

บทที่ 2

หลักการและทฤษฎีที่เกี่ยวข้อง

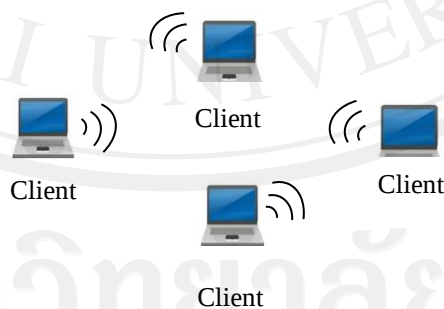
2.1 ระบบเครือข่ายท้องถิ่นแบบไร้สาย (Wireless Local Area Network)

ระบบเครือข่ายท้องถิ่นแบบไร้สาย (Wireless LAN) หมายถึงเทคโนโลยีที่ทำให้อุปกรณ์การสื่อสารตั้งแต่ 2 เครื่องขึ้นไปสามารถติดต่อสื่อสารกันได้เป็นระบบเครือข่ายโดยใช้คลื่นความถี่วิทยุในการรับและส่งข้อมูล

การทำงานของระบบเครือข่ายท้องถิ่นแบบไร้สายประกอบด้วยอุปกรณ์ในการส่งสัญญาณและกระจายสัญญาณเรียกว่า Access Point และอุปกรณ์คอมพิวเตอร์ลูกข่ายเรียกว่า Client ซึ่งติดตั้ง Wireless LAN card สำหรับการเชื่อมต่อกับอุปกรณ์กระจายสัญญาณโดยใช้คลื่นวิทยุ การเชื่อมต่อของระบบเครือข่ายท้องถิ่นแบบไร้สายมีรูปแบบใช้งานอยู่ 2 ลักษณะคือ แบบ Ad-hoc (Peer to Peer) และแบบ Infrastructure

2.1.1 รูปแบบ Ad-hoc (Peer to Peer)

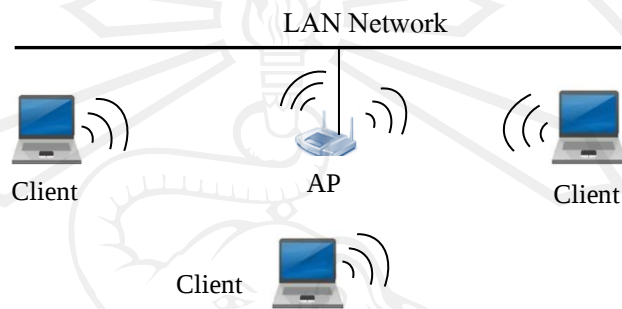
การเชื่อมต่อรูปแบบ Ad-Hoc หรือ IBSS (Independent Basic Service Set) เป็นการเชื่อมต่อระบบเครือข่ายท้องถิ่นแบบไร้สาย ที่ประกอบด้วยเครื่องคอมพิวเตอร์ตั้งแต่ 2 เครื่องขึ้นไป ที่ติดตั้งการ์ดแลนไร้สาย ที่ทำการสื่อสารกันโดยตรงไม่ต้องผ่านอุปกรณ์กระจายสัญญาณหรือ Access Point ดังแสดงรูปแบบในรูปที่ 2.1



รูปที่ 2.1 การเชื่อมต่อรูปแบบ Ad-hoc

2.1.2 รูปแบบ Infrastructure

การเชื่อมต่อรูปแบบ Infrastructure หรือ BSS (Basic Service Set) เป็นการเชื่อมต่อที่มีอุปกรณ์กระจายสัญญาณ (Access Point) เป็นตัวกลางทำหน้าที่รับ-ส่งข้อมูลจากเครื่องคอมพิวเตอร์บนเครือข่ายไร้สาย ไปสู่เครือข่ายมีสาย ซึ่งทำงานเหมือนอุปกรณ์สับในเครือข่ายคอมพิวเตอร์แบบมีสายดังแสดงรูปแบบในรูปที่ 2.2



รูปที่ 2.2 การเชื่อมต่อรูปแบบ Infrastructure

2.2 มาตรฐานเครือข่ายท้องถิ่นแบบไร้สาย IEEE 802.11

มาตรฐานเครือข่ายไร้สาย (Wireless LAN) IEEE 802.11 ได้รับการตีพิมพ์เผยแพร่ครั้งแรกเมื่อปี ค.ศ. 1997 [16] โดยสถาบัน IEEE (The Institute of Electronics and Electrical Engineers) ซึ่งมีข้อกำหนดระบุไว้ว่าผลิตภัณฑ์เครือข่ายไร้สายในส่วนของ Physical Layer นั้นมีความสามารถในการรับส่งข้อมูลที่มีความเร็ว 1, 2, 5.5, 11 และ 54 Mbps โดยใช้คลื่นวิทยุย่านความถี่ 2.4 GHz และ 5 GHz ส่วนในระดับชั้น MAC Layer นั้นได้กำหนดการทำงานแบบ Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) โดยมีกลไกในการเข้ารหัสข้อมูลก่อนแพร่กระจายสัญญาณไปบนอากาศ พร้อมกับมีการตรวจสอบผู้ใช้งาน โดยมีมาตรฐานย่อยที่สำคัญได้แก่ มาตรฐาน IEEE 802.11a, มาตรฐาน IEEE 802.11b, และ มาตรฐาน IEEE 802.11g

