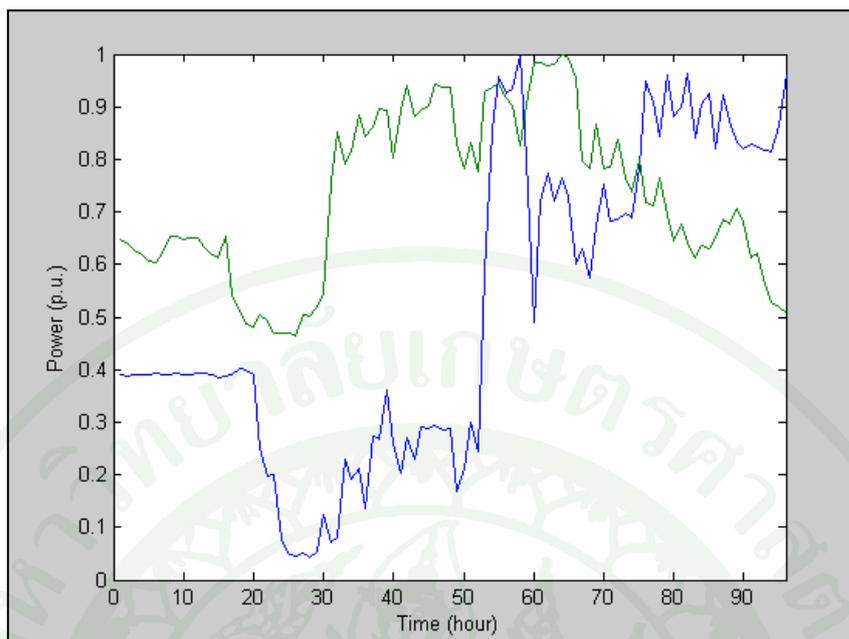
2. การบรรทัดฐานข้อมูล (Data Normalization)

หลังจากทำการจัดการข้อมูลเบื้องต้นแล้ว ก่อนจะเริ่มทำการจัดกลุ่ม (Clustering) นั้น จะต้องทำการบรรทัดฐานข้อมูล (Data Normalization) ก่อน เนื่องจากว่าผู้ใช้ไฟฟ้าแต่ละรายนั้น จะมีปริมาณการใช้ไฟฟ้าที่แตกต่างกัน ดังแสดงในภาพที่ 11 หากเราไม่มีการบรรทัดฐานข้อมูลก่อนที่จะทำการจัดกลุ่มข้อมูล เมื่อเรานำข้อมูลของผู้ใช้ไฟฟ้าแต่ละรายมารวมกันแล้วนำไปทำการจัดกลุ่มผลที่ได้ คือ จะมีจำนวนกลุ่มจำนวนมาก จนไม่สามารถนำไปใช้ประโยชน์ได้ เนื่องจากว่า ปริมาณการใช้ไฟฟ้าที่แตกต่างกัน ทำให้การจัดกลุ่มถือว่าอยู่คนละกลุ่มกัน แต่อันที่จริงแล้ว การจัดกลุ่มนั้นจะยึดเอารูปแบบ (Pattern) ของข้อมูลเป็นหลัก ว่าข้อมูลใดมีรูปแบบเดียวกัน หรือข้อมูลใดมีรูปแบบต่างกัน เพราะฉะนั้นจึงต้องทำการบรรทัดฐานข้อมูลก่อนเพื่อที่จะให้ข้อมูลทุกตัวอยู่ในช่วงเดียวกัน



ภาพที่ 11 ตัวอย่างข้อมูลโพรไฟล์ภาระที่ไม่ได้ทำการบรรทัดฐานข้อมูล

จากภาพที่ 11 ซึ่งเป็นตัวอย่างข้อมูลโพรไฟล์ภาระที่ไม่ได้ทำการบรรทัดฐานข้อมูล จะเห็นได้ว่าเมื่อนำข้อมูลโพรไฟล์ภาระของผู้ใช้ไฟฟ้า 2 ราย ที่มีปริมาณการใช้ไฟฟ้าแตกต่างกันมาพลอตกราฟ ช่วงของข้อมูลจะแตกต่างกันมาก ซึ่งจะส่งผลต่อการจัดกลุ่มข้อมูลต่อไป สังเกตข้อมูลตัวที่ 1 (กราฟเส้นสีเขียว) มีการใช้ไฟฟ้ามากกว่าข้อมูลตัวที่ 2 (กราฟเส้นสีน้ำเงิน) ดังนั้นจึงต้องทำการบรรทัดฐานข้อมูลโพรไฟล์ภาระ ดังแสดงในภาพที่ 12



ภาพที่ 12 ตัวอย่างข้อมูลโพรไฟล์ภาระที่ทำการบรรทัดฐานข้อมูลแล้ว

จากภาพที่ 12 ซึ่งเป็นตัวอย่างข้อมูลโพรไฟล์ภาระที่ทำการบรรทัดฐานข้อมูลแล้ว จะเห็นได้ว่าเมื่อทำการบรรทัดฐานข้อมูลโพรไฟล์ภาระของผู้ใช้ไฟฟ้า 5 ราย ที่มีปริมาณการใช้ไฟฟ้าแตกต่างกัน จากนั้นจึงนำมาพลอตกราฟ ทำให้ช่วงของข้อมูลอยู่ในช่วงเดียวกัน ซึ่งจะทำให้การจัดกลุ่มข้อมูลมีความถูกต้องมากยิ่งขึ้น

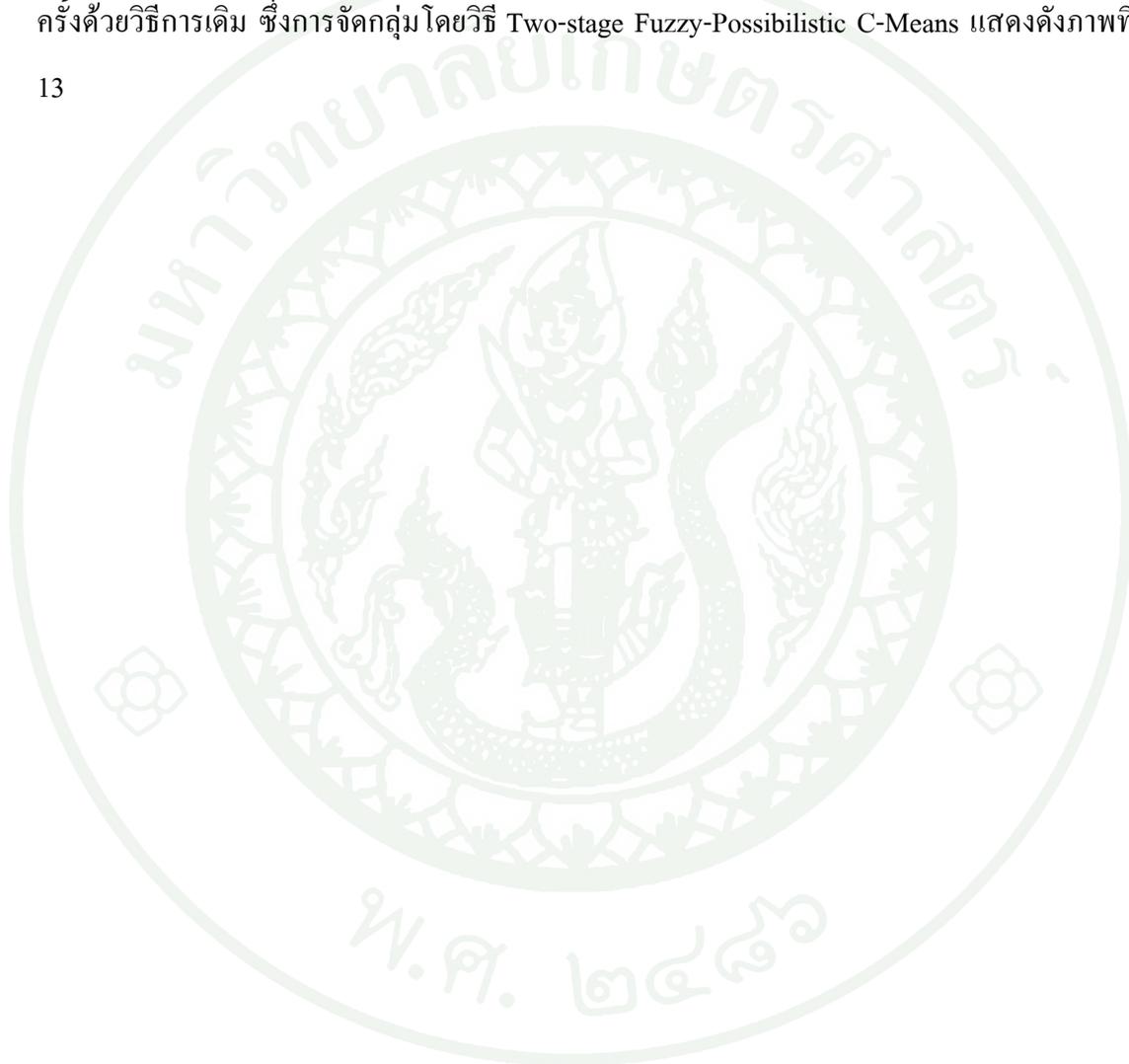
3. การจัดกลุ่ม (Clustering)

