

ผลของระยะห่างระหว่างอักษร ความหนาของเส้นรบกวน และเงื่อนไขการใช้สี ต่ออัตรายืนยันความ
เป็นมนุษย์และอัตราความทนทานของแคปทซ์ข้อความ



บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศทางธุรกิจ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2560
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

EFFECTS OF CHARACTER SPACING, THICKNESS OF DISTURBING LINE AND CONDITIONS
OF COLOR USAGE ON HUMAN AFFIRMATIVE RATES AND ROBUSTNESS RATES OF TEXT-
BASED CAPTCHA



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Information Technology in Business
Faculty of Commerce and Accountancy
Chulalongkorn University
Academic Year 2017
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

ผลของระยะห่างระหว่างอักษร ความหนาของเส้นรบกวน และเงื่อนไขการใช้สี ต่ออัตรายืนยันความเป็นมนุษย์และ อัตราความทนทานของแคปTCHAข้อความ

โดย

นางสาวศุภาพิชญ์ ชูชาติ

สาขาวิชา

เทคโนโลยีสารสนเทศทางธุรกิจ

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ผู้ช่วยศาสตราจารย์ ดร.ชัชพงศ์ ตั้งมณี

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโทบริหารธุรกิจ

บัญชี

..... คณบดีคณะพาณิชยศาสตร์และการ
(รองศาสตราจารย์ ดร.พลุ เดชะรินทร์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.จันทร์เจ้า มงคลนาวิน)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.ชัชพงศ์ ตั้งมณี)

..... กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.ปรวัฒน์ วิสูตรศักดิ์)

..... กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.พงษ์สิน ภูแสนคำ)

ศุภาพิชญ์ ชูชาติ : ผลของระยะห่างระหว่างอักษร ความหนาของเส้นรบกวน และเงื่อนไขการใช้สี ต่ออัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปต์ชาข้อความ (EFFECTS OF CHARACTER SPACING, THICKNESS OF DISTURBING LINE AND CONDITIONS OF COLOR USAGE ON HUMAN AFFIRMATIVE RATES AND ROBUSTNESS RATES OF TEXT-BASED CAPTCHA) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร.ชัชพงศ์ ตั้งมณี, หน้า.

แคปต์ชาข้อความ เป็นเครื่องมือที่ใช้ทดสอบว่าผู้ใช้งานเป็นมนุษย์จริงหรือไม่ ดังนั้นแคปต์ชาข้อความที่ดีควรป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ แต่มนุษย์ยังคงเข้าใจได้ง่าย การวัดความสามารถของแคปต์ชาข้อความคือ (1) การวัดอัตราการยืนยันความเป็นมนุษย์ อันหมายถึงอัตราที่ผู้ใช้สามารถระบุอักษรที่ปรากฏในแคปต์ชาข้อความได้ถูกต้อง และ (2) การวัดอัตราความทนทานของแคปต์ชาข้อความ อันหมายถึงอัตราที่แคปต์ชาข้อความสามารถป้องกันการโจมตีของโปรแกรมอัตโนมัติได้ ดังนั้นจึงน่าสนใจที่จะวิเคราะห์ผลของ (1) ระยะห่างระหว่างอักษร (2) ความหนาของเส้นรบกวน และ (3) สีของเส้นรบกวน ต่ออัตราการยืนยันความเป็นมนุษย์ และอัตราความทนทานของแคปต์ชาข้อความ

งานวิจัยนี้เป็นการวิจัยเชิงทดลองในห้องปฏิบัติการคอมพิวเตอร์กับหน่วยทดลองจำนวน 360 คน การวิเคราะห์ข้อมูลพบว่าที่ระดับนัยสำคัญทางสถิติ 0.05 ระยะห่างระหว่างอักษร และความหนาของเส้นรบกวนมีผลต่ออัตราการยืนยันความเป็นมนุษย์ ทว่าผลของสีของเส้นรบกวนต่ออัตราการยืนยันความเป็นมนุษย์ไม่มีนัยสำคัญ นอกจากนี้ยังพบว่าระยะห่างระหว่างอักษร และความหนาของเส้นรบกวนมีผลต่ออัตราความทนทานของแคปต์ชาข้อความ ทว่าผลของสีของเส้นรบกวนต่ออัตราความทนทานของแคปต์ชาข้อความไม่มีนัยสำคัญ ข้อสรุปที่ได้จากงานวิจัยนี้เป็นการต่อยอดองค์ความรู้สำหรับพัฒนาแคปต์ชาข้อความ อีกทั้งผู้พัฒนาแคปต์ชาข้อความสามารถใช้ข้อค้นพบนี้ในเลือกระยะห่างระหว่างอักษร และเส้นรบกวน เพื่อออกแบบแคปต์ชาข้อความให้ง่ายต่อการใช้งานของมนุษย์ และยากต่อการโจมตีของโปรแกรมอัตโนมัติ

สาขาวิชา เทคโนโลยีสารสนเทศทางธุรกิจ

ลายมือชื่อนิสิต

ปีการศึกษา 2560

ลายมือชื่อ อ.ที่ปรึกษาหลัก

5781588026 : MAJOR INFORMATION TECHNOLOGY IN BUSINESS

KEYWORDS: TEXT-BASED CAPTCHA / CHARACTER SPACING / THICKNESS OF DISTURBING LINE / CONDITIONS OF COLOR USAGE / HUMAN AFFIRMATIVE RATES / ROBUSTNESS RATES OF TEXT-BASED CAPTCHA

SUPAPICH CHOOCHART: EFFECTS OF CHARACTER SPACING, THICKNESS OF DISTURBING LINE AND CONDITIONS OF COLOR USAGE ON HUMAN AFFIRMATIVE RATES AND ROBUSTNESS RATES OF TEXT-BASED CAPTCHA. ADVISOR: ASST. PROF. CHATPONG TANGMANEE, Ph.D., pp.

Text-based CAPTCHA is a tool to assess if an online request is from a human user or from a software bot. Acceptable CAPTCHAs should be tolerant to automated software attack but still easy for humans to solve. The methods to measure the effectiveness of CAPTCHAs are (1) a human affirmative rate which means the rate through which users can identify the text that appears in CAPTCHAs correctly and (2) a robustness rate of CAPTCHAs which means the rate through which CAPTCHAs are safe from automated software attack. Therefore, it is interesting to examine factors related to text-based CAPTCHAs. This study examined the effects of (1) character spacing, (2) thickness of disturbing line, and (3) color of disturbing line on the human affirmative rates and the robustness rates of CAPTCHAs.

The study is based on a laboratory experiment. The analysis at 0.05 level indicates that the effects of character spacing and thickness of disturbing line are statistically significant on the human affirmative rate, but the effect of color of disturbing line is not significant. In addition, the effects of character spacing and thickness of disturbing line are statistically significant on the robustness rate of CAPTCHA, but the effect of color of disturbing line is not significant. In addition to extending knowledge of text-based CAPTCHA's design, developers could apply the findings to define character spacing and thickness of disturbing line to design text-based CAPTCHAs that are easy for human but difficult for automated software attack.

Field of Study: Information Technology in Student's Signature

Business

Advisor's Signature

Academic Year: 2017

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความช่วยเหลือของผู้ช่วยศาสตราจารย์ ดร. ชัชพงศ์ ตังมณี อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่กรุณาสละเวลาเพื่อให้คำปรึกษาและความรู้อันเป็นประโยชน์ในการทำวิทยานิพนธ์อย่างยิ่ง ตลอดจน ตรวจสอบและแก้ไขข้อบกพร่องต่างๆ รวมถึงช่วยแก้ไขปัญหาที่เกิดขึ้นระหว่างการดำเนินงานอีกด้วย ผู้วิจัยจึงขอขอบพระคุณไว้ ณ โอกาสนี้ และขอขอบพระคุณผู้ช่วยศาสตราจารย์ ดร. จันท์เจ้า มงคลนาวิน ประธานกรรมการวิทยานิพนธ์ อาจารย์ ดร. พงษ์สิน ภูแสนคำ และผู้ช่วยศาสตราจารย์ ดร. ประวิทย์ วิสูตรศักดิ์ กรรมการวิทยานิพนธ์ภายนอกมหาวิทยาลัย ที่ช่วยชี้แนะแนวทาง และให้ความรู้ ตลอดจนคำแนะนำเพื่อให้วิทยานิพนธ์ฉบับนี้สมบูรณ์มากยิ่งขึ้น