2.2.1 IEEE 802.11a

มาตรฐาน IEEE 802.11a ได้รับการตีพิมพ์และเผยแพร่เมื่อปี ค.ศ. 1999 [16] โดยใช้เทคโนโลยี Orthogonal Frequency Division Multiplexing (OFDM) เพื่อพัฒนาให้ผลิตภัณฑ์ไร้สายมีความสามารถในการรับส่งข้อมูลด้วยอัตราความเร็วสูงสุด 54 Mbps โดยใช้คลื่นวิทยุย่านความถี่ 5 GHz ซึ่งเป็นย่านความถี่ที่ไม่ได้รับอนุญาตให้ใช้งานโดยทั่วไปในประเทศไทย เนื่องจากสงวนไว้สำหรับกิจการทางด้านการทหาร

ข้อเสียของผลิตภัณฑ์มาตรฐาน IEEE 802.11a ก็คือมีรัศมีการใช้งานในระยะสั้น และมีราคาแพง ดังนั้นผลิตภัณฑ์ไร้สายมาตรฐาน IEEE 802.11a จึงได้รับความนิยมน้อย

2.2.2 IEEE 802.11b

มาตรฐาน IEEE 802.11b ได้รับการตีพิมพ์ และเผยแพร่ออกมาพร้อมกับมาตรฐาน IEEE 802.11a เมื่อปี ค.ศ. 1999 [16] โดยใช้เทคโนโลยี Complimentary Code Keying (CCK) ร่วมกับเทคโนโลยี Direct Sequence Spread Spectrum (DSSS) เพื่อให้สามารถรับส่งข้อมูลได้ด้วยอัตราความเร็วสูงสุดที่ 11 Mbps โดยใช้คลื่นสัญญาณวิทยุย่านความถี่ 2.4 GHz ซึ่งเป็นย่านความถี่ที่อนุญาตให้ใช้งานในแบบสาธารณะทางด้านวิทยาศาสตร์ อุตสาหกรรม และการแพทย์ ข้อดีของมาตรฐาน IEEE 802.11b ก็คือการสนับสนุนการใช้งานเป็นบริเวณกว้างกว่ามาตรฐาน IEEE 802.11a ผลิตภัณฑ์มาตรฐาน IEEE 802.11b เป็นที่รู้จักใน เครื่องหมายการค้า Wi-Fi ซึ่งกำหนดขึ้นโดย Wireless Ethernet Compatibility Alliance (WECA) โดยผลิตภัณฑ์ที่ได้รับเครื่องหมาย Wi-Fi ได้ผ่านการตรวจสอบ และรับรองว่าเป็นไปตามข้อกำหนดของมาตรฐาน IEEE 802.11b

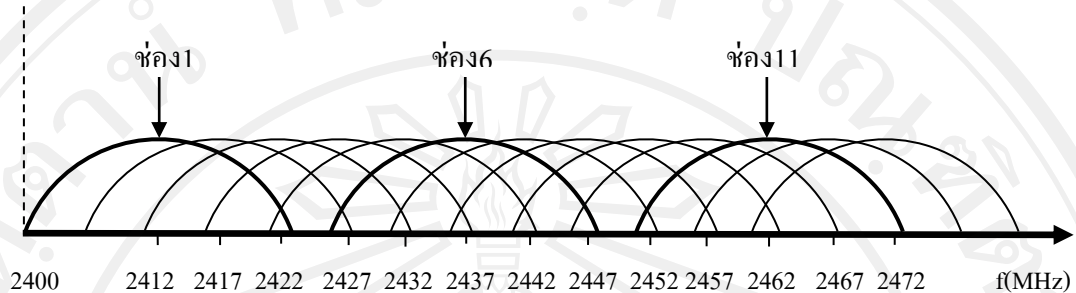
2.2.3 IEEE 802.11g

มาตรฐาน IEEE 802.11g ได้รับการตีพิมพ์ในปี ค.ศ 2003 [16] และได้เข้ามาทดแทนผลิตภัณฑ์ที่รองรับมาตรฐาน IEEE 802.11b เนื่องจากสนับสนุนอัตราความเร็วของการรับส่งข้อมูลในระดับ 54 Mbps โดยใช้เทคโนโลยี Orthogonal Frequency Division Multiplexing (OFDM) บนคลื่นสัญญาณวิทยุย่านความถี่ 2.4 GHz และให้รัศมีการทำงานที่มากกว่า IEEE 802.11a พร้อมความสามารถในการใช้งานร่วมกันกับมาตรฐาน IEEE 802.11b ได้ (Backward-Compatible)