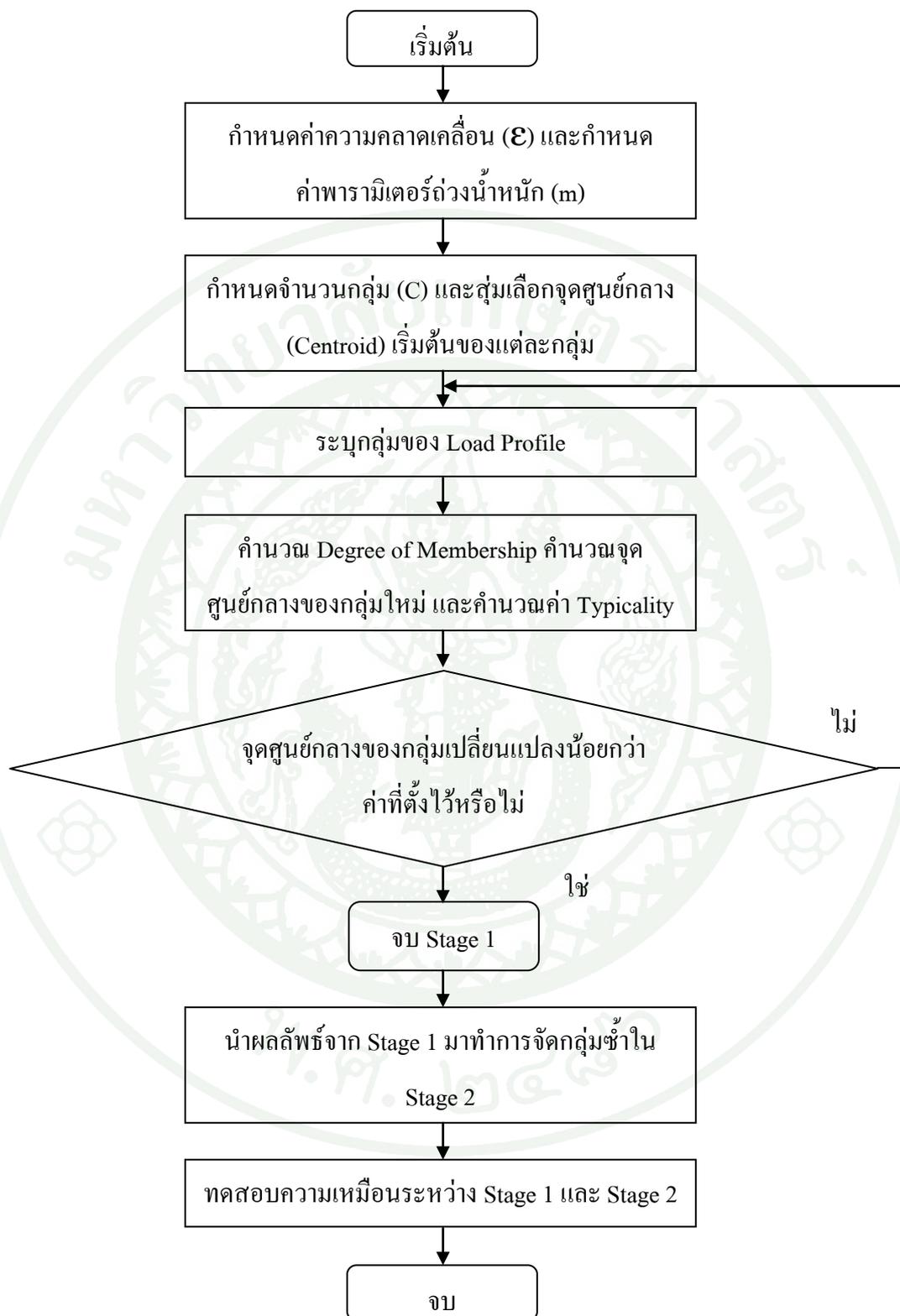
หลังจากทำการบรรทัดฐานข้อมูลโพรไฟล์ภาระทั้งหมดที่ต้องการนำมาจัดกลุ่มแล้ว เข้าสู่กระบวนการการจัดกลุ่ม (Clustering) ซึ่งในวิทยานิพนธ์นี้จะทำการจัดกลุ่มข้อมูลโพรไฟล์ภาระทั้งหมด 4 วิธี คือ Fuzzy C-Means Clustering, Fuzzy-Possibilistic C-Means Clustering, Two-stage Fuzzy C-Means Clustering และ Two-stage Fuzzy-Possibilistic C-Means Clustering โดยวิธีสุดท้ายคือ Two-stage Fuzzy-Possibilistic C-Means Clustering เป็นวิธีที่ได้ทำการเสนอในวิทยานิพนธ์นี้

Two-stage Fuzzy-Possibilistic C-Means

การจัดกลุ่มโดยใช้วิธี Two-stage Fuzzy-Possibilistic C-Means เป็นการจัดกลุ่มซ้ำ (Re-clustering) เพื่อให้ได้รูปแบบที่แท้จริงของข้อมูล ซึ่งกระบวนการทำนั้นเหมือนกับวิธี Fuzzy-Possibilistic C-Means แต่จะมีการจัดกลุ่มซ้ำอีกรอบ โดยผลที่ได้จากรอบแรกจะถูกจัดกลุ่มซ้ำอีกครั้งด้วยวิธีการเดิม ซึ่งการจัดกลุ่มโดยวิธี Two-stage Fuzzy-Possibilistic C-Means แสดงดังภาพที่

13





ภาพที่ 13 การจัดกลุ่มโดยวิธี Two-stage Fuzzy-Possibilistic C-Means

4. การตรวจสอบผลการจัดกลุ่ม (Cluster Validation)

หลังจากทำการจัดกลุ่มข้อมูลโพรไฟล์ภาวะแล้ว เข้าสู่กระบวนการการตรวจสอบผลการจัดกลุ่ม (Cluster Validation) ซึ่งในวิทยานิพนธ์นี้จะใช้ดัชนีที่ใช้ในการตรวจสอบทั้งหมด 3 ตัว คือ Partition Coefficient (PC), Classification Entropy (CE) และ Xie and Beni (XB)

5. สรุปผลการจัดกลุ่มข้อมูลโพรไฟล์ภาวะ

หลังจากทำการจัดกลุ่มโพรไฟล์ภาวะแล้วทำการสรุปผลการจัดกลุ่มเพื่อวิเคราะห์ว่าการจัดกลุ่มแต่ละวิธีนั้นมีข้อดีข้อเสียอย่างไร และวิธีใดให้ผลการจัดกลุ่มที่เหมาะสมกับข้อมูลโพรไฟล์ภาวะมากที่สุด

ผลและวิจารณ์

ผล

ในวิทยานิพนธ์นี้จะยกตัวอย่างการจัดกลุ่มโพรไฟล์ภาระ (Load Profile Clustering) ซึ่งข้อมูลที่นำมาใช้เป็นข้อมูลที่ได้มาจากเครื่องวัดแบบดิจิทัล ซึ่งเก็บค่าโพรไฟล์ภาระทุกๆ 15 นาที โดยเป็นข้อมูลการใช้ไฟฟ้าในรอบ 1 ปีของผู้ใช้ไฟฟ้า 11 ราย รวมทั้งสิ้นมีข้อมูลโพรไฟล์ภาระ = $11 \times 365 = 4,015$ ข้อมูล ดังแสดงในภาพที่ 14

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	dd/mm/yy	0:15	0:30	0:45	1:00	1:15	1:30	1:45	2:00	2:15	2:30	2:45	3:00
2	1/1/2011	6.396	6.36	6.42	6.42	6.408	6.516	6.504	6.468	6.528	6.468	6.468	6.552
3	2/1/2011	6.48	6.432	6.396	6.444	6.468	6.324	6.288	6.384	6.336	6.348	6.456	6.288
4	3/1/2011	6.324	6.42	6.372	7.164	7.14	6.852	6.936	6.456	6.84	6.684	6.828	6.792
5	4/1/2011	5.148	5.124	5.172	5.16	5.148	5.196	5.16	5.172	5.184	5.16	5.16	5.184
6	5/1/2011	10.98	11.304	12.18	10.332	10.284	10.2	10.08	10.164	10.176	10.152	10.944	12.024
7	6/1/2011	13.032	11.376	12.024	12.6	10.86	12.096	12.192	10.896	12.612	11.904	11.136	12.96
8	7/1/2011	11.784	11.532	10.452	10.332	10.452	10.356	10.596	12.312	10.86	10.212	13.236	12.42
9	8/1/2011	13.536	12.156	12.252	13.2	12.096	11.784	13.056	12.12	11.784	13.104	12.288	11.856
10	9/1/2011	12.252	11.16	11.232	11.244	11.124	10.764	11.184	12.312	12.552	11.508	11.688	12.48
11	10/1/2011	12.24	12.156	13.596	12.072	11.352	12.06	11.232	11.748	12.456	13.44	12.564	11.784
12	11/1/2011	13.308	12.876	11.94	11.892	12.06	12.408	11.22	12.876	12.672	11.568	12.084	12.612
13	12/1/2011	13.632	12.156	12.96	13.356	10.644	11.544	12.432	10.872	10.848	10.944	11.34	12.324
14	13/1/2011	13.212	12.576	11.196	13.068	12.42	11.496	12.936	12.54	10.344	11.712	11.148	11.496
15	14/1/2011	13.128	12.504	13.56	12.588	12.228	12.96	12.144	11.676	12.528	11.208	11.328	12.612
16	15/1/2011	11.616	12.468	13.152	12	11.376	12.996	11.988	12.036	12.996	11.244	12.204	13.26
17	16/1/2011	12.912	12.54	11.724	13.128	12.528	10.992	12.972	12.744	11.364	12.816	12.9	11.904
18	17/1/2011	14.916	13.524	13.092	12.552	12.252	12.204	12.144	13.716	14.22	13.488	14.916	14.76
19	18/1/2011	12.24	11.664	12	14.592	13.788	13.008	14.424	13.908	13.056	14.388	13.968	13.188
20	19/1/2011	15.276	14.448	14.016	15.012	14.484	13.416	14.964	14.184	13.956	15	13.416	12.42
21	20/1/2011	14.532	15.156	14.112	14.184	15	13.968	13.632	13.62	13.668	13.728	13.776	13.908
22	21/1/2011	15.444	15.624	14.568	13.56	13.392	12.564	12.516	14.088	13.476	12.528	14.34	14.916
23	22/1/2011	14.82	13.824	13.26	12.492	13.056	12.372	13.596	14.52	14.364	12.456	13.356	12.912
24	23/1/2011	14.208	13.128	14.436	12.624	13.14	12.792	12.6	13.08	14.004	14.508	13.308	14.22
25	24/1/2011	4.656	4.572	4.584	4.524	4.512	4.5	4.488	4.5	4.524	4.584	4.584	4.596

ภาพที่ 14 ข้อมูลโพรไฟล์ภาระในรอบ 1 ปีของผู้ใช้ไฟฟ้า 11 ราย

เมื่อได้ข้อมูลโพรไฟล์ภาระของผู้ใช้ไฟฟ้ามาแล้ว ทำการตัดวันหยุดต่างๆออกไป เนื่องจากว่าเป็นวันที่มีการใช้ไฟฟ้าไม่ปกติ และทำการตัดวันที่มีข้อมูลผิดปกติออกไป ทำให้มีข้อมูลที่สามารนำไปใช้ในการจัดกลุ่มทั้งสิ้น 2,662 ข้อมูล ดังแสดงในภาพที่ 15

