



ใบรับรองวิทยานิพนธ์
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมไฟฟ้า)

ปริญญา

วิศวกรรมไฟฟ้า	วิศวกรรมไฟฟ้า
สาขา	ภาควิชา
เรื่อง	การวิเคราะห์กลุ่มข้อมูลลักษณะการใช้ไฟฟ้าโดยใช้วิธีการทำงานร่วมกันของเคมีนและเทคนิคการขยายพื้นที่บนพื้นฐานของข่ายงานประสาทแบบก่อดำด้วยตนเอง A Combined Method of K-Means and Self-Organizing Map based Region-Growing Technique for Electric Load Profile Clustering
นามผู้วิจัย	นายวรายุทธ พราหมณ์ทอง
ได้พิจารณาเห็นชอบโดย	
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	(อาจารย์พิสุทธิ รัชศักดิ์, Ph.D.)
หัวหน้าภาควิชา	(รองศาสตราจารย์วิชัย สุระพัฒน์, วศ.ม.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญญา ชีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ เดือน พ.ศ.

สืบสินธุ์ มหาวิทยาลัยเกษตรศาสตร์

วิทยานิพนธ์

เรื่อง

การวิเคราะห์กลุ่มข้อมูลลักษณะการใช้ไฟฟ้าโดยใช้วิธีการทำงานร่วมกันของเคมีน
และเทคนิคการขยายพื้นที่บนพื้นฐานของข่ายงานประสาทแบบก่อดำด้วยตนเอง

A Combined Method of K-Means and Self-Organizing Map based
Region-Growing Technique for Electric Load Profile Clustering

โดย

นายวราวุธ พราหมณ์ทอง

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์
เพื่อความสมบูรณ์แห่งปริญญาวิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมไฟฟ้า)

พ.ศ. 2557

วราวุธ พรหมณ์ทอง 2557: การวิเคราะห์กลุ่มข้อมูลลักษณะการใช้ไฟฟ้าโดยใช้
วิธีการทำงานร่วมกันของเคมีน และเทคนิคการขยายพื้นที่บนพื้นฐานของข่ายงาน
ประสาทแบบก่อดำด้วยตนเอง ปริญญาวิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมไฟฟ้า)
สาขาวิศวกรรมไฟฟ้า ภาควิชาวิศวกรรมไฟฟ้า อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก:
อาจารย์พิสุทธิ์ รพีศักดิ์, Ph.D. 53 หน้า

วิทยานิพนธ์นี้นำเสนอการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าเพื่อค้นหา
รูปแบบตัวอย่างโดยใช้วิธีแบบผสม การทำงานเริ่มด้วยการค้นหาตำแหน่งศูนย์กลางกลุ่มของ
ข้อมูลอาศัยขั้นตอนของวิธีเคมีน เพื่อใช้เป็นข้อมูลนำเข้าสู่การวิเคราะห์กลุ่มด้วยเทคนิคการขยาย
พื้นที่อาศัยความเชื่อมโยงของเซลล์บนข่ายงานประสาทที่จัดเตรียมไว้ และเมื่อนำผลจากการ
วิเคราะห์กลุ่มด้วยวิธีที่เสนอเปรียบเทียบกับผลการวิเคราะห์ของวิธีเคมีน และเทคนิคการขยาย
พื้นที่แบบดั้งเดิม โดยพิจารณาจากค่าประสิทธิภาพที่คำนวณได้ ประกอบด้วย ค่าความกระชับ
ภายในกลุ่ม และความต่างระหว่างกลุ่ม พบว่าวิธีการตามที่เสนอ ช่วยให้การวิเคราะห์กลุ่มข้อมูล
ของลักษณะการใช้ไฟฟ้าให้ค่าประสิทธิภาพที่ดีกว่า

ลายมือชื่อนิสิต

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

Warayuth Phramthong 2014: A Combined Method of K-Means and Self-Organizing Map based Region-Growing Technique for Electric Load Profile Clustering.
Master of Engineering (Electrical Engineering), Major Field: Electrical Engineering,
Department of Electrical Engineering. Thesis Advisor: Mr. Pisut Raphisak, Ph.D.
53 pages.

This thesis proposes a combined clustering method to generate typical daily load profile. The process analysis starts with K-means in order to provide the cluster centers for selecting the suitable neurons for forming clusters by the Region-Growing technique based on neighborhood relation achieved by Self-Organizing Map. In addition, for the cluster validation process, the compactness of the obtained clusters and the separation between clusters have been used to measure and compare results from different methods: the proposed method, K-means and traditional Region-Growing respectively. After comparison, the result of the proposed method has a better performance.

Student's signature

Thesis Advisor's signature

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จได้ด้วยความอนุเคราะห์จาก อ.พิสุทธิ์ รพีศักดิ์ อาจารย์ที่ปรึกษา วิทยานิพนธ์หลัก ซึ่งเป็นทั้งผู้ให้ความรู้ คำแนะนำและคำปรึกษา จนทำให้วิทยานิพนธ์นี้สำเร็จมาได้ด้วยดี

ขอขอบพระคุณบิดา มารดา และทุกคนในครอบครัว ที่เป็นกำลังใจและให้การสนับสนุนมา โดยตลอดในทุกๆ ด้าน ขอขอบคุณเพื่อนๆ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ที่ให้ กำลังใจและอยู่เคียงข้างกันมาโดยตลอด ขอขอบพระคุณอาจารย์ภาควิชาวิศวกรรมไฟฟ้าทุกท่าน ที่ได้อบรมสั่งสอนและมอบความรู้อันเป็นประโยชน์เพื่อนำไปใช้ในอนาคต ขอขอบคุณพี่ๆ และ เพื่อนๆ ในโครงการวิจัยและพัฒนาความชำนาญด้านไฟฟ้ากำลังที่สนับสนุนรวมถึงให้คำแนะนำในการเรียนและการเขียนวิทยานิพนธ์ฉบับนี้จนสำเร็จ

สุดท้ายนี้ ข้าพเจ้าขอขอบพระคุณ โครงการจัดการด้านการใช้ไฟฟ้าให้มีประสิทธิภาพกับ ผู้ใช้ไฟฟ้าที่มีตัวประกอบการใช้ไฟฟ้า (Load Factor) ค่า ของการไฟฟ้านครหลวง ที่สนับสนุน ข้อมูลที่ใช้ในการทำวิทยานิพนธ์นี้จนสำเร็จ

วรายุทธ พราหมณ์ทอง

มิถุนายน 2557

สารบัญ

	หน้า
สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(3)
คำนำ	1
วัตถุประสงค์	2
การตรวจเอกสาร	3
อุปกรณ์และวิธีการ	21
อุปกรณ์	21
วิธีการ	21
ผลและวิจารณ์	34
ผล	34
วิจารณ์	48
สรุปและข้อเสนอแนะ	49
สรุป	49
ข้อเสนอแนะ	49
เอกสารและสิ่งอ้างอิง	50
ประวัติการศึกษาและการทำงาน	53

สารบัญตาราง

ตารางที่		หน้า
1	รายละเอียดของข่างานประสาทจากการจัดเตรียมตามเงื่อนไขที่กำหนด	32
2	ผลการคำนวณค่าประสิทธิภาพด้วยดัชนี WSJ เมื่อทำการวิเคราะห์กลุ่มตามขั้นตอนของวิธีแบบเป็นลำดับชั้น ในช่วง 2-10 กลุ่ม	34
3	ค่าประสิทธิภาพของดัชนี WSJ และขนาดของข่างานประสาทที่มีประสิทธิภาพต่อการวิเคราะห์ด้วยจำนวนกลุ่มตามที่กำหนด ตั้งแต่ 2-10 กลุ่ม	37
4	จำนวนกลุ่มที่เหมาะสม และค่าประสิทธิภาพ เมื่อทำการวิเคราะห์กลุ่มด้วยวิธีแบบเป็นลำดับชั้นของแต่ละข่างานประสาทที่จัดเตรียมไว้	37
5	ค่าประสิทธิภาพของดัชนี WSJ และขนาดข่างานประสาทที่สอดคล้องกับผลการวิเคราะห์ที่มีประสิทธิภาพตามจำนวนกลุ่มที่กำหนด ในช่วง 2-10 กลุ่ม	40
6	จำนวนกลุ่มที่เหมาะสม และค่าประสิทธิภาพ เมื่อทำการวิเคราะห์กลุ่มด้วยวิธีขยายพื้นที่บนพื้นฐานข่างานประสาทที่มีขนาดต่างกันตามที่จัดเตรียมไว้	40
7	ค่าประสิทธิภาพของดัชนี WSJ เมื่อทำการวิเคราะห์กลุ่มด้วยวิธีเคมีนในช่วง 2-10 กลุ่ม	42
8	ค่าประสิทธิภาพของดัชนี WSJ และขนาดข่างานประสาทที่สอดคล้องกับผลการวิเคราะห์ที่มีประสิทธิภาพตามจำนวนกลุ่มที่กำหนด ในช่วง 2-10 กลุ่ม	44
9	จำนวนกลุ่มที่เหมาะสม และค่าประสิทธิภาพ เมื่อทำการวิเคราะห์กลุ่มด้วยเคมีนร่วมกับเทคนิคขยายพื้นที่อาศัยข่างานประสาทตามที่จัดเตรียมไว้	44
10	สรุปผลการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้า	46
11	สัดส่วนของสมาชิกคิดเป็นเปอร์เซ็นต์จากการวิเคราะห์กลุ่มด้วยวิธีที่ต่างกัน	46

สารบัญภาพ

ภาพที่		หน้า
1	ลักษณะการใช้ไฟฟ้ารวมของระบบ และลักษณะการใช้ไฟฟ้าแยกตามประเภท ผู้ใช้ไฟฟ้า	4
2	เส้นโค้งภาระแสดงลักษณะการใช้ไฟฟ้าของผู้ใช้ไฟฟ้ารายหนึ่ง ณ วันที่ 11/06/2012	5
3	ตัวอย่างเส้นโค้งแสดงความยาวนานในการใช้ไฟฟ้าของผู้ใช้ไฟฟ้ารายหนึ่งในเวลา 1 ปี	6
4	การกระจายทางสถิติของตัวประกอบการใช้ไฟฟ้าก่อน-หลังการปรับบรรทัดฐาน	9
5	หลักการพื้นฐานของการวิเคราะห์กลุ่มข้อมูล	10
6	ตัวอย่างการวิเคราะห์กลุ่มตามขั้นตอนของวิธีแบบเป็นลำดับชั้น	11
7	แผนผังลำดับงานแสดงขั้นตอนการวิเคราะห์กลุ่มของวิธีแบบเป็นลำดับชั้น	12
8	แผนผังลำดับงานแสดงขั้นตอนการวิเคราะห์กลุ่มข้อมูลด้วยวิธีเคมีน	14
9	สถาปัตยกรรมของข่ายงานประสาทแบบก่อดำด้วยตัวเอง หรือ SOM	15
10	การปรับตำแหน่งเข้าหาข้อมูลของเซลล์ประสาท BMU และเซลล์ประสาท ใกล้เคียง	16
11	แผนผังกระบวนการเรียนรู้ของข่ายงานประสาทแบบก่อดำด้วยตนเอง	18
12	ลำดับขั้นตอนการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้า	21
13	ตัวอย่างลักษณะการใช้ไฟฟ้าที่มีการสูญหายของค่าการใช้ไฟฟ้าบางช่วงเวลา	22
14	ตัวอย่างลักษณะการใช้ไฟฟ้าที่มีค่าการใช้ไฟฟ้าเป็น 0 ต่อเนื่องนานผิดปกติ	23
15	ลักษณะการใช้ไฟฟ้าก่อนผ่านขั้นตอนการปรับค่าให้เป็นบรรทัดฐาน	23
16	ลักษณะการใช้ไฟฟ้าหลังจากผ่านขั้นตอนการปรับค่าให้เป็นบรรทัดฐาน	24
17	โครงสร้างการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้า	25
18	ขั้นตอนการคัดเลือกเซลล์ประสาทจากเงื่อนไขระยะการเชื่อมโยงต่ำสุดเฉพาะที่	27
19	ขั้นตอนการวิเคราะห์กลุ่มด้วยเทคนิคการขยายพื้นที่อาศัยความเชื่อมโยงของ เซลล์บนข่ายงานประสาทแบบดั้งเดิม	28

สารบัญญภาพ (ต่อ)

ภาพที่		หน้า
20	ขั้นตอนการวิเคราะห์กลุ่มข้อมูลด้วยวิธีการทำงานร่วมกันของเคมีน และเทคนิคการขยายพื้นที่อาศัยการเชื่อมโยงของเซลล์บนข่ายงานประสาท	29
21	ข่ายงานประสาทในแบบตาราง (Rectangular) และแบบหกเหลี่ยม (Hexagonal)	30
22	ผลการวิเคราะห์กลุ่มในรูปโครงสร้างแบบแผนภาพต้นไม้ลำดับชั้นจำนวน 10 กลุ่ม	35
23	ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 10 กลุ่ม หลังจากทำการวิเคราะห์ด้วยวิธีแบบเป็นลำดับชั้น	35
24	ค่าประสิทธิภาพของดัชนี WSJ เมื่อทำการวิเคราะห์กลุ่มด้วยขั้นตอนของวิธีแบบเป็นลำดับชั้นในช่วง 2-10 กลุ่ม บนข่ายงานประสาทที่มีขนาดต่างกัน	36
25	ผลการวิเคราะห์กลุ่มในรูปแบบแผนภาพต้นไม้ลำดับชั้นจำนวน 9 กลุ่ม บนข่ายงานประสาทขนาด 41×11 เซลล์ประสาท	38
26	ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 9 กลุ่ม เมื่อทำการวิเคราะห์ที่ตามขั้นตอนของวิธีแบบเป็นลำดับชั้นร่วมกับข่ายงานประสาทขนาด 41×11 เซลล์ประสาท	38
27	เปรียบเทียบผลการคำนวณค่าประสิทธิภาพด้วยดัชนี WSJ ในช่วง 2-10 กลุ่ม ด้วยเทคนิคการขยายพื้นที่บนข่ายงานประสาทที่มีขนาดต่างกัน	39
28	ผลการวิเคราะห์กลุ่มในรูปของแผนภาพต้นไม้ลำดับชั้นบนข่ายงานประสาทขนาด 40×10 เซลล์ประสาท จำนวน 3 กลุ่ม	41
39	ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 3 กลุ่ม จากการวิเคราะห์ด้วยเทคนิคการขยายพื้นที่ร่วมกับข่ายงานประสาทที่มีขนาด 40×10 เซลล์ประสาท	42
30	ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 3 กลุ่มจากการวิเคราะห์ด้วยวิธีเคมีน	43
31	ค่าประสิทธิภาพของดัชนี WSJ จากการวิเคราะห์กลุ่มด้วยวิธีการทำงานร่วมกันของเคมีน และเทคนิคการขยายพื้นที่ บนข่ายงานประสาทที่มีขนาดต่างกัน	43

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
32	ผลการวิเคราะห์กลุ่มในรูปถ่ายงานประสาธจำนวน 3 กลุ่ม จากการทำงานร่วมกันของเคมีน และเทคนิคการขยายพื้นที่โดยใช้ข่ายงานประสาธขนาด 20×20 เซลล์ประสาธ	45
33	ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 3 กลุ่ม จากการทำงานร่วมกันของเคมีนกับเทคนิคการขยายพื้นที่บนข่ายงานประสาธขนาด 20×20 เซลล์ประสาธ	45

การวิเคราะห์กลุ่มข้อมูลลักษณะการใช้ไฟฟ้าโดยใช้วิธีการทำงานร่วมกันของเคมีน และเทคนิคการขยายพื้นที่บนพื้นฐานของข่ายงานประสาทแบบก่อดตัวด้วยตนเอง

A Combined Method of K-Means and Self-Organizing Map based Region-Growing Technique for Electric Load Profile Clustering

คำนำ

การศึกษาลักษณะการใช้ไฟฟ้าทำให้ทราบพฤติกรรมการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าในแต่ละช่วงเวลาได้อย่างชัดเจน และอำนวยความสะดวกต่อการวิเคราะห์แผนงานทางไฟฟ้าเป็นอย่างมาก ตัวอย่างเช่น ช่วยให้ผู้ดำเนินธุรกิจไฟฟ้ามีแนวทางที่เหมาะสมในการพัฒนาประสิทธิภาพด้านการให้บริการทางไฟฟ้า ทางฝ่ายผู้ใช้ไฟฟ้า สามารถนำลักษณะการใช้ไฟฟ้าหลังจากทำการศึกษามาปรับปรุงแก้ไข ซึ่งนอกจากจะทำให้รายจ่ายค่าไฟฟ้าลดลงแล้ว ยังช่วยให้เกิดการใช้ไฟฟ้าที่มีประสิทธิภาพได้อีกด้วย และประกอบกับปัจจุบัน การใช้ไฟฟ้ามีแนวโน้มเพิ่มสูงขึ้นเรื่อย ๆ ดังนั้น ถ้าผู้ใช้ไฟฟ้าหันมาให้ความสำคัญกับการปรับพฤติกรรมการใช้ไฟฟ้ามากขึ้น จะส่งผลดีคือ ทำให้เกิดการประหยัดพลังงานไฟฟ้าในภาพรวมของประเทศได้อย่างสูงสุด

การวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าเป็นการศึกษาในลักษณะของการค้นหา กลุ่มธรรมชาติ ซึ่งพิจารณาได้จากความคล้ายคลึงกันของข้อมูล ผลลัพธ์ของการวิเคราะห์กลุ่มจะช่วยให้ทราบว่าลักษณะการใช้ไฟฟ้ามีรูปแบบเป็นอย่างไรบ้าง ซึ่งช่วยอำนวยความสะดวกให้กับทั้ง ทางฝ่ายผู้ดำเนินธุรกิจไฟฟ้าและผู้ใช้ไฟฟ้า สามารถกำหนดทิศทางการบริหารจัดการทาง ไฟฟ้าได้อย่างเหมาะสม

วิทยานิพนธ์ฉบับนี้นำเสนอการวิเคราะห์กลุ่มของลักษณะการใช้ไฟฟ้าจากผู้ใช้ไฟฟ้ากลุ่ม อุตสาหกรรมหนึ่งในเวลา 1 ปี โดยจะนำความรู้จากการประมวลของข่ายงานประสาทแบบก่อดตัว ด้วยตนเอง (Self-Organizing Map) มาใช้เป็นพื้นฐานสำหรับการวิเคราะห์กลุ่มตามขั้นตอนของวิธี ขยายพื้นที่ (Region-Growing) โดยจะเริ่มจากการค้นหาตำแหน่งเซลล์อ้างอิงที่อยู่ใกล้กับตำแหน่ง ศูนย์กลางหลังจากทำการวิเคราะห์กลุ่มด้วยเคมีน จากนั้นจะทำการเชื่อมโยงกับเซลล์ประสาทที่ เหลือเพื่อสร้างพื้นที่กลุ่ม การทำงานจะสิ้นสุดเมื่อพื้นที่ของแต่ละกลุ่มไม่สามารถขยายออกไปได้อีก

วัตถุประสงค์

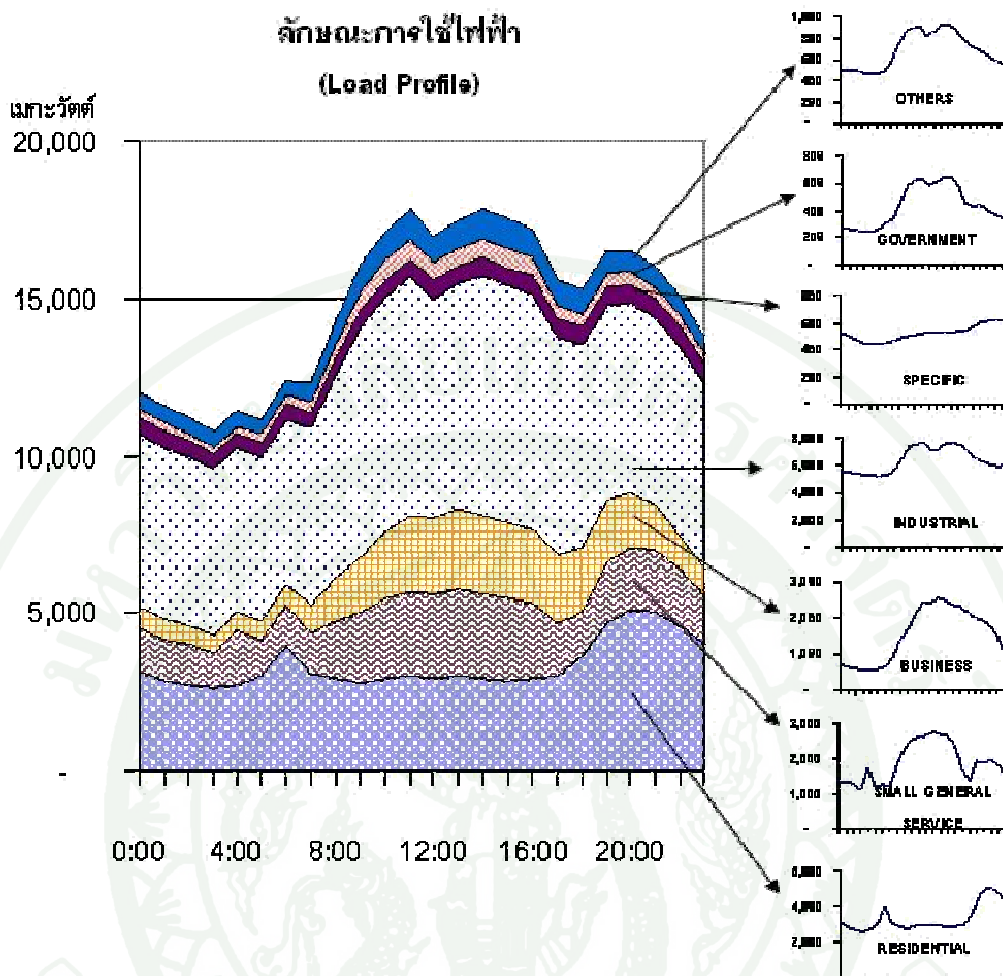
1. ศึกษาทฤษฎีการเรียนรู้ของช่างงานประสาธแบบก่อตัวด้วยตนเอง
2. วิเคราะห์หลักขณะการใช้ไฟฟ้ารูปแบบตัวแทนด้วย ด้วยวิธีการจัดกลุ่มรูปแบบต่าง ๆ
3. นำเสนอเทคนิคการวิเคราะห์กลุ่มตามขั้นตอนของวิธีการขยายพื้นที่ อาศัยคุณสมบัติความเชื่อมโยงระหว่างเซลล์บนช่างงานประสาธ
4. ทำการเปรียบเทียบผลการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าในแต่ละวิธีการที่มีขั้นตอนต่างกัน เพื่อหาผลลัพธ์ที่เหมาะสม

การตรวจเอกสาร

ลักษณะการใช้ไฟฟ้า

ลักษณะการใช้ไฟฟ้า (Load Profile) เป็นข้อมูลที่สะท้อนพฤติกรรมการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าตามเวลาที่กำหนด อาจจะเป็น 1 ชั่วโมง 30 นาที หรือ 15 นาที เป็นต้น และจากกิจกรรมที่มีความต่างกันของผู้ใช้ไฟฟ้าแต่ละประเภท ส่งผลให้พฤติกรรมการใช้ไฟฟ้ามีลักษณะต่างกันด้วย ตัวอย่างเช่น ผู้ใช้ไฟฟ้าภาคครัวเรือนจะมีลักษณะการใช้ไฟฟ้ารูปแบบที่เปลี่ยนไปตามเวลา และฤดูกาลอย่างสม่ำเสมอ ลักษณะการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าในภาคธุรกิจ จะเปลี่ยนตามนโยบายการดำเนินธุรกิจ และในภาคอุตสาหกรรมจะมีลักษณะการใช้ไฟฟ้าขึ้นอยู่กับประเภทของอุตสาหกรรม กำลังการผลิตของเครื่องจักร ตารางการทำงาน และปัจจัยด้านอื่น ๆ เป็นต้น (Chang and Lu, 2003: 974) ในภาพที่ 1 แสดงลักษณะการใช้ไฟฟ้าโดยรวมในระบบ ประกอบด้วยลักษณะการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าหลาย ๆ ประเภท