2.3 ช่องความถี่ที่ใช้ในระบบเครือข่ายท้องถิ่นแบบ ไร้สายตามมาตรฐานของ IEEE 802.11 b/g

ช่องความถี่ตามมาตรฐานของ IEEE 802.11 b/g (Wi-Fi) ประกอบด้วยช่องสัญญาณทั้งหมด 14 ช่องสัญญาณคือช่องความถี่ที่ 1 ถึง ช่องความถี่ที่ 14 แต่ละช่องสัญญาณมี bandwidth ขนาด 22 MHz และความถี่กลางห่างกัน 5 MHz [17] โดยเริ่มจากความถี่ 2400 MHz ถึง 2495 MHz โดยค่าความถี่กลางของช่องสัญญาณที่ n สามารถคำนวณได้ตามสมการที่ (2.1) และแสดงแผนที่สเปกตรัมของช่องสัญญาณดังรูปที่ 2.3

$$f_{cn} = \begin{cases} 2412 + (n-1) \times 5, & 1 \leq n \leq 13 \\ 2484, & n = 14 \end{cases} \text{ (MHz)} \quad (2.1)$$



รูปที่ 2.3 แผนทีสเปกตรัมของช่องสัญญาณ IEEE 802.11b/g [17]

ตามมาตรฐานของ IEEE 802.11 b/g มีการแบ่งช่วงของช่องสัญญาณให้แต่ละช่องสามารถซ้อนทับกันได้ ซึ่งสำหรับการใช้งานจริงแล้วช่องสัญญาณที่ 16 และ 11 เป็นช่อง ที่นิยมใช้งาน เพราะไม่มีการซ้อนทับกันระหว่างช่องสัญญาณเหล่านี้

2.4 Received Signal Strength (RSS)

RSS เป็นค่าที่ใช้บอกถึงระดับความแรงหรือกำลังของสัญญาณที่ได้รับ มีหน่วยเป็น dBm (decibels milliwatt) โดยค่า RSS นี้จะแปรผันตรงกับความแรงของสัญญาณ นั่นคือ ถ้าค่า RSS มีค่ามากแสดงว่าสัญญาณที่ได้รับมีความแรงสูง นั่นคือตัวส่งและตัวรับอยู่ใกล้กันและในทางกลับกัน หากค่า RSS มีค่าน้อยแสดงว่าสัญญาณที่ได้รับมีความแรงต่ำเพราะตัวส่งและตัวรับอยู่ไกลกัน นอกจากนี้ค่า RSS จะมีความแรงแปรไปตามระยะห่างของตัวส่งและตัวรับแล้ว สภาพแวดล้อมและสิ่งรอบข้างก็เป็นปัจจัยหนึ่งที่สามารถส่งผลกระทบต่อความแรงของสัญญาณได้เช่นว่า ช่องทางการกระจายสัญญาณระหว่างตัวส่งและตัวรับถูกบดบังหรือมีการรบกวนจากสิ่งกีดขวางต่างๆ ซึ่งสามารถหาค่า RSS หรือ P_r [16] ได้จากสมการที่ (2.2) ดังนี้

$$P_r(\text{dBm}) = P_t(\text{dBm}) - L_T(\text{dB}) \quad (2.2)$$

โดยที่ P_r คือค่าความแรงสัญญาณที่รับได้ที่ตัวรับ (dBm)

P_t คือค่าความแรงสัญญาณที่ตัวส่ง (dBm)

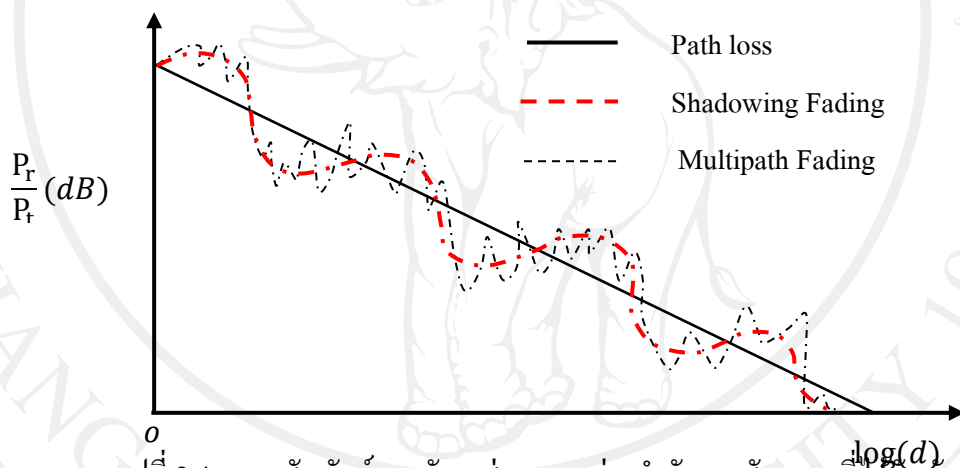
L_T คือกำลังสูญเสียทั้งหมด (dB)