ผู้วิจัยขอขอบพระคุณอาจารย์ทุกท่านที่ให้ความกรุณาช่วยเหลือในการติดต่อหน่วยทดลองเพื่อเก็บข้อมูลในงานวิจัยนี้ ขอขอบพระคุณ อาจารย์ ดร. อรุณี กำลิ่ง รองศาสตราจารย์ ดร. ณัฐพล อัสสระรัตน์ ผู้ช่วยศาสตราจารย์ ดร. ดนุพล หุ่นโสภณ ผู้ช่วยศาสตราจารย์ ดร. ธัญลักษณ์ วิจิตรสารวงค์ อาจารย์ ดร. นภมณี เตพละกุล อาจารย์ ดร. ปริญญา มณีโรจน์ อาจารย์ ดร. วศธร ชุตติภิญโญ อาจารย์ ดร. กฤษณา วิสมิตะนันท์ ผู้ช่วยศาสตราจารย์ ดร. ธารทัศน์ โมกขมรรคกุล อาจารย์ ดร. วรพงษ์ ลีวัฒนกิจ และอาจารย์ ดร. ณัตติฤดี เจริญรักษ์ นอกจากนี้ขอขอบคุณเจ้าหน้าที่ศูนย์คอมพิวเตอร์ที่ 9 คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ที่ช่วยเหลือในการเก็บข้อมูลของงานวิจัยเป็นอย่างดี รวมทั้งหน่วยทดลองทุกท่านที่สละเวลาเพื่อมาให้ข้อมูลในห้องปฏิบัติการคอมพิวเตอร์

ที่สำคัญอย่างยิ่งต้องขอขอบพระคุณคุณพ่อ คุณแม่ ที่มอบกำลังใจและคอยเป็นห่วงตลอดมา รวมทั้งเพื่อนๆทุกคนที่ให้ความช่วยเหลือจนทำให้งานวิจัยนี้สำเร็จไปได้ด้วยดี

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญรูปภาพ.....	ซ
สารบัญตาราง.....	ฅ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	11
1.3 ตัวแปรสำคัญที่ศึกษา.....	12
1.4 คำจำกัดความของการวิจัย.....	14
1.5 ขอบเขตของการวิจัย.....	14
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	15
บทที่ 2 การทบทวนวรรณกรรม.....	16
2.1 ความนำ.....	16
2.2 รูปแบบของแคปTCHAข้อความ.....	16
2.3 การยืนยันความเป็นมนุษย์ของแคปTCHAข้อความ.....	22
2.4 โปรแกรมแคปTCHAโซลเวอร์ (CAPTCHA Solver).....	23
2.5 การป้องกันโปรแกรมอัตโนมัติโจมตีแคปTCHAข้อความ.....	25
2.6 ระยะห่างระหว่างอักษร (Spacing).....	29
2.7 สิ่งรบกวนของแคปTCHAข้อความ (Noise).....	32
2.8 วรรณกรรมที่เกี่ยวข้อง: สรุปรูป.....	35

บทที่ 3 ระเบียบวิธีวิจัย	36
3.1 ความนำ	36
3.2 ประชากรและหน่วยทดลอง	36
3.3 การทดสอบสมมติฐาน	37
3.4 เครื่องมือเก็บข้อมูล	41
3.5 ประเด็นความเชื่อถือได้ (Reliability) และความถูกต้อง (Validity)	51
3.6 ขั้นตอนการเก็บข้อมูล	52
3.7 เทคนิคที่ใช้วิเคราะห์ข้อมูลที่เก็บได้	53
บทที่ 4 ผลการวิเคราะห์ข้อมูล	56
4.1 ความนำ	56
4.2 ลักษณะทางประชากรของหน่วยทดลอง	56
4.3 อัตราการยืนยันความเป็นมนุษย์	58
4.4 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อแคปท์ชาข้อความแสดงด้วย ระยะห่างระหว่างอักขระที่แตกต่างกัน	62
4.5 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อแคปท์ชาข้อความมีเส้นรบกวนที่มี ความหนาแตกต่างกัน	64
4.6 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อแคปท์ชาข้อความแสดงด้วยสีของ เส้นรบกวนที่แตกต่างกัน	66
4.7 อัตราความทนทานของแคปท์ชาข้อความ	67
4.8 การเปรียบเทียบอัตราความทนทานของแคปท์ชาข้อความ เมื่อแคปท์ชาข้อความแสดง ด้วยระยะห่างระหว่างอักขระที่แตกต่างกัน	71
4.9 การเปรียบเทียบอัตราความทนทานของแคปท์ชาข้อความ เมื่อแคปท์ชาข้อความมีเส้น รบกวนที่มีความหนาแตกต่างกัน	73

4.10 การเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อแคปท์ซาข้อความแสดงสี ของเส้นรบกวนที่แตกต่างกัน	75
4.11 สรุปผลการวิเคราะห์อัตราการยืนยันความเป็นมนุษย์ และอัตราความทนทานของแคปท์ ซาข้อความ	76
4.12 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ จำแนกตามข้อมูลส่วนบุคคลของหน่วย ทดลอง	79
4.13 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ จำแนกตามตัวอักษร	81
4.14 การเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ จำแนกตามตัวอักษร.....	87
บทที่ 5 สรุปและอภิปรายผล.....	95
5.1 ความนำ.....	95
5.2 การทดลองและลักษณะของหน่วยทดลอง.....	95
5.3 อัตราการยืนยันความเป็นมนุษย์กับระยะห่างระหว่างอักษรของแคปท์ซาข้อความ	95
5.4 อัตราความทนทานของแคปท์ซาข้อความกับระยะห่างระหว่างอักษรของแคปท์ซาข้อความ .	96
5.5 อัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปท์ซาข้อความกับ ระยะห่างระหว่างอักษรของแคปท์ซาข้อความ	97
5.6 อัตราการยืนยันความเป็นมนุษย์กับความหนาของเส้นรบกวนของแคปท์ซาข้อความ	97
5.7 อัตราความทนทานของแคปท์ซาข้อความกับความหนาของเส้นรบกวนของแคปท์ซา ข้อความ	98
5.8 อัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปท์ซาข้อความกับความ หนาของเส้นรบกวนของแคปท์ซาข้อความ	99
5.9 อัตราการยืนยันความเป็นมนุษย์กับสีของเส้นรบกวนของแคปท์ซาข้อความ.....	99
5.10 อัตราความทนทานของแคปท์ซาข้อความกับสีของเส้นรบกวนของแคปท์ซาข้อความ	100
5.11 อัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปท์ซาข้อความกับสีของ เส้น.....	100
รบกวนของแคปท์ซาข้อความ.....	100

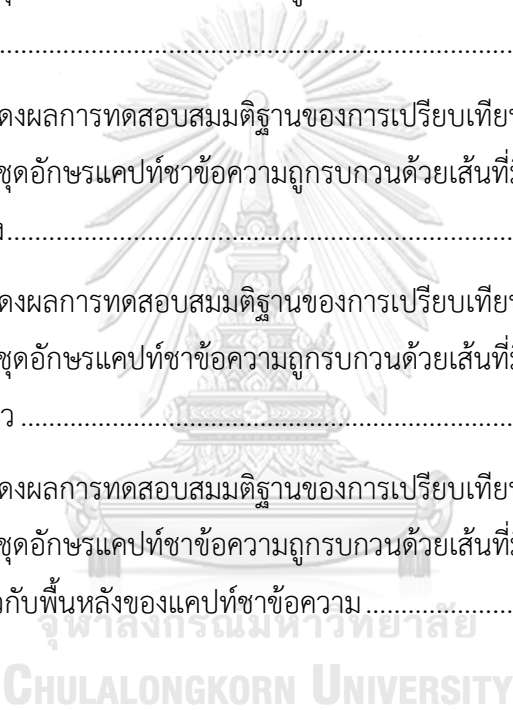
5.12 อัตราการยืนยันความเป็นมนุษย์กับตัวอักษรในชุดอักษรแคปทชาข้อความ	101
5.13 การนำงานวิจัยไปใช้ (contribution).....	101
5.14 ข้อจำกัดของงานวิจัยและข้อเสนอแนะ.....	104
รายการอ้างอิง	106
ภาคผนวก.....	109
ภาคผนวก ก ใบงาน	109
ภาคผนวก ข สมการคำนวณค่า z.....	110
ภาคผนวก ค ค่าระยะห่างระหว่างอักษรแบบขีด (0 พอยต์).....	120
ภาคผนวก ง การเปรียบเทียบจำแนกตามคู่อักษร.....	124
ภาคผนวก จ การกำหนด Filter ของโปรแกรม GSA Captcha Breaker	185
ประวัติผู้เขียนวิทยานิพนธ์	205

