

บทที่ 2

ข้อมูลและแหล่งข้อมูล

2.1 นิยามค่าสถิติความผิดพลาด

ข้อมูลที่จะใช้ในการวิจัยนี้จะอยู่ในรูปของสัญลักษณ์ที่จะแทนค่าความผิดพลาดที่เกิดขึ้น ณ เครื่องรับโดยที่ให้สัญลักษณ์ '0' แทนบิตข้อมูลที่ได้รับได้ถูกต้องและสัญลักษณ์ '1' แทนบิตข้อมูลที่เกิดผิดพลาด ดังนั้นขบวนข้อมูลที่ได้จะอยู่ในรูปของขบวนสัญลักษณ์ ดังตัวอย่างต่อไปนี้

.....00010001000001100000000000000100000000000110100000000.....

เพื่อให้การวิเคราะห์พฤติกรรมของความผิดพลาดที่เกิดขึ้นในช่องสื่อสารมีความถูกต้องและแม่นยำ งานวิจัยนี้จะใช้ค่าสถิติความผิดพลาดที่นิยามไว้โดย Brayer, K. and Cardinale, O. [2] ซึ่งมีค่าสถิติความผิดพลาดที่เป็นประโยชน์ต่อการวิจัยดังนี้

2.1.1 อัตราบิตผิดพลาด (Bit Error Rate)

อัตราบิตผิดพลาด คือ อัตราส่วนของจำนวนบิตผิดพลาดทั้งหมดของขบวนข้อมูลต่อจำนวนบิตทั้งหมดของขบวนข้อมูลเช่น ขบวนข้อมูลจำนวน 10,000 บิตมีบิตผิดพลาด 3 บิต ค่าอัตราบิตผิดพลาดคือ 3×10^{-4}

2.1.2 ค่าการกระจายของช่วงข้อมูลระหว่างบิตผิดพลาด (Gap Distribution)

ช่วงข้อมูลระหว่างบิตผิดพลาด (Gap) คือ ช่วงของบิตข้อมูลที่อยู่ระหว่างบิตผิดพลาดสองบิตเช่น ช่วงข้อมูล100000001..... ช่วงข้อมูลระหว่างบิตผิดพลาดคือ 7 และค่าการกระจายของช่วงข้อมูลระหว่างบิตผิดพลาดถูกนิยามให้เป็น ค่าความถี่สะสมของช่วงข้อมูลระหว่างบิตผิดพลาดเช่น ค่าการกระจายของช่วงข้อมูลระหว่างบิตผิดพลาดที่มีความยาว 100 บิตคือ ค่าความถี่สะสมของช่วงข้อมูลระหว่างบิตผิดพลาดที่มีความยาวน้อยกว่าหรือเท่ากับ 100 บิต

สำหรับกรณีช่องสื่อสารแบบสมมาตรสองสถานะ (Binary Symmetric Channel) ซึ่งมีความผิดพลาดแบบสุ่ม มีอัตราบิตผิดพลาดเท่ากับ p และให้ q เท่ากับ $1-p$ แล้ว ความน่าจะเป็นของช่วงข้อมูลระหว่างบิตผิดพลาดที่มีความยาว m บิตคือ

$$P(0^m) = q^m \quad (2.1)$$

และมีค่าการกระจายของช่วงข้อมูลระหว่างบิตผิดพลาดที่มีความยาว m บิตคือ

$$P_g(m) = 1 - q^m \quad (2.2)$$

2.1.3 ค่าการกระจายของช่วงปราศจากความผิดพลาด (Error-Free Run Distribution)