ภาควิชาวิศวกรรมไฟฟ้า มหาวิทยาลัยเกษตรศาสตร์ [วฟ. มก.] (2555: 6) การผลิตและการส่งจ่ายไฟฟ้าในระบบ ถูกกำหนดโดยผู้ใช้ไฟฟ้าซึ่งถือว่าเป็นโหลดของระบบ และปกติการใช้ไฟฟ้ามิได้คงที่อยู่ทุกขณะ โดยจะมีค่าสูงบ้างต่ำบ้างในระยะเวลาหนึ่ง ๆ ซึ่งเป็นปัญหาสำคัญต่อฝ่ายการผลิต และการส่งจ่ายไฟฟ้าเพราะสิ่งสำคัญคือ จะต้องพยายามรักษาระดับแรงดันในการส่งจ่ายไฟฟ้าไปยังผู้ใช้ไฟฟ้าให้อยู่ในระดับที่เหมาะสม ไม่สูงหรือต่ำเกินไป ไม่เช่นนั้นจะเป็นเหตุให้เกิดความเสียหายกับระบบได้ ดังนั้นการศึกษาความเปลี่ยนแปลงจากลักษณะการใช้ไฟฟ้า ช่วยให้ฝ่ายผู้ผลิตไฟฟ้าสามารถวางแผนเพื่อใช้ในการเดินเครื่องกำเนิดแต่ละเครื่องได้อย่างเหมาะสมกับการเปลี่ยนแปลงในแต่ละช่วงเวลา โดยวิธีดังกล่าวเมื่อมองในแง่ของเศรษฐศาสตร์จะช่วยให้เกิดความประหยัด เพราะในช่วงที่ระบบมีความต้องการใช้ไฟฟ้าปริมาณน้อย ๆ ก็ไม่จำเป็นต้องเดินเครื่องกำเนิดไฟฟ้าทุกเครื่อง ซึ่งจะช่วยลดปริมาณการใช้เชื้อเพลิงสำหรับผลิตไฟฟ้าที่เกินความจำเป็นและช่วยให้ต้นทุนจากการผลิตไฟฟ้าลดลงได้ นอกจากนี้ในด้านการสร้าง และการขยายระบบไฟฟ้า ตลอดจนการจัดหาพลังงานทดแทนและการสำรองไว้ล่วงหน้าเพื่อรองรับกับความต้องการกำลังไฟฟ้าที่เพิ่มสูงขึ้น ต้องใช้การพยากรณ์การขยายตัวของการใช้ไฟฟ้าในอนาคตเข้ามาช่วย ซึ่งถ้าหากนำลักษณะการใช้ไฟฟ้าที่เก็บเป็นสถิติประกอบการพยากรณ์จะทำให้ผลลัพธ์มีความน่าเชื่อถือมากขึ้น จากตัวอย่างที่กล่าวมาข้างต้น จะเห็นได้ว่าลักษณะการใช้ไฟฟ้าเป็นข้อมูลที่เป็นประโยชน์ และมีความสำคัญต่อการปฏิบัติงานในระบบไฟฟ้ากำลังเป็นอย่างมาก



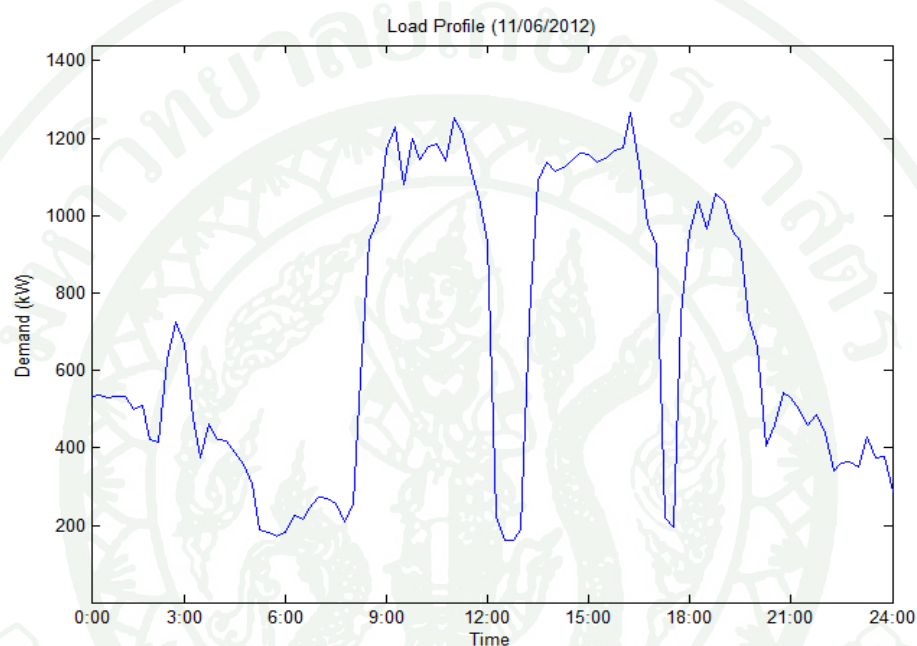
ภาพที่ 1 ลักษณะการไฟฟ้ารวมของระบบ และลักษณะการใช้ไฟฟ้าแยกตามประเภทผู้ใช้ไฟฟ้า

ที่มา: สำนักงานนโยบายและแผนพลังงาน กระทรวงพลังงาน [สนพ. พน.] (2547)

การเก็บตัวอย่างลักษณะการใช้ไฟฟ้าหรือติดตามพฤติกรรมของผู้ใช้ไฟฟ้า สามารถทำได้โดยติดตั้งระบบการอ่านเครื่องวัดหน่วยไฟฟ้าแบบอัตโนมัติ (Automatic Meter Reading-System, AMR System) ให้กับผู้ใช้ไฟฟ้า สำหรับทางฝ่ายผู้ให้บริการจำหน่ายไฟฟ้าในประเทศไทย ได้มีการติดตาม และเก็บตัวอย่างลักษณะการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าทุก ๆ 15 นาที ดังนั้นลักษณะการใช้ซึ่งเป็นข้อมูลของแต่ละวันจะประกอบด้วยค่าการใช้กำลังไฟฟ้าจำนวน 96 ค่า

การแสดงตัวอย่างลักษณะการใช้ไฟฟ้า

การศึกษารูปแบบลักษณะการใช้ไฟฟ้าที่ช่วยให้สามารถทำความเข้าใจได้อย่างรวดเร็ว และง่าย คือ ใช้การแสดงรูปแบบด้วยเส้นโค้งภาระ (Load Curve) และเส้นโค้งแสดงความยาวนานในการใช้ไฟฟ้า (Load Duration Curve) ตามตัวอย่างในภาพที่ 2 และ 3



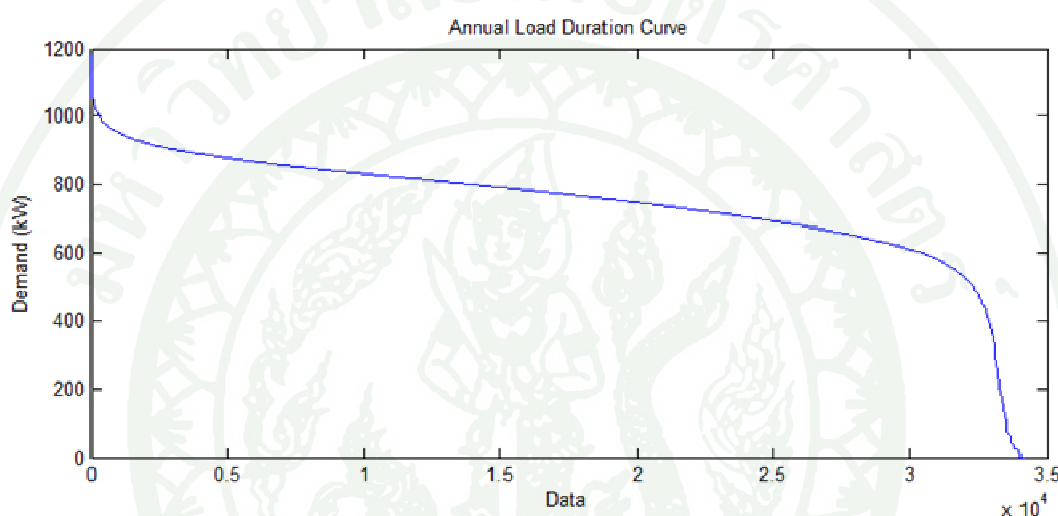
ภาพที่ 2 เส้นโค้งภาระแสดงลักษณะการใช้ไฟฟ้าของผู้ใช้ไฟฟ้ารายหนึ่ง ณ วันที่ 11/06/2012

เส้นโค้งภาระ หรือ Load Curve คือเส้นแสดงลักษณะการใช้ไฟฟ้าในช่วงเวลาที่พิจารณา เช่น เส้นโค้งภาระรายวัน (Daily Load Curve) ซึ่งใช้แสดงลักษณะการใช้ไฟฟ้า หรือค่าการใช้กำลังไฟฟ้าทั้งหมดที่บันทึกไว้ทั้งหมดในช่วงเวลาหนึ่งวัน เป็นต้น การแสดงข้อมูลด้วยเส้นโค้งภาระมีความสำคัญต่อการดำเนินงานทางไฟฟ้าหลาย ๆ ด้าน เช่น การวางแผนผลิตไฟฟ้า การวางมาตรการเพื่อปรับปรุงพฤติกรรมการใช้ไฟฟ้าให้มีประสิทธิภาพมากขึ้น การจำแนกประเภทผู้ใช้ไฟฟ้า สำหรับปริมาณทางไฟฟ้าที่สำคัญซึ่งปรากฏบนเส้นโค้งภาระ ประกอบด้วย

ค่าความต้องการกำลังไฟฟ้าเฉลี่ย (Average Demand) คือค่าเฉลี่ยของผลรวมค่าความการใช้กำลังไฟฟ้าทั้งหมดที่นำมาแสดงด้วยเส้นโค้งภาระ หรือจากลักษณะการใช้ไฟฟ้าตามระยะเวลาที่พิจารณา

ค่าความต้องการกำลังไฟฟ้าสูงสุด (Maximum Demand) หรือบางครั้งเรียกว่า ค่ายอดสูงสุด (Peak Demand) คือ ค่าการใช้กำลังไฟฟ้าสูงสุดที่ปรากฏบนเส้นโค้งภาระตามช่วงเวลาที่พิจารณา

ตัวประกอบการใช้ไฟฟ้า (Load Factor) คือ อัตราส่วนของค่าความต้องการกำลังไฟฟ้าเฉลี่ยต่อค่ายอดสูงสุด และเป็นดัชนีที่บ่งบอกว่า ลักษณะการใช้ไฟฟ้าเป็นรูปแบบที่มีประสิทธิภาพหรือไม่



ภาพที่ 3 ตัวอย่างเส้นโค้งแสดงความยาวนานในการใช้ไฟฟ้าของผู้ใช้ไฟฟ้ารายหนึ่งในเวลา 1 ปี

เส้นโค้งแสดงความยาวนานในการใช้ไฟฟ้า หรือ Load Duration Curve คือเส้นโค้งภาระที่นำมาแสดงใหม่ในลักษณะเรียงลำดับ เริ่มต้นจากค่าการใช้กำลังไฟฟ้าสูงสุดไปสู่ค่าการใช้กำลังไฟฟ้าต่ำสุดตลอดช่วงเวลาที่สนใจ ทำให้สามารถประมาณสัดส่วนของค่าความต้องการกำลังไฟฟ้าต่ำสุดหรือการใช้โหลดฐาน (Base Load) และโหลดสูงสุด (Peak Load) ซึ่งจะนำไปใช้ประโยชน์ในด้านการจัดหาเชื้อเพลิงจากแหล่งต่าง ๆ สำหรับใช้เดินเครื่องกำเนิดเพื่อผลิตไฟฟ้าได้อย่างเหมาะสม รวมทั้งสามารถใช้เป็นข้อมูลประกอบการพยากรณ์ตลอดจนใช้หาขนาดของเครื่องกำเนิด และหม้อแปลงไฟฟ้าเพื่อรองรับความต้องการไฟฟ้าของระบบที่มีแนวโน้มเพิ่มขึ้นในอนาคต (วฟ. มก., 2555: 8)

ประโยชน์ของการศึกษาลักษณะการใช้ไฟฟ้า

การศึกษาลักษณะการใช้ไฟฟ้ามีความสำคัญต่อการดำเนินงานทางไฟฟ้าเป็นอย่างมาก กล่าวคือ ทำให้ทราบพฤติกรรมการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าที่เปลี่ยนไปในแต่ละช่วงเวลาอย่างชัดเจน หรือช่วยให้การจัดการทางไฟฟ้าเป็นไปอย่างเหมาะสม และมีประสิทธิภาพ ประโยชน์ของการศึกษาลักษณะการใช้ไฟฟ้าจึงมีมากมาย ดังตัวอย่างต่อไปนี้ (สนพ. พน., 2541)

การจัดทำอัตราค่ากระแสไฟฟ้าได้เหมาะสม และยุติธรรมต่อผู้ใช้ไฟฟ้า เนื่องจากต้นทุนค่าไฟฟ้าจะแปรผันตามราคาเชื้อเพลิงสำหรับใช้เดินเครื่องกำเนิดไฟฟ้า ค่าการลงทุนในส่วนของการจัดเตรียมระบบส่งกำลังไฟฟ้าเพื่อรองรับความต้องการไฟฟ้าที่มีมากขึ้น รวมถึงค่าใช้จ่ายอื่น ๆ เพราะฉะนั้นผู้ใช้ไฟฟ้ายิ่งใช้ไฟฟ้าในปริมาณมาก และมีค่าความต้องการสูงสุดเกิดขึ้นในช่วงเวลาเดียวกันกับระบบไฟฟ้า ทำให้ผู้ให้บริการจำหน่ายไฟฟ้ามีการลงทุนสูง ควรที่จะถูกเรียกเก็บค่าไฟฟ้ามากตามไปด้วย ในทางตรงกันข้าม สำหรับผู้ใช้ไฟฟ้ายิ่งมีผลให้เกิดการลงทุนน้อย ผู้ใช้ไฟฟ้ายิ่งนั้นก็ควรที่จะถูกเรียกเก็บค่าไฟฟ้าน้อยลงด้วย

การพยากรณ์ความต้องการไฟฟ้า จากผลการศึกษาลักษณะการใช้ไฟฟ้า ทำให้ทราบรายละเอียดของผู้ใช้ไฟฟ้าแต่ละรายว่า มีความต้องการกำลังไฟฟ้าน้อยในแต่ละช่วงเวลาอย่างไร และถ้าหากนำข้อมูลที่มีการเก็บบันทึกในลักษณะสถิติมาประกอบการพยากรณ์ความต้องการไฟฟ้า จะทำให้ผลลัพธ์มีความน่าเชื่อถือ สามารถอธิบายได้อย่างชัดเจน และสามารถนำไปใช้เป็นข้อมูลประกอบการวางแผนขยายกำลังการผลิตไฟฟ้า ระบบส่งและระบบจำหน่ายให้เพียงพอกับความต้องการไฟฟ้าที่คาดว่าจะเกิดขึ้นในอนาคตต่อไป

ประโยชน์ด้านอื่น ๆ เช่น สามารถเพิ่มรายได้ให้กับผู้ดำเนินธุรกิจไฟฟ้า จากการให้บริการข้อมูลลักษณะการใช้ไฟฟ้าให้กับผู้ใช้ไฟฟ้า หรือในด้านอุตสาหกรรมการผลิตและจำหน่ายไฟฟ้า ที่ไม่มีการผูกขาด หรือเปิดแข่งขันกันอย่างเสรี ลักษณะการใช้ไฟฟ้าถือเป็นข้อมูลสำคัญต่อการวางกลยุทธ์เพื่อสร้างความได้เปรียบเหนือคู่แข่ง (Anuar and Zakaria, 2011: 139) สำหรับทางด้านผู้ใช้ไฟฟ้า สามารถวิเคราะห์พฤติกรรมของตนจากลักษณะการใช้ไฟฟ้า พร้อมทั้งกำหนดแนวทางการปรับปรุงให้มีการใช้ไฟฟ้าที่มีประสิทธิภาพได้ โดยประโยชน์ที่จะได้รับคือ ใช้จ่ายด้านค่าไฟฟ้าลดลง เนื่องจากการลดปริมาณการใช้ไฟฟ้ารวมทั้งการย้ายไปใช้ในช่วงเวลาอื่นส่งผลให้สัดส่วนของกำลังการผลิตและความต้องการอยู่ในเกณฑ์ที่สมดุล ช่วยให้ทางฝ่ายผู้ดำเนินธุรกิจสามารถ

ชะลอการลงทุนในด้านการสร้างโรงจักรไฟฟ้า และการขยายระบบส่งจ่ายไฟฟ้าเพิ่มเติม ซึ่งถือเป็นส่วนหนึ่งของต้นทุนการผลิต ที่จะนำไปใช้ในเรียกเก็บค่าไฟฟ้าในภายหลัง

การจัดการเบื้องต้นกับข้อมูล

ข้อมูลอาจได้รับความเสียหายในระหว่างการบันทึก และการถ่ายโอนเพื่อนำมาใช้ ดังนั้นก่อนนำข้อมูลมาทำการวิเคราะห์ จะต้องมีการจัดการเบื้องต้นกับข้อมูล (Data preprocessing) เพื่อปรับข้อมูลให้อยู่ในรูปแบบที่ช่วยให้การค้นหาคำถามที่มีความถูกต้อง วิธีการจัดการเบื้องต้นประกอบด้วย การทำความสะอาดข้อมูล และการบรรทัดฐานข้อมูล

การทำความสะอาดข้อมูล (Data Cleaning) เป็นขั้นตอนที่ใช้คัดแยกข้อมูลที่ผิดปกติออกจากการพิจารณา สำหรับกรณีของลักษณะการใช้ไฟฟ้า ความผิดปกติที่ตรวจพบโดยส่วนใหญ่จะเกี่ยวกับการบันทึกค่าการใช้ไฟฟ้าซ้ำซ้อน ค่าการใช้ไฟฟ้ามีค่าเป็น 0 ต่อเนื่องนานผิดปกติ และการสูญหายของค่าการใช้ไฟฟ้าในช่วงระยะเวลาหนึ่ง ข้อมูลที่ผิดปกติเหล่านี้เป็นสาเหตุสำคัญที่ทำให้การจัดกลุ่มเกิดผิดพลาดได้

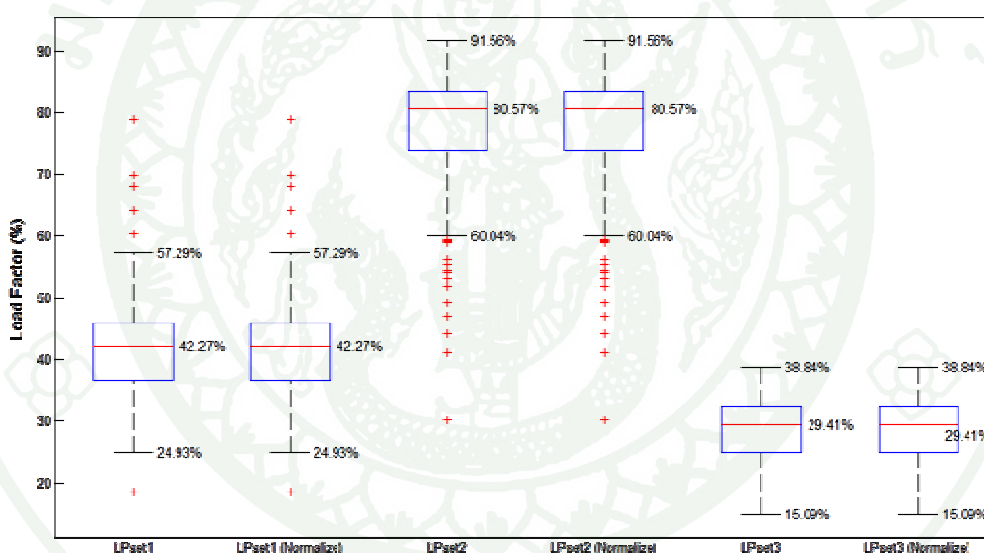
การบรรทัดฐานข้อมูล (Data normalization) เป็นขั้นตอนการปรับข้อมูลเพื่อให้มีค่าอยู่ในช่วงระดับเดียวกัน วิธีการบรรทัดฐานข้อมูลมีหลายวิธี แต่ในวิทยานิพนธ์จะทำการปรับค่าตามสมการที่ 1

$$x_{d-norm} = \frac{x_d}{x_{max}} \quad (1)$$

โดยที่ x_d คือ ค่าการใช้ไฟฟ้า ณ เวลา d
 x_{max} คือ ค่ายอดสูงสุดจากภายในช่วงที่พิจารณา

สมการที่ 1 เป็นการปรับค่ายอดสูงสุดของลักษณะการใช้ไฟฟ้าภายในช่วงเวลาที่พิจารณาให้มีค่าเท่ากับ 1 และจะมีการปรับระดับค่าการใช้ไฟฟ้าต่ำสุดสองคี่งกับค่ายอดสูงสุด เพื่อช่วยให้มองเห็นโหลดฐาน (Base Load) จากลักษณะการใช้ไฟฟ้าได้อย่างชัดเจน

Gerbec *et al.* (2003: 2) จุดประสงค์ของการบรรทัดฐานข้อมูล เนื่องจากปริมาณการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าแต่ละรายมีค่าต่างกัน ซึ่งอาจส่งผลกระทบต่อความถูกต้องของการวิเคราะห์บนพื้นฐานการพิจารณาความคล้ายเป็นหลักได้ ตัวอย่างเช่น ลักษณะการใช้ไฟฟ้าที่มีรูปแบบคล้ายกัน แต่ปริมาณการใช้ไฟฟ้ามีค่าต่างกันมาก มีโอกาสสูงที่จะถูกจัดลงต่างกลุ่ม แม้จะมีรูปแบบที่ซ้ำกันหรือกรณีที่สามารถจัดลงในกลุ่มเดียวกันได้ แต่รูปแบบของตัวแทนกลุ่มอาจไม่ครอบคลุมกับสมาชิกบางส่วนภายในกลุ่มได้มากพอ ไม่สามารถนำไปใช้ประโยชน์ได้ และในบางครั้งอาจถูกมองว่าเป็นข้อมูลที่ผิดปกติ (Outlier) เนื่องจากมีค่าต่างจากข้อมูลค่าอื่น ๆ อย่างชัดเจน ในภาพที่ 4 แสดงการกระจายทางสถิติของตัวประกอบการใช้ไฟฟ้าก่อน และหลังจากทำการบรรทัดฐานลักษณะการใช้ไฟฟ้าของผู้ใช้ไฟฟ้าจำนวน 3 ราย ในระยะเวลา 1 ปี จะเห็นได้ว่าการบรรทัดฐานข้อมูลลักษณะการใช้ไฟฟ้าตามสมการที่ 1 ไม่ส่งผลกระทบต่อค่าตัวประกอบการใช้ไฟฟ้า