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	dd/mm/yy	0.15	0.30	0.45	1.00	1.15	1.30	1.45	2.00	2.15	2.30	2.45	3.00
2	4/1/2011	5.148	5.124	5.172	5.16	5.148	5.196	5.16	5.172	5.184	5.16	5.16	5.184
3	5/1/2011	10.98	11.304	12.18	10.332	10.284	10.2	10.08	10.164	10.176	10.152	10.944	12.024
4	6/1/2011	13.032	11.376	12.024	12.6	10.86	12.096	12.192	10.896	12.612	11.904	11.136	12.96
5	7/1/2011	11.784	11.532	10.452	10.332	10.452	10.356	10.596	12.312	10.86	10.212	13.236	12.42
6	10/1/2011	12.24	12.156	13.596	12.072	11.352	12.06	11.232	11.748	12.456	13.44	12.564	11.784
7	11/1/2011	13.308	12.876	11.94	11.892	12.06	12.408	11.22	12.876	12.672	11.568	12.084	12.612
8	12/1/2011	13.632	12.156	12.96	13.356	10.644	11.544	12.432	10.872	10.848	10.944	11.34	12.324
9	13/1/2011	13.212	12.576	11.196	13.068	12.42	11.496	12.936	12.54	10.344	11.712	11.148	11.496
10	14/1/2011	13.128	12.504	13.56	12.588	12.228	12.96	12.144	11.676	12.528	11.208	11.328	12.612
11	17/1/2011	14.916	13.524	13.092	12.552	12.252	12.204	12.144	13.716	14.22	13.488	14.916	14.76
12	18/1/2011	12.24	11.664	12	14.592	13.788	13.008	14.424	13.908	13.056	14.388	13.968	13.188
13	19/1/2011	15.276	14.448	14.016	15.012	14.484	13.416	14.964	14.184	13.956	15	13.416	12.42
14	20/1/2011	14.532	15.156	14.112	14.184	15	13.968	13.632	13.62	13.668	13.728	13.728	13.908
15	21/1/2011	15.444	15.624	14.568	13.56	13.392	12.564	12.516	14.088	13.476	12.528	14.34	14.916
16	24/1/2011	4.656	4.572	4.584	4.524	4.512	4.5	4.488	4.5	4.524	4.584	4.584	4.596
17	25/1/2011	14.148	14.076	13.44	12.984	12.96	12.708	13.68	14.016	14.04	14.316	14.412	14.364
18	26/1/2011	12.012	11.616	11.64	13.524	13.944	13.92	13.92	13.656	11.652	11.652	11.688	11.64
19	27/1/2011	11.916	13.044	13.26	11.184	10.44	12.588	13.044	12.984	13.056	12.06	12.084	12.084
20	28/1/2011	15.288	15.54	14.508	14.292	11.808	12.588	14.484	14.964	14.856	15.6	15.372	13.884
21	31/1/2011	14.1	14.04	13.812	12.324	11.1	13.38	13.896	13.476	13.128	11.496	12.156	12.756
22	1/2/2011	17.028	16.176	15.444	14.844	14.868	16.416	16.68	16.092	16.044	17.004	15.468	13.284
23	2/2/2011	15.912	15.876	16.944	16.704	13.836	13.68	16.248	16.26	15.756	16.764	16.74	15.636
24	3/2/2011	16.284	15.636	14.952	16.128	13.8	13.956	16.92	16.56	15.816	16.788	16.956	15.78
25	4/2/2011	17.592	17.076	16.656	17.004	17.496	16.08	15.48	14.916	13.704	16.56	17.34	15.912

ภาพที่ 15 ข้อมูลโพรไฟล์ภาระที่ทำการตัดวันหยุดและวันที่มีข้อมูลผิดปกติออกแล้ว

ทำการบรรทัดฐานข้อมูล (Data Normalization) เพื่อให้ข้อมูลทั้งหมดอยู่ในช่วงเดียวกัน
ข้อมูลโพรไฟล์ภาระที่ทำการบรรทัดฐานแล้วแสดงในภาพที่ 16

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	dd/mm/yy	0.15	0.30	0.45	1.00	1.15	1.30	1.45	2.00	2.15	2.30	2.45	3.00
2	4/1/2011	0.389646	0.387829	0.391462	0.390554	0.389646	0.393279	0.390554	0.391462	0.392371	0.390554	0.390554	0.392371
3	5/1/2011	0.225647	0.232306	0.250308	0.21233	0.211344	0.209618	0.207152	0.208878	0.209125	0.208631	0.224908	0.247102
4	6/1/2011	0.25583	0.223322	0.236042	0.24735	0.213192	0.237456	0.23934	0.213899	0.247585	0.233687	0.21861	0.254417
5	7/1/2011	0.243371	0.238166	0.215861	0.213383	0.215861	0.213879	0.218835	0.254275	0.224287	0.210905	0.273358	0.256506
6	10/1/2011	0.260005	0.258221	0.28881	0.256436	0.241142	0.256181	0.238593	0.249554	0.264593	0.285496	0.266888	0.250319
7	11/1/2011	0.272482	0.263636	0.244472	0.243489	0.246929	0.254054	0.22973	0.263636	0.259459	0.236855	0.24742	0.258231
8	12/1/2011	0.277819	0.247738	0.264123	0.272194	0.216923	0.235265	0.253363	0.22157	0.221081	0.223037	0.231108	0.251162
9	13/1/2011	0.275871	0.262591	0.233776	0.272864	0.259334	0.24004	0.270108	0.261839	0.215986	0.24455	0.232774	0.24004
10	14/1/2011	0.272682	0.259721	0.281655	0.261466	0.253988	0.269192	0.252243	0.242522	0.260219	0.232802	0.235294	0.261964
11	17/1/2011	0.327623	0.297048	0.287559	0.275698	0.269109	0.268055	0.266737	0.301265	0.312335	0.296257	0.327623	0.324196
12	18/1/2011	0.253416	0.241491	0.248447	0.302112	0.285466	0.269317	0.298634	0.28795	0.270311	0.297888	0.289193	0.273043
13	19/1/2011	0.300874	0.284566	0.276058	0.295675	0.285275	0.26424	0.294729	0.279367	0.274876	0.295438	0.26424	0.244623
14	20/1/2011	0.291737	0.304264	0.283305	0.284751	0.301132	0.280414	0.273669	0.273428	0.274392	0.275596	0.27656	0.27921
15	21/1/2011	0.328233	0.332058	0.309615	0.288192	0.284621	0.267024	0.266004	0.299413	0.286407	0.266259	0.304769	0.317011
16	24/1/2011	0.088182	0.086591	0.086818	0.085682	0.085455	0.085227	0.085	0.085227	0.085682	0.086818	0.086818	0.087045
17	25/1/2011	0.262408	0.261073	0.249277	0.240819	0.240374	0.2357	0.253728	0.25996	0.260405	0.265524	0.267305	0.266414
18	26/1/2011	0.235308	0.227551	0.228021	0.264927	0.273155	0.272685	0.272685	0.267513	0.228256	0.228256	0.228961	0.228021
19	27/1/2011	0.231361	0.253262	0.257456	0.217148	0.202703	0.244408	0.253262	0.252097	0.253495	0.234157	0.234623	0.234623
20	28/1/2011	0.291333	0.296135	0.276469	0.272353	0.225017	0.239881	0.276012	0.285159	0.283101	0.297279	0.292934	0.264578
21	31/1/2011	0.300819	0.299539	0.294675	0.262929	0.236815	0.285458	0.296467	0.287506	0.280082	0.245264	0.259345	0.272145
22	1/2/2011	0.333333	0.316655	0.302326	0.29058	0.29105	0.321353	0.326521	0.315011	0.314071	0.332864	0.302795	0.260042
23	2/2/2011	0.317909	0.31719	0.338528	0.333733	0.276433	0.273316	0.324622	0.324862	0.314793	0.334932	0.334452	0.312395
24	3/2/2011	0.320198	0.307456	0.294007	0.317131	0.271354	0.274422	0.332704	0.325625	0.310996	0.330109	0.333412	0.310288
25	4/2/2011	0.335086	0.325257	0.317257	0.323886	0.333257	0.306286	0.294857	0.284114	0.261029	0.315429	0.330286	0.303086

ภาพที่ 16 ข้อมูลโพรไฟล์ภาระที่ทำการบรรทัดฐานแล้ว

เมื่อทำการบรรทัดฐานข้อมูลแล้ว จะเริ่มกระบวนการจัดกลุ่ม (Clustering) ทั้ง 4 วิธี ดังนี้

1. Fuzzy C-Means (FCM)

ทำการจัดกลุ่มโดยเลือกใช้ค่า Weighting Parameter (m) = 1.5 และใช้จำนวนกลุ่ม (C) ตั้งแต่ 2-8 กลุ่ม ค่าดัชนี Partition Coefficient (PC) และ Classification Entropy (CE) แสดงดังตารางที่ 1 และ 2 ตามลำดับ

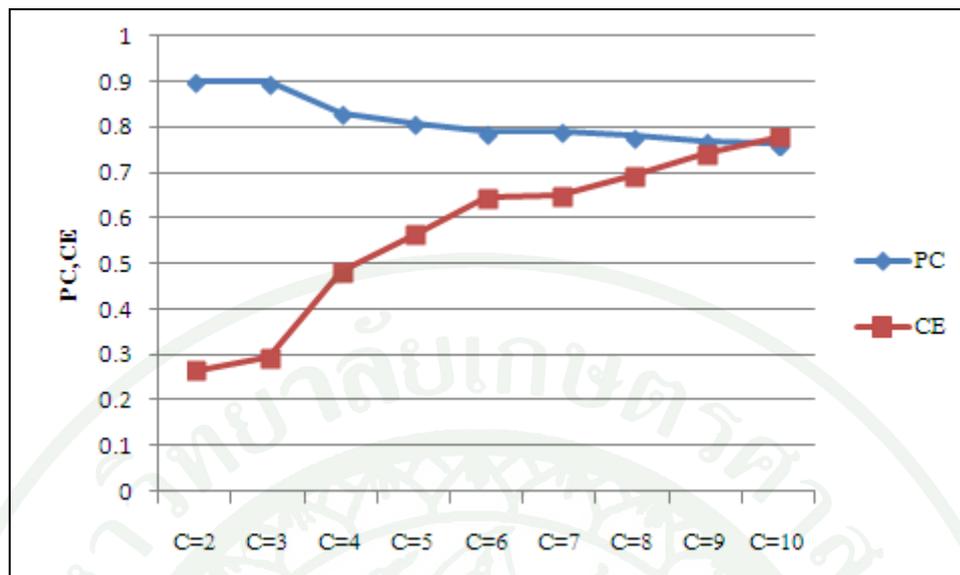
ตารางที่ 1 ค่าดัชนี Partition Coefficient เมื่อทำการจัดกลุ่มโดยวิธี FCM