ค่าการกระจายของช่วงปราศจากความผิดพลาดคือค่าความถี่สะสมของความน่าจะเป็นของช่วงปราศจากความผิดพลาดที่เกิดหลังจากบิตผิดพลาดและมีความยาวอย่างน้อย x บิต และสามารถคำนวณได้จากสมการ

$$P(0^x|1) = [1 - P_g(x-1)]P(0^1|1) \quad (2.3)$$

โดยที่

$P(0^x|1)$ คือ N_g/N_c

N_g คือ จำนวนทั้งหมดของช่วงข้อมูลระหว่างบิตผิดพลาด

N_c คือ จำนวนทั้งหมดของบิตผิดพลาด

X คือ ความยาวของช่วงปราศจากความผิดพลาด

สำหรับกรณีช่องสื่อสารแบบสมมาตรสองสถานะที่มีความผิดพลาดแบบสุ่มและมีอัตราบิตผิดพลาดเฉลี่ยคือ p และให้ $q = 1-p$ แล้วค่าการกระจายของช่วงปราศจากความผิดพลาดคือ

$$P(0^x|1) = q^x \quad (2.4)$$

2.1.4 ค่าสถิติความผิดพลาดแบบบล็อกข้อมูล (Block Error Statistics)

ค่าสถิติความผิดพลาดแบบบล็อกข้อมูลถูกนิยามให้เป็นความน่าจะเป็นของช่วงข้อมูลที่มีความยาว n บิตและมีบิตผิดพลาด m บิต สำหรับกรณีช่องสื่อสารแบบสมมาตรสองสถานะที่มีความผิดพลาดแบบสุ่มและมีอัตราบิตผิดพลาดเฉลี่ยคือ p ค่าสถิติความผิดพลาดแบบบล็อกข้อมูลคือ

$$P(m,n) = \binom{n}{m} p^m (1-p)^{n-m} \quad (2.5)$$

2.1.5 ค่าการกระจายของเบิร์สต์ (Burst Distribution)

นิยามของเบิร์สต์คือช่วงข้อมูลที่เริ่มต้นด้วยบิตผิดพลาดซึ่งไม่ใช่ของเบิร์สต์ที่เกิดขึ้นล่วงหน้าก่อนและจบด้วยบิตผิดพลาดที่ไม่ใช่ของเบิร์สต์ที่จะเกิดตามมา โดยที่ความยาวของแต่ละเบิร์สต์นั้นจะถูกกำหนดด้วยเงื่อนไขของค่ากำหนดความหนาแน่นของบิตผิดพลาด (Specified Error Density, Δ_0) ซึ่งความหนาแน่นของบิตผิดพลาดเท่ากับอัตราส่วนของจำนวนบิตผิดพลาดทั้งหมดที่มีในช่วงข้อมูลต่อจำนวนบิตทั้งหมดของช่วงข้อมูลนั้น และเบิร์สต์ที่พิจารณาจะต้องมีความหนาแน่นของบิตผิดพลาดเท่ากับหรือมากกว่าค่ากำหนดความหนาแน่นของบิตผิดพลาด

ตัวอย่างการหาความยาวของเบิร์สต์นั้นมีดังนี้ สมมุติค่ากำหนดความหนาแน่นของบิตผิดพลาดมีค่า 0.3 และพิจารณาช่วงข้อมูล

.....000011001000100000000001000.....