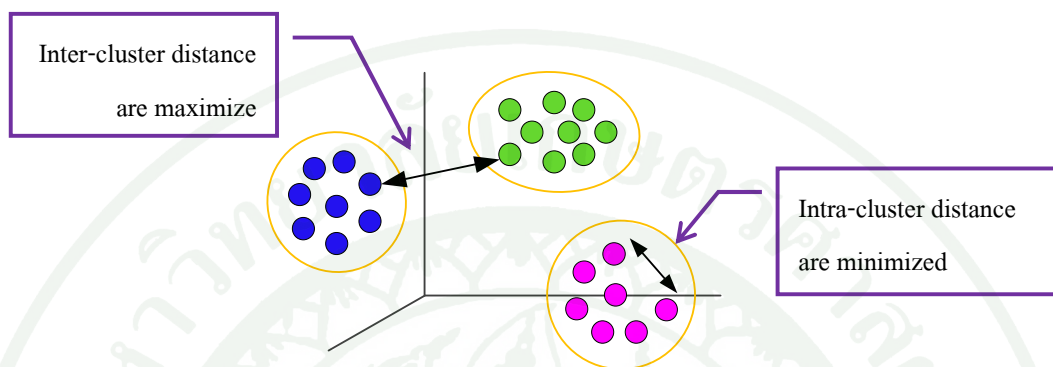


ภาพที่ 4 การกระจายทางสถิติของตัวประกอบการใช้ไฟฟ้าก่อน-หลังการปรับบรรทัดฐาน

การวิเคราะห์กลุ่มข้อมูล

การวิเคราะห์กลุ่มข้อมูล เป็นกระบวนการหนึ่งสำหรับใช้ค้นหารูปแบบตัวแทนหรือลักษณะเฉพาะทางข้อมูลที่สำคัญ ซึ่งซ่อนอยู่ท่ามกลางข้อมูลจำนวนมาก หลักการวิเคราะห์คือข้อมูลที่มีลักษณะคล้ายกัน (Similarity) หรือใกล้เคียงกัน (Proximity) จะถูกจัดไว้ในกลุ่มเดียวกัน โดยที่ระยะห่างระหว่างสมาชิกภายในกลุ่มเดียวกัน (Intra-cluster) จะต้องมีค่าน้อย และระยะห่างระหว่างสมาชิกจากต่างกลุ่ม (Inter-cluster) ต้องมีค่ามาก ดังแสดงในภาพที่ 5 ทุกวันนี้มีการนำเสนอ

เทคนิคใหม่ ๆ หรือพัฒนาต่อยอดจากวิธีการดั้งเดิม เพื่อนำมาใช้วิเคราะห์กลุ่มข้อมูลมีประสิทธิภาพมากขึ้น และสำหรับวิทยานิพนธ์นี้ จะนำเสนอขั้นตอนการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าของวิธีการพื้นฐานต่าง ๆ รวมทั้งวิธีการแบบผสมผสาน โดยมีรายละเอียดดังต่อไปนี้



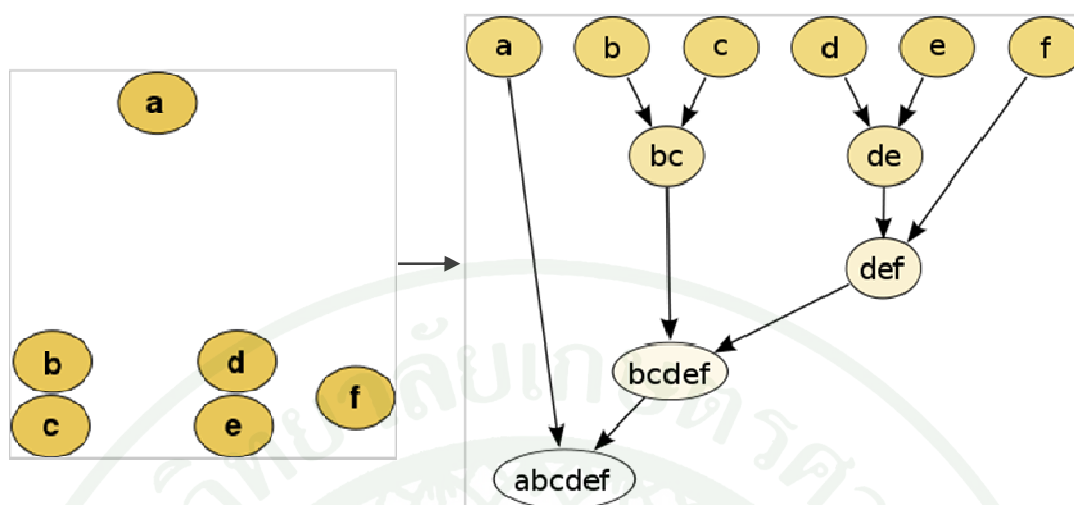
ภาพที่ 5 หลักการพื้นฐานของการวิเคราะห์กลุ่มข้อมูล

ทฤษฎีการวิเคราะห์กลุ่มข้อมูลด้วยวิธีแบบเป็นลำดับขั้น

การวิเคราะห์กลุ่มข้อมูลด้วยวิธีแบบเป็นลำดับขั้น (Hierarchical Clustering) เป็นวิธีจัดกลุ่มข้อมูลอาศัยหลักการเชื่อมโยงข้อมูลที่มีลักษณะคล้ายเข้าไว้ด้วยกันเป็นลำดับขั้นคล้ายกับต้นไม้ ดังตัวอย่างในภาพที่ 6 สามารถแบ่งวิธีการทำงานออกเป็น 2 ลักษณะ ตามการเชื่อมโยงข้อมูล ได้แก่ Agglomerative Hierarchical Clustering และ Divisive Hierarchical Clustering

Agglomerative Hierarchical Clustering (Bottom-Up) วิธีนี้เริ่มต้นจากการกำหนดกลุ่มย่อยตามจำนวนข้อมูลที่น่าวิเคราะห์จัดกลุ่ม จากนั้นทำการรวมข้อมูลครั้งละ 2 กลุ่มย่อยที่มีความคล้ายหรือระยะการเชื่อมโยงใกล้กันมากที่สุดเข้าไว้ด้วยกัน เพื่อสร้างกลุ่มใหม่ที่มีขนาดใหญ่ขึ้น และดำเนินการรวมกลุ่มต่อ ๆ ไป จนกระทั่งเหลือกลุ่มหลักเพียงกลุ่มเดียว

Divisive Hierarchical Clustering (Top-Down) ลักษณะการทำงานจะตรงข้ามกับวิธี Agglomerative กล่าวคือ วิธีนี้จะเริ่มจากกลุ่มข้อมูลเพียงกลุ่มเดียว แล้วแบ่งย่อยออกเป็นหลายกลุ่ม จนกระทั่งได้กลุ่มย่อยที่มีจำนวนเท่ากับข้อมูลที่น่าวิเคราะห์ แต่วิธีนี้ไม่ได้รับความนิยมเท่ากับวิธี Agglomerative เนื่องจากใช้เวลาในการคำนวณมากกว่า



ภาพที่ 6 ตัวอย่างการวิเคราะห์กลุ่มตามขั้นตอนของวิธีแบบเป็นลำดับชั้น

วิธีคำนวณระยะการเชื่อมโยง (Linkage Method) ระหว่างกลุ่มที่เป็นที่รู้จัก และถูกใช้ร่วมกับการจัดกลุ่มแบบลำดับชั้นอยู่บ่อยครั้ง แบ่งออกได้ 3 วิธีตามสมการที่ 2-4

$$d_{mn}(i, p) = \min(\text{dist}(x_{ij}, x_{pq})) \quad (2)$$

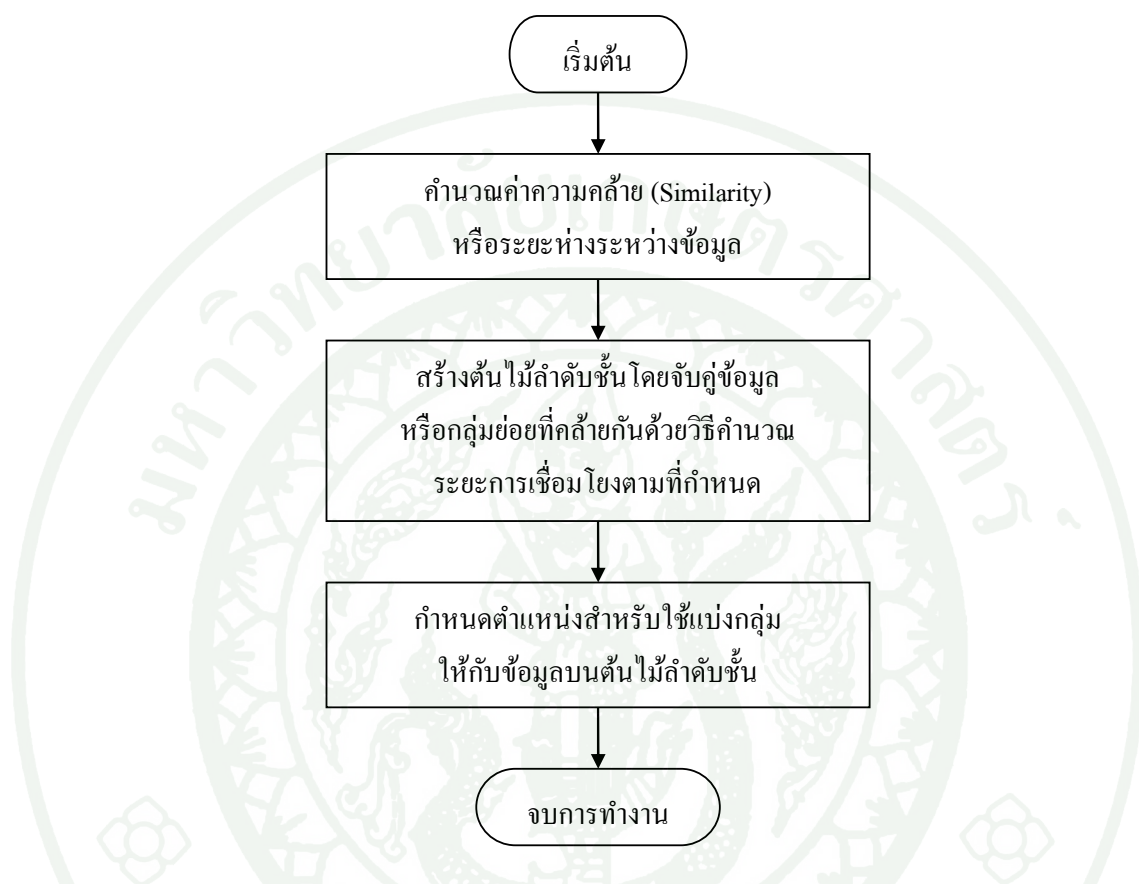
$$d_c(i, p) = \max(\text{dist}(x_{ij}, x_{pq})) \quad (3)$$

$$d_a(i, p) = \frac{1}{n_i n_p} \sum_{j=1}^{n_i} \sum_{q=1}^{n_p} d(x_{ij}, x_{pq}) \quad (4)$$

โดยที่ x_{ij} คือ ค่าของข้อมูลที่เป็นสมาชิกของกลุ่ม i ตำแหน่งที่ j
 x_{pq} คือ ตำแหน่งข้อมูลที่เป็นสมาชิกของกลุ่ม p ตำแหน่งที่ q
 $dist$ คือ ระยะห่างระหว่างจุดข้อมูลที่พิจารณา

สมการที่ 2 เป็นการคำนวณระยะห่างระหว่าง 2 กลุ่มที่ใกล้กันมากที่สุด เรียกว่า Single linkage สำหรับสมการที่ 3 เรียกว่า Complete linkage เป็นวิธีคำนวณซึ่งตรงข้ามกับสมการที่ 2 นั่นคือ เป็นการคำนวณหาระยะระหว่างสมาชิกของทั้ง 2 กลุ่มที่ห่างกันมากที่สุด แต่วิธีนี้ไม่เป็นที่ยอมรับเนื่องจากได้รับผลกระทบจากตัวรบกวนมาก ทำให้การคำนวณใช้เวลานาน และสมการที่ 4 คือ

Average linkage เป็นวิธีการที่ผสมระหว่าง 2 วิธีแรก จากนั้นจึงเลือกค่าเฉลี่ยของระยะทางระหว่าง 2 กลุ่ม สำหรับขั้นตอนการวิเคราะห์กลุ่มของวิธีแบบเป็นลำดับชั้น สรุปได้ดังภาพที่ 7



ภาพที่ 7 แผนผังลำดับงานแสดงขั้นตอนการวิเคราะห์กลุ่มของวิธีแบบเป็นลำดับชั้น

ทฤษฎีการวิเคราะห์กลุ่มข้อมูลด้วยวิธีเคมีน

วิธีเคมีน (K-Means) เป็นการวิเคราะห์กลุ่มข้อมูลจัดอยู่ในประเภทของการจัดกลุ่มแบบแบ่งส่วน (Partition Clustering) เป็นวิธีการหนึ่งที่ได้รับคามนิยม และถูกนำไปใช้ประกอบงานวิจัยแขนงต่าง ๆ อย่างแพร่หลาย ตัวอย่างเช่น การรู้จำแบบ (Pattern Recognition) รวมถึงการประมวลผลจากรูปภาพ (Image Processing) ตลอดจนนำไปใช้ในด้านการวินิจฉัยโรค สำหรับกระบวนการจัดกลุ่มด้วยวิธีเคมีนสรุปได้ดังนี้ (Balasko *et al.*, 2005)

กำหนดให้ $X = \{x_1, x_2, \dots, x_N \in \mathcal{R}^d\}$ แทนข้อมูลจำนวน N รูปแบบ ซึ่งถูกกำหนดให้สามารถจัดกลุ่มได้จำนวน c กลุ่ม โดยที่ $V = \{v_1, v_2, \dots, v_c \in \mathcal{R}^d\}$ คือลักษณะเฉพาะหรือจุดศูนย์กลางกลุ่ม แต่ละข้อมูลรวมทั้งจุดศูนย์กลางกลุ่มจะประกอบด้วยสมาชิกหรือตัวแปรจำนวน d ค่า การทำงานของเคมีนเริ่มจากระบุจำนวนกลุ่ม และสุ่มเลือกตำแหน่งศูนย์กลางเริ่มต้นของแต่ละกลุ่ม จากนั้นจะเข้าสู่กระบวนการปรับตำแหน่งของศูนย์กลาง โดยอาศัยการทำซ้ำเพื่อให้ได้ผลลัพธ์ที่เหมาะสม ซึ่งประเมินจากการลดลงของฟังก์ชันเป้าหมายหรือค่าความผิดพลาดโดยรวมระหว่างข้อมูลกับจุดศูนย์กลางกลุ่ม (J) ในเกณฑ์ที่ยอมรับได้ โดยคำนวณได้จากสมการที่ 5

$$J = \sum_{i=1}^c \sum_{j \in A_i} \|x_j - v_i\|^2 \quad (5)$$

$\|x_j - v_i\|$ เป็นระยะห่างระหว่างข้อมูลตำแหน่ง j กับจุดศูนย์กลางของกลุ่ม i จากการวัดระยะแบบยูคลิดีเนียน (Euclidean) จุดศูนย์กลางกลุ่มในแต่ละรอบจะถูกปรับตำแหน่งใหม่โดยใช้การคำนวณจากสมการที่ 6

$$v_i = \frac{\sum_{j=1}^{N_i} x_j}{N_i}, j \in A_i \quad (6)$$

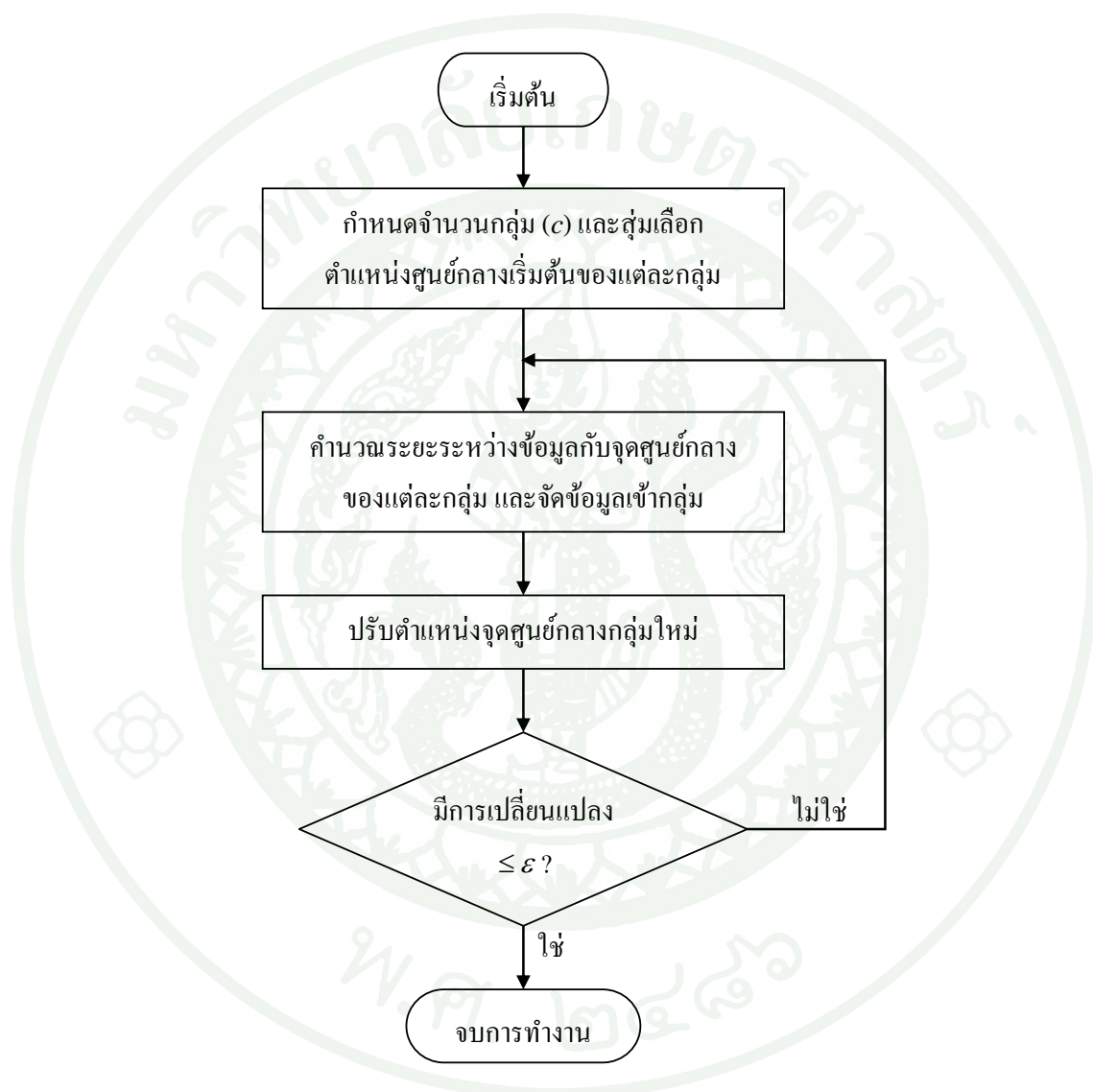
N_i คือจำนวนข้อมูลที่เป็นสมาชิกของกลุ่ม i (A_i) และเนื่องจากขั้นตอนของเคมีนทำงานในลักษณะของกระบวนการทำซ้ำ ดังนั้นจะต้องกำหนดเงื่อนไขหยุดการทำงาน โดยกระบวนการจัดกลุ่มจะสิ้นสุดลงเมื่อฟังก์ชันเป้าหมายเปลี่ยนแปลงไปจากรอบก่อนหน้าน้อยกว่าค่าความคลาดเคลื่อนที่กำหนดไว้ ($\leq \epsilon$) สำหรับลำดับขั้นตอนการทำงานด้วยเทคนิคการจัดกลุ่มวิธีเคมีน อธิบายได้ด้วยแผนผังลำดับการทำงานตามภาพที่ 8

ข้อดีของวิธีเคมีน

- 1) เป็นวิธีที่มีกระบวนการทำงานเข้าใจได้ง่าย ไม่ซับซ้อน และได้รับความนิยมจึงถูกนำไปประยุกต์ใช้กับงานวิจัยได้หลากหลายแขนง ช่วยให้การสืบค้นรูปแบบสำคัญทำได้ง่ายขึ้น
- 2) สามารถจำแนกกลุ่มข้อมูลที่มีความคล้ายได้ดี
- 3) ใช้เวลาในการประมวลผลที่รวดเร็ว และสามารถจัดการกับข้อมูลที่มีจำนวนมากได้ดี

ข้อสังเกตของวิธีเคมิน

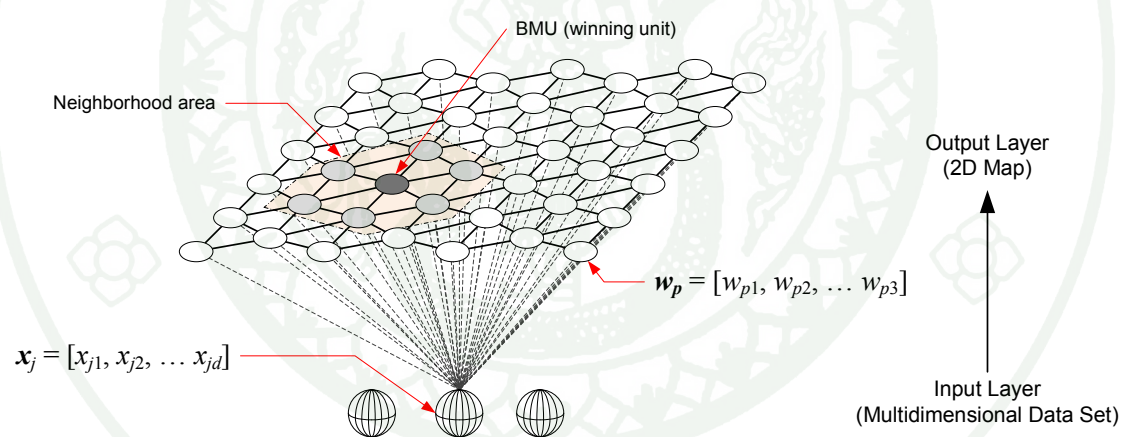
- 1) ต้องระบุจำนวนกลุ่มให้กับชุดข้อมูล ซึ่งส่วนใหญ่จะไม่ทราบจำนวนกลุ่มของข้อมูล
- 2) การวัดระยะห่างอ้างอิงจากจุดศูนย์กลางของแต่ละกลุ่ม ดังนั้นลักษณะของกลุ่มจะถูกพิจารณาเป็นทรงกลม ซึ่งถือเป็นข้อจำกัดกรณีที่กลุ่มของข้อมูลมีลักษณะเป็นรูปทรงอื่น ๆ