	C=2	C=3	C=4	C=5	C=6	C=7	C=8	C=9	C=10
PC	0.9002	0.8999	0.8316	0.81	0.789	0.7912	0.7809	0.77	0.7639

ตารางที่ 2 ค่าดัชนี Classification Entropy เมื่อทำการจัดกลุ่มโดยวิธี FCM

	C=2	C=3	C=4	C=5	C=6	C=7	C=8	C=9	C=10
CE	0.2653	0.2937	0.4814	0.5634	0.6433	0.6481	0.692	0.742	0.777

จากตารางที่ 1 และ 2 หลังจากทำการจัดกลุ่มโดยใช้วิธี Fuzzy C-Means แล้วใช้ดัชนี Partition Coefficient และ Classification Entropy ในการตรวจสอบผลการจัดกลุ่ม พบว่าดัชนี Partition Coefficient มีค่าลดลงเมื่อค่า m เพิ่มขึ้น ในขณะที่ดัชนี Classification Entropy มีค่าเพิ่มขึ้นเมื่อค่า m เพิ่มขึ้น ดังนั้นดัชนีทั้ง 2 ค่านี้จะไม่ถูกนำมาใช้ในการเลือกจำนวนกลุ่มที่เหมาะสม เพราะค่าที่ได้มีแนวโน้มเปลี่ยนแปลงไปในทิศทางเดียวอย่างต่อเนื่อง ดังแสดงในภาพที่ 17



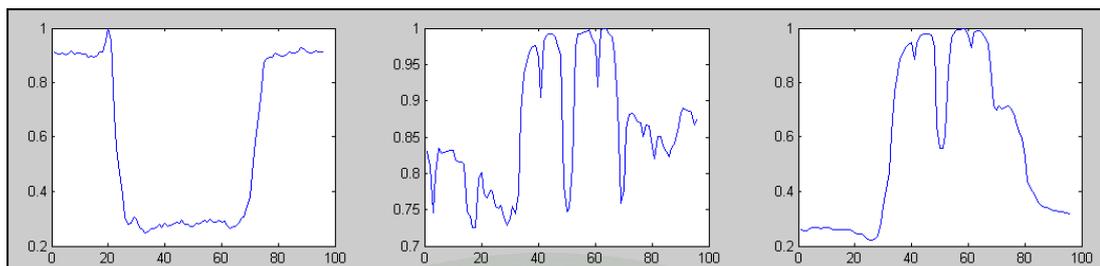
ภาพที่ 17 ค่าดัชนี PC และ CE เมื่อทำการจัดกลุ่มโดยวิธี FCM ที่ค่า C ต่างกัน

ทำการจัดกลุ่มโดยวิธี FCM ใช้ค่า $m = 1.5$ และใช้จำนวนกลุ่ม (C) ตั้งแต่ 3-8 กลุ่ม จากนั้นใช้ดัชนี Xie and Beni (XB) ในการตรวจสอบผลการจัดกลุ่ม ดังแสดงในตารางที่ 3

ตารางที่ 3 ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี FCM

	C=3	C=4	C=5	C=6	C=7	C=8
XB	0.201675	0.519063	0.482612	0.615795	0.615242	0.548368

จากผลการจัดกลุ่มโดยวิธี FCM เมื่อใช้ค่า $m = 1.5$ ได้จำนวนกลุ่มที่เหมาะสมคือ 3 กลุ่ม มี Typical Load Profile ดังแสดงในภาพที่ 18 และมีจำนวนสมาชิกในแต่ละกลุ่มดังแสดงในตารางที่ 4



ภาพที่ 18 Typical Load Profile ของการจัดกลุ่มโดยวิธี FCM

ตารางที่ 4 จำนวนสมาชิกของกลุ่มจากการจัดกลุ่มโดยวิธี FCM

	จำนวนสมาชิก
กลุ่มที่ 1	428
กลุ่มที่ 2	950
กลุ่มที่ 3	1,284
รวม	2,662

2. Fuzzy-Possibilistic C-Means (FPCM)

ทำการจัดกลุ่มโดยเลือกใช้ค่า Weighting Parameter (m) = 1.5 และใช้จำนวนกลุ่ม (C) ตั้งแต่ 2-10 กลุ่ม ค่าดัชนี Partition Coefficient (PC) และ Classification Entropy (CE) แสดงดังตารางที่ 5 และ 6 ตามลำดับ

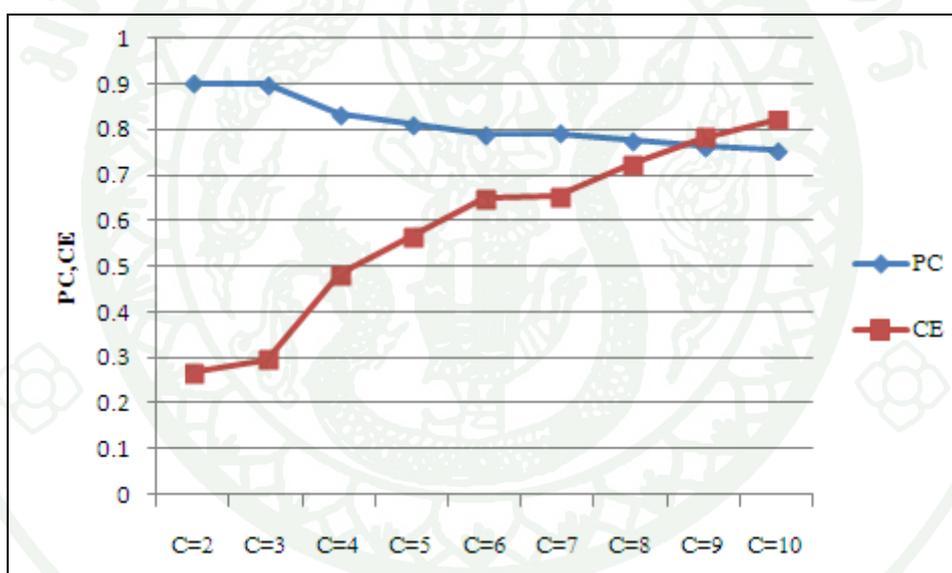
ตารางที่ 5 ค่าดัชนี Partition Coefficient เมื่อทำการจัดกลุ่มโดยวิธี FPCM

	C=2	C=3	C=4	C=5	C=6	C=7	C=8	C=9	C=10
PC	0.901	0.9008	0.8327	0.8111	0.7903	0.7926	0.7757	0.7631	0.7556

ตารางที่ 6 ค่าดัชนี Classification Entropy เมื่อทำการจัดกลุ่มโดยวิธี FPCM

	C=2	C=3	C=4	C=5	C=6	C=7	C=8	C=9	C=10
CE	0.2655	0.2951	0.4837	0.5671	0.6483	0.6547	0.725	0.7832	0.8222

จากตารางที่ 5 และ 6 หลังจากทำการจัดกลุ่มโดยใช้วิธี Fuzzy-Possibilistic C-Means แล้วใช้ดัชนี Partition Coefficient และ Classification Entropy ในการตรวจสอบผลการจัดกลุ่ม พบว่าดัชนี Partition Coefficient มีค่าลดลงเมื่อค่า m เพิ่มขึ้น ในขณะที่ดัชนี Classification Entropy มีค่าเพิ่มขึ้นเมื่อค่า m เพิ่มขึ้น ดังนั้นดัชนีทั้ง 2 ค่านี้จะไม่ถูกนำมาใช้ในการเลือกจำนวนกลุ่มที่เหมาะสม เพราะค่าที่ได้มีแนวโน้มเปลี่ยนแปลงไปในทิศทางเดียวอย่างต่อเนื่อง ดังแสดงในภาพที่ 19



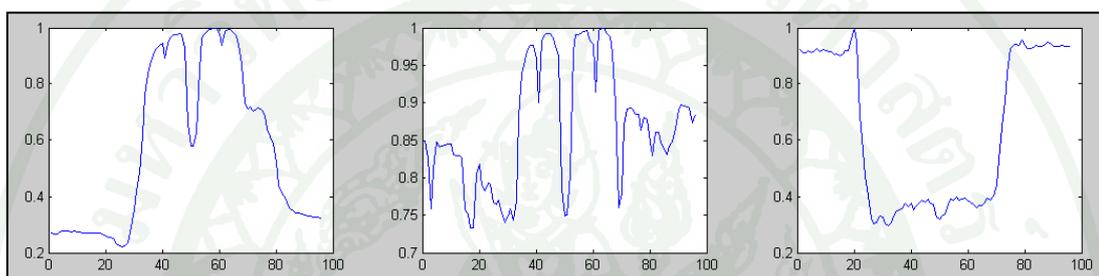
ภาพที่ 19 ค่าดัชนี PC และ CE เมื่อทำการจัดกลุ่มโดยวิธี FPCM ที่ค่า C แตกต่างกัน

ทำการจัดกลุ่มโดยวิธี FPCM ใช้ค่า $m = 1.5$ และใช้จำนวนกลุ่ม (C) ตั้งแต่ 3-8 กลุ่ม จากนั้นใช้ดัชนี Xie and Beni (XB) ในการตรวจสอบผลการจัดกลุ่ม ดังแสดงในตารางที่ 7

ตารางที่ 7 ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี FPCM

	C=3	C=4	C=5	C=6	C=7	C=8
XB	0.201987	0.520172	3.835799	3.91791	4.128782	6.140045

จากผลการจัดกลุ่มโดยวิธี FPCM เมื่อใช้ค่า $m = 1.5$ ได้จำนวนกลุ่มที่เหมาะสมคือ 3 กลุ่ม มี Typical Load Profile ดังแสดงในภาพที่ 20 และมีจำนวนสมาชิกในแต่ละกลุ่มดังแสดงในตารางที่ 8



ภาพที่ 20 Typical Load Profile ของการจัดกลุ่มโดยวิธี FPCM

ตารางที่ 8 จำนวนสมาชิกของกลุ่มจากการจัดกลุ่มโดยวิธี FPCM

	จำนวนสมาชิก
กลุ่มที่ 1	1,315
กลุ่มที่ 2	967
กลุ่มที่ 3	380
รวม	2,662

3. Two-stage Fuzzy C-Means (Two-stage FCM)

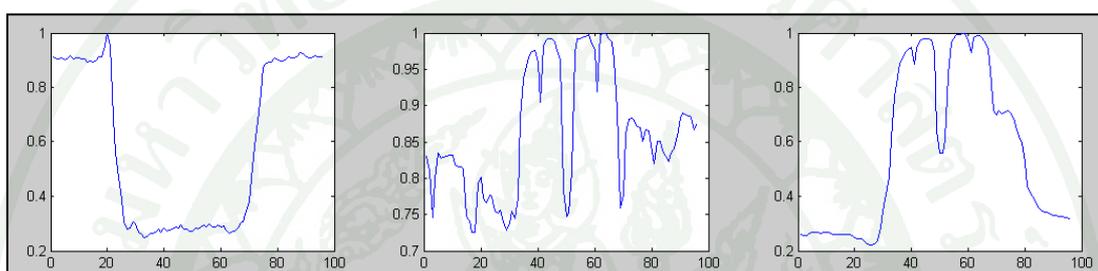
Stage ที่ 1

ทำการจัดกลุ่มโดยเลือกใช้ค่า $m = 1.5$ และใช้จำนวนกลุ่ม (C) ตั้งแต่ 3-8 กลุ่ม จากนั้นใช้ดัชนี Xie and Beni (XB) ในการตรวจสอบผลการจัดกลุ่ม ดังแสดงในตารางที่ 9