2.5 เฟดดิ้งของช่องสัญญาณไร้สาย (Wireless Channel Fading)

ในระบบการสื่อสารไร้สายเมื่อสัญญาณถูกส่งผ่านช่องสัญญาณ คลื่นสัญญาณต้องเดินทางผ่านอุปสรรคและสิ่งกีดขวาง ส่งผลให้ค่ากำลังของสัญญาณที่ได้รับมีระดับเปลี่ยนแปลงไปตามเวลาเรียกปรากฏการณ์นี้ว่า การเกิดเฟดดิ้งของช่องสัญญาณ โดยสามารถแบ่งประเภทเฟดดิ้งของช่องสัญญาณไร้สายออกเป็น 2 ลักษณะคือ Large-Scale Fading และ Small-Scale Fading [18] รูปที่ 2.4 แสดงความสัมพันธ์ของอัตราส่วนระหว่างกำลังของสัญญาณที่ได้รับกับกำลังของสัญญาณที่ส่ง ต่อระยะทาง (d) ใน log สเกลเมื่อมีผลกระทบจากเฟดดิ้งลักษณะต่างๆ

2.5.1 Large-Scale Fading

Large-Scale Fading เกิดขึ้นเนื่องจาก 2 สาเหตุคือ การสูญเสียตามระยะทาง (Path Loss) และการบดบัง (Shadowing)



รูปที่ 2.4 ความสัมพันธ์ของอัตราส่วน ระหว่างกำลังของสัญญาณที่ได้รับกับกำลังของสัญญาณที่ส่ง ต่อระยะทาง (d) ใน log สเกล [18]

2.5.1.1 Path Loss

การสูญเสียตามระยะทาง (Path Loss) เป็นการลดทอนของสัญญาณที่แปรผันตามระยะทางระหว่างเครื่องส่งและเครื่องรับสัญญาณ โดยทั่วไปค่าเฉลี่ยกำลังของสัญญาณจะลดลงตามค่า $d^{-\alpha}$ หรือ กฎของการสูญเสียกำลังตามระยะทางโดยที่ d คือ ระยะทางระหว่างเครื่องส่ง และเครื่องรับ α คือตัวชี้กำลังการสูญเสียโดยมีค่าคงที่ที่แตกต่างกันไปตามสภาพ แวกดล้อมของบริเวณพื้นที่การรับส่งสัญญาณ เมื่อสภาพแวดล้อมเป็นอากาศว่าง (Free Space) α จะมีค่าเท่ากับ 2 [16] โดยมีกฎการสูญเสียกำลังตามระยะทางเป็นไปดังสมการที่ (2.3)

$$P_r = \frac{P_0}{d^\alpha} \quad (2.3)$$

P_r คือกำลังของสัญญาณที่ได้รับ (Watt)

P_0 คือกำลังของสัญญาณที่ได้รับในระยะ 1 m จากเครื่องส่ง (Watt)

d คือระยะทางระหว่างเครื่องส่งและเครื่องรับ (m)

α คือตัวชี้กำลังการสูญเสีย (Path loss exponent)

สมการที่ (2.4) แสดงความสัมพันธ์ระหว่างระยะทางและกำลังของสัญญาณในหน่วยเดซิเบล (Decibels; dB)

$$10\log_{10}P_r = 10\log_{10}P_0 - 10\alpha\log_{10}d \quad (2.4)$$

โดย $10\log_{10}P_r$ คือกำลังของสัญญาณที่ได้รับ (dB)

$10\log_{10}P_0$ คือกำลังของสัญญาณที่ได้รับในระยะ 1 m (dB)

ถ้ากำหนดให้ L_0 หมายถึง Path loss ในหน่วย dB ที่ระยะ 1m แสดงด้วยสมการ (2.5) ดังนี้

$$L_0 = 10\log_{10}P_t - 10\log_{10}P_0 \quad (2.5)$$

Path loss (L_P) ที่เกิดขึ้นระหว่างตัวส่งและตัวรับตามระยะทาง d แสดงด้วยสมการ (2.6) ดังนี้

$$L_P = L_0 + 10\alpha\log_{10}d \quad (2.6)$$

ตารางที่ 2.1 แสดงตัวอย่างของค่า L_0 และ α ของ Path Loss ในสถานการณ์ Line-of-sight (LOS) และ Non-line-of-sight (NLOS) ที่วัดได้จากการทดลองจริงในสภาพแวดล้อมต่างๆ

ตารางที่ 2.1 ตัวอย่าง Path loss ที่วัดในระบบย่านความถี่ 2.4 GHz และ 5 GHz [16]