สารบัญรูปภาพ

ภาพที่ 1. 1 ตัวอย่างภาพอักษรพิมพ์เล็ก a ของแบบอักษร Arial	4
ภาพที่ 1. 2 ตัวอย่างการวัดค่าระยะห่างระหว่างอักษรพิมพ์เล็ก b กับ w ของแบบอักษร Arial	5
ภาพที่ 1. 3 ตัวอย่างเส้นรบกวนสีเดียวกับอักษรในแคปต์ซาข้อความ	5
ภาพที่ 1. 4 ตัวอย่างการกำหนดระยะห่างระหว่างอักษร b กับ w โดยกำหนดค่า kerning เท่ากับลบ 44 Ps	8
ภาพที่ 1. 5 ตัวอย่างการกำหนดระยะห่างระหว่างอักษร g t s x k และ y โดยกำหนดค่า tracking เท่ากับลบ 156 Ps.....	9
ภาพที่ 1. 6 แสดงตัวอย่างตัวอักษร a และเส้นรบกวนที่มีขนาด 6 พอยต์เท่ากัน	9
ภาพที่ 1. 7 ตัวอย่างการกำหนดความหนาของเส้นรบกวนเป็น 3 พิกเซล ด้วยโปรแกรม Adobe Photoshop CS3.....	10
ภาพที่ 1. 8 ตัวอย่างการกำหนดสีของเส้นรบกวนเป็นสีเดียวกับอักษรด้วยโปรแกรม Adobe Photoshop CS3.....	11
ภาพที่ 1. 9 ตัวอย่างแคปต์ซาข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าปกติ	12
ภาพที่ 1. 10 ตัวอย่างแคปต์ซาข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าชิด	12
ภาพที่ 1. 11 ตัวอย่างแคปต์ซาข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าทับซ้อน	12
ภาพที่ 1. 12 ตัวอย่างแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล	12
ภาพที่ 1. 13 ตัวอย่างแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล	12
ภาพที่ 1. 14 ตัวอย่างแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล	13
ภาพที่ 1. 15 ตัวอย่างแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับอักษร และมีขนาด 1 พิกเซล	13
ภาพที่ 1. 16 ตัวอย่างแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับพื้นหลัง และมีขนาด 3 พิกเซล	13
ภาพที่ 2. 1 ตัวอย่างแคปต์ซาข้อความจาก authorize.net	17

ภาพที่ 2. 2 ตัวอย่างแคปต์ชาข้อความจาก baidu.com	18
ภาพที่ 2. 3 ตัวอย่างแคปต์ชาข้อความจาก captcha.net	18
ภาพที่ 2. 4 ตัวอย่างแคปต์ชาข้อความจาก digg.com	18
ภาพที่ 2. 5 ตัวอย่างแคปต์ชาข้อความจาก ebay.com.....	19
ภาพที่ 2. 6 ตัวอย่างแคปต์ชาข้อความจาก google.com	19
ภาพที่ 2. 7 ตัวอย่างแคปต์ชาข้อความจาก mail.ru	20
ภาพที่ 2. 8 ตัวอย่างแคปต์ชาข้อความจาก live.com	20
ภาพที่ 2. 9 ตัวอย่างแคปต์ชาข้อความจาก recaptcha.net	20
ภาพที่ 2. 10 ตัวอย่างแคปต์ชาข้อความจาก skyrock.com	21
ภาพที่ 2. 11 ตัวอย่างแคปต์ชาข้อความจาก slashdot.org	21
ภาพที่ 2. 12 ตัวอย่างแคปต์ชาข้อความจาก worldofwarcraft.com.....	22
ภาพที่ 2. 13 ตัวอย่างแคปต์ชาข้อความจาก yahoo.com	22
ภาพที่ 2. 14 (ก) แคปต์ชาข้อความต้นฉบับ (ข) แคปต์ชาข้อความที่โปรแกรมอัตโนมัติไม่สามารถ แบ่งส่วน (Segment) เส้นของตัวอักษรได้.....	28
ภาพที่ 2. 15 (ก) แคปต์ชาข้อความต้นฉบับ (ข) แคปต์ชาข้อความที่โปรแกรมอัตโนมัติวิเคราะห์ แล้วทำให้เส้นของตัวอักษรบางส่วนขาดหายไป.....	28
ภาพที่ 2. 16 (ก) แคปต์ชาข้อความต้นฉบับ (ข) แคปต์ชาข้อความที่โปรแกรมอัตโนมัติไม่สามารถ ระบุค่าระยะห่างระหว่างอักขรได้อย่างถูกต้อง.....	29
ภาพที่ 2. 17 (ก) แคปต์ชาข้อความต้นฉบับ (ข) แคปต์ชาข้อความที่โปรแกรมอัตโนมัติไม่สามารถ เพิ่มเส้นของตัวอักษรได้อย่างถูกต้อง.....	29
ภาพที่ 2. 18 ตัวอย่างภาพอักขร (Microsoft Corporation, 2014; วงศกร รัตติรุจิเสวต, 2557)...	30
ภาพที่ 2. 19 ตัวอย่างแคปต์ชาข้อความที่ได้ปรับระยะห่างระหว่างอักขรจนชิดกัน.....	31
ภาพที่ 2. 20 ตัวอย่างแคปต์ชาข้อความที่ได้ปรับระยะห่างระหว่างอักขรจนทับซ้อน.....	31
ภาพที่ 2. 21 ตัวอย่างแคปต์ชาข้อความที่ใช้ภาพพื้นหลังที่ซับซ้อน	32
ภาพที่ 2. 22 ตัวอย่างแคปต์ชาข้อความที่เพิ่มเส้นรบกวน.....	32

ภาพที่ 2. 23 ตัวอย่างแคปทซาข้อความที่มีตัวอักษรทับซ้อนกัน	33
ภาพที่ 2. 24 ตัวอย่างรูปสามเหลี่ยมที่ไม่มีเส้นเชื่อมต่อจนสมบูรณ์ แต่มนุษย์สามารถรับรู้ได้จาก รูปร่างที่คล้ายกับรูปสามเหลี่ยม (Nuttanont Hongwarittorn & Suttikiat Meelap, 2015).....	35
ภาพที่ 3. 1 ตัวอย่างแคปทซาข้อความที่มีระยะห่างระหว่างอักขรเป็นค่าปกติ.....	42
ภาพที่ 3. 2 ตัวอย่างแคปทซาข้อความที่มีระยะห่างระหว่างอักขรเป็นค่า 0 พอยต์.....	43
ภาพที่ 3. 3 ตัวอย่างแคปทซาข้อความที่มีระยะห่างระหว่างอักขรเป็นค่าลบ 3.75 พอยต์	43
ภาพที่ 3. 4 ตัวอย่างการกำหนดความหนาของเส้นรบกวนเป็น 1 พิกเซล ด้วยโปรแกรม Adobe Photoshop CS3.....	44
ภาพที่ 3. 5 ตัวอย่างเส้นรบกวนที่มีความหนา 1 พิกเซล พาดข้ามอักขร bwexzj.....	44
ภาพที่ 3. 6 ตัวอย่างเส้นรบกวนที่มีความหนา 3 พิกเซล พาดข้ามอักขร wycblz	44
ภาพที่ 3. 7 ตัวอย่างเส้นรบกวนที่มีความหนา 5 พิกเซล พาดข้ามอักขร mhxfnz.....	44
ภาพที่ 3. 8 ตัวอย่างการกำหนดสีของเส้นรบกวนให้เป็นสีดำด้วยโปรแกรม Adobe Photoshop CS3.....	45
ภาพที่ 3. 9 ตัวอย่างเส้นรบกวนขนาด 1 พิกเซล สีดำพาดข้ามอักขร bwexzj.....	45
ภาพที่ 3. 10 ตัวอย่างการกำหนดสีของเส้นรบกวนให้เป็นสีขาวด้วยโปรแกรม Adobe Photoshop CS3.....	46
ภาพที่ 3. 11 ตัวอย่างเส้นรบกวนขนาด 3 พิกเซล สีขาวพาดข้ามอักขร rphbnx	46
ภาพที่ 3. 12 ส่วนของเว็บไซต์สำหรับตรวจสอบการเข้าสู่ระบบ.....	49
ภาพที่ 3. 13 ส่วนของเว็บไซต์สำหรับตอบคำถามข้อมูลส่วนบุคคล	50
ภาพที่ 3. 14 แบบทดสอบตาบอดสีจากเว็บไซต์ enchroma.com	50
ภาพที่ 3. 15 กรอบความคิดของงานวิจัยนี้.....	53
ภาพที่ ข. 1 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการยืนยันความเป็น มนุษย์ เมื่อชุดอักขรแคปทซาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ และขีด.....	110