ตารางที่ 9 ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 1

	C=3	C=4	C=5	C=6	C=7	C=8
XB	0.201675	0.519063	0.482612	0.615795	0.615242	0.548368

จากผลการจัดกลุ่มโดยวิธี Two-stage FCM เมื่อใช้ค่า $m = 1.5$ ได้จำนวนกลุ่มที่เหมาะสมคือ 3 กลุ่ม มี Typical Load Profile ดังแสดงในภาพที่ 21



ภาพที่ 21 Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 1

จากผลการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 1 ตั้งสมมติฐานว่า ตัวแทนในแต่ละกลุ่มนั้นมีกลุ่มย่อย (Sub-cluster) แฝงอยู่ จึงอาจทำให้ไม่ได้ตัวแทนของกลุ่มที่ต้องการ เพราะฉะนั้นจึงทำการจัดกลุ่มซ้ำ (Re-clustering) เพื่อให้ได้มาซึ่งรูปแบบที่แท้จริงของโพรไฟล์ภาระ

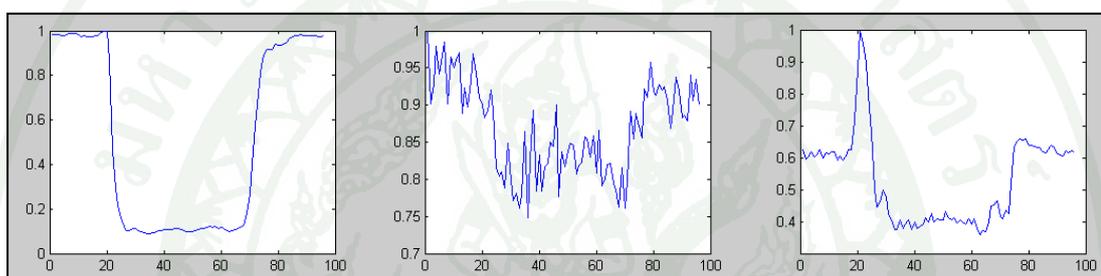
Stage ที่ 2

นำผลการจัดกลุ่มที่ได้จาก stage ที่ 1 มาทำการจัดกลุ่มซ้ำ โดยใช้ค่า $m = 1.5$ และใช้จำนวนกลุ่ม (C) ตั้งแต่ 3-8 กลุ่ม จากนั้นใช้ดัชนี Xie and Beni (XB) ในการตรวจสอบผลการจัดกลุ่ม ดังแสดงในตารางที่ 10-12

ตารางที่ 10 ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1

	C=3	C=4	C=5	C=6	C=7	C=8
XB	1.124219	2.688878	2.373001	2.221316	2.021712	1.873714

จากตารางที่ 10 การจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1 เมื่อใช้ค่า $m = 1.5$ ได้จำนวนกลุ่มที่เหมาะสมคือ 3 กลุ่ม มี Typical Load Profile ดังแสดงในภาพที่ 22

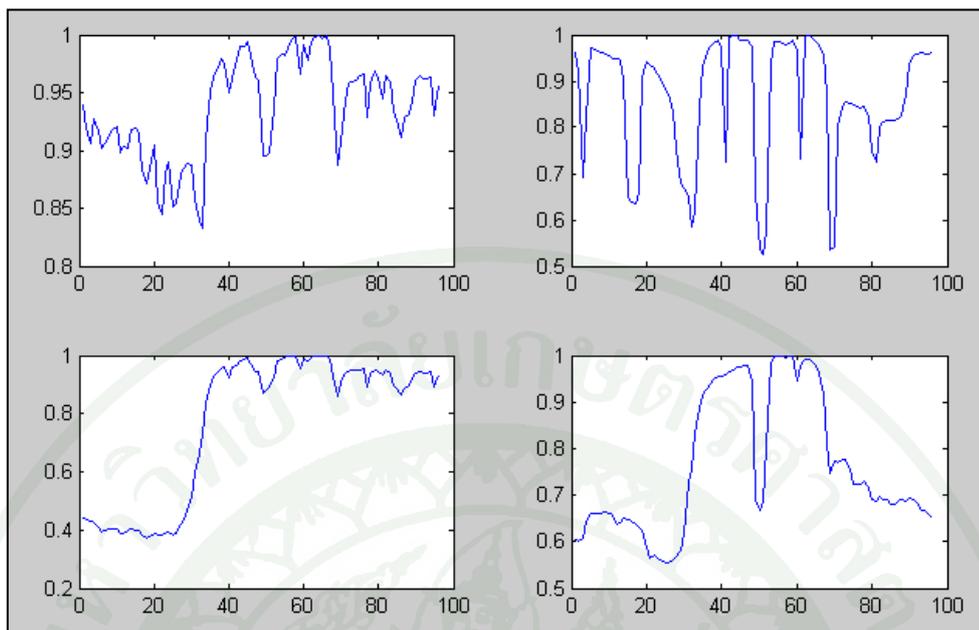


ภาพที่ 22 Typical Load Profile ของการ จัดกลุ่ม โดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1

ตารางที่ 11 ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2

	C=3	C=4	C=5	C=6	C=7	C=8
XB	0.240943	0.182789	0.278497	0.28133	0.240103	0.234117

จากตารางที่ 11 การ จัดกลุ่ม โดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2 ได้จำนวนกลุ่มที่เหมาะสมคือ 4 กลุ่ม และมี Typical Load Profile ดังแสดงในภาพที่ 23

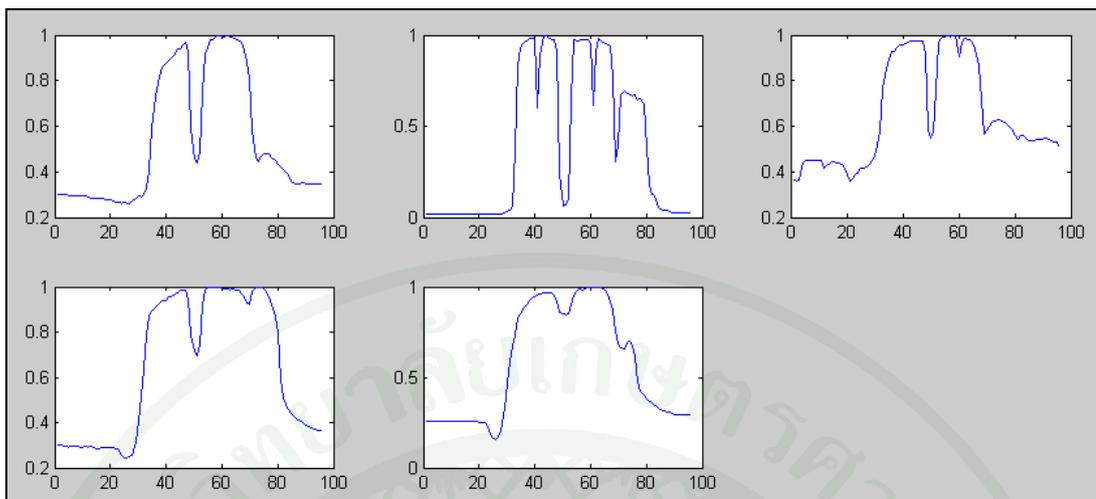


ภาพที่ 23 Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2

ตารางที่ 12 ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3

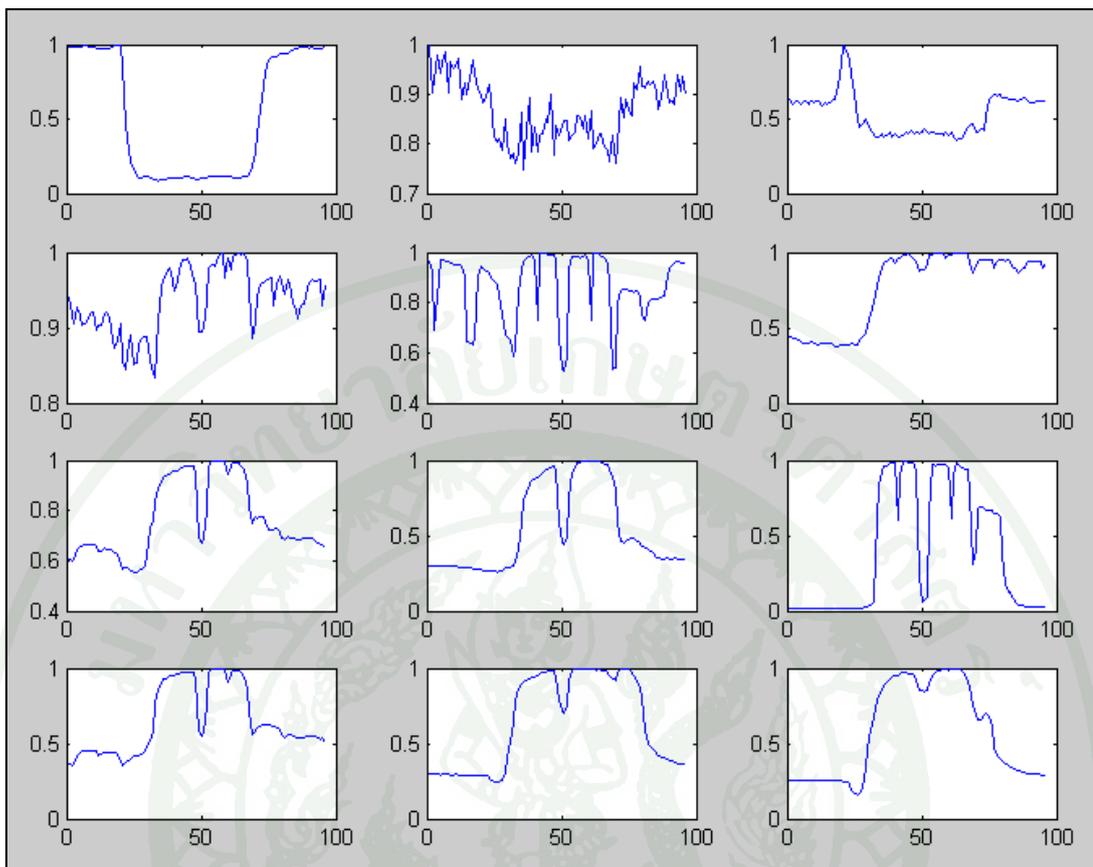
	C=3	C=4	C=5	C=6	C=7	C=8
XB	0.744384	0.729037	0.494416	0.757949	0.93546	0.917943