ภาพที่ 8 แผนผังลำดับงานแสดงขั้นตอนการวิเคราะห์กลุ่มข้อมูลด้วยวิธีเคมิน

ทฤษฎีการเรียนรู้ของข่ายงานประสาทแบบก่อดำด้วยตนเอง

ข่ายงานประสาทแบบก่อดำด้วยตนเอง (Self – Organizing Map) หรือ SOM ถูกนำเสนอขึ้นในปี 1995 โดย Teuvo Kohonen เป็นข่ายงานประสาทเทียมที่มีลักษณะการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) สำหรับใช้วิเคราะห์โครงสร้างข้อมูล โดยใช้แนวคิดคือ ข้อมูลที่มีลักษณะคล้ายกันในหลายมิติ ก็จะมีลักษณะที่คล้ายกันใน 2 มิติด้วย ทำให้สามารถนำไปใช้อธิบายความสัมพันธ์ของข้อมูลที่ประกอบขึ้นมาจากหลายตัวแปรได้ สำหรับสถาปัตยกรรมของ SOM สามารถแยกได้เป็น 2 ชั้น ตามภาพที่ 9 ประกอบด้วยชั้นนำเข้า (Input Layer) หรือชั้นของข้อมูลจริง สำหรับใช้ในกระบวนการเรียนรู้ และชั้นผลลัพธ์ (Output Layer) หรือชั้นข่ายงานประสาทในรูปของแผนภาพขนาด 1-2 มิติ การประมวลผลด้วย SOM ถูกนำมาใช้ร่วมกับการวิจัยอย่างกว้างขวางในหลาย ๆ ด้าน เช่น การแสดงภาพจำลองข้อมูล การรู้จำรูปแบบ การวิเคราะห์สภาวะการทำงานในอุตสาหกรรมและทางการแพทย์ งานควบคุมกระบวนการผลิต

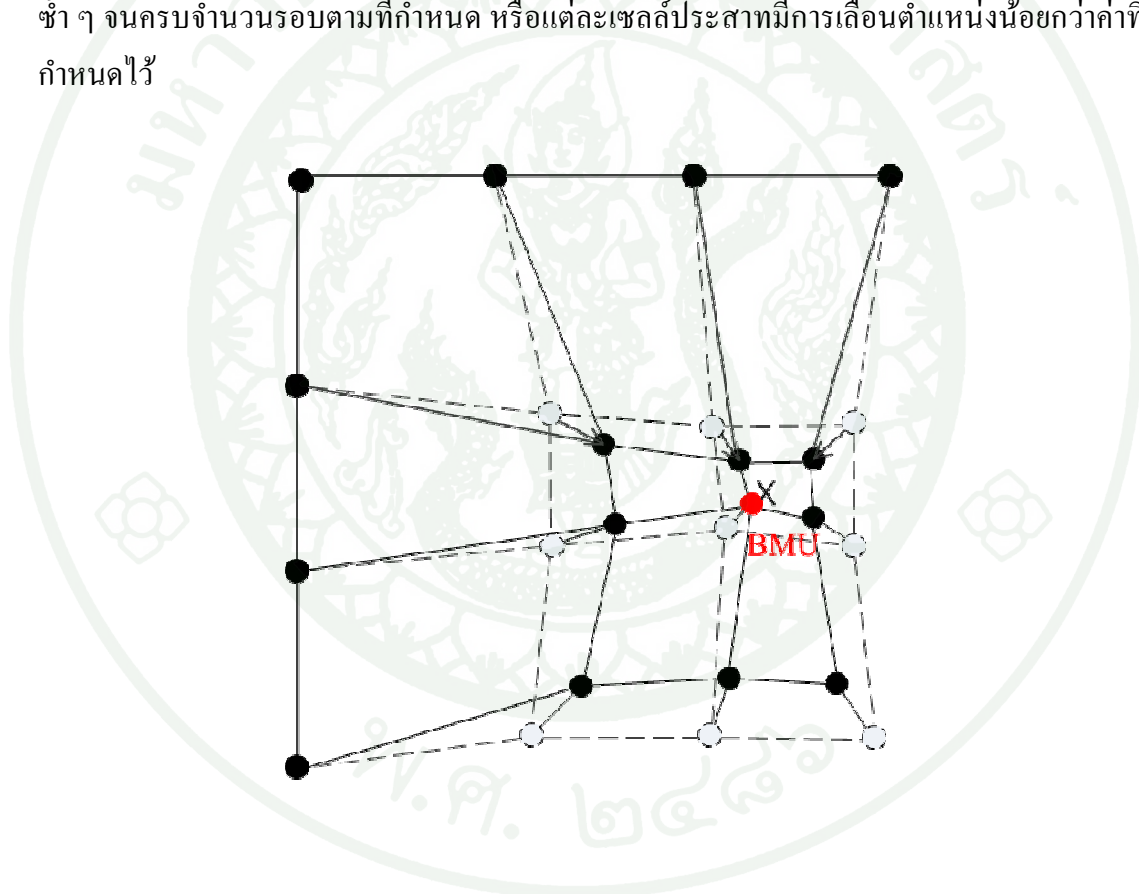


ภาพที่ 9 สถาปัตยกรรมของข่ายงานประสาทแบบก่อดำด้วยตัวเอง หรือ SOM

Vesanto *et al.* (2000) กระบวนการเรียนรู้ของ SOM เป็นการปรับข่ายงานประสาทเพื่อให้ครอบคลุมกับโครงสร้างข้อมูลที่นำมาฝึกสอนตามจำนวนรอบที่กำหนด ในแต่ละรอบของการเรียนรู้ จะมีการสุ่มเลือกข้อมูลนำเข้า $x(t)$ จากชุดข้อมูลจำนวนหนึ่ง สำหรับใช้เปรียบเทียบกับเวกเตอร์น้ำหนักของแต่ละเซลล์ประสาท ซึ่งจะมีจำนวนตัวแปรเท่ากับข้อมูล เพื่อหา Best Matching Unit (BMU) หรือตำแหน่งของเซลล์ประสาทที่อยู่ใกล้กับข้อมูลที่สุ่มเลือกมากที่สุด การเปรียบเทียบจะใช้วิธีวัดระยะแบบยูคลิเดียน ตามสมการที่ 7

$$\|x(t) - BMU\| = \min_p \{\|x(t) - w_p(t)\|\} \quad (7)$$

จากนั้นเวกเตอร์น้ำหนักของเซลล์ประสาท BMU จะถูกปรับค่าเพื่อให้เข้าใกล้ข้อมูลที่ถูกลูกเลือกให้มากยิ่งขึ้น ในการปรับค่าจะพิจารณาจากความแตกต่างระหว่าง BMU กับข้อมูลที่ใช้เปรียบเทียบ ผลจากความเชื่อมโยง (Neighborhood relation) ส่งผลให้เวกเตอร์น้ำหนักของเซลล์ประสาทใกล้เคียงมีการปรับค่าด้วยเช่นกัน ตามภาพที่ 10 เพื่อเพิ่มโอกาสให้กับข้อมูลใหม่ที่ถูกลูกเข้ามาในรอบถัดไป ซึ่งมีลักษณะใกล้เคียงกับข้อมูลเดิม สามารถที่จะมีเซลล์ประสาท BMU ใหม่ใกล้เคียงกับ BMU ตำแหน่งเดิมจากการลูกลูกก่อนหน้าได้ กระบวนการปรับข่ายงานประสาทจะเกิดขึ้นซ้ำ ๆ จนครบจำนวนรอบตามที่กำหนด หรือแต่ละเซลล์ประสาทมีการเลื่อนตำแหน่งน้อยกว่าค่าที่กำหนดไว้



ภาพที่ 10 การปรับตำแหน่งเข้าหาข้อมูลของเซลล์ประสาท BMU และเซลล์ประสาทใกล้เคียง

ที่มา: Vesano and Himberg (2000)

การปรับค่าเวกเตอร์น้ำหนักให้กับเซลล์ประสาทตำแหน่ง และเซลล์ประสาทใกล้เคียง จะใช้การคำนวณจากสมการที่ 8

$$\mathbf{w}_p(t+1) = \mathbf{w}_p(t) + \alpha(t) h_{cp}(t) [\mathbf{x}(t) - \mathbf{w}_p(t)] \quad (8)$$

$\mathbf{w}_p(t)$ คือ เวกเตอร์น้ำหนักของเซลล์ประสาท t คือรอบการเรียนรู้ในปัจจุบัน $\alpha(t)$ เป็น อัตราการเรียนรู้ ซึ่งจะมีการปรับลดค่าลงโดยอ้างอิงจากจำนวนรอบตามสมการที่ 9

$$\alpha(t) = \alpha_0 \times \left(\frac{T-t}{T} \right) \quad (9)$$

โดยที่ T คือจำนวนรอบของกระบวนการเรียนรู้ตามที่กำหนด $h_{cp}(t)$ เป็นฟังก์ชันค่าน้ำหนักสำหรับใช้ปรับขนาดเซลล์ประสาทใกล้เคียง ตามสมการที่ 10

$$h_{cp}(t) = \exp\left(-\|r_c - r_p\| / 2\sigma^2(t)\right) \quad (10)$$

$\|r_c - r_p\|$ คือระยะห่างทางโครงสร้างระหว่างเซลล์ประสาท BMU กับเซลล์ประสาทตำแหน่ง p ใด ๆ บนข่ายงานประสาท $\sigma(t)$ เป็นรัศมีระยะเซลล์ประสาทข้างเคียง ซึ่งตามปกติจะมีการลดค่าลงตามจำนวนรอบที่เพิ่มขึ้น ดังแสดงในสมการที่ 11

$$\sigma(t+1) = 1 + (\sigma(t) - 1) \times \left(\frac{T-t}{T} \right) \quad (11)$$

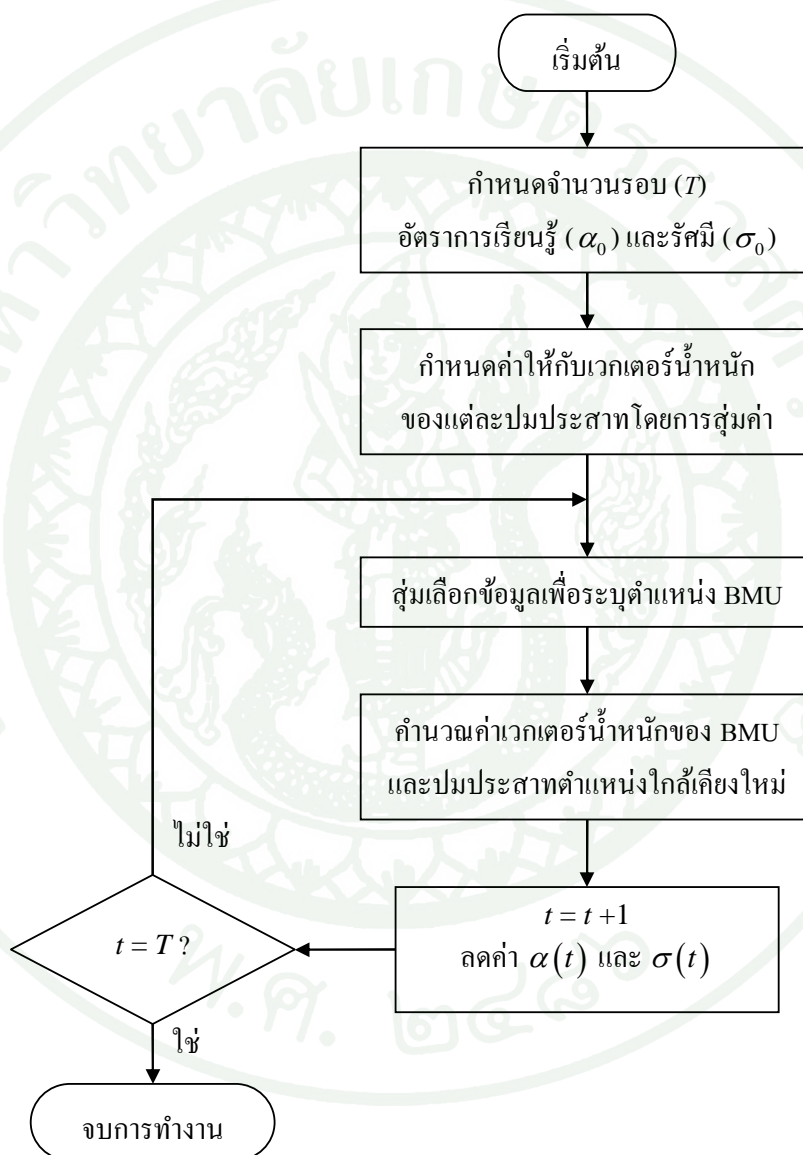
กระบวนการเรียนรู้ของ SOM เพื่อปรับข่ายงานสำหรับใช้แสดงโครงสร้าง และความถี่ของข้อมูล อธิบายได้ด้วยแผนผังลำดับการทำงานตามภาพที่ 11

จุดเด่นของการประมวลข้อมูลโดยใช้ข่ายงานประสาทแบบก่อดัวด้วยตนเอง

- 1) สามารถจัดการกับข้อมูลจำนวนมากได้
- 2) เป็นวิธีที่ใช้แสดงรายละเอียดของข้อมูลที่ประกอบขึ้นมาจากหลายตัวแปร ทำให้สามารถแสดงลักษณะของข้อมูลชนิดหลายตัวแปรให้อยู่ในรูปของแผนภาพที่มีความซับซ้อนน้อยลง ช่วยให้การศึกษาค้นคว้าข้อมูลได้รับความสะดวกมากขึ้น
- 3) สามารถเชื่อมโยงกับโครงข่ายภายนอก และจัดการตัวเองได้

ข้อสังเกตจากการเรียนรู้ตามขั้นตอนของข่ายงานประสาทแบบก่อดำด้วยตนเอง

- 1) ใช้เวลาในการประมวลผลสูง
- 2) การใช้ตัวแปรต่างกันก็จะให้ผลต่างกัน ซึ่งในปัจจุบัน การกำหนดค่าของตัวแปรที่เหมาะสมกำลังอยู่ในระหว่างการศึกษา



ภาพที่ 11 แผนผังกระบวนการเรียนรู้ของข่ายงานประสาทแบบก่อดำด้วยตนเอง

การวัดประสิทธิภาพการจัดกลุ่ม

Balasko *et al.* (2005: 13) ปัญหาหนึ่งของการวิเคราะห์กลุ่มข้อมูลคือ การเลือกจำนวนกลุ่มให้เหมาะสม เนื่องจากข้อมูลที่นำมาวิเคราะห์กลุ่มโดยส่วนใหญ่ไม่ทราบจำนวนกลุ่มที่แน่ชัด กรณีที่กำหนดจำนวนกลุ่มน้อยไปทำให้ลักษณะเด่นของข้อมูลแสดงออกมาได้ไม่ครอบคลุมมากพอ หรือในกรณีที่มีจำนวนกลุ่มมากไป รูปแบบตัวแทนของบางกลุ่มอาจซ้ำซ้อนกัน หรือได้กลุ่มที่ไม่มี ความสำคัญ ในทางปฏิบัติจะมีการแก้ปัญหาโดยทำการวิเคราะห์ตามจำนวนกลุ่มที่แตกต่างกัน ภายในช่วงที่กำหนด จากนั้นจะทำการตรวจสอบผลการวิเคราะห์กลุ่ม (Cluster validation) โดยการ คำนวณค่าประสิทธิภาพ และนำมาเปรียบเทียบเพื่อเลือกจำนวนกลุ่มที่เหมาะสมกับชุดข้อมูลมากที่สุด สำหรับวิทยานิพนธ์นี้มีการคำนวณประสิทธิภาพด้วยวิธีการของดัชนี WSJ

Sun *et al.* (2004) ได้นำเสนอดัชนีวัดประสิทธิภาพ WSJ สำหรับใช้เปรียบเทียบเพื่อเลือก จำนวนกลุ่มที่เหมาะสม การคำนวณประกอบด้วยการคำนวณหาค่าการกระจายภายในกลุ่ม (Scattering) และการคำนวณหาค่าความแตกต่างหรือความชัดเจนในการแบ่งกลุ่ม (Separation) กรณีที่ WSJ มีค่าน้อยแสดงว่าผลการวิเคราะห์ตามจำนวนกลุ่มที่กำหนดมีประสิทธิภาพ

ค่าการกระจายแสดงถึงความกระชับของกลุ่มข้อมูล ซึ่งคำนวณได้จากความแปรปรวนของกลุ่ม และความแปรปรวนรวมของชุดข้อมูล ตามสมการที่ 12

$$\begin{array}{ll} \text{ค่าความแปรปรวนภายในกลุ่ม} & \text{ค่าความแปรปรวนของชุดข้อมูล} \\ \sigma_i = \frac{1}{N_i} \sum_{j=1}^{N_i} [x_j - v_i]^2 & \sigma_x = \frac{1}{N} \sum_{j=1}^N [x_j - \bar{x}]^2 \end{array} \quad (12)$$

\bar{x} คือตำแหน่งศูนย์กลางของข้อมูล หลังจากการคำนวณค่าความแปรปรวนของแต่ละกลุ่ม และความแปรปรวนของชุดข้อมูลโดยรวมทั้งหมดแล้ว สามารถคำนวณหาค่าการกระจายตัว (Scat) หลังจากการจัดกลุ่มได้จากสมการที่ 13

$$Scat = \frac{1/c \sum_{i=1}^c (\sigma_i^T \cdot \sigma_i)^{1/2}}{(\sigma_x^T \cdot \sigma_x)^{1/2}} \quad (13)$$

สำหรับการคำนวณค่าความแตกต่างระหว่างกลุ่ม (*Sep*) จะอาศัยผลการวัดระยะระหว่างตำแหน่งศูนย์กลางกลุ่ม ตามสมการที่ 14

$$Sep = \frac{D_{\max}^2}{D_{\min}^2} \sum_{i=1}^c \left(\sum_{i'=1}^c \|v_i - v_{i'}\| \right)^{-1}, i \neq i' \quad (14)$$

โดยที่ $D_{\min} = \min_{i \neq i'} \|v_i - v_{i'}\|$ และ $D_{\max} = \max_{i \neq i'} \|v_i - v_{i'}\|$ ในส่วนของการคำนวณประสิทธิภาพโดยรวมจะพิจารณาผลรวมจากทั้งสองส่วน ได้แก่ ค่าการกระจายภายในกลุ่ม และค่าความแตกต่างระหว่างกลุ่ม เมื่อ c_{\max} คือจำนวนกลุ่มสูงสุดภายในช่วงการจัดกลุ่มตามที่กำหนด ดังนั้นสามารถคำนวณประสิทธิภาพตามวิธีการของดัชนี WSJ ได้จากสมการที่ 15

$$WSJ = Scat + \frac{Sep}{Sep(c_{\max})} \quad (15)$$

อุปกรณ์และวิธีการ

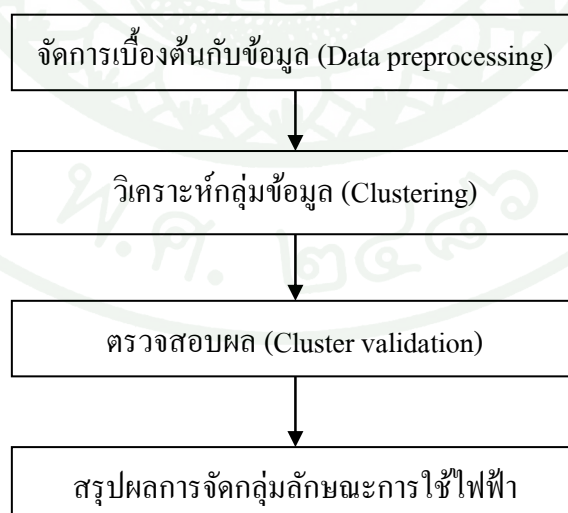
อุปกรณ์

อุปกรณ์ที่จำเป็นสำหรับทำวิทยานิพนธ์ ประกอบด้วย

1. เครื่องคอมพิวเตอร์โน้ตบุ๊ก ระบบปฏิบัติการ Window 7 Ultimate 32 bit
2. โปรแกรม Microsoft Office (Excel, Word)
3. โปรแกรม MATLAB R2010b

วิธีการ

การค้นหาลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนด้วยการวิเคราะห์กลุ่มข้อมูล ด้วยแต่ละวิธีที่ต่างกันจะประกอบด้วยขั้นตอนหลักตามภาพที่ 12 ขั้นตอนแรกเป็นการจัดการเบื้องต้นกับข้อมูล (Data preprocessing) เพื่อให้พร้อมต่อการนำไปใช้ในขั้นตอนการวิเคราะห์กลุ่มข้อมูล (Clustering) จากนั้นจะทำการตรวจสอบผลการวิเคราะห์ (Cluster validation) ด้วยการวัดประสิทธิภาพการจัดกลุ่ม ซึ่งจะอาศัยเงื่อนไขการตรวจสอบตามเกณฑ์ที่ระบุไว้ สำหรับใช้เลือกผลการวิเคราะห์ที่เหมาะสม เพื่อนำไปสู่การสรุปผลการจัดกลุ่มลักษณะการใช้ไฟฟ้าในขั้นตอนสุดท้าย



ภาพที่ 12 ลำดับขั้นตอนการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้า

การจัดการเบื้องต้นกับข้อมูล

การจัดการเบื้องต้นกับข้อมูล เริ่มต้นจากการสำรวจ และคัดแยกลักษณะการใช้ไฟฟ้าที่เป็นปัญหาออกจากการวิเคราะห์ เพราะอาจส่งผลกระทบต่อความถูกต้องในการวิเคราะห์กลุ่มข้อมูลได้ ภาพที่ 13 แสดงตัวอย่างของปัญหาอย่างหนึ่ง เป็นกรณีที่มีการสูญหายของค่าการใช้กำลังไฟฟ้า บางส่วนหรือตลอดทั้งวัน และภาพที่ 14 เป็นปัญหาของค่าการใช้ไฟฟ้าที่ผิดปกติ จากภาพเป็นการบันทึกค่าการใช้ไฟฟ้าเป็น 0 ติดต่อกัน ซึ่งต่างไปจากช่วงเวลาเดียวกันของวันอื่น ๆ อย่างชัดเจน สำหรับลักษณะการใช้ไฟฟ้าของวันที่ตรวจพบค่าการใช้กำลังไฟฟ้าที่เป็นปัญหาเหล่านี้ไม่ว่าจะเป็นเพียงบางช่วงเวลาหรือตลอดทั้งวัน จะถูกตัดออกและไม่นำมาใช้ร่วมกับการวิเคราะห์กลุ่ม

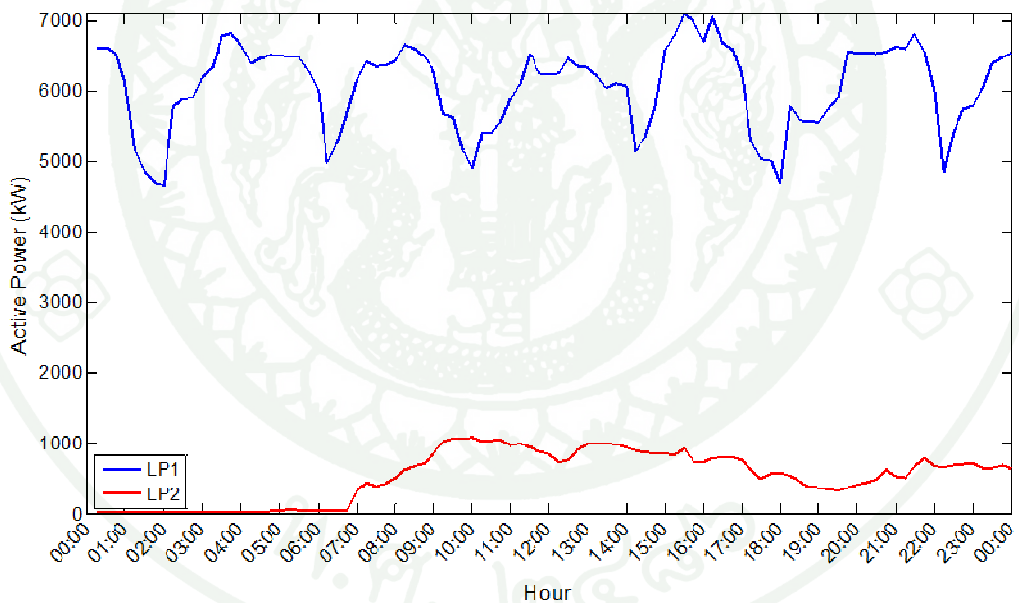
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	dd/mm/yy	0.15	0.30	0.45	1.00	1.15	1.30	1.45	2.00	2.15	2.30	2.45	3.00	3.15	3.30
2	16/9/2011	79.74	81.36	89.46	79.38	79.74	86.04	76.32	80.64	77.22	76.68	76.32	77.94	77.94	75.78
3	19/9/2011	35.1	32.4	32.4	39.42	32.58	32.4	31.86	32.4	32.22	32.22	32.4	31.86	32.4	32.22
4	20/9/2011	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	22.32	22.14	21.6	20.7	20.34	20.52	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	3.24	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	3.6	3.24	3.42	3.24	3.42	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 13 ตัวอย่างลักษณะการใช้ไฟฟ้าที่มีการสูญหายของค่าการใช้ไฟฟ้าบางช่วงเวลา