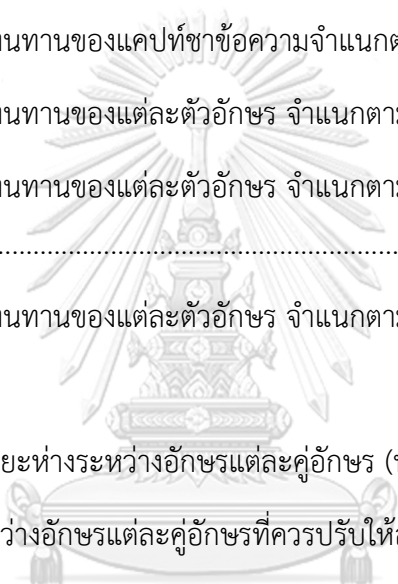
ภาพที่ ข. 13 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของ แคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักษรเป็นขีดและทับซ้อน แบบทางเดียว	116
ภาพที่ ข. 14 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของ แคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 3 พิกเซล	117
ภาพที่ ข. 15 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของ แคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 5 พิกเซล	117
ภาพที่ ข. 16 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของ แคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล แบบสองทาง.....	118
ภาพที่ ข. 17 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของ แคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล แบบทางเดียว	118
ภาพที่ ข. 18 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของ แคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ ชาข้อความ และสีเดียวกับพื้นหลังของแคปต์ชาข้อความ	119
 CHULALONGKORN UNIVERSITY	
ภาพที่ จ. 1 สร้าง SDK ใหม่สำหรับแคปต์ชาข้อความแต่ละลักษณะ.....	185
ภาพที่ จ. 2 เลือกโพลเดอรรูปแคปต์ชาข้อความเพื่อบรรจุรูปแคปต์ชาข้อความลงในโปรแกรม.....	185
ภาพที่ จ. 3 ตรวจสอบตัวอักษรที่เป็นไปได้สำหรับคำตอบของแคปต์ชาข้อความ	186
ภาพที่ จ. 4 วิเคราะห์ Filter ที่เหมาะสมที่สุดสำหรับรูปแบบของแคปต์ชาข้อความนั้น	186
ภาพที่ จ. 5 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 1	187
ภาพที่ จ. 6 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 2	187
ภาพที่ จ. 7 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 3	188
ภาพที่ จ. 8 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 4	188

ภาพที่ จ. 33 Filter ที่เหมาะสมสำหรับแคปซูลชาข้อความรูปแบบที่ 16 (3).....	201
ภาพที่ จ. 34 Filter ที่เหมาะสมสำหรับแคปซูลชาข้อความรูปแบบที่ 17 (1).....	201
ภาพที่ จ. 35 Filter ที่เหมาะสมสำหรับแคปซูลชาข้อความรูปแบบที่ 17 (2).....	202
ภาพที่ จ. 36 Filter ที่เหมาะสมสำหรับแคปซูลชาข้อความรูปแบบที่ 17 (3).....	202
ภาพที่ จ. 37 Filter ที่เหมาะสมสำหรับแคปซูลชาข้อความรูปแบบที่ 18 (1).....	203
ภาพที่ จ. 38 Filter ที่เหมาะสมสำหรับแคปซูลชาข้อความรูปแบบที่ 18 (2).....	203
ภาพที่ จ. 39 Filter ที่เหมาะสมสำหรับแคปซูลชาข้อความรูปแบบที่ 18 (3).....	204



สารบัญตาราง

ตารางที่ 2. 1 ความหมายของส่วนประกอบต่าง ๆ ของภาพอักษร	30
ตารางที่ 3. 1 ค่าระยะห่างด้านซ้าย และระยะห่างด้านขวาของตัวอักษรพิมพ์เล็ก แบบอักษร Arial (มีหน่วยเป็นพอยต์ ทั้งนี้การคำนวณค่าระยะห่างจากหน่วย U เป็นหน่วย พอยต์ อยู่ในสูตรที่ 1.2 ของบทที่ 1).....	42
ตารางที่ 3. 2 รูปแบบแคปต์ซาข้อความที่พัฒนามีจำนวน 18 เงื่อนไข.....	46
ตารางที่ 4. 1 ข้อมูลส่วนบุคคลของหน่วยทดลองจำนวน 360 คน.....	57
ตารางที่ 4. 2 ข้อมูลปัญหาตาบอดสี จำแนกตามผลการทดสอบตาบอดสีจากแบบทดสอบ	58
ตารางที่ 4. 3 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามรูปแบบแคปต์ซาข้อความ โดยมีจำนวนผู้ตอบแคปต์ซาข้อความรูปแบบละ 20 คน.....	59
ตารางที่ 4. 4 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามระยะห่างระหว่างอักษร	61
ตารางที่ 4. 5 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามความหนาของเส้นรบกวน	62
ตารางที่ 4. 6 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามสีของเส้นรบกวน.....	62
ตารางที่ 4. 7 อัตราความทนทานของแคปต์ซาข้อความ จำแนกตามรูปแบบแคปต์ซาข้อความ โดยจำนวนครั้งที่ GSA ตอบแคปต์ซาข้อความรูปแบบละ 20 ครั้ง	67
ตารางที่ 4. 8 อัตราความทนทานของแคปต์ซาข้อความ จำแนกตามระยะห่างระหว่างอักษร	70
ตารางที่ 4. 9 อัตราความทนทานของแคปต์ซาข้อความ จำแนกตามความหนาของเส้นรบกวน	70
ตารางที่ 4. 10 อัตราความทนทานของแคปต์ซาข้อความ จำแนกตามสีของเส้นรบกวน.....	71
ตารางที่ 4. 11 สรุปผลการวิเคราะห์ระยะห่างระหว่างอักษรของแคปต์ซาข้อความ	76
ตารางที่ 4. 12 สรุปผลการวิเคราะห์ความหนาของเส้นรบกวนที่พาดข้ามชุดอักษรแคปต์ซาข้อความ	77
ตารางที่ 4. 13 สรุปผลการวิเคราะห์สีของเส้นรบกวนที่พาดข้ามชุดอักษรแคปต์ซาข้อความ.....	79

ตารางที่ 4. 14 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามข้อมูลส่วนบุคคลของหน่วยทดลอง.....	79
ตารางที่ 4. 15 อัตราการยืนยันความเป็นมนุษย์จำแนกตามตัวอักษร	81
ตารางที่ 4. 16 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามระยะห่างระหว่าง อักษร	83
ตารางที่ 4. 17 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามขนาดความหนา ของเส้นรบกวน (พิกเซล).....	84
ตารางที่ 4. 18 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามสีของเส้นรบกวน ...	86
ตารางที่ 4. 19 อัตราความทนทานของแคปทาชข้อความจำแนกตามตัวอักษร	88
ตารางที่ 4. 20 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามระยะห่างระหว่างอักษร	89
ตารางที่ 4. 21 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามขนาดความหนาของเส้น รบกวน (พิกเซล).....	91
ตารางที่ 4. 22 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามสีของเส้นรบกวน	93
ตารางที่ ค. 1 ผลรวมของระยะห่างระหว่างอักษรแต่ละคู่อักษร (หน่วย Ps)	120
ตารางที่ ค. 2 ระยะห่างระหว่างอักษรแต่ละคู่อักษรที่ควรปรับให้ลดลงเพื่อให้อึดติดกันพอดี (หน่วย Ps).....	122
 จุฬาลงกรณ์มหาวิทยาลัย CHULALONGKORN UNIVERSITY	
ตารางที่ ง. 1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และ บางคู่อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง.....	124
ตารางที่ ง. 2 อัตราความทนทานของแคปทาชข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษร ซ้ำ และบางคู่อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง	154