จากตารางที่ 12 การจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3 ได้จำนวนกลุ่มที่เหมาะสมคือ 5 กลุ่ม และมี Typical Load Profile ดังแสดงในภาพที่ 24



ภาพที่ 24 Typical Load Profile ของการจัดกลุ่ม โดยวิธี Two-stage FCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3

จากผลการจัดกลุ่ม โดยวิธี Two-stage Fuzzy C-Means ให้ผลการจัดกลุ่มทั้งสิ้น 12 กลุ่ม ดังแสดงในภาพที่ 25 และมีจำนวนสมาชิกของกลุ่มดังแสดงในตารางที่ 13



ภาพที่ 25 Typical Load Profile ของการจัดกลุ่ม โดยวิธี Two-stage Fuzzy C-Means

ตารางที่ 13 จำนวนสมาชิกของกลุ่มจากการจัดกลุ่มโดยวิธี Two-stage Fuzzy C-Means

	จำนวนสมาชิก
กลุ่มที่ 1	165
กลุ่มที่ 2	170
กลุ่มที่ 3	93
กลุ่มที่ 4	367
กลุ่มที่ 5	245
กลุ่มที่ 6	64
กลุ่มที่ 7	274
กลุ่มที่ 8	287

ตารางที่ 13 (ต่อ)

	จำนวนสมาชิก
กลุ่มที่ 9	235
กลุ่มที่ 10	233
กลุ่มที่ 11	287
กลุ่มที่ 12	242
รวม	2,662

4. Two-stage Fuzzy-Possibilistic C-Means (Two-stage FPCM)

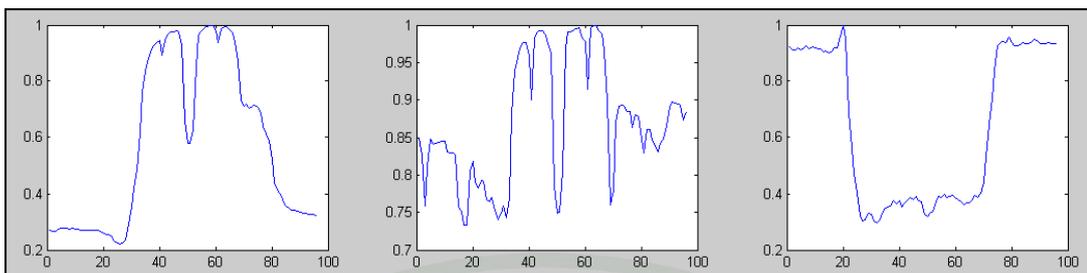
Stage ที่ 1

ทำการจัดกลุ่มโดยเลือกใช้ค่า $m = 1.5$ และใช้จำนวนกลุ่ม (C) ตั้งแต่ 3-8 กลุ่ม จากนั้นใช้ดัชนี Xie and Beni (XB) ในการตรวจสอบผลการจัดกลุ่ม ดังแสดงในตารางที่ 14

ตารางที่ 14 ค่าดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 1

	C=3	C=4	C=5	C=6	C=7	C=8
XB	0.201987	0.520172	3.835799	3.91791	4.128782	6.140045

จากตารางที่ 14 การจัดกลุ่มโดยวิธี Two-stage FPCM เมื่อใช้ค่า $m = 1.5$ ได้จำนวนกลุ่มที่เหมาะสมคือ 3 กลุ่ม และมี Typical Load Profile ดังแสดงในภาพที่ 26



ภาพที่ 26 Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 1

จากผลการจัดกลุ่มโดยวิธี Two-stage FCM ใน stage ที่ 1 ตั้งสมมติฐานว่า ตัวแทนในแต่ละกลุ่มนั้นมีกลุ่มย่อย (Sub-cluster) แฝงอยู่ จึงอาจทำให้ไม่ได้ตัวแทนของกลุ่มที่ต้องการ เพราะฉะนั้นจึงทำการจัดกลุ่มซ้ำ (Re-clustering) เพื่อให้ได้มาซึ่งรูปแบบที่แท้จริงของโพรไฟล์ภาระ

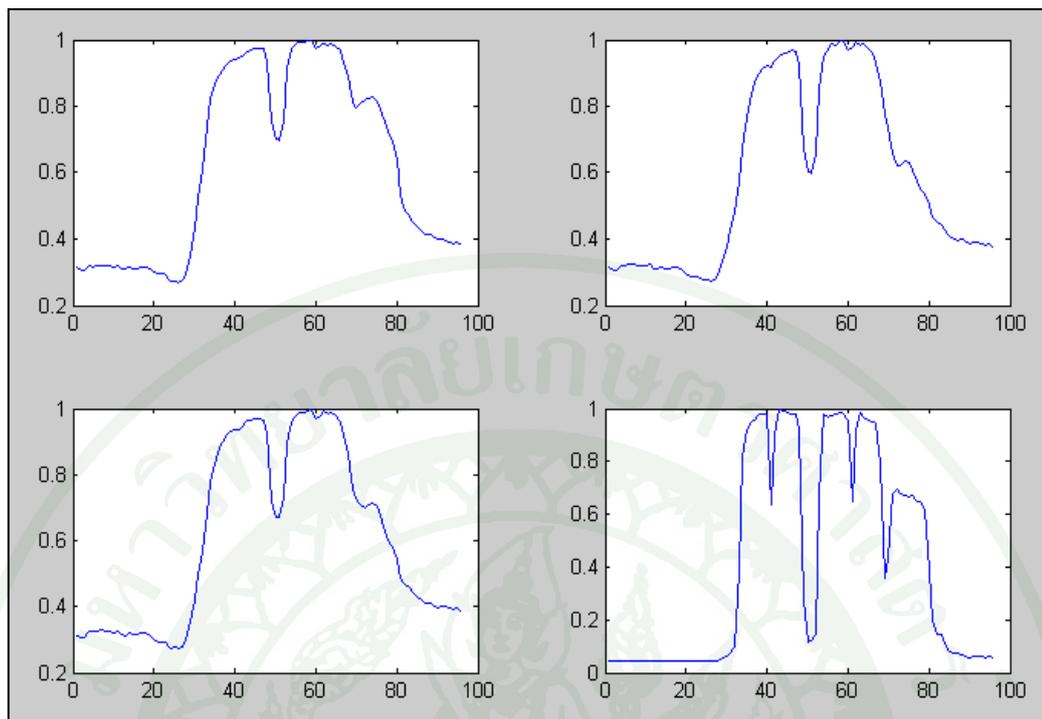
Stage ที่ 2

นำผลการจัดกลุ่มที่ได้จาก stage ที่ 1 มาทำการจัดกลุ่มซ้ำ โดยใช้ค่า $m = 1.5$ และใช้จำนวนกลุ่ม (C) ตั้งแต่ 3-8 กลุ่ม จากนั้นใช้ดัชนี Xie and Beni (XB) ในการตรวจสอบผลการจัดกลุ่ม ดังแสดงในตารางที่ 15-17

ตารางที่ 15 ดัชนี Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1

	C=3	C=4	C=5	C=6	C=7	C=8
XB	1.668689	0.850057	1.58761	1.777704	1.788927	2.936067

จากตารางที่ 15 การจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1 เมื่อใช้ค่า $m = 1.5$ ได้จำนวนกลุ่มที่เหมาะสมคือ 4 กลุ่ม และมี Typical Load Profile ดังแสดงในภาพที่

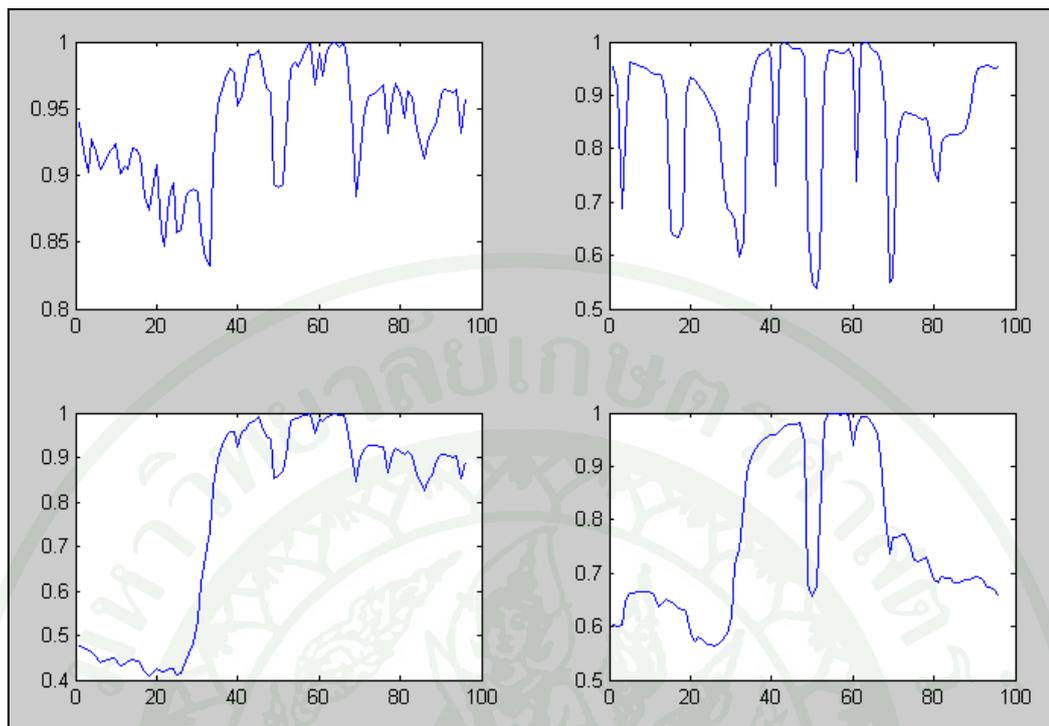


ภาพที่ 27 Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 1

ตารางที่ 16 คำนวณ Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2

	C=3	C=4	C=5	C=6	C=7	C=8
XB	0.254329	0.193857	1.852	2.136573	1.841469	3.580387

จากตารางที่ 16 การจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2 ได้จำนวนกลุ่มที่เหมาะสมคือ 4 กลุ่ม และมี Typical Load Profile ดังแสดงในภาพที่ 28