หลังจากทำการสำรวจ และตัดลักษณะการใช้ไฟฟ้าในวันที่ตรวจพบค่าการใช้ไฟฟ้าที่มีปัญหาออกก่อนการนำไปวิเคราะห์ ขั้นตอนต่อมาจะทำการบรรทัดฐานข้อมูล (Data normalization) เพื่อให้ข้อมูลมีค่าอยู่ในช่วงเดียวกัน วิธีการที่ใช้คือ จะทำการปรับค่ายอดสูงสุด (Peak Demand) ของลักษณะการใช้ไฟฟ้าแต่ละวันให้มีค่าเท่ากับ 1 ในภาพที่ 15 และ 16 แสดงการเปรียบเทียบตัวอย่างลักษณะการใช้ไฟฟ้าก่อนและหลังการบรรทัดฐานข้อมูลตามลำดับ

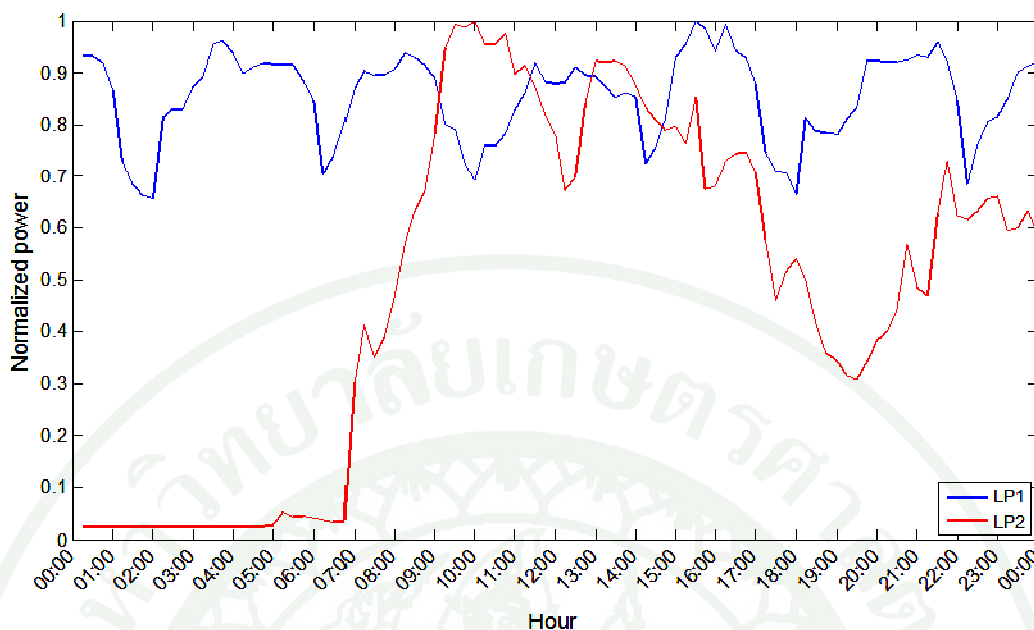
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	dd/mm/yy	0.15	0.30	0.45	1.00	1.15	1.30	1.45	2.00	2.15	2.30	2.45	3.00	3.15	3.30
2	2/6/2011	16.116	15.996	16.152	16.812	16.14	15.824	16.344	16.116	15.96	13.872	13.32	14.544	14.148	15.816
3	3/6/2011	15.936	16.044	16.524	16.584	15.888	16.14	16.656	16.164	15.996	13.992	15.204	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 14 ตัวอย่างลักษณะการใช้ไฟฟ้าที่มีค่าการใช้ไฟฟ้าเป็น 0 ต่อเนื่องนานผิดปกติ



ภาพที่ 15 ลักษณะการใช้ไฟฟ้าก่อนผ่านขั้นตอนการปรับค่าให้เป็นบรรทัดฐาน

จากภาพที่ 15 เป็นตัวอย่างลักษณะการใช้ไฟฟ้าจำนวน 2 รูปแบบภายในฐานข้อมูลชุดเดียวกัน จะเห็นได้ว่าช่วงของปริมาณการใช้ไฟฟ้า LP1 และ LP2 มีค่าต่างกันมาก ซึ่งจะส่งผลต่อการจัดกลุ่มข้อมูล เนื่องจากในการวิเคราะห์จะให้ความสนใจกับรูปแบบการใช้ไฟฟ้า เป็นสำคัญ ดังนั้นจึงต้องทำบรรทัดฐานลักษณะการใช้ไฟฟ้าก่อนนำไปวิเคราะห์ ดังในภาพที่ 16

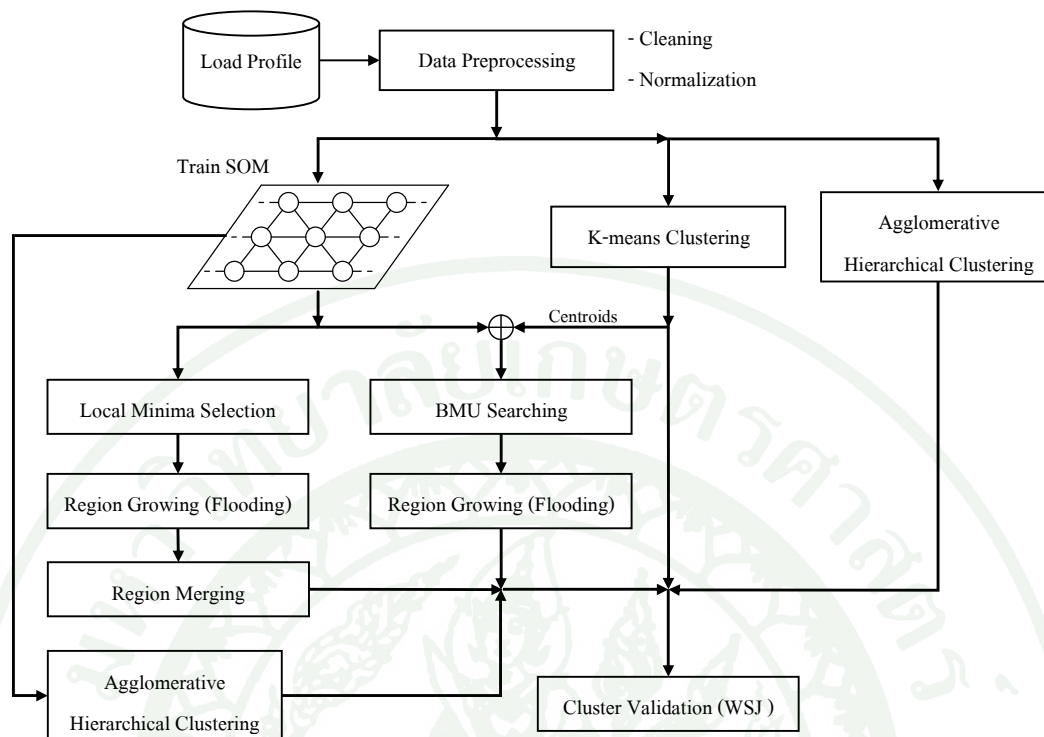


ภาพที่ 16 ลักษณะการใช้ไฟฟ้าหลังจากผ่านขั้นตอนการปรับค่าให้เป็นบรรทัดฐาน

ในภาพที่ 16 เป็นเส้นโค้งการแสดงผลการใช้ไฟฟ้าที่ทำการบรรทัดฐานข้อมูลแล้ว จะเห็นได้ว่าค่ายอดสูงสุดของลักษณะการใช้ไฟฟ้า LP1 และ LP2 ถูกปรับให้เท่ากับ 1 เหมือนกัน ผลจากการปรับค่าการใช้ไฟฟ้าให้อยู่ในช่วงเดียวกัน จะทำให้การจัดกลุ่มลักษณะการใช้ไฟฟ้าได้รูปแบบตัวแทนที่ครอบคลุมกับสมาชิกภายในกลุ่มมากขึ้น

ขั้นตอนการวิเคราะห์กลุ่ม

ในวิทยานิพนธ์นี้จะทำการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าอาศัยขั้นตอนของวิธีการต่าง ๆ กัน เพื่อเปรียบเทียบประสิทธิภาพ โดยประกอบด้วย ขั้นตอนของวิธีเคมีน ขั้นตอนของวิธีแบบเป็นลำดับชั้น ขั้นตอนของวิธีแบบเป็นลำดับชั้นบนข่ายงานประสาท ขั้นตอนของเทคนิคการขยายพื้นที่ (Region-Growing) บนพื้นฐานของข่ายงานประสาท และขั้นตอนการทำงานร่วมกันแบบผสมระหว่างเคมีนกับเทคนิคการขยายพื้นที่บนพื้นฐานของข่ายงานประสาท ดังแสดงในภาพที่ 17



ภาพที่ 17 โครงสร้างการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้า

Vesanto *et al.* (2006) เสนอวิธีจัดกลุ่มใหม่ โดยใช้ประโยชน์จากคุณสมบัติทางโครงสร้างของข่ายงานประสาท การทำงานจะอยู่ในลักษณะของการขยายพื้นที่ (Region-Growing) อาศัยความสัมพันธ์ที่เชื่อมโยงกัน โดยตรงระหว่างเซลล์บนข่ายงานประสาท จากกระบวนการเรียนรู้ในระหว่างการปรับข่ายงานประสาทเข้าหาโครงสร้างของข้อมูลที่น่ามาฝึกสอน สำหรับขั้นตอนการจัดกลุ่มด้วยวิธีขยายพื้นที่บนพื้นฐานข่ายงานประสาทมีรายละเอียดดังต่อไปนี้

1. ระบุตำแหน่งเริ่มต้นสำหรับใช้สร้างกลุ่มหรือพื้นที่ เพื่อเริ่มการเชื่อมต่อของเซลล์ประสาทหนึ่งเข้ากับอีกเซลล์ประสาทหนึ่ง สำหรับเซลล์ประสาทตำแหน่งพื้นที่ว่างบนข่ายงานประสาท (Sample Hit = 0) จะถูกกำหนดให้เป็นบริเวณขอบของกลุ่ม เพื่อใช้จำกัดการแผ่ขยายพื้นที่จากกลุ่มใด ๆ ที่มีการเชื่อมโยงของเซลล์ประสาทใกล้เคียงเข้ามา
2. คำนวณระยะห่างระหว่างเซลล์ประสาทที่เหลือกับทุก ๆ สมาชิกในแต่ละกลุ่ม โดยประยุกต์จากวิธีคำนวณระยะแบบค่าใกล้เคียงกันมากที่สุด (Single Linkage) ตามสมการที่ 16

$$d(A_i, j) = \min \{ \text{dist}(w_k, w_j) \}, i=1, 2 \dots c \quad (16)$$

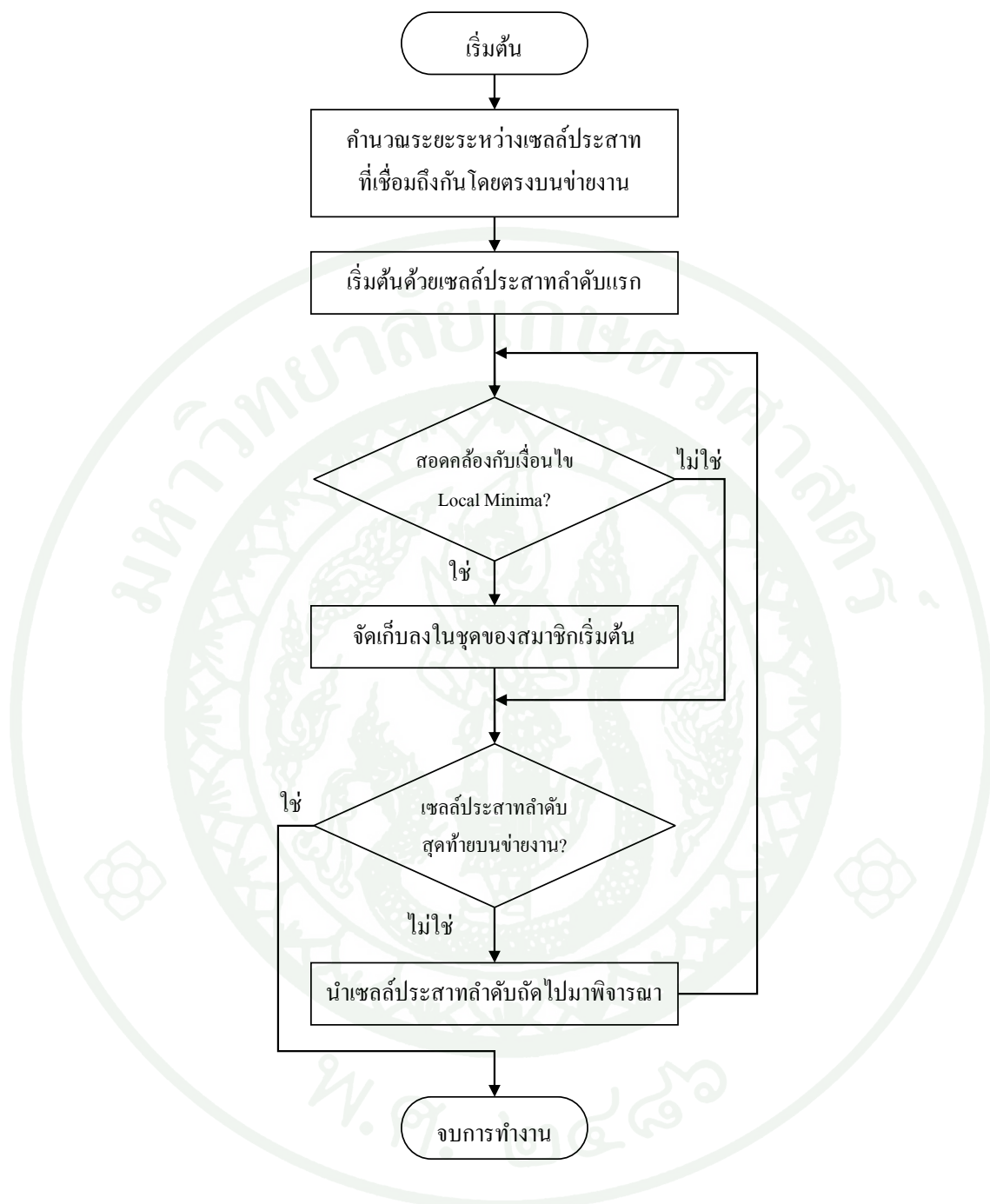
n_i คือ จำนวนเซลล์ประสาททั้งหมดที่เป็นสมาชิกของกลุ่ม i และ j แทนตำแหน่งของเซลล์ประสาทใด ๆที่กำลังรอการจัดเข้ากลุ่ม

3. ขยายพื้นที่ โดยทำการเชื่อมต่อเซลล์ประสาทกับกลุ่มที่มีระยะการเชื่อมโยงเฉลี่ยใกล้กันมากที่สุดเข้าไว้ด้วยกัน แต่ต้องเป็นไปตามข้อบังคับด้านความต่อเนื่อง (Continuity Constraint) คือภายในกลุ่มที่จะต้องมีความเชื่อมโยงถึงกันโดยตรงกับตำแหน่งของเซลล์ประสาทที่กำลังจะเชื่อมต่อเข้ากับกลุ่มเพิ่มเติม เพื่อให้แน่ใจว่าการขยายของพื้นที่จะเป็นไปในลักษณะเชื่อมโยงจากเซลล์ที่อยู่ใกล้กันอย่างต่อเนื่อง

4. ทำซ้ำตั้งแต่ขั้นตอนที่ 2 จนกระทั่งแต่ละกลุ่มไม่สามารถขยายพื้นที่ออกไปได้อีก

5. ในกรณีที่พบเซลล์ประสาทบางส่วนตกค้าง ซึ่งยังไม่ได้รับการจัดเข้ากลุ่มใดๆ เช่น ตำแหน่งของเซลล์ประสาทที่ถูกใช้เพื่อจำกัดการขยายขอบเขตพื้นที่จากขั้นตอนที่ 1 จะต้องกำหนดกลุ่มให้กับเซลล์ประสาทเหล่านี้ โดยจะถูกจัดลงในกลุ่มเดียวกันกับกลุ่มของเซลล์ประสาทข้างเคียงที่เชื่อมถึงกัน (Neighboring map unit) และอยู่ใกล้กันมากที่สุด

ในการจัดกลุ่มตามวิธีการดั้งเดิม จำนวนพื้นที่ย่อยบนขำงานประสาทจะถูกจัดแบ่งออกเป็นส่วน ๆ เท่ากับจำนวนเซลล์ประสาทเริ่มต้นตามที่กำหนด ซึ่งอยู่ในตำแหน่งที่สอดคล้องกับเงื่อนไขของระยะการเชื่อมโยงต่ำสุดเฉพาะที่ (Local Minima) โดยมีขั้นตอนในการคัดเลือกคือ เริ่มจากการคำนวณหาค่าระยะเฉลี่ยเฉพาะเซลล์ประสาทที่เชื่อมโยงถึงกัน จากนั้นจะทำการสำรวจโดยเริ่มต้นจากเซลล์ประสาทตำแหน่งแรกจนครบทุกเซลล์บนขำงานประสาท เซลล์ประสาทตำแหน่งใดที่มีค่าระยะจากการคำนวณในขั้นตอนแรกต่ำกว่าทุกเซลล์ประสาทที่เชื่อมถึงกัน โดยตรง จะได้รับการคัดเลือก และนำเข้าสู่กระบวนการจัดกลุ่มต่อไป ลำดับขั้นตอนการคัดเลือกสมาชิกเริ่มต้นตามเงื่อนไขของระยะการเชื่อมโยงต่ำสุดเฉพาะที่ เพื่อใช้เป็นข้อมูลนำเข้าให้กับการวิเคราะห์กลุ่มข้อมูลด้วยกระบวนการขยายพื้นที่ แสดงไว้ในภาพที่ 18



ภาพที่ 18 ขั้นตอนการคัดเลือกเซลล์ประสาทจากเงื่อนไขระยะการเชื่อมโยงต่ำสุดเฉพาะที่

หลังจากทำการจัดแบ่งพื้นที่ย่อยบนขั้วงานประสาทแล้ว สุดท้ายจะเป็นขั้นตอนการสร้างกลุ่มหลัก โดยทำการรวมกลุ่มพื้นที่ย่อยต่าง ๆ (Merge Region) ที่มีลักษณะคล้ายเข้าไว้ด้วยกัน เพื่อแก้ปัญหาจากการกระจายของข้อมูลที่มีลักษณะคล้ายกันเป็นช่วงกว้าง ทำให้เซลล์ประสาทที่อยู่ห่างไกลออกไปยังคงมีลักษณะคล้ายกัน ทำให้มีโอกาที่จะถูกเลือกซ้ำได้ วิธีการรวมกลุ่มของพื้นที่

ย่อย จะใช้ขั้นตอนของวิธีแบบเป็นลำดับชั้นชนิด Agglomerative สำหรับขั้นตอนการวิเคราะห์กลุ่ม ด้วยเทคนิคการขยายพื้นที่อาศัยการเชื่อมโยงของเซลล์บนข่ายงานประสาทในแบบพื้นฐาน สรุปได้ ดังภาพที่ 19



ภาพที่ 19 ขั้นตอนการวิเคราะห์กลุ่มด้วยเทคนิคการขยายพื้นที่อาศัยความเชื่อมโยงของเซลล์บน ข่ายงานประสาทแบบดั้งเดิม

และสำหรับวิธีที่เสนอในบทความนี้ เป็นการทำงานแบบร่วมกัน โดยในขั้นตอนการจัดกลุ่มจะใช้กระบวนการแผ่ขยายพื้นที่จากการเชื่อมโยงของเซลล์บนข่ายงานประสาท ตามวิธีดั้งเดิม ที่ได้เสนอไปก่อนหน้านี้ แต่จะมีการปรับเปลี่ยนเงื่อนไขในส่วนของการค้นหาชุดของสมาชิกเริ่มต้น โดยวิธีปรับปรุงใหม่จะใช้นำเข้าตำแหน่งศูนย์กลางจากผลการจัดกลุ่มด้วยวิธีเคมีน จากนั้นจะสำรวจหาตำแหน่งเริ่มต้นที่เหมาะสม อาศัยวิธีการทำงานเช่นเดียวกับการค้นหาตำแหน่ง BMU โดยมองว่า ข้อมูลที่ถูกสุ่มเลือกในที่นี้ถูกแทนที่ด้วยตำแหน่งศูนย์กลางข้อมูลของวิธีเคมีน ในกรณีที่ BMU ของศูนย์กลางข้อมูลตรงกับเซลล์ประสาทตำแหน่งพื้นที่ว่าง ก็ให้เลือก BMU ลำดับถัดไป เซลล์แรกที่ไม่ใช่ตำแหน่งพื้นที่ว่างเข้ามาแทนที่ ในภาพที่ 20 เป็นการสรุปภาพรวมขั้นตอนการจัดกลุ่มของวิธีการทำงานแบบผสมผสานตามที่นำเสนอ



ภาพที่ 20 ขั้นตอนการวิเคราะห์กลุ่มข้อมูลด้วยวิธีการทำงานร่วมกันของเคมีน และเทคนิคการขยายพื้นที่อาศัยความเชื่อมโยงของเซลล์บนข่ายงานประสาท

การตรวจสอบผลการวิเคราะห์กลุ่ม

หลังจากทำการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าแล้ว ขั้นตอนถัดมาคือการตรวจสอบผลการจัดกลุ่ม (Cluster Validation) ซึ่งในวิทยานิพนธ์นี้จะใช้การตรวจสอบด้วยการคำนวณค่าประสิทธิภาพการจัดกลุ่มด้วยดัชนี WSI โดยวิธีวิเคราะห์กลุ่มที่ให้ค่าประสิทธิภาพจากการคำนวณต่ำกว่า แสดงว่ามีการจัดข้อมูลเข้ากลุ่มได้ดีกว่า

การใช้ข่ายงานประสาทเข้ามาช่วยวิเคราะห์ เป็นการทำงานกับค่าเวกเตอร์นำหนักของเซลล์ประสาทแทนการวิเคราะห์โดยตรงจากชุดข้อมูล แต่ในด้านการตรวจสอบผลการจัดกลุ่มเป็นการวัดค่าประสิทธิภาพการเข้ากลุ่มของชุดข้อมูล ดังนั้นจะต้องมีการระบุกลุ่มที่ชัดเจนให้กับแต่ละข้อมูลจะทำการวัดประสิทธิภาพ ซึ่งตามปกติหลังจากเสร็จสิ้นกระบวนการเรียนรู้ ข้อมูลฝึกสอนจะกระจายลงไปในแต่ละเซลล์ประสาทตามรูปแบบเวกเตอร์นำหนักที่คล้ายมากที่สุด ข้อมูลที่มีลักษณะคล้ายกันจะถูกจัดวางในตำแหน่งเซลล์ประสาทเดียวกันหรือเซลล์ประสาทที่ใกล้เคียงกัน