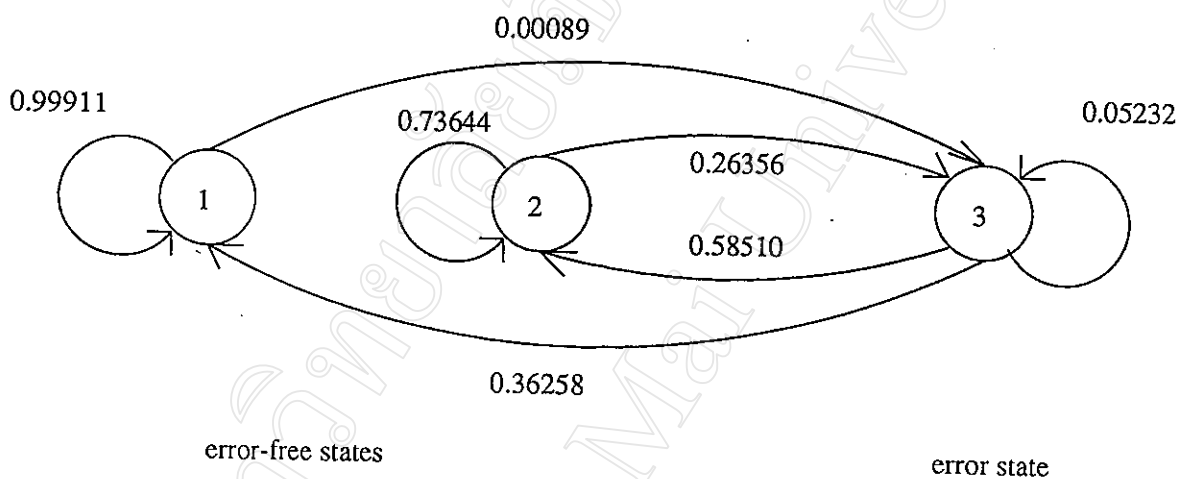
เบิร์สต์จะเริ่มจากบิตผิดพลาด “1” ตัวที่หนึ่ง ถ้านับรวมเอาบิตผิดพลาดตัวที่สองค่าความหนาแน่นของบิตผิดพลาดจะเท่ากับ 1 ซึ่งมีค่ามากกว่าค่ากำหนดความหนาแน่นของบิตผิดพลาด ถ้านับรวมเอาบิตผิดพลาดตัวที่สามค่าความหนาแน่นของบิตผิดพลาดจะเท่ากับ 3/5 หรือ 0.6 ซึ่งมีค่ามากกว่าค่ากำหนดความหนาแน่นของบิตผิดพลาด ถ้านับรวมเอาบิตผิดพลาดตัวที่สี่ค่าความหนาแน่นของบิตผิดพลาดจะเท่ากับ 4/9 หรือ 0.44 ซึ่งมีค่ามากกว่าค่ากำหนดความหนาแน่นของบิตผิดพลาด ถ้านับรวมเอาบิตผิดพลาดตัวที่ห้าค่าความหนาแน่นของบิตผิดพลาดจะเท่ากับ 5/20 หรือ 0.25 ซึ่งมีค่าน้อยกว่าค่ากำหนดความหนาแน่นของบิตผิดพลาด ดังนั้นเบิร์สต์จึงจบลงที่บิตผิดพลาดตัวที่สี่และมีความยาวเท่ากับ 9 บิต และช่วงของเบิร์สต์ก็คือ ...110010001..... และค่าการกระจายของเบิร์สต์คือค่าความถี่สะสมของเบิร์สต์ความยาวใด ๆ

2.1.6 ค่าการกระจายของช่วงห่างของเบิร์สต์ (Burst Interval Distribution)

ช่วงห่างของเบิร์สต์ (Burst Interval) ก็คือช่วงของข้อมูลระหว่างเบิร์สต์สองเบิร์สต์ซึ่งจะเริ่มต้นด้วยบิตถูกต้อง “0” ที่เกิดตามหลังบิตผิดพลาดตัวสุดท้ายของเบิร์สต์ที่ผ่านมา และจบด้วยบิตถูกต้อง “0” ที่เกิดขึ้นก่อนบิตผิดพลาดตัวแรกของเบิร์สต์ตัวต่อไป และแม้ว่าช่วงห่างของเบิร์สต์จะสามารถมีบิตผิดพลาดเกิดขึ้นได้แต่ความหนาแน่นของบิตผิดพลาดจะต้องมีค่าน้อยกว่าค่ากำหนดความหนาแน่นของบิตผิดพลาดที่ถูกเลือกไว้สำหรับการหาความยาวของเบิร์สต์ และค่าการกระจายของช่วงห่างของเบิร์สต์ก็คือค่าความถี่สะสมของช่วงห่างของเบิร์สต์ความยาวใด ๆ