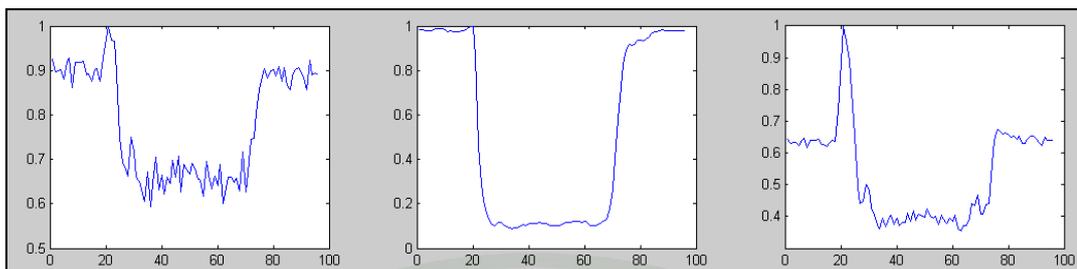


ภาพที่ 28 Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 2

ตารางที่ 17 คำนวณ Xie and Beni เมื่อทำการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3

	C=3	C=4	C=5	C=6	C=7	C=8
XB	2.740781	3.295582	7.452309	7.211834	8.27914	4.339771

จากตารางที่ 17 การจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3 ได้จำนวนกลุ่มที่เหมาะสมคือ 3 กลุ่ม และมี Typical Load Profile ดังแสดงในภาพที่ 29



ภาพที่ 29 Typical Load Profile ของการจัดกลุ่มโดยวิธี Two-stage FPCM ใน stage ที่ 2 ของกลุ่มหลักที่ 3

จากผลการจัดกลุ่มซ้ำสามารถจัดกลุ่มได้ทั้งสิ้น 11 กลุ่ม แต่เมื่อทำการพิจารณาพบว่าบางกลุ่มนั้นมีลักษณะการใช้ไฟฟ้าในลักษณะคล้ายกันหรือลักษณะใกล้เคียงกัน สามารถรวมอยู่ในกลุ่มเดียวกันได้ จึงมีการใช้ Goodness of fit มาเป็นเครื่องมือในการตัดสินใจว่าลักษณะการใช้ไฟฟ้ากลุ่มใดบ้างที่มีลักษณะคล้ายกันหรือต่างกัน โดยเลือกใช้วิธี Kolmogorov-Smirnov Two-Sample Test ดังแสดงในตารางที่ 18-20

ตารางที่ 18 ค่า D ที่ใช้ในการทดสอบสมมติฐานของกลุ่มหลักที่ 1

	กลุ่มย่อยที่ 1	กลุ่มย่อยที่ 2	กลุ่มย่อยที่ 3	กลุ่มย่อยที่ 4
กลุ่มย่อยที่ 1	0	0.1875	0.2292	0.5417
กลุ่มย่อยที่ 2	0.1875	0	0.2188	0.5417
กลุ่มย่อยที่ 3	0.2292	0.2188	0	0.5313
กลุ่มย่อยที่ 4	0.5417	0.5417	0.5313	0

จากตารางที่ 18 เมื่อพิจารณาที่ค่า $D < D_{critical}$ สามารถสรุปได้ว่า กลุ่มย่อยที่ 1 เหมือนกับกลุ่มย่อยที่ 2, กลุ่มย่อยที่ 1 เหมือนกับกลุ่มย่อยที่ 3 และกลุ่มย่อยที่ 2 เหมือนกับกลุ่มย่อยที่ 3 ดังนั้นในกลุ่มหลักที่ 1 จะเหลือกลุ่มย่อย 2 กลุ่ม

ตารางที่ 19 ค่า D ที่ใช้ในการทดสอบสมมติฐานของกลุ่มหลักที่ 2

	กลุ่มย่อยที่ 1	กลุ่มย่อยที่ 2	กลุ่มย่อยที่ 3	กลุ่มย่อยที่ 4
กลุ่มย่อยที่ 1	0	0.375	0.3542	0.6771
กลุ่มย่อยที่ 2	0.375	0	0.3125	0.4375
กลุ่มย่อยที่ 3	0.3542	0.3125	0	0.3333
กลุ่มย่อยที่ 4	0.6771	0.4375	0.3333	0

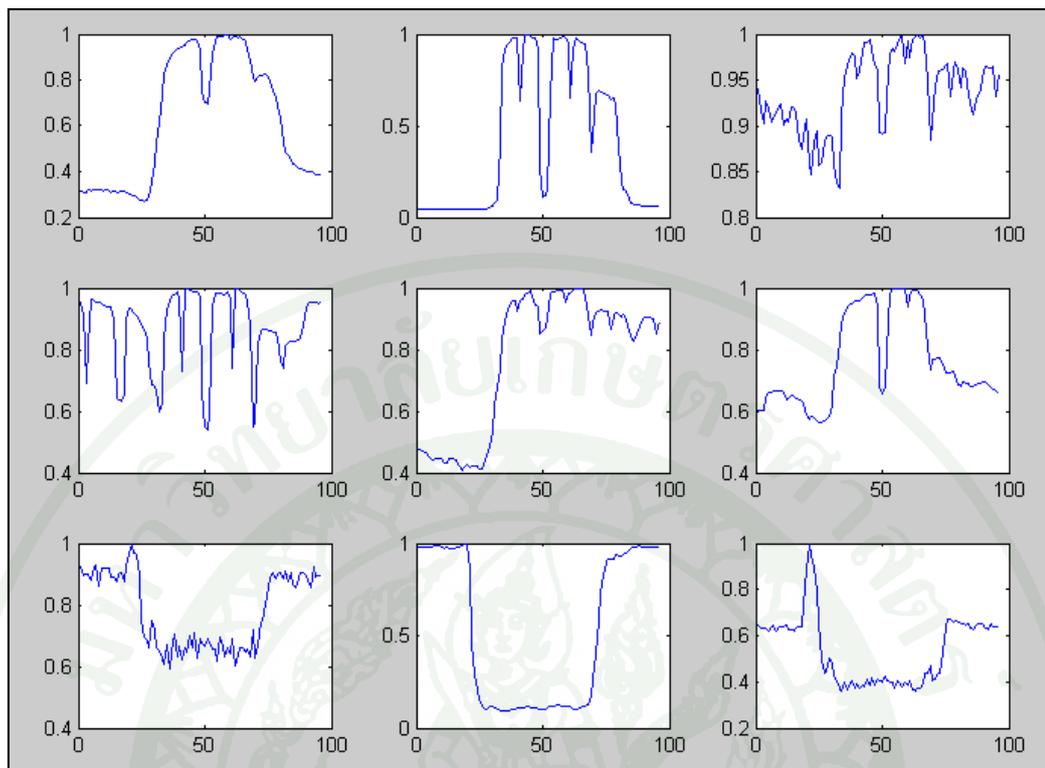
จากตารางที่ 19 เมื่อพิจารณาที่ค่า $D < D_{critical}$ ไม่มีกลุ่มย่อยใดมีลักษณะเหมือนกัน ดังนั้นในกลุ่มหลักที่ 2 มี 4 กลุ่มย่อย

ตารางที่ 20 ค่า D ที่ใช้ในการทดสอบสมมติฐานของกลุ่มหลักที่ 3

	กลุ่มย่อยที่ 1	กลุ่มย่อยที่ 2	กลุ่มย่อยที่ 3
กลุ่มย่อยที่ 1	0	0.5313	0.9479
กลุ่มย่อยที่ 2	0.5313	0	0.5
กลุ่มย่อยที่ 3	0.9479	0.5	0

จากตารางที่ 20 เมื่อพิจารณาที่ค่า $D < D_{critical}$ ไม่มีกลุ่มย่อยใดมีลักษณะเหมือนกัน ดังนั้นในกลุ่มหลักที่ 3 มี 3 กลุ่มย่อย

จากผลการทดสอบลักษณะความคล้ายกันดังตารางที่ 18-20 สามารถสรุปได้ว่า เมื่อทำการจัดกลุ่มโดยใช้วิธี Two-stage FPCM แล้ว สามารถจัดกลุ่มได้ทั้งสิ้น 9 กลุ่ม และมี Typical Load Profile ดังแสดงในภาพที่ 30 และมีจำนวนสมาชิกในแต่ละกลุ่มดังแสดงในตารางที่ 21



ภาพที่ 30 Typical Load Profile ของการจัดกลุ่ม โดยวิธี Two-stage FPCM

ตารางที่ 21 จำนวนสมาชิกของกลุ่มจากการจัดกลุ่ม โดยวิธี Two-stage FPCM

	จำนวนสมาชิก
กลุ่มที่ 1	1,078
กลุ่มที่ 2	237
กลุ่มที่ 3	367
กลุ่มที่ 4	245
กลุ่มที่ 5	66
กลุ่มที่ 6	289
กลุ่มที่ 7	132
กลุ่มที่ 8	164
กลุ่มที่ 9	84
รวม	2,662

5. สรุปผลการจัดกลุ่มข้อมูลโพรไฟล์ภาระ

การจัดกลุ่มโพรไฟล์ภาระโดยใช้วิธีข้างต้น 4 วิธี คือ Fuzzy C-Means, Fuzzy-Possibilistic C-Means, Two-stage Fuzzy C-Means และ Two-stage Fuzzy-Possibilistic C-Means ได้ผลการจัดกลุ่มดังแสดงในตารางที่ 22

ตารางที่ 22 ผลการจัดกลุ่มโพรไฟล์ภาระ

ลำดับที่	วิธีการจัดกลุ่ม	จำนวนกลุ่ม
1	Fuzzy C-Means	3
2	Fuzzy-Possibilistic C-Means	3
3	Two-stage Fuzzy C-Means	12
4	Two-stage Fuzzy-Possibilistic C-Means	9