ความถี่ f_c	สภาพแวดล้อม	สถานการณ์	L_0	ตัวชี้กำลังการสูญเสีย (α)
2.4 GHz	Indoor office	LOS	41.5 dB	2
		NLOS	37.7 dB	3.3
5.1 GHz	Meeting room	LOS	46.6 dB	2.22
		NLOS	61.6 dB	2.22

2.5.1.2 Shadowing หรือ Shadow Fading

การถูกบดบัง (Shadowing) เกิดจากการที่สัญญาณถูกบดบังไปชั่วขณะจากสิ่งกีดขวางต่างๆ ส่งผลให้กำลังเฉลี่ยของสัญญาณที่รับได้มีการเปลี่ยนแปลงที่ขึ้นกับขนาดและชนิดของสิ่งกีดขวาง ซึ่งการเปลี่ยนแปลงของสัญญาณในลักษณะนี้สามารถจำลองได้ด้วยการใช้ตัวแปรสุ่มที่มีการกระจายแบบล็อกนอร์มอล (Log-normal Distribution) ซึ่งอธิบายถึงปรากฏการณ์บดบัง (Shadowing) แบบสุ่มที่เกิดขึ้น ณ ตำแหน่งต่างๆ ตามสมการที่ (2.7) [16]

$$p_{LN}(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right), \quad x > 0 \quad (2.7)$$

โดย $p_{LN}(x)$ คือค่าความน่าจะเป็นของสัญญาณที่มีค่าเท่ากับ x
 x คือตัวแปรสุ่มแบบเกาส์ (Gaussian random variables)
 σ คือค่าเบี่ยงเบนมาตรฐาน (Standard deviation)
 σ^2 คือค่าความแปรปรวน (Variance)
 μ คือค่าเฉลี่ย (Mean)

สำหรับกำลังของสัญญาณที่รับได้ P_r จะเปลี่ยนแปลงไปตาม Path Loss L_P และผลกระทบจาก Shadow Fading X_σ ดังในสมการที่ (2.8)

$$P_r = P_t - L_P - X_\sigma \quad (2.8)$$

สมการที่ (2.6) สามารถใช้ในกรณีที่พิจารณาพื้นที่ภายใน (floor) เดียวกัน หากอาคารมีหลายชั้นการหาค่า Path Loss หรือ Indoor Path Loss สามารถพิจารณาโดยใช้รูปแบบการสูญเสียกำลังสัญญาณภายในอาคารดังสมการที่ (2.9)

$$L_T = L_P + L_f(n) + X_\sigma \quad (2.9)$$

โดย L_T คือการสูญเสียกำลังสัญญาณโดยรวม (Total Path Loss)
 L_P คือ Path Loss จากสมการที่ (2.6)
 $L_f(n)$ คือการลดทอนของสัญญาณที่ผ่าน n ชั้นของอาคาร
 X_σ คือผลกระทบของ Shadow Fading

การลดทอนของสัญญาณจะมากหรือน้อยขึ้นอยู่กับชนิดของพาร์ทิชันที่สัญญาณเดินทางผ่านดังตัวอย่างในตารางที่ 2.2

ตารางที่ 2.2 การลดทอนของสัญญาณที่ความถี่ 2.4 GHz เมื่อเดินทางผ่านพาร์ทิชันต่างๆ [16]

ประเภทพาร์ทิชัน	การลดทอน (dB)
Window in brick wall	2
Metal frame, glass wall in to building	6
Office wall	6
Metal door in office wall	6
Cinder wall	4
Metal door in Brick wall	12.4
Brick wall next to metal door	3

2.5.2 Small-Scale Fading หรือ Multipath Fading

คือปรากฏการณ์ที่ค่าสัญญาณที่ได้รับได้เปลี่ยนแปลงไปเนื่องจากมีจำนวนเส้นทางเดินทางของสัญญาณมากกว่าหนึ่งเส้นทางหรือมี Multipath propagation เกิดขึ้น เพราะคลื่นสัญญาณต้องเดินทางผ่านสิ่งกีดขวางรอบข้างส่งผลให้คลื่นสัญญาณเกิดการสะท้อน (Reflection) การหักเห (Diffraction) และการกระเจิง (Scattering) ในหลายทิศทาง โดยทั่วไปแล้วสัญญาณเหล่านี้มักจะเดินทางมาถึงเครื่องรับด้วยเวลาที่แตกต่างกันและมีทั้งขนาดและเฟสที่ต่างกันทำให้สัญญาณบางช่วงเกิดการเสริมกันและบางช่วงเกิดการหักล้างกัน [16]

มีแบบจำลองการเปลี่ยนแปลงของสัญญาณที่ได้รับได้ตามลักษณะการกระจายแบบ Rician (Rician distribution) [19] และการกระจายแบบ Rayleigh (Rayleigh distribution) [19]