2.2 ข้อมูล และค่าสถิติความผิดพลาด

เนื่องจากมีข้อจำกัดในด้านเครื่องมือทดลองเพื่อให้ได้ขบวนการข้อมูลของช่องสื่อสารที่มีความผิดพลาดแบบเบรสต์ จึงจะอาศัยแบบจำลองลูกโซ่มาร์คอฟแบบแยกส่วนที่ Fritchman, B.D. [7] เสนอไว้แบบกรณีทั่วไป โดยแบ่งกลุ่มสถานะของบิตออกเป็นสองกลุ่ม กลุ่มแรกเป็นสถานะที่บิตไม่ผิดพลาด (Error-free States) มี k สถานะ ส่วนอีกกลุ่มเป็นสถานะที่บิตผิดพลาด (Error States) มี n สถานะ โดยที่ขบวนการข้อมูลเพื่อนำมาใช้ในการวิจัยนี้จะถูกจำลองด้วยแบบจำลองลูกโซ่มาร์คอฟแบบแยกส่วนที่มีค่า k เท่ากับสองหรือมีสถานะที่บิตไม่ผิดพลาดสองสถานะ และค่า n เท่ากับหนึ่งหรือมีสถานะที่บิตผิดพลาดหนึ่งสถานะมีลักษณะดังรูป 2.1



รูป 2.1 แบบจำลองลูกโซ่มาร์คอฟแบบแยกส่วน

แบบจำลองในรูป 2.1 นั้น เป็นแบบจำลองลูกโซ่มาร์คอฟแบบแยกส่วนที่ใช้ในการจำลองลักษณะการเกิดความผิดพลาดของช่องสื่อสารความถี่สูงที่มีค่าอัตราบิตผิดพลาดเฉลี่ยจริงคือ 1.789×10^{-3} ที่ Tsai, S. [24] เคยนำเสนอไว้ในการจำลองแบบช่องสื่อสารความถี่สูง โดยแบบจำลองจะผลิตลักษณะของข้อมูลสองสัญลักษณ์คือสัญลักษณ์ “0” แทนบิตที่ถูกต้องที่เครื่องรับ และสัญลักษณ์ “1” แทนบิตที่เกิดความผิดพลาดในการรับข้อมูลที่เครื่องรับ แต่ละบิตที่ผลิตขึ้นมาจะอยู่ในแต่ละสถานะ โดยที่บิตต่อไปจะเป็นสัญลักษณ์ใดในสถานะใดนั้นขึ้นอยู่กับความน่าจะเป็นส่งผ่าน (Transition Probabilities) ในการเปลี่ยนสถานะหรือคงอยู่สถานะเดิม

ความน่าจะเป็นส่งผ่านจากสถานะ i ไปสู่สถานะ j เขียนแทนด้วยรูปของเซต

$\{p_{ij} : i = 1,2,3 : j = 1,2,3\}$ และความน่าจะเป็นส่งผ่านของแบบจำลองในรูป 2.1 เขียนอยู่ในรูปของเมตริกซ์ส่งผ่านได้ดังนี้

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 0.99911 & 0 & 0.00089 \\ 0 & 0.73644 & 0.26356 \\ 0.36258 & 0.58510 & 0.05232 \end{bmatrix}$$

ค่าอัตราบิตผิดพลาดเฉลี่ยของแบบจำลองถูกใช้มาร์คอฟแบบแยกส่วนสามารถคำนวณได้จากสมการดังต่อไปนี้ [14]

$$p_e = [1 + \sum_{v=1}^{N-1} (p_{Nv} / p_{vN})] - 1 \quad (2.6)$$

โดยที่ p_e คือ ค่าอัตราบิตผิดพลาดเฉลี่ย

N คือ จำนวนสถานะทั้งหมด และ N เป็นเลขของสถานะที่บิตผิดพลาด

p_{Nv} คือ ความน่าจะเป็นส่งผ่านจากสถานะ N ไปสู่สถานะ v

p_{vN} คือ ความน่าจะเป็นส่งผ่านจากสถานะ v ไปสู่สถานะ N

ดังนั้นอาศัยค่าความน่าจะเป็นส่งผ่านของแบบจำลองแทนค่าลงในสมการที่ (2.6) จากการคำนวณจะได้ค่าอัตราบิตผิดพลาดเฉลี่ยของแบบจำลองมีค่า 2.435×10^{-3}