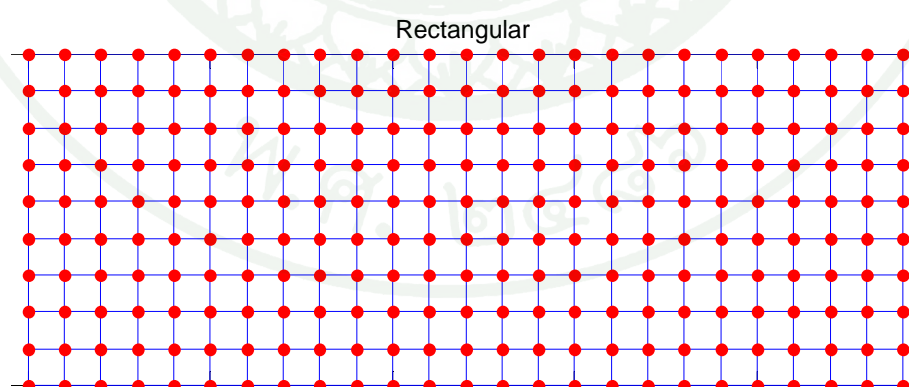
นั่นก็หมายความว่า ถ้าข้อมูลเป็นสมาชิกของเซลล์ประสาทตำแหน่งใด ให้ถือว่าอยู่ในกลุ่มเดียวกันกับเซลล์ประสาทที่ ตำแหน่งนั้นด้วยเช่นกัน

สรุปผลการจัดกลุ่มลักษณะการใช้ไฟฟ้า

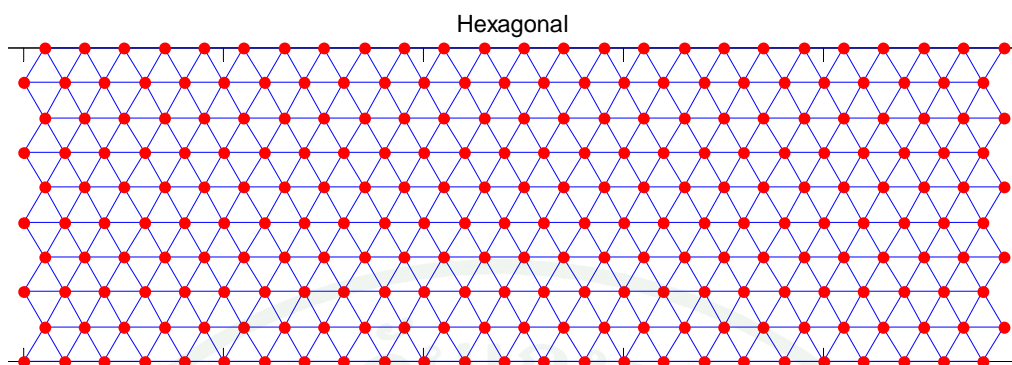
หลังจากทำการจัดกลุ่มและผ่านขั้นตอนการตรวจสอบผลการจัดกลุ่มลักษณะการใช้ไฟฟ้าแล้ว จะนำไปสู่การสรุปผลเพื่อวิเคราะห์ว่าการจัดกลุ่มแต่ละวิธีนั้นมีข้อดีข้อเสียอย่างไร และวิธีใดให้ผลการจัดกลุ่มที่เหมาะสมกับลักษณะการใช้ไฟฟ้ามากที่สุด โดยนำค่าประสิทธิภาพจากการจัดกลุ่มของแต่ละวิธีการมาเปรียบเทียบ

การเตรียมข่ายงานประสาท

การเตรียมข่ายงานประสาทประกอบการวิเคราะห์ห้กลุ่ม จะต้องมีการกำหนดโครงสร้างการเชื่อมต่อระหว่างเซลล์ และจำนวนเซลล์ประสาท ซึ่งขึ้นอยู่กับผู้ออกแบบว่าต้องการแสดงรายละเอียดให้ปรากฏบนข่ายงานประสาทมากน้อยเพียงใด สำหรับในส่วนของโครงสร้างการเชื่อมต่อระหว่างเซลล์ประสาท จะขึ้นอยู่กับลักษณะฟังก์ชันโครงสร้าง (Topology Function) โดยมีอยู่ 2 ลักษณะคือ ฟังก์ชันโครงสร้างในแบบของตาราง (Rectangular Topology Function) และฟังก์ชันโครงสร้างในแบบหกเหลี่ยม (Hexagonal Topology Function) ดังในภาพที่ 21



ภาพที่ 21 ข่ายงานประสาทในแบบตาราง (Rectangular) และแบบหกเหลี่ยม (Hexagonal)



ภาพที่ 21 (ต่อ)

จากภาพที่ 21 โครงสร้างการเชื่อมต่อในแบบของตาราง แต่ละเซลล์ประสาทจะเชื่อมต่อถึงกันโดยตรงกับเซลล์ตำแหน่งอื่น ๆ ได้สูงสุดจำนวน 4 เซลล์ประสาท ซึ่งจะแตกต่างจากโครงสร้างการเชื่อมต่อในแบบหกเหลี่ยมที่มีการเชื่อมต่อโดยตรงกับเซลล์ข้างเคียงได้สูงสุดจำนวน 6 เซลล์ประสาท ในวิทยานิพนธ์นี้ได้เลือกใช้โครงสร้างแบบหกเหลี่ยมเพราะมีการเชื่อมโยงกันเซลล์ประสาทด้วยจำนวนที่มากกว่า

การกำหนดจำนวนเซลล์ประสาทเพื่อให้เหมาะสมกับชุดข้อมูลฝึกสอนในปัจจุบันยังไม่มีกฎเกณฑ์ที่ตายตัว และสำหรับในวิทยานิพนธ์นี้มีการเตรียมข่ายงานประสาทประกอบการวิเคราะห์กลุ่มที่มีขนาดต่างกัน แบ่งได้เป็น 2 กรณี คือ

กรณีที่ 1 ทำการสร้างข่ายงานประสาท กำหนดให้มีจำนวนเซลล์เรียงตัวเท่ากันในทุก ๆ ด้าน โดยปรับขนาดตั้งแต่ 14×14 ไปจนถึง 30×30 แต่ละข่ายงานประสาทจะมีขนาดต่างกันด้านละ 2 เซลล์ประสาท

กรณีที่ 2 กำหนดขนาดของข่ายงานประสาทด้วยอัตราส่วนด้านกว้างต่อด้านยาว ซึ่งสัมพันธ์กับค่าลักษณะเฉพาะ (Eigenvalue) จากเมตริกซ์ความแปรปรวนร่วม (Covariance Matrix) หรือ $\text{cov}(X)$ ของข้อมูล ซึ่งคำนวณได้จากสมการที่ 17 (Tsekouras and Salis, 2008)

$$\text{cov}(X) = \frac{1}{N-1} \sum_{j=1}^N (\mathbf{x}_j - \bar{\mathbf{x}}) \cdot (\mathbf{x}_j - \bar{\mathbf{x}})^T \quad (17)$$

\bar{x} เป็นตำแหน่งศูนย์กลางหรือค่าเฉลี่ยข้อมูล ขนาดของเมตริกซ์ที่คำนวณได้จะเท่ากับ $d \times d$ เมื่อ d คือจำนวนตัวแปรของแต่ละข้อมูล สำหรับจำนวนเซลล์ที่เรียงตัวในแต่ละด้านของข่ายงานประสาท (n_x, n_y) แปรตามสัดส่วนค่าลักษณะเฉพาะของเมตริกซ์ความแปรปรวนร่วม ที่มีค่าสูงสุดสองค่าแรก (λ_1, λ_2 โดยที่ $\lambda_1 > \lambda_2$) ตามสมการที่ 18

$$n_x \sim \sqrt{munits/ratio}, n_y \sim munits/R \text{ โดยที่ } ratio = \sqrt{\lambda_1/\lambda_2} \quad (18)$$

$munits$ คือจำนวนเซลล์ประสาทตามที่กำหนด สำหรับในกรณีนี้ 2 จะกำหนดจำนวนเซลล์ประสาทภายในช่วง เริ่มตั้งแต่ 200 จนถึง 900 เซลล์ประสาท โดยทำการปรับจำนวนเซลล์เพิ่มให้กับข่ายงานประสาทลำดับถัดไป ครั้งละ 50 เซลล์ประสาท

การสร้างจัดเตรียมโดยใช้กระบวนการเรียนรู้ของข่ายงานประสาทแบบก่อดำด้วยตนเองในวิทยานิพนธ์นี้ จะดำเนินการภายใต้เงื่อนไขที่มีการควบคุมประกอบด้วย จำนวนรอบการเรียนรู้ (Epochs) เท่ากับ 300 รอบ อัตราการเรียนรู้ $\alpha(0)$ มีค่าเริ่มต้นเท่ากับ 0.5 ซึ่งเป็นค่าตามที่โปรแกรมกำหนด (Vesanto *et al.*, 2000) และค่าเริ่มต้นของรัศมี $\sigma(0)$ กำหนดให้มีค่าเท่ากับครึ่งหนึ่งของด้านที่มีเซลล์ประสาทเรียงตัวกันเป็นจำนวนมากว่า (Tsekouras and Salis, 2008) สำหรับรายละเอียดของทุกข่ายงานประสาทที่สร้างขึ้นจากเงื่อนไขที่กำหนดไว้ ได้แก่ ขนาดของงานประสาท จำนวนเซลล์ทั้งหมดบนข่ายงานประสาท และจำนวนเซลล์ในตำแหน่งพื้นที่ว่าง ที่จะนำมาใช้วิเคราะห์กลุ่มข้อมูลประกอบการทำวิทยานิพนธ์นี้ แสดงไว้ในตารางที่ 1

ตารางที่ 1 รายละเอียดของข่ายงานประสาทจากการจัดเตรียมตามเงื่อนไขที่กำหนด

ลำดับที่	ขนาดของข่ายงาน	จำนวนเซลล์ประสาททั้งหมด	จำนวนเซลล์ตำแหน่งพื้นที่ว่าง
1	29×7	203	0
2	31×8	248	0
3	33×9	297	2
4	35×10	350	3
5	40×10	400	8
6	41×11	451	6

ตารางที่ 1 (ต่อ)

ลำดับที่	ขนาดของข่ายงาน	จำนวนเซลล์ประสาททั้งหมด	จำนวนเซลล์ตำแหน่งพื้นที่ว่าง
7	42×12	504	16
8	46×12	552	23
9	46×13	598	25
10	50×13	650	38
11	50×14	700	38
12	54×14	756	42
13	53×15	795	54
14	14×14	196	0
15	16×16	256	1
16	18×18	324	2
17	20×20	400	4
18	22×22	484	17
19	24×24	576	19
20	26×26	676	24
21	28×28	784	39
22	57×15	855	56
23	56×16	896	61
24	30×30	900	62

ผลและวิจารณ์

ผล

ลักษณะการใช้ไฟฟ้ารายวันที่นำมาวิเคราะห์เป็นของผู้ใช้ไฟฟ้าในกลุ่มอุตสาหกรรมประเภทหนึ่งจำนวน 14 รายในช่วงระยะเวลา 1 ปี โดยเลือกใช้เฉพาะวันจันทร์-เสาร์ เพราะเป็นช่วงเวลาที่ทางภาคอุตสาหกรรมมีการใช้ไฟฟ้าที่ส่งผลให้การใช้ไฟฟ้าโดยรวมของระบบมีค่าสูง ทั้งนี้มีลักษณะการใช้ไฟฟ้าที่ผ่านการคัดเลือกตามขั้นตอนการจัดการเบื้องต้นจำนวน 3,795 ข้อมูล สำหรับผลการตรวจสอบการวิเคราะห์ด้วยดัชนี WSJ ของแต่ละวิธีการในช่วงจำนวนกลุ่ม (c) ตั้งแต่ 2 ถึง 10 กลุ่ม มีรายละเอียดดังต่อไปนี้

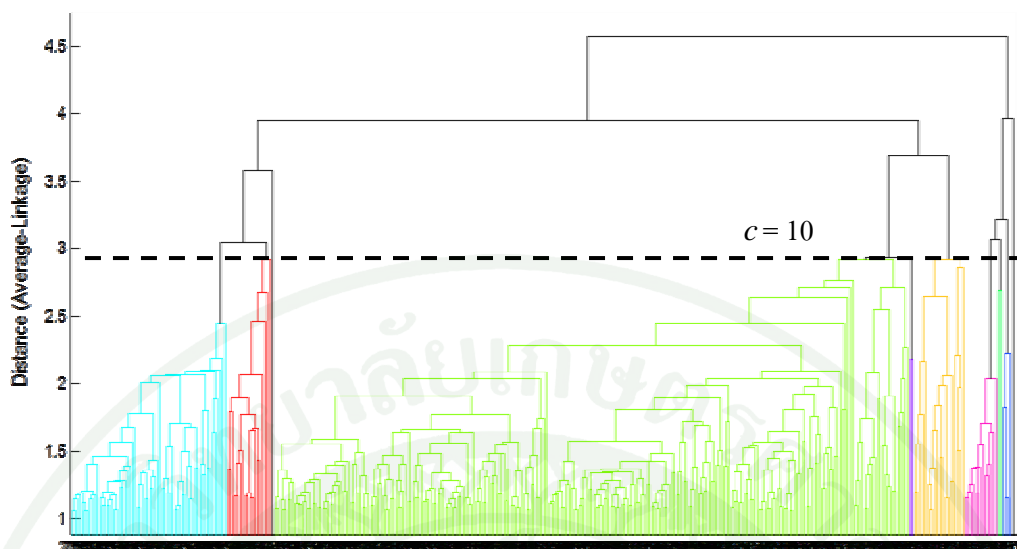
ผลการวิเคราะห์กลุ่มตามขั้นตอนของวิธีแบบเป็นลำดับขั้น

ผลการวิเคราะห์กลุ่มข้อมูลลักษณะการใช้ไฟฟ้าตามขั้นตอนของวิธีแบบเป็นลำดับขั้น และคำนวณระยะด้วยวิธีการเชื่อมโยงแบบค่าเฉลี่ย (Average-Linkage) โดยใช้การคำนวณด้วยดัชนี WSJ แสดงไว้ในตารางที่ 2

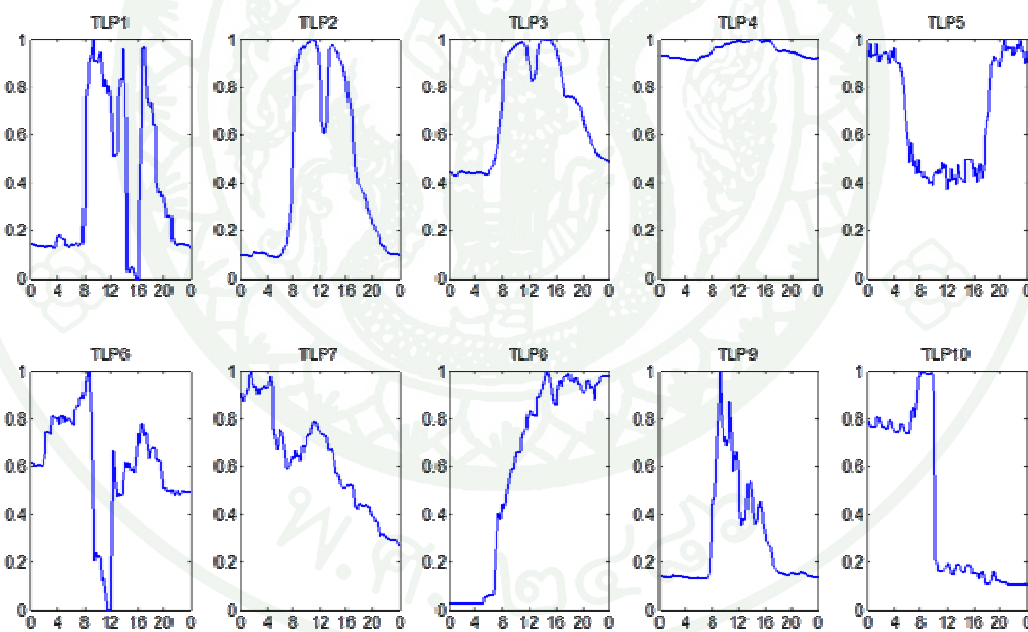
ตารางที่ 2 ผลการคำนวณค่าประสิทธิภาพด้วยดัชนี WSJ เมื่อทำการวิเคราะห์กลุ่มตามขั้นตอนของวิธีแบบเป็นลำดับขั้น ในช่วง 2-10 กลุ่ม

	c=2	c=3	c=4	c=5	c=6	c=7	c=8	c=9	c=10
WSJ	8.1641	3.8966	2.2515	1.7279	1.6787	1.6237	1.4639	1.0752	1.0386

จากผลการวิเคราะห์กลุ่มข้อมูลลักษณะการใช้ไฟฟ้าในช่วง 2-10 กลุ่ม ตามตารางที่ 2 เมื่อเปรียบเทียบค่าประสิทธิภาพ พบว่า จำนวนกลุ่มที่มีค่าประสิทธิภาพซึ่งคำนวณด้วยดัชนี WSJ ต่ำสุดคือ 10 กลุ่ม ดังนั้นสรุปได้ว่า เมื่อทำการวิเคราะห์ตามขั้นตอนของวิธีแบบเป็นลำดับขั้น สามารถวิเคราะห์กลุ่มของลักษณะการใช้ไฟฟ้าได้จำนวน 10 กลุ่ม โดยมีผลการวิเคราะห์ในรูปแบบแผนภาพต้นไม้ลำดับขั้นตามภาพที่ 22 และมีลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนของแต่ละกลุ่มแสดงไว้ในภาพที่ 23



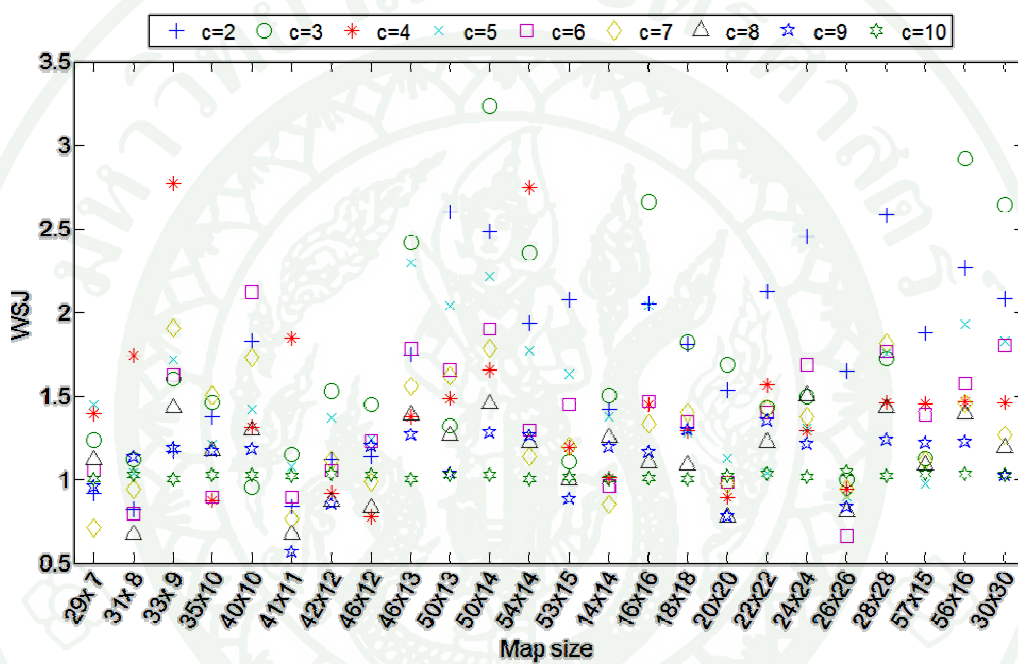
ภาพที่ 22 ผลการวิเคราะห์กลุ่มในรูปแบบโครงสร้างแบบแผนภาพต้นไม้ลำดับชั้นจำนวน 10 กลุ่ม



ภาพที่ 23 ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 10 กลุ่ม หลังจากทำการวิเคราะห์ด้วยวิธีแบบเป็นลำดับชั้น

ผลการวิเคราะห์กลุ่มตามขั้นตอนของวิธีแบบเป็นลำดับชั้นบนข่ายงานประสาทที่จัดเตรียมไว้

ดำเนินการวิเคราะห์กลุ่มข้อมูลลักษณะการใช้ไฟฟ้าจากข่ายงานประสาทที่เตรียมไว้ตามตารางที่ 1 โดยขั้นตอนของวิธีแบบเป็นลำดับชั้น และคำนวณระยะด้วยวิธีการเชื่อมโยงแบบค่าเฉลี่ย (Average-Linkage) ผลการวัดประสิทธิภาพจากการคำนวณด้วยดัชนี WSJ บนข่ายงานประสาทที่มีขนาดต่างกัน แสดงไว้ในภาพที่ 24



ภาพที่ 24 ค่าประสิทธิภาพของดัชนี WSJ เมื่อทำการวิเคราะห์กลุ่มข้อมูลด้วยขั้นตอนของวิธีแบบเป็นลำดับชั้นในช่วง 2-10 กลุ่ม บนข่ายงานประสาทที่มีขนาดต่างกัน

จากผลการวิเคราะห์กลุ่มตามขั้นตอนของวิธีแบบเป็นลำดับชั้น โดยใช้ร่วมข่ายงานประสาทที่มีขนาดต่างกัณดังแสดงในภาพที่ 24 สามารถแยกพิจารณาได้เป็น 2 ลักษณะ ตามตารางที่ 3 และ 4 ตามลำดับ ในตารางที่ 3 คือการพิจารณาเทอมของค่าประสิทธิภาพที่คำนวณได้ ในแต่ละช่วงของจำนวนกลุ่มตั้งแต่ 2-10 กลุ่ม โดยจะแสดงเฉพาะค่าประสิทธิภาพ และขนาดของข่ายงานประสาทที่ช่วยให้การวิเคราะห์ในแต่ละช่วงจำนวนกลุ่มมีประสิทธิภาพมากที่สุด (msize) เทียบกับข่ายงานอื่น ๆ และสำหรับตารางที่ 4 เป็นการสรุปผลการวิเคราะห์ของแต่ละข่ายงานประสาท ในเทอมของจำนวนกลุ่มที่เหมาะสม (c^*) และค่าประสิทธิภาพตามวิธีคำนวณของดัชนี WSJ