วิจารณ์

การจัดกลุ่มโพรไฟล์ภาระ (Load Profile Clustering) โดยใช้วิธีที่แตกต่างกัน 4 วิธี คือ Fuzzy C-Means, Fuzzy-Possibilistic C-Means, Two-stage Fuzzy C-Means และ Two-stage Fuzzy-Possibilistic C-Means ให้ผลการจัดกลุ่มที่แตกต่างกัน คือ วิธี Fuzzy C-Means และ Fuzzy-Possibilistic C-Means ให้ผลการจัดกลุ่มเป็น 3 กลุ่ม ซึ่งเมื่อนำรูปแบบของทั้ง 3 กลุ่มมาพิจารณา พบว่าไม่ครอบคลุมโพรไฟล์ภาระในบางรูปแบบ เนื่องจากว่าทั้ง 2 วิธีนี้เป็นการจัดกลุ่มเพียงครั้งเดียว ส่วนวิธี Two-stage Fuzzy C-Means และ Two-stage Fuzzy-Possibilistic C-Means ให้ผลการจัดกลุ่มเป็น 12 และ 9 กลุ่มตามลำดับ ซึ่งเมื่อนำรูปแบบของทั้ง 12 และ 9 กลุ่มมาพิจารณา พบว่าครอบคลุมโพรไฟล์ภาระในทุกรูปแบบ เนื่องจากว่า 2 วิธีนี้เป็นการจัดกลุ่มถึง 2 ครั้ง แต่วิธี Two-stage Fuzzy-Possibilistic C-Means จะมีข้อดีกว่าวิธี Two-stage Fuzzy C-Means ตรงที่สามารถจัดการกับข้อมูลที่เป็น outlier หรือ noise ได้ดีกว่า และเมื่อทำการประยุกต์ใช้ทฤษฎี Goodness of Fit เข้ากับผลจากการจัดกลุ่มโดยวิธี Two-stage Fuzzy-Possibilistic C-Means ทำให้ได้ผลการจัดกลุ่มที่มีความกระชับมากยิ่งขึ้น

ผลที่ได้จากการทดลองนี้เป็นไปตามเป้าหมายที่ตั้งไว้คือ วิธี Two-stage Fuzzy C-Means และ Two-stage Fuzzy-Possibilistic C-Means ซึ่งเป็นการจัดกลุ่ม 2 ครั้งนั้นจะให้ผลการจัดกลุ่มที่ครอบคลุมมากกว่าวิธี Fuzzy C-Means และ Fuzzy-Possibilistic C-Means ซึ่งเป็นการจัดกลุ่มเพียงครั้งเดียว เนื่องจากการจัดกลุ่ม 2 ครั้งนั้นทำให้ได้รูปแบบบางรูปแบบที่ซ่อนอยู่ในการจัดกลุ่มครั้งแรก

การทดลองนี้มีประโยชน์คือ ทั้งการไฟฟ้าฯ หรือผู้ใช้ไฟฟ้า สามารถที่จะนำวิธีการจัดกลุ่มโพรไฟล์ภาระดังที่นำเสนอไปใช้ต่อยอดในการวางแผนต่างๆ ได้ เช่น การไฟฟ้าฯ สามารถนำวิธีการจัดกลุ่มโพรไฟล์ภาระไปใช้ เพื่อวางแผนเรื่องการคิดอัตราค่าไฟฟ้า การวางแผนการผลิตไฟฟ้า หรือการวางแผนด้านการส่งจ่ายไฟฟ้า ส่วนผู้ใช้ไฟฟ้า สามารถนำวิธีการจัดกลุ่มโพรไฟล์ภาระไปใช้ในการวางแผนการใช้ไฟฟ้า วางแผนการผลิตต่างๆ เพื่อให้ใช้ไฟฟ้าอย่างมีประสิทธิภาพมากที่สุดและลดค่าไฟฟ้าต่อเดือนลงได้

สรุปและข้อเสนอแนะ

สรุป

วิทยานิพนธ์นี้นำเสนอวิธีการจัดกลุ่มโพรไฟล์ภาระ (Load Profile Clustering) โดยใช้วิธีการแตกต่างกันออกไป 4 วิธี ได้แก่ Fuzzy C-means, Fuzzy-possibilistic C-means, Two-stage fuzzy C-means และ Two-stage fuzzy-possibilistic C-means ซึ่งผลการจัดกลุ่มโพรไฟล์ภาระโดยใช้วิธี Two-stage fuzzy-possibilistic C-means พบว่าในขั้นที่ 1 สามารถจัดกลุ่มโพรไฟล์ภาระได้ 3 กลุ่ม เมื่อนำโพรไฟล์ภาระมาจัดกลุ่มซ้ำเพื่อหารูปแบบที่แท้จริงของโพรไฟล์ภาระและใช้ Goodness of Fit เป็นเครื่องมือในการตรวจสอบความคล้ายกัน สามารถจัดกลุ่มได้ทั้งหมด 9 กลุ่ม ซึ่งเมื่อเปรียบเทียบกับวิธีการจัดกลุ่มรอบเดียวซึ่งได้จำนวนกลุ่มคือ 3 กลุ่ม พบว่าการจัดกลุ่มซ้ำสามารถแสดงถึงกลุ่มหลักที่แท้จริงของโพรไฟล์ภาระได้ชัดเจนกว่า เมื่อเปรียบเทียบผลการจัดกลุ่มระหว่างวิธีต่างๆ ข้างต้นกับวิธีที่นำเสนอในวิทยานิพนธ์นี้ ผลการเปรียบเทียบแสดงในตารางที่ 22 ซึ่งเห็นได้ว่าวิธี Two-stage fuzzy C-means สามารถจัดกลุ่มได้ 12 กลุ่ม แต่เมื่อพิจารณาพบว่ารูปแบบของบางกลุ่มมีลักษณะคล้ายกันมากสามารถที่จะจัดให้อยู่ในกลุ่มเดียวกันได้ เพราะฉะนั้น วิธี Two-stage fuzzy-possibilistic C-means ที่ได้นำเสนอในวิทยานิพนธ์นี้เป็นวิธีการจัดกลุ่มที่เหมาะสม คือ ให้ผลการจัดกลุ่มที่กระชับและครอบคลุมกว่า

ข้อเสนอแนะ

การจัดกลุ่มโพรไฟล์ภาระ (Load Profile Clustering) ที่ได้นำเสนอในวิทยานิพนธ์นี้ยังมีส่วนที่สามารถปรับปรุงให้ดีขึ้นได้อีก คือ ในขั้นตอน Pre-processing น่าจะมีการระบุวันหยุดที่แน่นอนของผู้ใช้ไฟฟ้าแต่ละรายไปเลย ว่าวันไหนเป็นวันหยุดจริงๆ เพราะผู้ใช้ไฟฟ้าแต่ละรายอาจมีวันหยุดที่แตกต่างกันออกไป ซึ่งถ้าหากสามารถทำในขั้นตอนนี้เพิ่มขึ้นได้ ก็จะช่วยเพิ่มความถูกต้องในการจัดกลุ่มมากขึ้น เนื่องจากว่า หากไม่ได้มีการระบุวันหยุดที่แท้จริง ทำให้เรานำข้อมูลวันหยุดบางวันมารวมในการจัดกลุ่มด้วย ผลการจัดกลุ่มที่ได้ ก็จะมีกลุ่มที่เป็นวันหยุดรวมอยู่ด้วย ซึ่งกลุ่มที่เป็นวันหยุดนั้น เป็นข้อมูลที่เราไม่สนใจเนื่องจากมีปริมาณการใช้ไฟฟ้าน้อยมาก แต่การที่จะระบุวันหยุดที่แน่นอนของผู้ใช้ไฟฟ้าแต่ละรายนั้น ก็ต้องใช้เวลาพอสมควรในการขอข้อมูลเพิ่มเติมจากผู้ใช้ไฟฟ้า

เอกสารและสิ่งอ้างอิง

- ชินกร น้อยคำยาง. 2552. การเปรียบเทียบผลการทดสอบสถิติพารามตริกและนอนพารามตริกในการทดสอบความแตกต่างระหว่างประชากรสองกลุ่มที่อิสระจากกัน. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ชำนาญ ห่อเกียรติ. 2550. ถาม-ตอบ ไฟฟ้ากำลัง. พิมพ์ครั้งที่ 2. บริษัท จรัสสนิทวงส์การพิมพ์ จำกัด, กรุงเทพฯ.
- นิภา ศรีไพโรจน์. 2533. สถิตินอนพารามตริก. พิมพ์ครั้งที่ 2. บริษัท โอเดียนส โตร์ จำกัด, กรุงเทพฯ.
- Beni, G. and X. L. Xie. 1991. A Validity Measure for Fuzzy Clustering. **IEEE Transaction on Pattern Analysis and Machine Intelligence** PAMI 13(8): 841-847.
- Bezdek, J. C. and J. D. Harris. 1978. Fuzzy partitions and relations; an axiomatic basis for clustering. **Fuzzy Sets and Systems** 1: 111-127.
- Gašperic, S. , D. Gerbec, F. Gubina and I. Šmon. 2002. An Approach to Customers Daily Load Profile Determination. **Power Engineering Society Summer Meeting IEEE** 1: 587-591.
- Hyun, S. R. and W. O. Kyung. 1996. A Validity Measure for Fuzzy Clustering and Its Use in Selecting Optimal Number of Clusters, pp. 1020-1025. *In* 5th **IEEE International Conference on Fuzzy Systems**. Seoul, Korea.
- Kumar, V. , M. Steinbach and P. Tan. 2005. **An Introduction to Data Mining**. Available Source: <http://www-users.cs.umn.edu/~kumar/dmbook/index.php>, April 4, 2012.

Lo, K. L. , M. H. Sohod and Z. Zakaria. 2005. Determination of Consumers' Load Profiles based on Two-stage Fuzzy C-Means, pp. 212-217. *In Proceedings of the 5th WSEAS International Conference on Power Systems and Electromagnetic Compatibility.* 23-25 August 2005, Corfu, Greece.

Martin, T. , C. Zhang and Y. Zhou. 2008. A Validity Index for Fuzzy and Possibilistic C-means Algorithm, pp. 877-882. *In Proceedings of IPMU'08.* 22-27 June 2008, Malaga, Spain.

Sathiracheewin, S. and V. Surapatana. 2011. Daily Typical Load Clustering of Residential Customers, pp. 797-800. *In 8th Electrical Engineering/Electronic, Computer Telecommunications and Information Technology (ECTI) Association of Thailand Conference.* Khon Kaen University, Thailand.

ประวัติการศึกษา และการทำงาน

ชื่อ	นายกรัณย์ ตันทวีวงศ์
เกิดวันที่	1 พฤศจิกายน 2530
สถานที่เกิด	จังหวัดภูเก็ต
ประวัติการศึกษา	วิศวกรรมศาสตรบัณฑิต (วิศวกรรมไฟฟ้า) มหาวิทยาลัยเกษตรศาสตร์
ตำแหน่งปัจจุบัน	-
สถานที่ทำงานปัจจุบัน	-
ผลงานดีเด่นและ/หรือรางวัลทางวิชาการ	-
ทุนการศึกษาที่ได้รับ	-