2.5.2.1 Rician distribution

การกระจายแบบ Rician ใช้เพื่ออธิบายลักษณะการเปลี่ยนแปลงขนาดของสัญญาณตามเวลา เกี่ยวกับสถานการณ์เมื่อมีองค์ประกอบ LOS เป็นสำคัญ ซึ่งมีค่าฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability density function: PDF) [19] ดังแสดงในสมการที่ (2.10)

$$p(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2 + A^2}{2\sigma^2}\right) I_0\left(\frac{Ar}{\sigma^2}\right), (A \geq 0; r \geq 0) \quad (2.10)$$

และ
$$I_0\left(\frac{Ar}{\sigma^2}\right) = \frac{1}{2\pi} \int_0^{2\pi} \exp\left(\frac{Arcos\theta}{\sigma^2}\right) d\theta \quad (2.11)$$

I_0 คือ Modified Bessel Function อันดับที่สอง

$$I_0(x) = \sum_{i=0}^{\infty} \left(\frac{x^i}{i!.2^i}\right) \quad (2.12)$$

การกระจาย Rician โดยทั่วไปจะอธิบายอยู่ในรูปของ K พารามิเตอร์ หรือ Rician factor และมีการแสดงดังสมการที่ (2.13)

$$K = 10 \log_{10} \left(\frac{A^2}{2\sigma^2}\right) \text{ dB} \quad (2.13)$$

ถ้า A มีค่าเข้าใกล้ 0 จะทำให้ค่าพารามิเตอร์ K มีค่าเข้าใกล้ $-\infty$ dB ซึ่งทำให้ขนาดหรือความกว้างความหนาแน่นความน่าจะเป็นของการกระจาย Rician มีค่าลดลงไปเรื่อยๆ และจะกลายเป็นการกระจายแบบ Rayleigh ดังแสดงในรูปที่ 2.5

2.5.2.2 Rayleigh distribution

การกระจายแบบ Rayleigh ใช้เพื่ออธิบายลักษณะการเปลี่ยนแปลงขนาดของสัญญาณตามเวลา เกี่ยวกับสถานการณ์เมื่อไม่มีองค์ประกอบ LOS ในทางสถิติของสัญญาณที่เกิดการเฟดแบบ Rayleigh จะมีค่าการกระจายซึ่งมีค่าฟังก์ชันความหนาแน่นความน่าจะเป็น (pdf) [19] ดังแสดงในสมการที่ (2.14)

$$p(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0 \quad (2.14)$$

โดย r คือแอมพลิจูดของการกระจายแบบ Rayleigh (Volts)

σ คือค่าเบี่ยงเบนมาตรฐาน (Standard deviation)

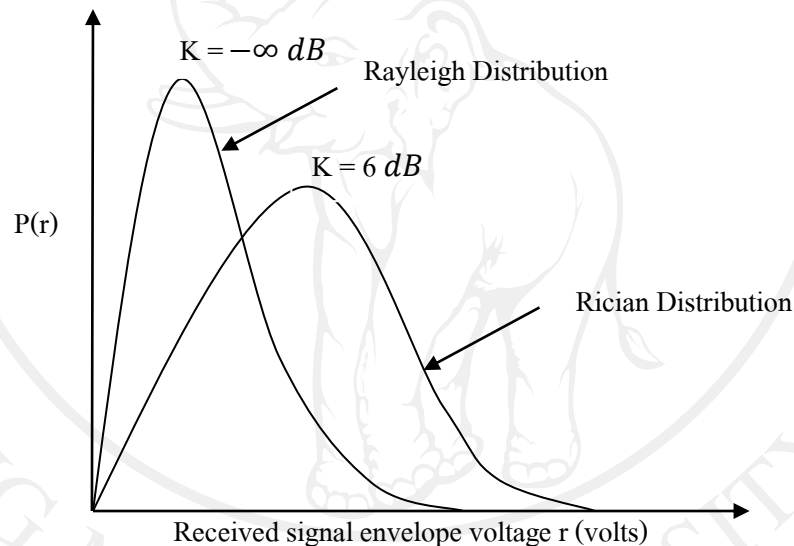
σ^2 คือค่าความแปรปรวน (Variance)

- ค่าเฉลี่ย (Mean) ของการกระจายแบบ Rayleigh ดังแสดงในสมการที่ (2.15)

$$r_{mean} = E[r] = \int_0^\infty r p(r) dr = \sigma \sqrt{\frac{\pi}{2}} = 1.25\sigma \quad (2.15)$$

- ความแปรปรวน (Variance) ของการกระจายแบบ Rayleigh ดังแสดงในสมการที่ (2.16)

$$r_{var} = E[r^2] - E^2[r] = \int_0^\infty r^2 p(r) dr - \frac{\sigma^2 \pi}{2} = \sigma^2 \left[2 - \frac{\pi}{2}\right] = 0.43\sigma^2 \quad (2.16)$$



รูปที่ 2.5 การกระจาย Rayleigh และ Rician [19]