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันอินเทอร์เน็ตได้เข้ามาเป็นส่วนสำคัญในการสื่อสารแลกเปลี่ยนข้อมูลผ่านเว็บไซต์ต่างๆ เว็บไซต์ส่วนหนึ่งต้องการให้ผู้ใช้สมัครเป็นสมาชิก สำหรับทำกิจกรรมต่างๆบนเว็บไซต์ได้ตามที่เจ้าของเว็บไซต์กำหนดไว้ อินเทอร์เน็ตจึงเป็นเทคโนโลยีที่ช่วยอำนวยความสะดวกให้กับผู้คนเป็นจำนวนมาก แต่ในทางกลับกันเป็นโอกาสให้กับผู้ไม่ประสงค์ดีที่พยายามก่อความเสียหายให้กับผู้อื่น หรือสร้างความเสียหายให้กับเจ้าของเว็บไซต์เช่นกัน ผู้ไม่ประสงค์ดีพยายามพัฒนาโปรแกรมคอมพิวเตอร์อัตโนมัติเพื่อส่งข้อความที่ผู้ใช้ไม่ได้ร้องขอเป็นจำนวนมาก ก่อให้เกิดความรำคาญต่อผู้ใช้ เนื่องจากเนื้อหาในข้อความเหล่านั้นไม่ได้มีความเกี่ยวข้องกับผู้ใช้ (Tangmanee C. & Sujarit – Apirak P., 2008)

แคปทชา (CAPTCHA) ย่อมาจาก Completely Automated Public Turing Test to Tell Computers and Humans Apart คือ โปรแกรมทดสอบแบบโต้ตอบชนิดหนึ่ง เพื่อทดสอบว่าผู้ใช้งานที่กำลังขอรับบริการเป็นมนุษย์จริงหรือไม่ (Luis von Ahn, Manuel Blum, & John Langford, 2004) การทดสอบนั้นสามารถจำแนกความเป็นมนุษย์ได้จากการตอบแบบทดสอบอย่างถูกต้อง โดยที่โปรแกรมคอมพิวเตอร์อัตโนมัติไม่สามารถตอบแบบทดสอบได้ถูกต้อง (Shirali-Shahreza M. H. & Shirali-Shahreza M., 2007) ในปัจจุบันแคปทชานั้นเป็นเทคโนโลยีมาตรฐานที่ถูกนำมาใช้อย่างแพร่หลายสำหรับป้องกันการโจมตีจากโปรแกรมคอมพิวเตอร์ที่ไม่ประสงค์ดี (Yan J. & El Ahmad A. S., 2008)

แคปทชาสามารถจำแนกได้เป็นสองกลุ่มหลักๆคือ (1) กลุ่มแคปทชาแบบรู้จำอักษรทางภาพ (OCR-based CAPTCHA) แคปทชากลุ่มนี้ได้แก่ แคปทชาข้อความ (Text-based CAPTCHA) โดยจะแสดงชุดอักษรที่มีลักษณะบิดเบือน และกำหนดให้ผู้ใช้ต้องพิมพ์ชุดอักษรที่แสดงอยู่ให้ถูกต้องเพื่อยืนยันความเป็นมนุษย์ และ (2) กลุ่มแคปทชาแบบไม่รู้จำอักษรทางภาพ (Non-OCR based CAPTCHA) แคปทชากลุ่มนี้ได้แก่ แคปทชารูปภาพ (Image-based CAPTCHA) โดยจะแสดงรูปภาพและกำหนดให้ผู้ใช้ทำตามคำสั่งให้ถูกต้องเพื่อยืนยันความเป็นมนุษย์ หรือแคปทชาเสียง (Audio-based CAPTCHA) โดยจะให้ผู้ใช้ฟังเสียง และกำหนดให้ผู้ใช้ต้องพิมพ์ชุดอักษรตามเสียงที่ได้ยินให้

ถูกต้องเพื่อยืนยันความเป็นมนุษย์ (Kurt Alfred Kluever & Richard Zanibbi, 2009; วงศกร รัตติ รุจิเสวต, 2557)

แคปต์ชาข้อความในรูปแบบของแคปต์ชาที่ได้รับความนิยมมากที่สุดเมื่อเทียบกับรูปแบบอื่น โดยถูกนำไปใช้อย่างแพร่หลายเพื่อทดสอบความสามารถในการรู้จำอักษรของมนุษย์ และการออกแบบแคปต์ชาข้อความให้มีความทนทานต่อการโจมตีจากโปรแกรมคอมพิวเตอร์ควรให้ความสำคัญกับการป้องกันการโจมตีด้านการแบ่งส่วน (Segmentation) เนื่องจากในขั้นตอนการรู้จำ (Recognition) นั้นโปรแกรมคอมพิวเตอร์สามารถรู้จำอักษรได้ถูกต้องแม่นยำในอัตราที่สูงมาก (Yan J. & El Ahmad A. S., 2008)

กระบวนการวิเคราะห์แคปต์ชาของโปรแกรมอัตโนมัติประกอบด้วยขั้นตอนหลักๆห้าขั้นตอน คือ (1) ขั้นตอนการประมวลผล (Pre-processing) ขั้นตอนนี้จะนำภาพพื้นหลัง หรือรูปภาพที่อื่นๆที่ไม่ใช่ชุดอักษรและสามารถลบส่วนออกไป (2) ขั้นตอนการแบ่งส่วน (Segmentation) ขั้นตอนนี้จะจำแนกชุดอักษรออกเป็นแต่ละตัวอักษร เป็นการประมวลผลว่าในแคปต์ชาข้อความนี้มีกี่ตัวอักษร (3) ขั้นตอนหลังจากการแบ่งส่วน (Post-segmentation) ขั้นตอนนี้จะปรับขนาดของรูปภาพอักษรที่ได้จากขั้นตอนการแบ่งส่วนให้มีขนาดปกติ (4) ขั้นตอนการรู้จำ (Recognition) ขั้นตอนนี้จะระบุว่ารูปภาพอักษรนั้นเป็นตัวอักษรใด และ (5) ขั้นตอนหลังจากการประมวลผล (Post-processing) ขั้นตอนนี้จะตรวจสอบความถูกต้องเพิ่มเติมหลังจากการระบุตัวอักษรในขั้นตอนรู้จำ เช่น การตรวจสอบการสะกดคำสำหรับแคปต์ชาที่ใช้คำที่มีความหมาย เป็นต้น (Bursztein E., Martin M., & Mitchell J. C., 2011)

แคปต์ชาข้อความควรง่ายต่อผู้ใช้งานเพื่อให้ผู้ใช้งานตอบได้ถูกต้อง อัตราการยืนยันความเป็นมนุษย์สามารถวัดจากสัดส่วนของจำนวนผู้ตอบแคปต์ชาข้อความได้ถูกต้องต่อจำนวนผู้ตอบแคปต์ชาข้อความทั้งหมด (Yan J. & El Ahmad A. S., 2008; มณีรัตน์ ชาตธีรังสรรค์, 2555) และแคปต์ชาข้อความต้องทนทานต่อการวิเคราะห์ของโปรแกรมอัตโนมัติเพื่อป้องกันการสร้างความเสียหายจากการประสงค์ร้าย อัตราความทนทานของแคปต์ชาข้อความสามารถวัดจากจำนวนครั้งการโจมตีแคปต์ชาข้อความของโปรแกรมอัตโนมัติที่ไม่ถูกต้องต่อจำนวนการโจมตีทั้งหมด (El Ahmand A. S., Yan J., & Tayara M., 2011) การออกแบบแคปต์ชาข้อความให้มีอัตราความทนทานการโจมตีจากโปรแกรมอัตโนมัติแบ่งออกเป็นสองประเด็น ได้แก่ (1) ป้องกันการโจมตีด้านการรู้จำ (anti-recognition) คือการป้องกันไม่ให้โปรแกรมอัตโนมัติสามารถระบุได้ว่ารูปภาพอักษรเป็นตัวอักษรใด และ (2) ป้องกันการโจมตีด้านการแบ่งส่วน (anti-segmentation) คือการป้องกันไม่ให้แคปต์ชานั้นถูกจำแนกออกเป็นอักษรแต่ละตัว ซึ่งทำให้โปรแกรมอัตโนมัติจะทราบถึงจำนวนตัวอักษรในชุดอักษรนั้น (Bursztein E. et al., 2011)