เพื่อให้มีขบวนการข้อมูลที่มีความหลากหลายสำหรับการวิเคราะห์ ผู้วิจัยจึงใช้แบบจำลองดังรูป 2.1 ผลิตขบวนการข้อมูลจำนวนสี่ชุด โดยเขียนแบบจำลองด้วยโปรแกรมคอมพิวเตอร์ภาษาซีบนเครื่องคอมพิวเตอร์ส่วนบุคคลและขบวนการข้อมูลแต่ละชุดมีจำนวนบิตทั้งหมดสี่ล้านบิต และให้ชื่อเพิ่มข้อมูลทั้งสี่ คือ “data1” “data2” “data3” และ “data4” ตามลำดับ ค่าสถิติความผิดพลาดต่างๆ ของข้อมูลทั้งสี่ชุดคือ ค่าการกระจายของช่วงข้อมูลระหว่างบิตผิดพลาด ค่าการกระจายของช่วงปราศจากความผิดพลาด ค่าการกระจายของเบรสต์ ค่าการกระจายของช่วงห่างของเบรสต์ แสดงดังกราฟรูป 2.2 - 2.5 ตามลำดับและข้อมูลสรุปเกี่ยวกับอัตราบิตผิดพลาดของข้อมูลทั้งสี่ชุดมีดังนี้

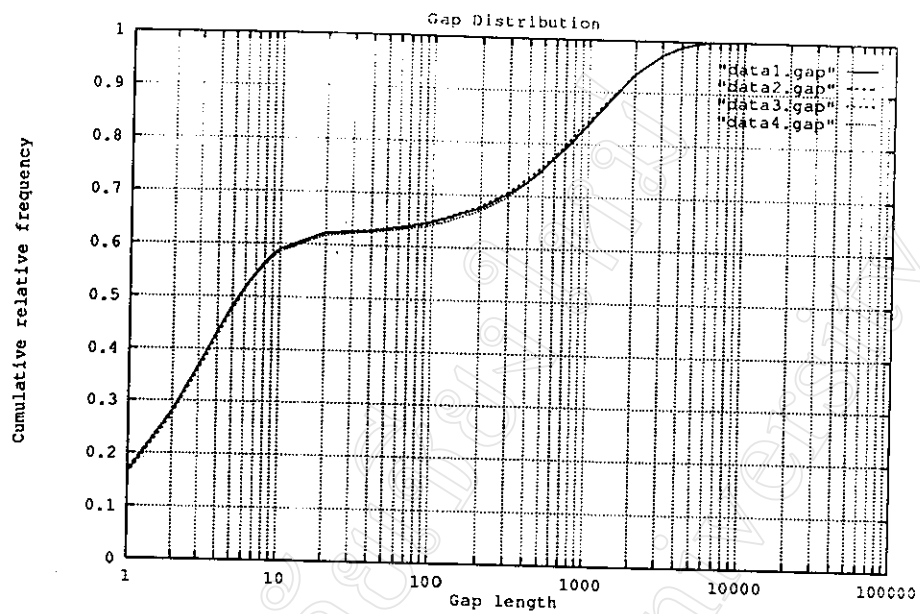
อัตราบิตผิดพลาดที่คำนวณได้จากแบบจำลอง คือ 2.435×10^{-3}