2.6 หลักการวิเคราะห์ข้อมูลทางสถิติ

ข้อมูลทางสถิติได้แบ่งออกเป็น 2 ลักษณะคือ ข้อมูลเชิงปริมาณ (Quantitative data) และ ข้อมูลเชิงคุณภาพ (Qualitative data) สำหรับในบทนี้จะได้อธิบายถึงการวัดค่าและการกระจายของข้อมูลเชิงปริมาณต่างๆ ใช้ในงานวิจัยนี้เท่านั้นคือ

2.6.1 ค่าเฉลี่ย (Mean)

หมายถึง ค่าที่ได้จากการนำข้อมูลหรือตัวอย่าง (Samples) ทั้งหมดมารวมกัน หาค่าด้วยจำนวนข้อมูลทั้งหมด ดังสมการที่ (2.17)

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.17)$$

โดย μ คือค่าเฉลี่ย

x_i คือค่า sample ที่ i

n คือจำนวน samples ทั้งหมด

2.6.2 ค่าความแปรปรวน (Variance)

หมายถึง ค่าเฉลี่ยของความแตกต่างยกกำลังสองของข้อมูลแต่ละตัวที่เบี่ยงเบนไปจากค่าเฉลี่ยของข้อมูลชุดนั้น ดังสมการที่ (2.18)

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 \quad (2.18)$$

โดย σ^2 คือค่าความแปรปรวน

x_i คือค่า sample ที่ i

n คือจำนวน samples ทั้งหมด

ถ้าใส่รากที่สองของค่าแปรปรวนเรียกว่า ค่าเบี่ยงเบนมาตรฐาน

(Standard deviation) มีการคำนวณ ดังสมการที่ (2.19)

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}} \quad (2.19)$$

โดย σ คือค่าเบี่ยงเบนมาตรฐาน

x_i คือค่า sample ที่ i

n คือจำนวน samples ทั้งหมด

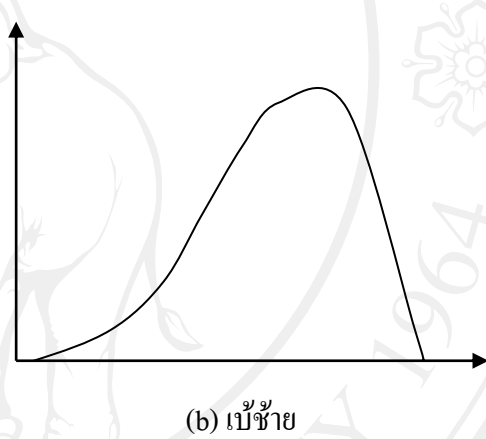
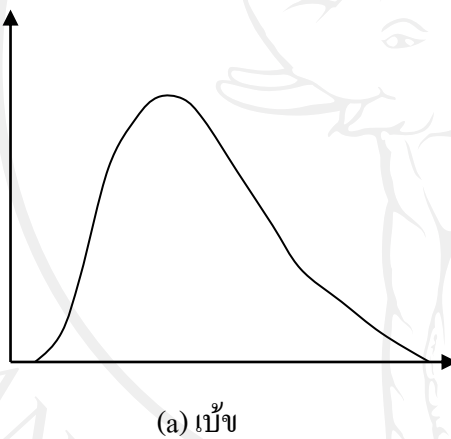
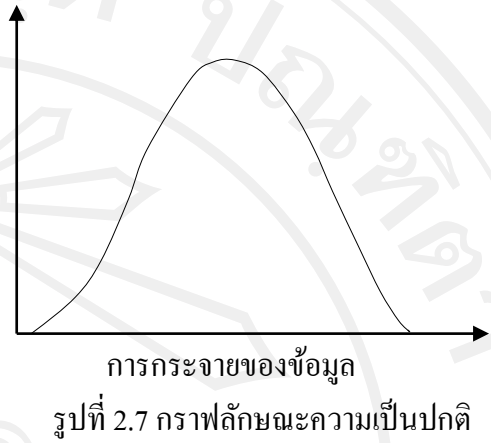
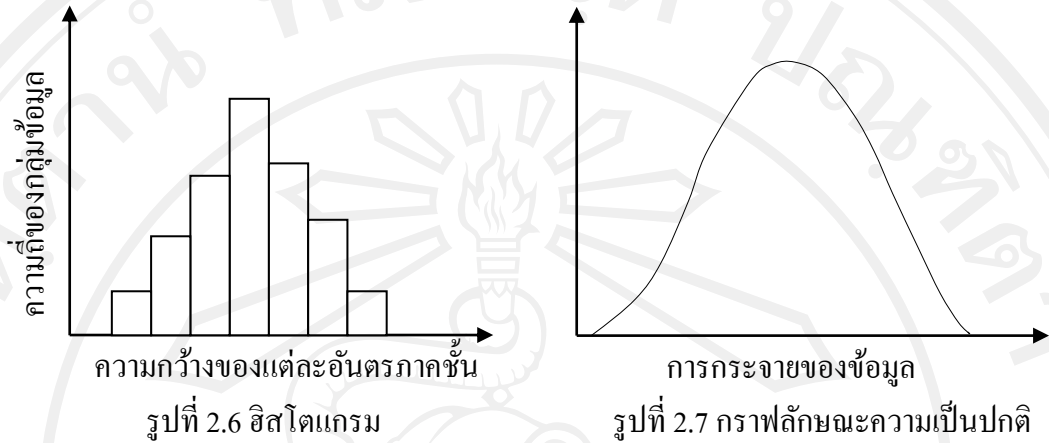
2.6.3 ฮิสโตแกรม (Histogram)