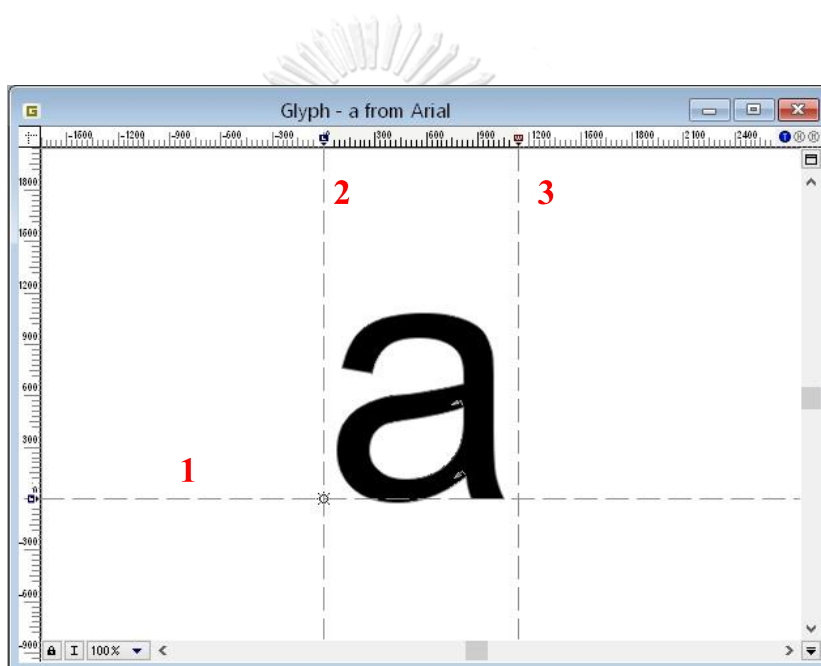
การออกแบบแคปซูลข้อความให้มีอัตราความทนทานสามารถป้องกันการโจมตีด้านการรู้จำ จะมุ่งเน้นไม่ให้โปรแกรมอัตโนมัติสามารถระบุได้ว่ารูปภาพอักษรของแคปซูลข้อความเป็นตัวอักษรใด (Bursztein E. et al., 2011) ได้ศึกษาคุณสมบัติป้องกันการโจมตีด้านการรู้จำด้วยการกำหนดแนวปฏิบัติว่าขนาดของตัวอักษรให้มีขนาดเล็ก ไม่ควรเป็นอักษรพิมพ์ใหญ่ ใช้การหมุนเอียงตัวอักษร และเลือกใช้แบบอักษรที่หลากหลาย และคุณสมบัติป้องกันการโจมตีด้านการแบ่งส่วนด้วยการเพิ่มเส้นรบกวนเข้าไปในชุดอักษรแคปซูลข้อความ (ตัวอย่างของเส้นรบกวนเป็นตามเส้นตรงทึบในภาพที่ 1.3) เพื่อให้โปรแกรมอัตโนมัติไม่สามารถจำแนกได้ว่าส่วนใดเป็นส่วนของอักษรรอยางชัดเจน และการลดระยะห่างระหว่างอักษรของชุดอักษรแคปซูลข้อความ

การวิจัยนี้ได้เลือกศึกษาแคปซูลข้อความ เนื่องจากแคปซูลข้อความนี้เป็นประเภทของแคปซูลข้อความที่ถูกนำมาใช้งานมากที่สุดและสามารถสร้างแบบทดสอบได้หลายรูปแบบ (Kolupave A. & Ogijenko J., 2008) โดยจะมุ่งเน้นที่การออกแบบแคปซูลข้อความเพื่อป้องกันการโจมตีด้านการแบ่งส่วน (anti-segmentation) จากโปรแกรมอัตโนมัติ เนื่องจากงานวิจัยในอดีตส่วนใหญ่ได้ศึกษาคุณลักษณะของแคปซูลข้อความเพื่อป้องกันการรู้จำ (anti-recognition) เช่น มณีรัตน์ ชาตธีรสรณ์ (2555) ได้ศึกษาคุณลักษณะด้านการหมุนเอียง คุณลักษณะด้านแบบอักษร และคุณลักษณะด้านจำนวนอักษร และพบว่าคุณลักษณะด้านการหมุนเอียงเท่านั้นที่สามารถป้องกันการโจมตีด้านการแบ่งส่วน ในขณะที่วงศกร รัตติรุจิเศวต (2557) ได้ศึกษาคุณลักษณะด้านระยะห่างระหว่างอักษร คุณลักษณะด้านตำแหน่งอักษร และ คุณลักษณะด้านจำนวนอักษร โดยคุณลักษณะที่ป้องกันการโจมตีด้านการแบ่งส่วน ได้แก่ คุณลักษณะด้านระยะห่างระหว่างอักษรและตำแหน่งอักษร เพื่อเพิ่มอัตราความทนทานของแคปซูลข้อความด้านการแบ่งส่วน และการออกแบบแคปซูลข้อความให้มีความทนทานต่อการโจมตีจากโปรแกรมคอมพิวเตอร์ควรให้ความสำคัญกับการป้องกันการโจมตีด้านการแบ่งส่วน (Segmentation) เนื่องจากในขั้นตอนการรู้จำ (Recognition) นั้นโปรแกรมคอมพิวเตอร์สามารถรู้จำอักษรได้ถูกต้องแม่นยำในอัตราที่สูงมาก (Yan J. & El Ahmad A. S., 2008)

นักวิจัยจำนวนหนึ่งได้สนใจตัวแปรที่มีผลการโจมตีแบ่งส่วน แต่อย่างน้อยเมื่อเทียบกับความสนใจในตัวแปรที่มีผลต่อการโจมตีด้านการรู้จำ (Bursztein E. et al., 2014) ดังนั้นจึงเป็นโอกาสให้วิเคราะห์ตัวแปรที่ส่งผลต่อการโจมตีการแบ่งส่วน โดยตัวแปรที่อยู่ในความสนใจของวิทยานิพนธ์ฉบับนี้ ประกอบด้วย (1) ระยะห่างระหว่างอักษร (2) ความหนาของเส้นรบกวน และ (3) สีของเส้นรบกวน

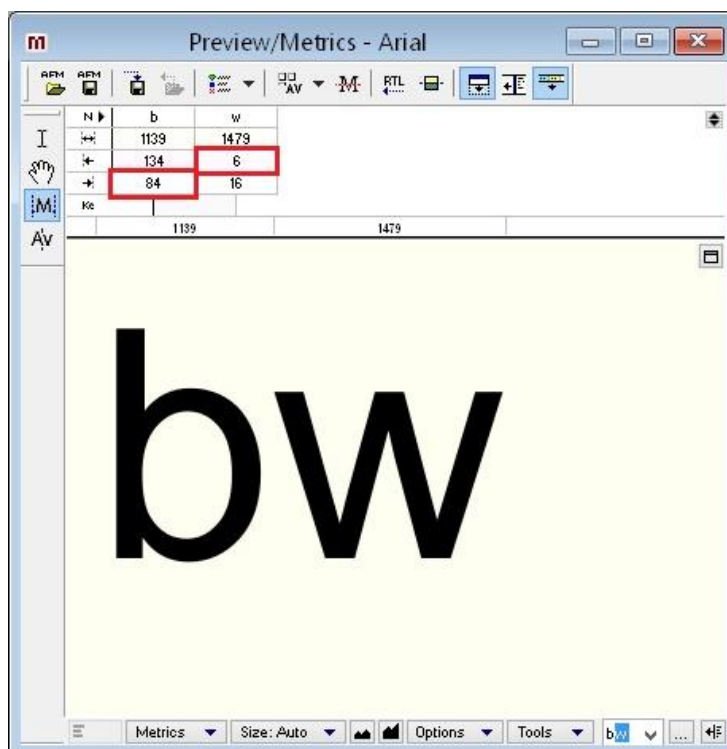
ระยะห่างระหว่างอักษร (Spacing) หมายถึง ระยะห่างด้านขวาของภาพอักษรแรกรวมกับระยะห่างด้านซ้ายของภาพตัวอักษรหลัง (วงศกร รัตติรุจิเศวต, 2557) โดยภาพอักษร (Glyph) หมายถึงภาพของตัวพิมพ์ (Type) ที่ใช้แสดงผลบนหน้าจอ (Microsoft Corporation, 2014; วงศกร