อัตราบิตผิดพลาดของขบวนการข้อมูล “data1” คือ 2.343×10^{-3}

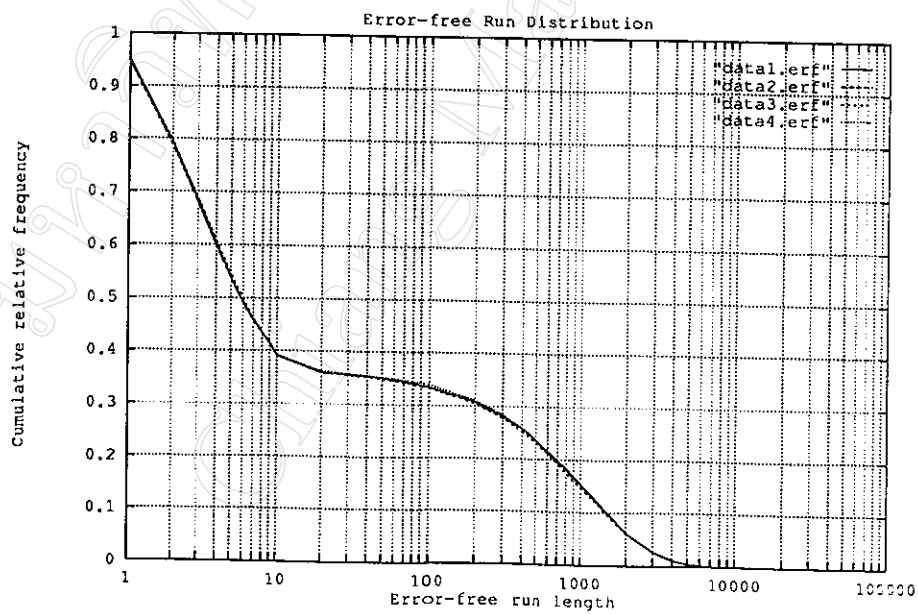
อัตราบิตผิดพลาดของขบวนการข้อมูล “data2” คือ 2.390×10^{-3}

อัตราบิตผิดพลาดของขบวนการข้อมูล “data3” คือ 2.440×10^{-3}

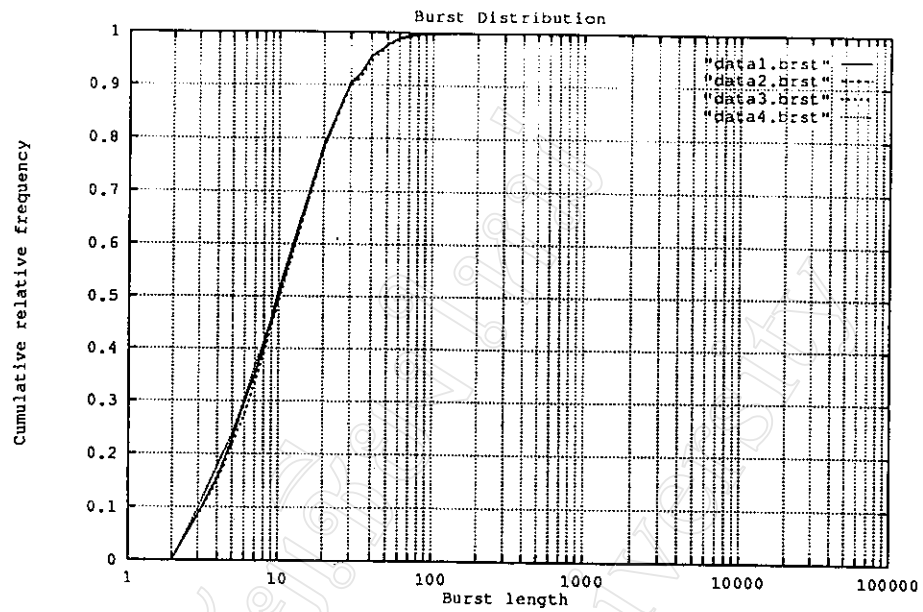
อัตราบิตผิดพลาดของขบวนการข้อมูล “data4” คือ 2.413×10^{-3}