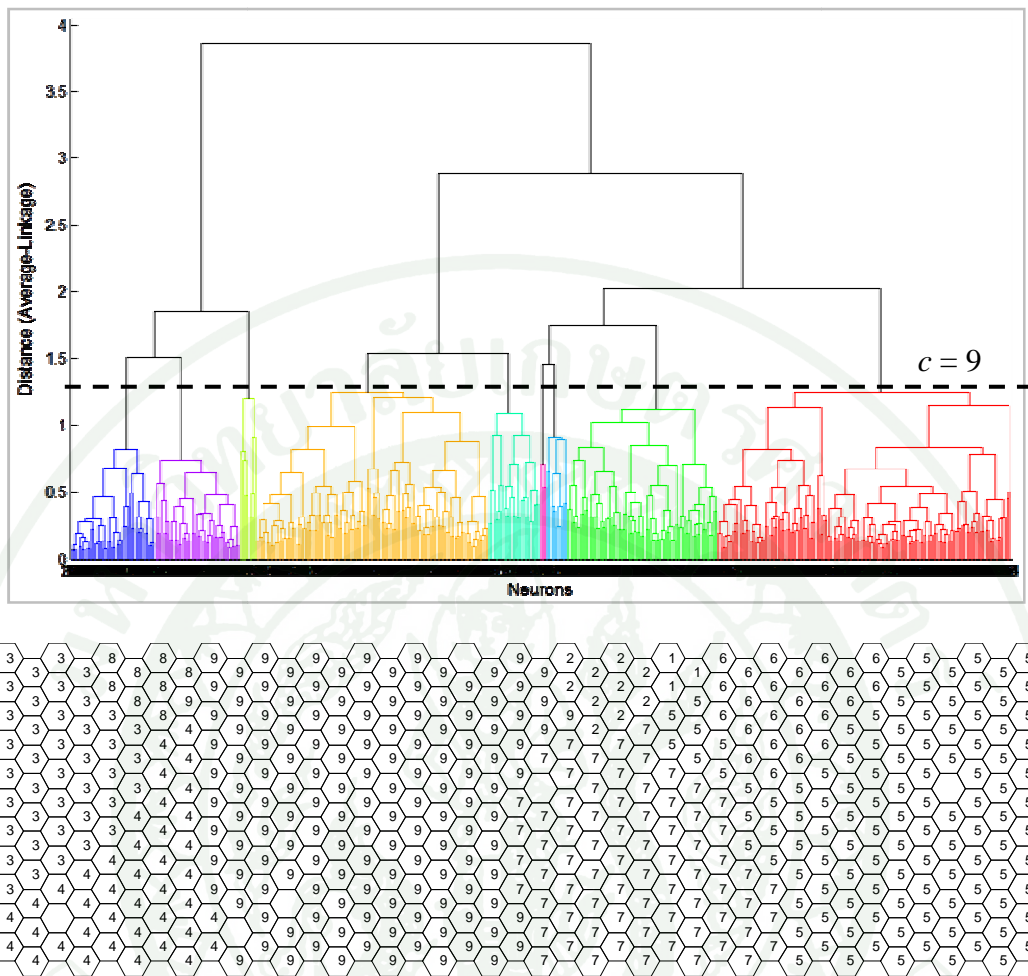
ตารางที่ 3 ค่าประสิทธิภาพของดัชนี WSJ และขนาดของข่ายงานที่มีประสิทธิภาพต่อการวิเคราะห์ด้วยจำนวนกลุ่มตามที่กำหนด ตั้งแต่ 2-10 กลุ่ม

	c = 2	c = 3	c = 4	c = 5	c = 6	c = 7	c = 8	c = 9	c = 10
WSJ	0.8199	0.9530	0.7773	0.9002	0.6646	0.7080	0.6721	0.5662	1.0003
msize	31×8	40×10	46×12	26×26	26×26	29×7	31×8	41×11	29×7

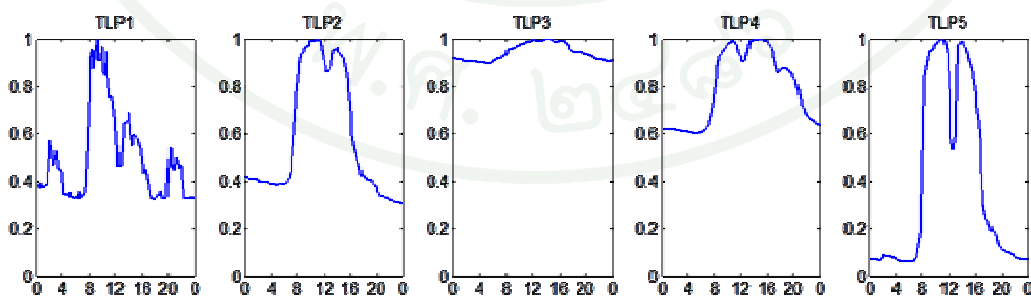
ตารางที่ 4 จำนวนกลุ่มที่เหมาะสม และค่าประสิทธิภาพการวิเคราะห์กลุ่มด้วยวิธีแบบเป็นลำดับชั้นของแต่ละข่ายงานประสาทที่จัดเตรียมไว้

ข่ายงานประสาท	c*	WSJ	ข่ายงานประสาท	c*	WSJ	ข่ายงานประสาท	c*	WSJ
29×7	7	0.7080	46×13	10	1.0017	20×20	9	0.7760
31×8	8	0.6721	50×13	10	1.0255	22×22	10	1.0352
33×9	10	1.0010	50×14	10	1.0283	24×24	10	1.0173
35×10	4	0.8773	54×14	10	1.0004	26×26	6	0.6646
40×10	3	0.9530	53×15	9	0.8831	28×28	10	1.0201
41×11	9	0.5662	14×14	7	0.8487	57×15	5	0.9698
42×12	9	0.8515	16×16	10	1.0062	56×16	10	1.0329
46×12	4	0.7773	18×18	10	1.0004	30×30	9	1.0219

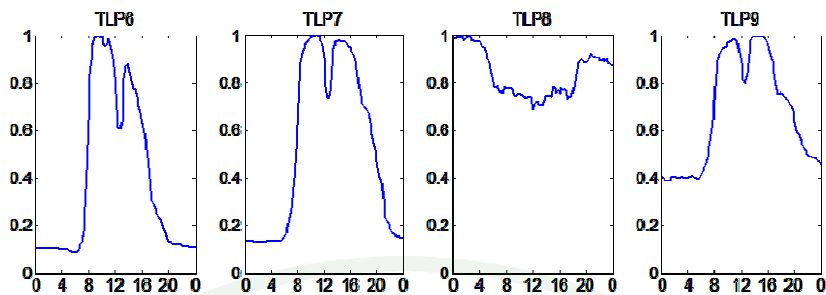
พิจารณาผลการวิเคราะห์โดยทำการเปรียบเทียบค่าประสิทธิภาพจากตารางที่ 3 และ 4 ตามลำดับ พบว่า เมื่อทำการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าตามขั้นตอนของวิธีแบบเป็นลำดับชั้นจำนวน 9 กลุ่ม กับข่ายงานประสาทที่มีขนาดเท่ากับ 41×11 เซลล์ประสาท จะได้ค่าประสิทธิภาพตามวิธีคำนวณของดัชนี WSJ มีค่าเท่ากับ 0.5662 ซึ่งต่ำสุดเมื่อเทียบกับข่ายงานประสาทที่เหลือ โดยมีผลการวิเคราะห์ในรูปแบบแผนภาพต้นไม้ลำดับชั้นบนข่ายงานประสาทขนาด 41×11 ตามภาพที่ 25 และมีลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนของแต่ละกลุ่มแสดงไว้ในภาพที่ 26



ภาพที่ 25 ผลการวิเคราะห์กลุ่มในรูปแบบแผนภาพต้นไม้ลำดับชั้นจำนวน 9 กลุ่ม บนข่ายงานประสาทที่มีขนาด 41×11 เซลล์ประสาท



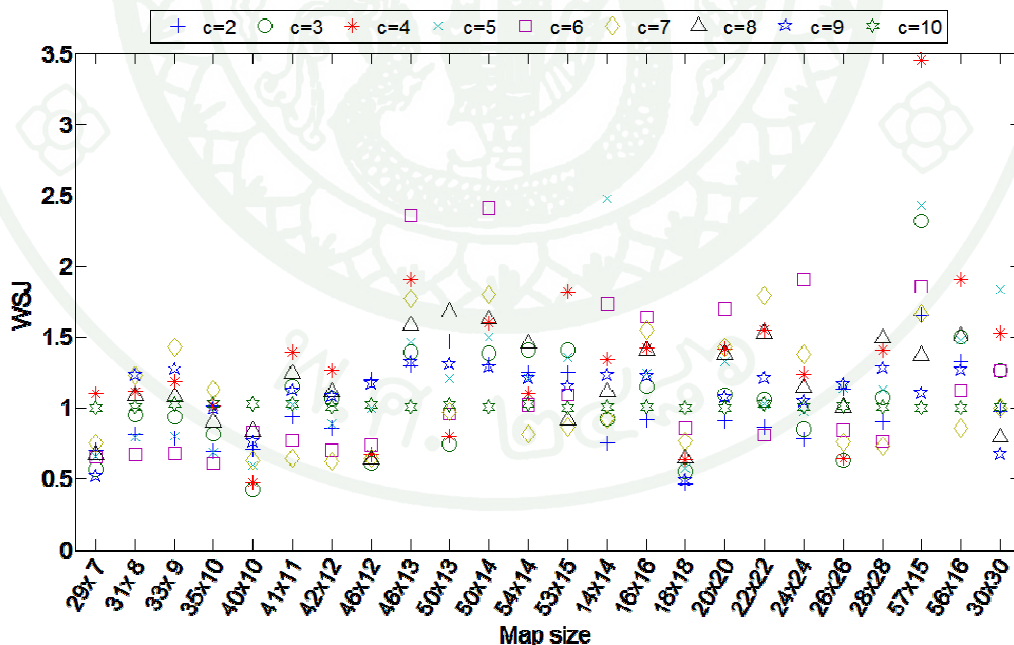
ภาพที่ 26 ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 9 กลุ่ม เมื่อทำการวิเคราะห์ตามขั้นตอนของวิธีแบบเป็นลำดับชั้นร่วมกับข่ายงานประสาทขนาด 41×11 เซลล์ประสาท



ภาพที่ 26 (ต่อ)

ผลการวิเคราะห์กลุ่มตามขั้นตอนของวิธีขยายพื้นที่ด้วยการเชื่อมโยงเซลล์ประสาทเข้าด้วยกัน

ดำเนินการวิเคราะห์กลุ่มของลักษณะการใช้ไฟฟ้าด้วยข่ายงานประสาทที่เตรียมไว้จากตารางที่ 1 ตามขั้นตอนของวิธีขยายพื้นที่แบบพื้นฐานอาศัยค่าระยะระหว่างเซลล์ประสาทที่เชื่อมโยงกัน โดยทำการคัดเลือกตำแหน่งเริ่มต้นจากระยะการเชื่อมโยงต่ำสุดเฉพาะที่ ผลการวัดประสิทธิภาพจากการคำนวณด้วยดัชนี WSJ ในแต่ละขนาดข่ายงานประสาท แสดงไว้ในภาพที่ 27



ภาพที่ 27 เปรียบผลการคำนวณค่าประสิทธิภาพด้วยดัชนี WSJ ในช่วง 2-10 กลุ่ม ด้วยเทคนิคการขยายพื้นที่บนข่ายงานประสาทที่มีขนาดต่างกัน

จากผลการวิเคราะห์กลุ่มตามขั้นตอนของวิธีขยายพื้นที่อาศัยการเชื่อมโยงระหว่างเซลล์บน ข่ายงานประสาทแบบพื้นฐาน โดยทำการคัดเลือกตำแหน่งเริ่มต้นจากระยะการเชื่อมโยงต่ำสุด เฉพาะที่ เมื่อวิเคราะห์กลุ่มร่วมกับข่ายงานประสาทที่มีขนาดต่างกันดังแสดงในภาพที่ 27 สามารถ แยกพิจารณาได้เป็นตารางที่ 5 และ 6 ตามลำดับ

ตารางที่ 5 ค่าประสิทธิภาพของดัชนี WSJ และขนาดข่ายงานประสาทที่สอดคล้องกับผลลัพธ์การ วิเคราะห์ที่มีประสิทธิภาพตามจำนวนกลุ่มที่กำหนด ในช่วง 2-10 กลุ่ม

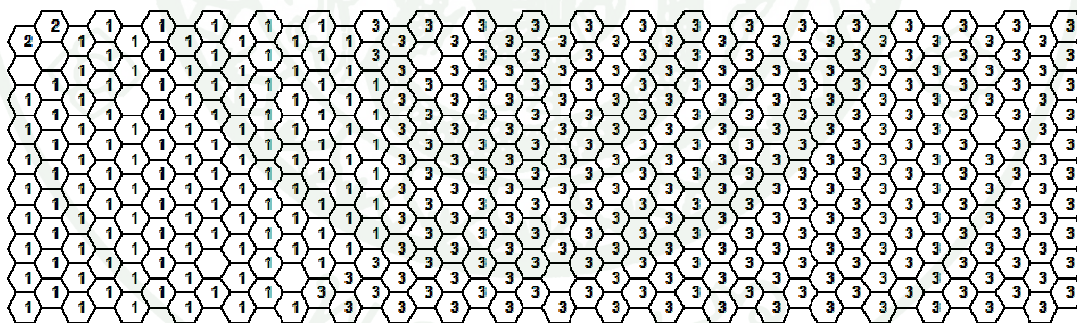
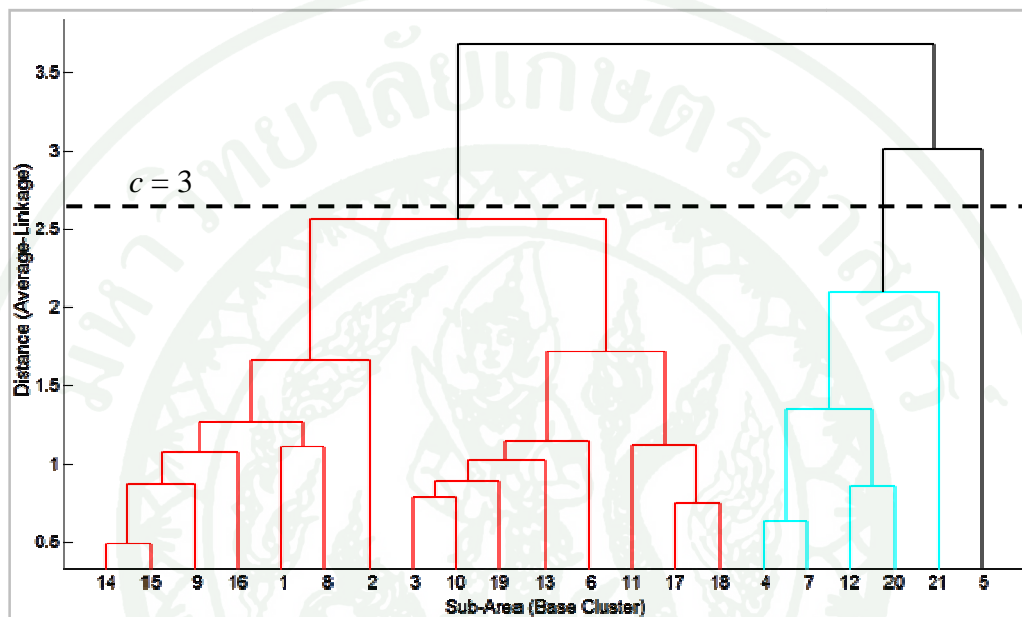
	c = 2	c = 3	c = 4	c = 5	c = 6	c = 7	c = 8	c = 9	c = 10
WSJ	0.4638	0.4240	0.4758	0.5713	0.6105	0.6245	0.6389	0.4872	1.0002
msize	18×18	40×10	40×10	18×18	35×10	42×12	46×12	18×18	18×18

ตารางที่ 6 จำนวนกลุ่มที่เหมาะสมและค่าประสิทธิภาพ เมื่อทำการวิเคราะห์กลุ่มด้วยวิธีขยายพื้นที่ บนพื้นฐานข่ายงานประสาทที่มีขนาดต่างกันตามที่จัดเตรียมไว้

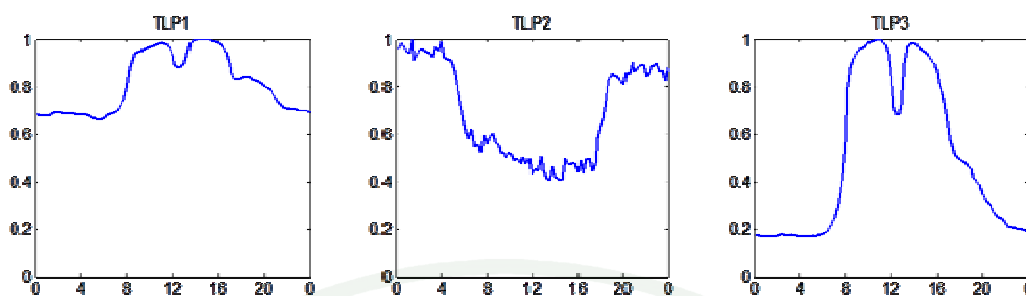
ข่ายงาน ประสาท	c*	WSJ	ข่ายงาน ประสาท	c*	WSJ	ข่ายงาน ประสาท	c*	WSJ
29×7	9	0.5191	46×13	10	1.0061	20×20	2	0.9114
31×8	6	0.6733	50×13	3	0.7455	22×22	6	0.8108
33×9	6	0.6772	50×14	10	1.0068	24×24	2	0.7773
35×10	6	0.6105	54×14	7	0.8184	26×26	3	0.6330
40×10	3	0.4240	53×15	7	0.8653	28×28	7	0.7276
41×11	7	0.6451	14×14	2	0.7522	57×15	10	1.0022
42×12	7	0.6245	16×16	2	0.9156	56×16	7	0.8608
46×12	3	0.6073	18×18	2	0.4638	30×30	9	0.6758

เปรียบเทียบประสิทธิภาพผลการวิเคราะห์ตามที่แสดงไว้ในตารางที่ 5 และ 6 ตามลำดับ พบว่า เมื่อทำการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าตามขั้นตอนของวิธีขยายพื้นที่อาศัย ความเชื่อมโยงระหว่างเซลล์บนข่ายงานประสาท โดยทำการสร้างพื้นที่ย่อยด้วยเงื่อนไขระยะเวลา เชื่อมโยงต่ำสุดเฉพาะที่ และทำการรวมกลุ่มของพื้นที่ย่อยจำนวน 3 กลุ่มด้วยขั้นตอนของวิธีแบบ

เป็นลำดับขั้น โดยใช้ข่ายงานประสาทที่มีขนาดเท่ากับ 40×10 เซลล์ประสาท จะได้ค่าประสิทธิภาพตามวิธีคำนวณของดัชนี WSJ มีค่าเท่ากับ 0.4240 ซึ่งต่ำสุดเมื่อเทียบกับข่ายงานประสาทที่เหลือ โดยมีการวิเคราะห์ในรูปแบบแผนภาพต้นไม้ลำดับขั้นบนข่ายงานประสาทขนาด 41×10 ตามภาพที่ 28 และมีลักษณะการใช้ไฟฟารูปแบบตัวแทนของแต่ละกลุ่มแสดงไว้ในภาพที่ 29



ภาพที่ 28 ผลการวิเคราะห์กลุ่มในรูปของแผนภาพต้นไม้ลำดับขั้นบนข่ายงานประสาทขนาด 40×10 เซลล์ประสาท จำนวน 3 กลุ่ม



ภาพที่ 29 ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 3 กลุ่ม จากการวิเคราะห์ด้วยเทคนิคการขยายพื้นที่ร่วมกับข่ายงานประสาทที่มีขนาด 40×10 เซลล์ประสาท

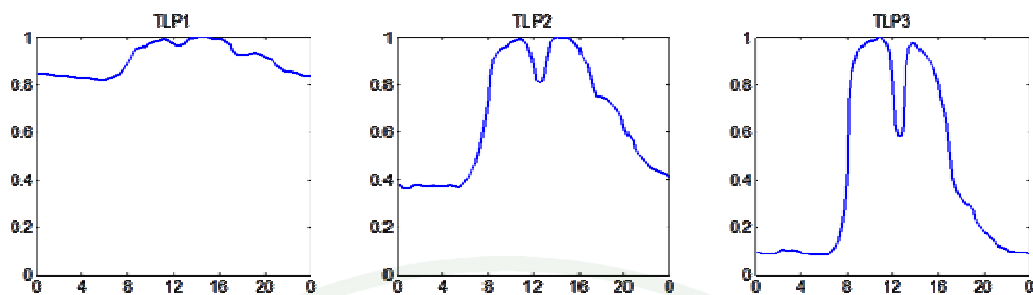
ผลการวิเคราะห์กลุ่มตามขั้นตอนของวิธีเคมิน

ดำเนินการวิเคราะห์กลุ่มข้อมูลลักษณะการใช้ไฟฟ้าตามขั้นตอนของวิธีเคมิน ผลจากการวิเคราะห์โดยใช้การคำนวณด้วยดัชนี WSJ แสดงไว้ในตารางที่ 7

ตารางที่ 7 ค่าประสิทธิภาพของดัชนี WSJ เมื่อทำการวิเคราะห์กลุ่มด้วยวิธีเคมินในช่วง 2-10 กลุ่ม

	c=2	c=3	c=4	c=5	c=6	c=7	c=8	c=9	c=10
WSJ	1.1091	1.0716	1.4373	2.1935	1.4612	1.6977	1.3566	1.2027	1.0826

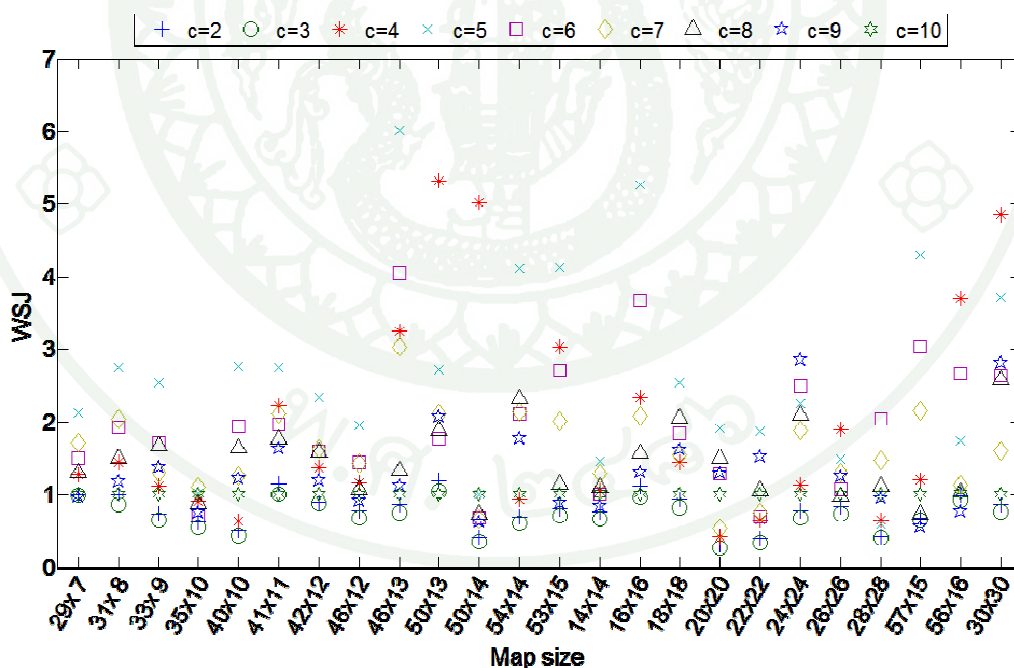
จากผลการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าในช่วง 2-10 กลุ่ม ตามตารางที่ 7 เมื่อเปรียบเทียบค่าประสิทธิภาพ พบว่า จำนวนกลุ่มที่มีค่าประสิทธิภาพซึ่งคำนวณด้วยดัชนี WSJ ต่ำสุดคือ 3 กลุ่ม ดังนั้นสรุปได้ว่า เมื่อทำการวิเคราะห์ตามขั้นตอนของวิธีเคมิน สามารถวิเคราะห์กลุ่มของลักษณะการใช้ไฟฟ้าได้จำนวน 3 กลุ่ม โดยมีลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนของแต่ละกลุ่มแสดงไว้ในภาพที่ 30



ภาพที่ 30 ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 3 กลุ่มจากการวิเคราะห์ด้วยวิธีเคมีน

ผลการวิเคราะห์กลุ่มด้วยวิธีแบบผสมระหว่างเคมีน และเทคนิคการขยายของพื้นที่

ผลการวิเคราะห์กลุ่มของลักษณะการใช้ไฟฟ้าในช่วง 2-10 กลุ่ม ด้วยขั้นตอนของวิธีแบบผสมระหว่างเคมีน และเทคนิคการขยายของพื้นที่ ในเทอมของค่าประสิทธิผลจากการคำนวณด้วยดัชนี WSJ แสดงไว้ในภาพที่ 31