รัตติรุจิเสวต, 2557) ภาพที่ 1.1 คือตัวอย่างภาพอักษร a ในแบบอักษร Arial ที่ได้จากโปรแกรม FontLab Studio เส้นประที่ใกล้หมายเลข 1 เรียกว่าเส้นบรรทัด (Baseline) เส้นประที่ใกล้หมายเลข 2 เรียกว่าเส้นด้านซ้าย และเส้นประที่ใกล้หมายเลข 3 เรียกว่าเส้นด้านขวา ภาพอักษรแต่ละตัวจะมี ค่าความกว้างของภาพอักษรคือค่าที่ได้จากการวัดเส้นด้านซ้ายจนถึงเส้นด้านขวาของภาพอักษร ค่าระยะห่างด้านซ้ายคือค่าที่ได้จากการวัดจากเส้นด้านซ้ายไปจนถึงด้านซ้ายของตัวอักษร และค่าระยะห่างด้านขวาคือค่าที่ได้จากการวัดจากด้านขวาของตัวอักษรไปจนถึงเส้นด้านขวา ดังนั้นการลดระยะห่างระหว่างอักษรคือการลดระยะห่างด้านขวาของภาพอักษรแรกและระยะห่างด้านซ้ายของภาพอักษรหลัง



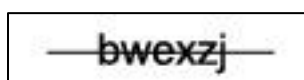
ภาพที่ 1.1 ตัวอย่างภาพอักษรพิมพ์เล็ก a ของแบบอักษร Arial

ภาพที่ 1.2 คือตัวอย่างการวัดค่าระยะห่างระหว่างอักษรของแบบอักษร Arial ด้วยโปรแกรม FontLab Studio โดยหน่วยในการวัด คือ U (Unit per EM) โดยกรอบสี่เหลี่ยมด้านมุมซ้ายบนแสดงค่าระยะห่างด้านขวาของอักษร b เท่ากับ 84 U และกรอบสี่เหลี่ยมใกล้เคียงกันแสดงค่าระยะห่างด้านซ้ายของอักษร w เท่ากับ 6 U



ภาพที่ 1. 2 ตัวอย่างการวัดค่าระยะห่างระหว่างอักษรพิมพ์เล็ก b กับ w ของแบบอักษร Arial

เพื่ออธิบายตัวแปรอิสระทั้งสามตัวให้เข้าใจตรงกัน ผู้วิจัยขอเริ่มจากความหมายของเส้นรอบวง เส้นรอบวง หมายถึง เส้นที่ถูกเพิ่มเข้าไปในลักษณะที่พาดข้ามชุดอักษรแคปซัลชาข้อความ (Bursztein E. et al., 2011) โดยมีวัตถุประสงค์เพื่อให้โปรแกรมอัตโนมัติเกิดความผิดพลาดในการจำแนกหรือแบ่งว่าส่วนใดเป็นส่วนใดเป็นเส้นของตัวอักษร เส้นรอบวงที่มีความหนามากขึ้นสามารถเพิ่มความทนทานของแคปซัลชาข้อความต่อโปรแกรมอัตโนมัติมากขึ้น แต่ในขณะเดียวกันก็ลดความสามารถในการยืนยันความเป็นมนุษย์เช่นกัน เมื่อเพิ่มเส้นรอบวงหนา 1-2 พิกเซล เข้าไปในลักษณะที่พาดข้ามชุดอักษรแคปซัลชาข้อความแล้ว มนุษย์ยังสามารถอ่านและตอบแคปซัลชาข้อความได้ถูกต้องโดยใช้เวลาในการตอบสั้น (Bursztein E. et al., 2014) โดยหลักการแล้วเส้นรอบวงไม่จำเป็นต้องเป็นเส้นตรง แต่การทบทวนวรรณกรรมในอดีตไม่พบการศึกษาอย่างเป็นระบบของเส้นรอบวงเลย ผู้วิจัยจึงเริ่มจากการใช้เส้นตรงเป็นเส้นรอบวง ทั้งนี้ภาพที่ 1.3 คือตัวอย่างเส้นรอบวงในลักษณะตรงพาดข้ามทุกตัวแคปซัลชาข้อความ



ภาพที่ 1. 3 ตัวอย่างเส้นรอบวงสีเดียวกับอักษรในแคปซัลชาข้อความ

ทั้งนี้ Bursztein et al. (2014) ได้วิจัยคุณลักษณะของเส้นรบกวนที่ควรใช้เพื่อเพิ่มความทนทานของแคปต์ซาข้อความ ได้แก่ (1) ประเภทของเส้นรบกวน (Line types) (2) จำนวนของเส้นรบกวน (Line counts) (3) ความหนาของเส้นรบกวน (Line widths) หรือ (4) ตำแหน่งของเส้นรบกวน (Line positions) และพบว่าเมื่อเพิ่มเส้นรบกวนประเภทเส้นตรงที่มีความหนา 1-2 พิกเซล จำนวน 1-3 เส้น เข้าไปในลักษณะพาดข้ามชุดอักษรแคปต์ซาข้อความแล้ว มนุษย์ยังสามารถเข้าใจได้ดีใกล้เคียงกับแคปต์ซาข้อความที่ไม่มีเส้นรบกวน แต่การเพิ่มความหนาของเส้นรบกวนมากขึ้นจะทำให้การยืนยันความเป็นมนุษย์ลดลง กล่าวคือความถูกต้องในการตอบตัวอักษรของมนุษย์จะลดลงและต้องใช้เวลาเพิ่มขึ้นในการอ่านเพื่อตอบให้ถูกต้อง นอกจากนี้การลดระยะห่างระหว่างอักษรของชุดอักษรแคปต์ซาข้อความในระยะห่างต่างๆ ได้พบว่าการลดระยะห่างระหว่างอักษรมากกว่า 5 พิกเซล จะทำให้การยืนยันความเป็นมนุษย์ลดลง นั่นหมายความว่าความถูกต้องในการตอบตัวอักษรของมนุษย์จะลดลง อย่างไรก็ตามยังไม่พบการวิเคราะห์อัตราความทนทานของแคปต์ซาข้อความต่อโปรแกรมอัตโนมัติ และยังไม่พบการพิสูจน์ว่าการเพิ่มเส้นรบกวนจะช่วยป้องกันการโจมตีด้านการแบ่งส่วนของแคปต์ซาข้อความ

เนื่องจากการลดระยะห่างระหว่างอักษรช่วยเพิ่มคุณสมบัติป้องกันการโจมตีการแบ่งส่วนของแคปต์ซาข้อความ ผู้วิจัยจึงได้เลือกวิเคราะห์การลดระยะห่างระหว่างอักษร (1) การลดระยะห่างระหว่างอักษรจนชิดกัน คือให้ระยะห่างด้านขวาของภาพอักษรแรกและระยะห่างด้านซ้ายของภาพอักษรหลังมีค่า 0 พอยต์ (ตัวอย่างการลดระยะห่างจนชิดกันในภาพที่ 1.4) และ (2) การลดระยะห่างระหว่างอักษรจนทับซ้อนกัน คือให้ระยะห่างด้านขวาของภาพอักษรแรกและระยะห่างด้านซ้ายของภาพอักษรหลังมีค่าลบ 3.75 พอยต์ (ตัวอย่างการลดระยะห่างเป็นทับซ้อนกันอยู่ในภาพที่ 1.4) การเลือกค่าลบ 3.75 พอยต์ เป็นเพราะระยะห่างระหว่างอักษรมีค่าลดลงน้อยกว่าลบ 3.75 พอยต์ หรือเท่ากับลบ 5 พิกเซล (1 พิกเซลเท่ากับ 0.75 พอยต์) จะทำให้ความสามารถในการยืนยันความเป็นมนุษย์ลดลง (Bursztein E. et al., 2011)

ผู้วิจัยเลือกกำหนดระยะห่างระหว่างอักษรด้วยโปรแกรม Adobe Photoshop CS3 มีหน่วยในการวัด คือ Ps (Photoshop) มีค่า 1/1000 em ดังนั้น สูตรความสัมพันธ์ระหว่างหน่วย Ps กับหน่วยพอยต์ ในงานวิจัยนี้เป็นตามสูตรที่ 1.1 ดังนี้

$$Ps = P \times \frac{1000}{24} \dots\dots\dots (1.1)$$

เมื่อ	Ps	คือ	ระยะห่างระหว่างอักขรมีหน่วยเป็น Ps
	P	คือ	ระยะห่างระหว่างอักขรมีหน่วยเป็นพอยต์
	FontSize	คือ	ขนาดของอักขรที่ใช้แสดงในแคปท์ซาข้อความ มีค่าคงที่คือ 24 พอยต์