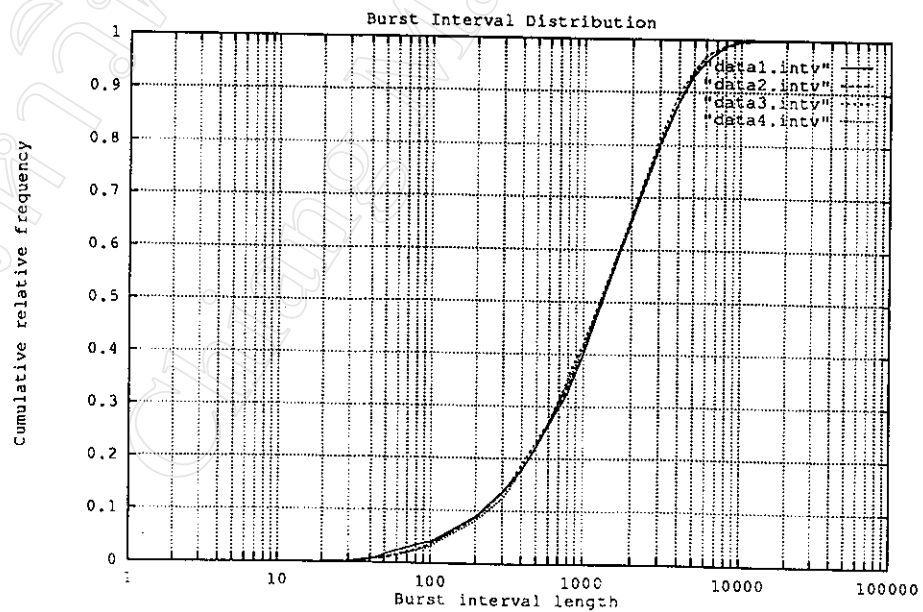
รูป 2.2 ค่าการกระจายของช่วงข้อมูลระหว่างบิตผิดพลาดของข้อมูลทั้ง 4 ชุด



รูป 2.3 ค่าการกระจายของช่วงปราศจากความผิดพลาดของข้อมูลทั้ง 4 ชุด



รูป 2.4 ค่าการกระจายของเบรสต์ของข้อมูลทั้ง 4 ชุด



รูป 2.5 ค่าการกระจายของช่วงห่างของเบรสต์ของข้อมูลทั้ง 4 ชุด

2.3 วิเคราะห์ผลการเปรียบเทียบค่าสถิติความผิดพลาดของข้อมูล

จากการเปรียบเทียบโดยการวาดกราฟพบว่าค่าสถิติความผิดพลาดของข้อมูลทั้งสี่ชุด มีความคล้ายคลึงกันมากซึ่งสามารถสรุปได้ว่าในเชิงสถิติข้อมูลทั้งสี่ชุดมีความเป็นหนึ่งเดียวกัน และถึงแม้ว่าค่าอัตราบิดผิดพลาดของข้อมูลทั้งสี่ชุดมีค่าไม่เท่ากับค่าอัตราบิดผิดพลาดของแบบจำลองที่ใช้ในการผลิตข้อมูลแต่ก็มีเปอร์เซ็นต์คลาดเคลื่อนของอัตราบิดผิดพลาดของข้อมูลจากเพิ่มข้อมูล “data1” “data2” “data3” และ “data4” เมื่อเปรียบเทียบกับค่าที่คำนวณได้จากแบบจำลองเพียง 3.78% 1.85% 0.20% และ 3.49% ตามลำดับ

จากการวิเคราะห์ผลการเปรียบเทียบค่าสถิติความผิดพลาดของข้อมูลทั้งสี่ชุดซึ่งแตกต่างกันเพียงเล็กน้อยและสามารถยอมรับได้ ผู้วิจัยจะนำไปใช้เป็นประโยชน์ในการจำลองแบบและการหาค่าพารามิเตอร์ต่างๆ ด้วยแบบจำลองที่จะได้เสนอต่อไป