ภาพที่ 31 ค่าประสิทธิภาพของดัชนี WSJ จากการวิเคราะห์กลุ่มด้วยวิธีการทำงานร่วมกันของเคมีน และเทคนิคการขยายพื้นที่กลุ่ม บนข่ายงานประสาทที่มีขนาดต่างกัน

จากการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าตามขั้นตอนของวิธีขยายพื้นที่จากการเชื่อมโยงเซลล์บนข่ายงานประสาท โดยคัดเลือกชุดของเซลล์ประสาทตำแหน่งเริ่มต้นอาศัยผลการวิเคราะห์จากวิธีเคมิน เมื่อวิเคราะห์กลุ่มร่วมกับข่ายงานประสาทที่มีขนาดต่างกันตามภาพที่ 31 สามารถแยกพิจารณาได้เป็นตารางที่ 8 และ 9 ตามลำดับ

ตารางที่ 8 ค่าประสิทธิภาพของดัชนี WSJ และขนาดของข่ายงานที่สอดคล้องกับผลการวิเคราะห์ ที่มีประสิทธิภาพตามจำนวนกลุ่มที่กำหนด ในช่วง 2-10 กลุ่ม

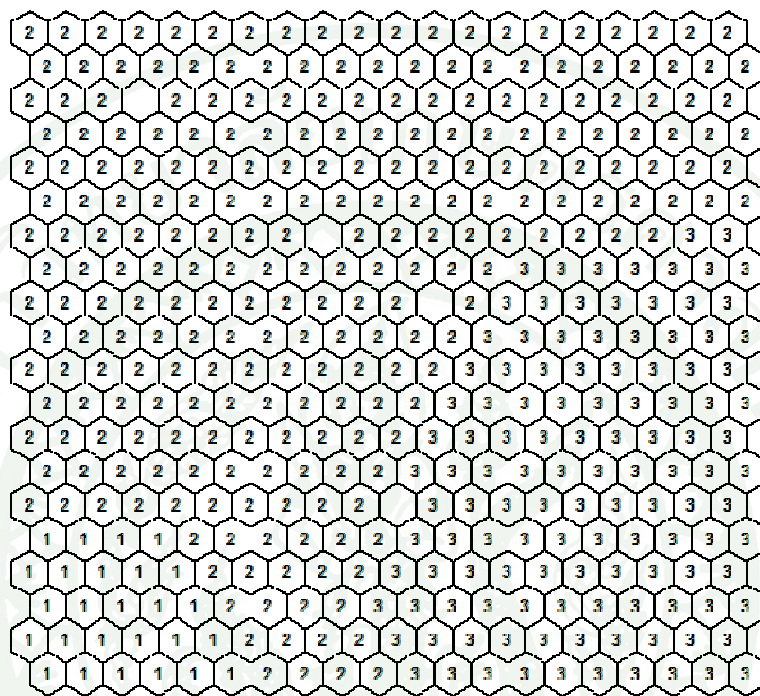
	c = 2	c = 3	c = 4	c = 5	c = 6	c = 7	c = 8	c = 9	c = 10
WSJ	0.3046	0.2559	0.4244	0.5836	0.6688	0.5269	0.7251	0.5512	1.0016
msize	20x20	20x20	20x20	28x28	50x14	20x20	50x14	50x14	33x 9

ตารางที่ 9 จำนวนกลุ่มที่เหมาะสม และค่าประสิทธิภาพ เมื่อทำการวิเคราะห์กลุ่มด้วยเคมินร่วมกับเทคนิคขยายพื้นที่อาศัยข่ายงานประสาทตามที่จัดเตรียมไว้

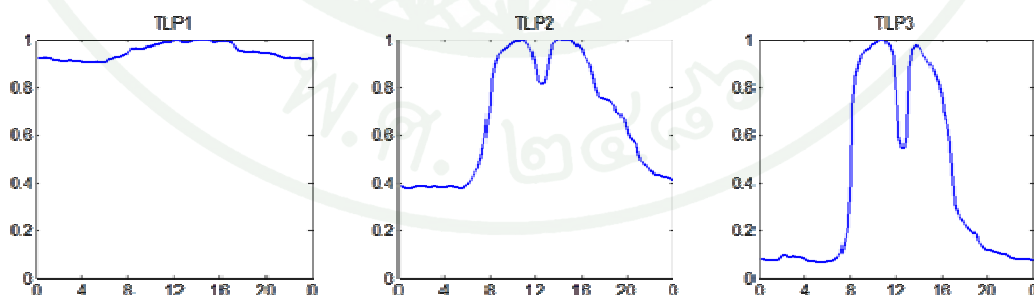
ข่ายงาน ประสาท	c*	WSJ	ข่ายงาน ประสาท	c*	WSJ	ข่ายงาน ประสาท	c*	WSJ
29×7	9	0.9547	46×13	3	0.7114	20×20	3	0.2559
31×8	3	0.8635	50×13	10	1.0034	22×22	3	0.3380
33×9	3	0.6468	50×14	3	0.3440	24×24	3	0.6758
35×10	3	0.5409	54×14	3	0.6004	26×26	3	0.7343
40×10	3	0.4383	53×15	3	0.6291	28×28	3	0.4102
41×11	3	0.9999	14×14	3	0.6671	57×15	3	0.7890
42×12	3	0.8779	16×16	3	0.9524	56×16	9	0.7595
46×12	3	0.6698	18×18	3	0.8133	30×30	3	0.7554

พิจารณาผลการคำนวณค่าประสิทธิภาพตามที่แสดงไว้ในตารางที่ 8 และ 9 ตามลำดับ พบว่า เมื่อทำการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าตามขั้นตอนของวิธีขยายพื้นที่จากการเชื่อมโยงเซลล์บนข่ายงานประสาท โดยคัดเลือกชุดของเซลล์ประสาทตำแหน่งเริ่มต้นอาศัยผลการวิเคราะห์จากวิธีเคมิน โดยใช้ข่ายงานประสาทที่มีขนาดเท่ากับ 20×20 เซลล์ประสาท จะได้ค่า

ประสิทธิภาพตามวิธีคำนวณของดัชนี WSI มีค่าเท่ากับ 0.2559 ซึ่งต่ำสุดเมื่อเทียบกับข่ายงานประสาทที่เหลือ สำหรับผลการวิเคราะห์ในรูปของข่ายงานประสาทขนาด 20×20 เซลล์ประสาทแสดงไว้ในภาพที่ 32 และมีลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนของแต่ละกลุ่มตามภาพที่ 33



ภาพที่ 32 ผลการวิเคราะห์กลุ่มในรูปข่ายงานประสาทจำนวน 3 กลุ่ม จากการทำงานร่วมกันของเคมีน และเทคนิคการขยายพื้นที่โดยใช้ข่ายงานประสาทขนาด 20×20 เซลล์ประสาท



ภาพที่ 33 ลักษณะการใช้ไฟฟ้ารูปแบบตัวแทนจำนวน 3 กลุ่ม จากการทำงานร่วมกันของเคมีนกับเทคนิคขยายพื้นที่บนข่ายงานประสาทขนาด 20×20 เซลล์ประสาท

สรุปผลการจัดกลุ่มลักษณะการใช้ไฟฟ้า

สรุปจำนวนกลุ่มที่เหมาะสม และผลการคำนวณค่าประสิทธิภาพด้วยวิธีของดัชนี WSJ หลังจากการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าในแต่ละวิธีการตามที่น่าเสนอภายในวิทยานิพนธ์นี้ แสดงไว้ในตารางที่ 10 และมีสัดส่วนของสมาชิกหลังจากการวิเคราะห์กลุ่มคิดเป็นเปอร์เซ็นต์แสดงไว้ในตารางที่ 11

ตารางที่ 10 สรุปผลการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้า

ลำดับ ที่	วิธีการ	จำนวนกลุ่ม ที่เหมาะสม	ค่าดัชนีWSJ
1	Agglomerative Hierarchical Clustering (Data Space)	10	1.0386
2	Agglomerative Hierarchical Clustering (Map Space)	9	0.5662
3	Region-Growing based on the Self-Organizing Map	3	0.4240
4	K-Means	3	1.0716
5	K-means combined to Region-Growing based on the Self-Organizing Map	3	0.2559

ตารางที่ 11 สรุปสัดส่วนของสมาชิกคิดเป็นเปอร์เซ็นต์จากการวิเคราะห์กลุ่มด้วยวิธีที่ต่างกัน

	จำนวนกลุ่ม	สัดส่วนสมาชิก
Agglomerative Hierarchical Clustering (Data Space)	10	กลุ่มที่ 1: 0.05%
		กลุ่มที่ 2: 44.74%
		กลุ่มที่ 3: 41.79%
		กลุ่มที่ 4: 12.12%
		กลุ่มที่ 5: 0.37%
		กลุ่มที่ 6: 0.05%
		กลุ่มที่ 7: 0.11%
		กลุ่มที่ 8: 0.03%

ตารางที่ 11 (ต่อ)

	จำนวนกลุ่ม	สัดส่วนสมาชิก
Agglomerative Hierarchical Clustering (Data Space)	10	กลุ่มที่ 9: 0.71% กลุ่มที่ 10: 0.03%
Agglomerative Hierarchical Clustering (Map Space)	9	กลุ่มที่ 1: 0.66% กลุ่มที่ 2: 1.84% กลุ่มที่ 3: 11.99% กลุ่มที่ 4: 7.09% กลุ่มที่ 5: 26.35% กลุ่มที่ 6: 4.35% กลุ่มที่ 7: 14.99% กลุ่มที่ 8: 1.34% กลุ่มที่ 9: 31.38%
Region-Growing based on the Self-Organizing Map	3	กลุ่มที่ 1: 33.81% กลุ่มที่ 2: 0.50% กลุ่มที่ 3: 65.69%
K-Means	3	กลุ่มที่ 1: 18.79% กลุ่มที่ 2: 41.29% กลุ่มที่ 3: 39.92%
K-Means combined to Region-Growing based on the Self-Organizing Map	3	กลุ่มที่ 1: 11.38% กลุ่มที่ 2: 57.60% กลุ่มที่ 3: 31.01%

วิจารณ์

ผลการวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าด้วยขั้นตอนวิธีจัดกลุ่มที่ต่างกัน ถ้าเปรียบเทียบในเชิงประสิทธิภาพการจัดกลุ่มจากผลการคำนวณความเป็นปึกแผ่นภายในกลุ่ม และความต่างระหว่างกลุ่มด้วยดัชนี WSI พบว่า ผลการวิเคราะห์โดยใช้การนำเข้าสู่รูปแบบตัวแทนจากการวิเคราะห์ด้วยวิธีเคมีนมาช่วยในจัดเรียงกลุ่มใหม่อาศัยขั้นตอนของวิธีขยายพื้นที่ด้วยการเชื่อมโยงเซลล์ประสาทเข้าด้วยกัน มีค่าประสิทธิภาพการจัดกลุ่มที่ดีกว่าวิธีอื่น ๆ และเมื่อพิจารณาผลการคำนวณค่าประสิทธิภาพจากการวิเคราะห์โดยดำเนินการบนกับข่ายงานประสาท จะเห็นได้ว่า ถ้าข่ายประสาทมีขนาดที่เหมาะสม จะมีส่วนทำให้การวิเคราะห์กลุ่มมีประสิทธิภาพมากขึ้น

ผลที่ได้จากการทดลองพบว่าเป็นไปตามเป้าหมายที่ตั้งไว้ คือ การวิเคราะห์กลุ่มของลักษณะการตามขั้นตอนของวิธีการขยายพื้นที่โดยการเชื่อมโยงเซลล์ประสาทเข้าด้วยกัน จากตำแหน่งเริ่มต้นซึ่งใช้การคัดเลือกตามขั้นตอนของวิธีเคมีน เมื่อใช้ร่วมกับข่ายงานประสาทที่มีขนาดเหมาะสม จะช่วยให้ผลการจัดกลุ่มมีประสิทธิภาพดีขึ้นกว่าผลการจัดกลุ่มลักษณะการใช้ไฟฟ้าด้วยวิธีเคมีนขั้นตอนเดียว โดยสังเกตได้จากผลการคำนวณค่าประสิทธิภาพตามด้วยดัชนี WSI มีค่าลดลงตามที่แสดงไว้ในตารางที่ 9 ในขณะเดียวกันเมื่อนำไปเปรียบเทียบกับผลจากการวิเคราะห์กลุ่มด้วยวิธีการขยายพื้นที่แบบพื้นฐานที่นำเสนอก่อนหน้านี้ พบว่าผลการวิเคราะห์กลุ่มด้วยเงื่อนไขที่ปรับปรุงใหม่ตามที่เสนอให้ค่า WSI ต่ำกว่า เช่นเดียวกัน อีกทั้งยังมีจำนวนกลุ่มที่กระชับมากกว่า

ประโยชน์ที่ได้รับจากการวิจัย นอกจากจะได้แนวทางในการจัดกลุ่มวิธีใหม่แล้ว ผลจากการจัดกลุ่มสามารถนำไปใช้ประโยชน์ต่อยอด หรือช่วยอำนวยความสะดวกต่อแผนงานต่าง ๆ ทางไฟฟ้าได้ เช่น ทางผู้ดำเนินธุรกิจไฟฟ้า สามารถนำผลการจัดกลุ่มไปใช้ประกอบการจัดทำอัตราค่ากระแสไฟฟ้า การวางแผนการผลิตรวมทั้งการส่งจ่ายไฟฟ้า และนำผลการวิเคราะห์ทางสถิติไปช่วยในด้านการพยากรณ์ความต้องการไฟฟ้า ส่วนทางด้านผู้ใช้ไฟฟ้า สามารถนำผลการจัดกลุ่มลักษณะการใช้ไฟฟ้ามาวิเคราะห์หาแนวทางปรับปรุงการใช้ไฟฟ้า เพื่อควบคุมค่าใช้จ่าย และเป็นการส่งเสริมให้เกิดการใช้ไฟฟ้าที่มีประสิทธิภาพ

สรุปและข้อเสนอแนะ

สรุป

วิทยานิพนธ์นี้นำเสนอวิธีวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้า บนพื้นฐานการเรียนรู้ของข่ายงานประสาทแบบก่อดำด้วยตนเอง โดยจะนำคุณสมบัติด้านการเชื่อมโยงระหว่างเซลล์ในข่ายงานใช้ร่วมกับการวิเคราะห์กลุ่มด้วยเทคนิคขยายพื้นที่ ซึ่งจะเริ่มดำเนินการแผ่ขยายกลุ่มจากตำแหน่งเซลล์ที่ตรงกับตำแหน่งศูนย์กลางข้อมูลตามขั้นตอนการวิเคราะห์ของวิธีเคมิน ข้อดีคือมีการระบุตำแหน่งอ้างอิงสำหรับใช้ในขั้นตอนขยายพื้นที่ที่ได้ใกล้เคียงกับโครงสร้างข้อมูลจริง และมีการทำงานที่ซับซ้อนน้อยลงเมื่อเทียบกับเทคนิคขยายพื้นที่แบบดั้งเดิมเพราะไม่ต้องทำการรวมพื้นที่ย่อยซ้ำอีกครั้ง เมื่อทำการเปรียบเทียบประสิทธิภาพผลการจัดกลุ่ม โดยพิจารณาจากค่าการคำนวณของดัชนี WSJ พบว่า วิธีวิเคราะห์กลุ่มแบบผสมตามที่เสนอให้ผลการจัดกลุ่มซึ่งสะท้อนถึงความกระชับหรือครอบคลุมสมาชิกได้ดีขึ้นกว่าการวิเคราะห์ด้วยเคมินเพียงขั้นตอนเดียว แสดงให้เห็นว่า ข่ายงานประสาทมีส่วนช่วยให้ข้อมูลสามารถจัดเข้ากลุ่มได้อย่างถูกต้องมากขึ้น

ข้อเสนอแนะ

การวิเคราะห์กลุ่มข้อมูลของลักษณะการใช้ไฟฟ้าด้วยวิธีต่าง ๆ ตามที่ได้นำเสนอในวิทยานิพนธ์นี้ยังมีส่วนที่สามารถปรับปรุงให้ดีขึ้นได้อีก โดยเริ่มจากขั้นตอนการจัดการเบื้องต้นกับข้อมูล ควรจะมีการระบุวันหยุดที่ชัดเจน เพราะผู้ใช้ไฟฟ้าแต่ละรายจะมีวันหยุดที่ต่างกัน ถ้าหากสามารถเพิ่มเติมรายละเอียดในส่วนนี้ได้ ก็จะช่วยให้การจัดกลุ่มมีความถูกต้องมากยิ่งขึ้น เนื่องจากในช่วงวันหยุดจะมีปริมาณการใช้ไฟฟ้าน้อยเมื่อเทียบกับวันทำงาน หากไม่ได้มีการระบุวันหยุดจริงของผู้ใช้ไฟฟ้าแต่ละราย รูปแบบตัวแทนและผลการจัดกลุ่มอาจเกิดการคลาดเคลื่อนเพราะมีข้อมูลของวันหยุดปะปนอยู่ อย่างไรก็ตามการระบุวันหยุดที่แน่นอนของผู้ใช้ไฟฟ้านั้นต้องใช้เวลาพอสมควรในการขอข้อมูลเพิ่มเติมจากผู้ใช้ไฟฟ้า และสำหรับการจัดเตรียมข่ายงานประสาทสำหรับใช้ร่วมกับเทคนิคการขยายพื้นที่ พบว่าต้องให้เวลากับกระบวนการส่วนนี้พอสมควร แต่ในปัจจุบันเริ่มมีงานวิจัยในลักษณะของการสร้างข่ายงานประสาทในลักษณะของการแบ่งเซลล์แบบอัตโนมัติ เพื่อให้มีโครงสร้างที่ครอบคลุมกับข้อมูล ซึ่งถ้าหากมีการพัฒนาอย่างต่อเนื่องจนนำไปสู่ข้อสรุป จะช่วยให้วิธีจัดกลุ่มตามที่เสนอสามารถตอบสนองด้านการใช้งานได้ดียิ่งขึ้น

เอกสารและสิ่งอ้างอิง

การไฟฟ้านครหลวง. 2555. ร่างขอบเขตของงาน (Terms of Reference: TOR). แหล่งที่มา:

http://www.mea.or.th/upload/download/file_b6597a13b68cd0f274f676065d5701c7.pdf,
14 พฤษภาคม 2557.

ภาควิชาวิศวกรรมไฟฟ้า มหาวิทยาลัยเกษตรศาสตร์ และ ฝ่ายเศรษฐกิจพลังไฟฟ้า การไฟฟ้านครหลวง. 2555. แนวทางการศึกษาเบื้องต้นโครงการจัดการด้านการใช้ไฟฟ้าให้มีประสิทธิภาพกับผู้ใช้ไฟฟ้าที่มีตัวประกอบการใช้ไฟฟ้า (Load Factor) ต่ำ.

สำนักงานนโยบายและแผนพลังงาน กระทรวงพลังงาน. 2541. การศึกษาลักษณะการใช้ไฟฟ้า.

แหล่งที่มา: <http://www.eppo.go.th/vrs/VRS39-05-power.html>, 1 สิงหาคม 2556.

สำนักงานนโยบายและแผนพลังงาน กระทรวงพลังงาน. 2547. การพยากรณ์ความต้องการไฟฟ้า.

แหล่งที่มา: <http://www.eppo.go.th/load/load2004-jan04rev>, 1 สิงหาคม 2556.

Anuar, N. and Z. Zakaria. 2011. Determination of fuzziness parameter in load profiling via Fuzzy C-Means, pp. 139-142. **IEEE Control and System Graduate Research Colloquium (ICSGRC)**. 27-28 June 2011, Universiti Teknologi Mara, Shah Alam, Malaysia.

Balasko, B., J. Abonyi and B. Feil. 2005. **Fuzzy Clustering and Data Analysis Toolbox**.

Department of Process Engineering, University of Veszprem, Hungary.

Chang, R.F. and C.N. Lu. 2003. Load Profiling and Its Applications in Power Market, pp. 974-

978. *In IEEE Power Engineering Society General Meeting*. 13-17 July 2003,

Toronto, Canada.

- Chicco G., R. Napoli and F. Piglione. 2006. Application of Clustering and Self Organizing Maps to Classify Electricity Customers. **IEEE Transaction on Power System** 21 (4): 1672-1682.
- Gerbec, D., S. Gasperic and F. Gubina. 2003. Determination and Allocation of Typical Load Profile to the Eligible Consumers, pp. 1-5. *In Proceedings of the IEEE Bologna Power Tech Conference*. 23-26 June 2013, Faculty of Engineering, University of Bologna, Bologna, Italy
- Heenández, L., C. Baladrón, J. M. Aguiar, B. Carro, and A. Sánchez-Esguevillas. 2012. Classification and Clustering of Electricity Demand Patterns in Industrial Parks. **Energies** 5 (12): 5215-5228.
- Kim, Y. I., S. J. Kang, J. M. Ko and S. H. Choi. 2011. A Study for Clustering Method to Generate Typical Load Profiles for Smart Grid, pp. 1102-1109. *In 8th International Conference on Power Electronics (ECCE Asia)*. The Shilla Jeju, Korea.
- Sathiracheewin, S. and V. Surapatana. 2011. Daily Typical Load Clustering of Residential Customers, pp. 797-800. *In 8th Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology Association of Thailand Conference*. Khon Kaen University, Thailand.
- Sun H. J., S.G. Wang and Q.S. Jiang. 2004. FCM-Based Model Selection Algorithms for Determining the Number of Clusters. **Pattern Recognition** 37: 2027-2037.
- Tarek K. M. and B. Farouk. 2011. Kohonen Maps Combined to Fuzzy C-means, a Two Level Clustering Approach. Application to Electricity Load Data, pp. 541-558. *In* J. I. Mwasiagi., ed. **Self Organizing Maps - Applications and Novel Algorithm Design**.

- Tsekouras J. G. and A. D. Salis. 2008. Load Time-Series Classification Based on Pattern Recognition Method, pp. 361-432. In P. Y. Yin., ed. **Pattern Recognition Techniques, Technology and Application**. Austria.
- Vesanto, J. and E. Alhoniemi. 2000a. Clustering of the Self-Organizing Map. **IEEE Transactions on Neural Networks** 11 (3): 586-600.
- Vesanto, J., J. Himberg, E. Alhoniemi, and J. Parhankangas. 2000b. **SOM Toolbox for Matlab 5**. Helsinki University of Technology, Finland.
- Vesanto, J. and M. Sulkava. 2002. Distance matrix based clustering of the Self-Organizing Map, pp. 951-956. **12th International Conference on Artificial Neural Networks (ICANN 2002)**. 27-30 August 2002, Madrid, Spain.
- Yu, I. H., J. K. Lee, J. M. Ko and S. I. Kim. 2005. A Method for Classification of Electricity Demands using Load Profile Data, pp. 164-168. *In* **4th Annual ACIS International Conference on Computer and Information Science (ICIS'05)**.

ประวัติการศึกษา และการทำงาน

ชื่อ	นายวรายุทธ พราหมณ์ทอง
เกิดวันที่	1 ตุลาคม 2530
สถานที่เกิด	จังหวัดสุราษฎร์ธานี
ประวัติการศึกษา	วิศวกรรมศาสตรบัณฑิต (วิศวกรรมไฟฟ้า) มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่ จังหวัดสงขลา
ตำแหน่งปัจจุบัน	-
สถานที่ทำงานปัจจุบัน	-
ผลงานดีเด่นและ/หรือรางวัลทางวิชาการ	-
ทุนการศึกษาที่ได้รับ	-