ฮิสโตแกรมเป็นการกระจายของข้อมูลอยู่ในรูปความถี่ด้วยกราฟแท่ง โดยความสูงของแต่ละแท่งที่เหลี่ยมจะแทนด้วยความถี่ของกลุ่มข้อมูลในแต่ละอันตรภาคชั้น ส่วนความกว้างของฐานแท่งที่เหลี่ยมจะแทนด้วยความกว้างของแต่ละอันตรภาคชั้น [21] ดังรูปที่ 2.6

2.6.4 ความเบ้ (Skewness)

ลักษณะความเบ้จะมีความแตกต่างจากลักษณะการกระจายข้อมูลแบบปกติ (Normal distribution) [21] คือ การกระจายแบบปกติข้อมูลจะมีการกระจายที่สมมาตร (symmetry) ดังในรูปที่ 2.7 เนื่องจากค่าเฉลี่ย (Mean) มัธยฐาน (Median) และฐานนิยม (Mode) มีค่าเท่ากัน แต่ถ้าข้อมูลมีการกระจายที่ไม่สมมาตร (Asymmetry) โดยมีลักษณะเบ้ไปด้านใดด้านหนึ่ง เนื่องจากค่าเฉลี่ย มัธยฐานและฐานนิยมมีค่าไม่เท่ากัน ถ้าปลายด้านหนึ่งของเส้นโค้งลาดยาวไปทางขวาจะเรียกว่า เบ้ขวา (Skewed to the right) หรือความเบ้เป็นบวก (Positive skewness) ดังแสดงในรูปที่ 2.8(a)และถ้า

ปลายด้านหนึ่งของเส้นโค้งลาดยาวไปท้ายซ้ายจะเรียกว่า เบ้ซ้าย (Skewed to the left) หรือความเบ้เป็นลบ (Negative skewness) ดังแสดงในรูปที่ 2.8(b)

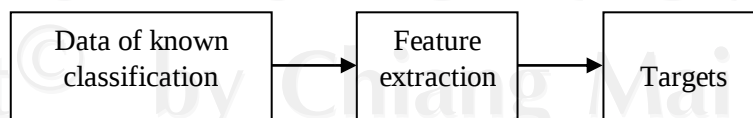


รูปที่ 2.8 กราฟลักษณะความเบ้ของข้อมูล

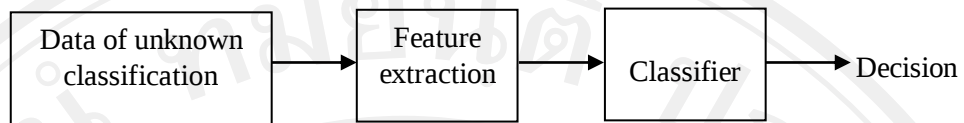
2.7 การเรียนรู้แบบมีผู้สอน (Supervised learning)

การเรียนรู้แบบมีผู้สอน เป็นเทคนิคหนึ่งของการเรียนรู้ของวิธีการรู้จำ [22] โดยอาศัยข้อมูลที่ทราบชนิดหรือคุณลักษณะแล้วเข้าสู่ขั้นตอนการฝึกฝน (Training phase) และเพื่อดำเนินการจำแนกข้อมูลที่ยังไม่ทราบชนิดในขั้นตอนทดสอบ (Testing phase) ดังแสดงแผนผังขั้นตอนการทำงานในรูปที่ 2.9

• Training phase



• **Testing phase**



รูปที่ 2.9 แผนผังขั้นตอนการทำงานการรู้จำ

การวัดประสิทธิภาพโดยวิธีการ 2-fold cross validation คือทำการแบ่งข้อมูลออกเป็นสองส่วนเท่าๆกัน (Training set และ Testing set) ซึ่งมีการทำไขว้กันระหว่างชุดข้อมูลฝึกฝนและชุดข้อมูลทดสอบเพื่อหาค่าเฉลี่ยความถูกต้อง (Average accuracy) สำหรับการคำนวณค่าความถูกต้องสามารถหาได้ดังสมการที่ (2.20)

$$\text{ค่าความถูกต้อง (\%)} = \frac{\text{จำนวนแถวข้อมูลที่ทำนายถูกต้อง}}{\text{จำนวนแถวข้อมูลทั้งหมดที่ทำนาย}} \times 100 \quad 2.20$$