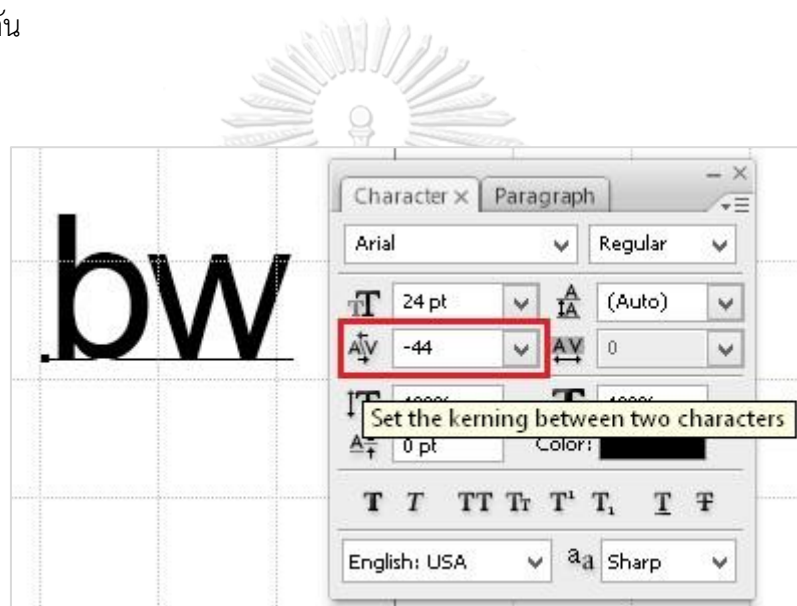
เนื่องจากการออกแบบภาพอักขรจะใช้หน่วย U กำหนดขนาด โดยหนึ่งภาพอักขรจะมีขนาดไม่เกิน 2048 U เรียกว่า UPM (Unit per EM) (Fontlab Ltd, 2006; วงศกร รัตติรุจิเสวต, 2557) ผู้วิจัยต้องแปลงหน่วย U เป็นหน่วยพอยต์ (Point) เพื่อนำไปใช้คำนวณหาค่าในหน่วย Ps สูตรความสัมพันธ์ระหว่างหน่วย U กับหน่วยพอยต์ ตามสูตรที่ 1.2 ดังนี้ (Microsoft Corporation, 2014; วงศกร รัตติรุจิเสวต, 2557)

$$P = U \times \frac{24}{2048} \dots\dots\dots (1.2)$$

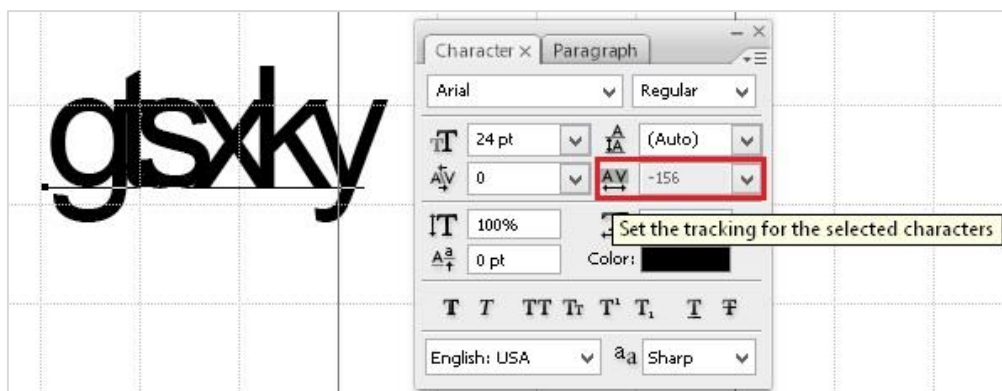
เมื่อ	P	คือ	ระยะห่างระหว่างอักขรมีหน่วยเป็นพอยต์
	U	คือ	ระยะห่างระหว่างอักขรมีหน่วยเป็น U
	FontSize	คือ	ขนาดของอักขรที่ใช้แสดงในแคปท์ซาข้อความ มีค่าคงที่คือ 24 พอยต์
	UPM	คือ	จำนวนหน่วย U สูงสุดของภาพอักขร มีค่าคงที่คือ 2048 U

จากภาพที่ 1.2 ระยะห่างด้านขวาของอักขร b คือ 84 U และระยะห่างด้านซ้ายของอักขร w คือ 6 U รวมระยะห่างระหว่างอักขร b กับ w เท่ากับ $84+6 = 90$ U เมื่อนำมาแปลงเป็นหน่วยพอยต์ โดยใช้สูตรที่ 1.2 จะได้ $90 \times 24 / 2048 = 1.0546875$ พอยต์ และจากหน่วยพอยต์ เมื่อนำมาแปลงเป็นหน่วย Ps โดยใช้สูตรที่ 1.1 จะได้ $1.0546875 \times 1000 / 24 = 43.9453125$ Ps หรือโดยประมาณคือ 44 Ps

เนื่องจากระยะห่างด้านขวาและระยะห่างด้านซ้ายของอักษรแต่ละตัวมีค่าไม่เท่ากัน การกำหนดระยะห่างระหว่างอักษรให้ชิดกัน จึงต้องกำหนดจากค่า kerning ของอักษร โดยที่ kerning คือการกำหนดช่องว่างระหว่างด้านหน้าของตัวอักษร โดยสามารถกำหนดเป็นค่าแตกต่างกันไปตามระยะห่างด้านขวาของอักษรตัวแรกกับระยะห่างด้านซ้ายของอักษรตัวหลังแต่ละคู่ ในกรอบสี่เหลี่ยมของภาพที่ 1.4 คือการกำหนด kerning ระหว่างอักษร b กับ w ให้เป็นค่าลบ 44 Ps เพื่อให้อักษรชิดกัน ตัวอย่างในภาพที่ 1.5 คือการกำหนดระยะห่างระหว่างอักษรให้ทับซ้อนกัน โดยข้อมูลในกรอบสี่เหลี่ยมของภาพ 1.5 คือการกำหนด tracking เป็นลระยะห่างจนมีค่าเท่ากับลบ 3.75 พอยต์ หรือลบ 156 Ps โดยที่ tracking คือการกำหนดระยะห่างระหว่างอักษรของทุกคู่อักษรในชุดอักษรนั้นให้เป็นค่าเดียวกัน

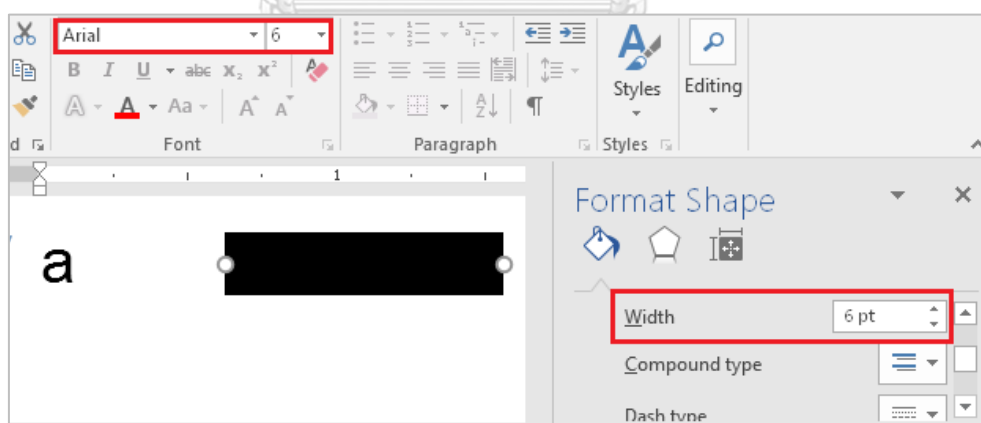


ภาพที่ 1.4 ตัวอย่างการกำหนดระยะห่างระหว่างอักษร b กับ w โดยกำหนดค่า kerning เท่ากับลบ 44 Ps



ภาพที่ 1. 5 ตัวอย่างการกำหนดระยะห่างระหว่างอักษร g t s x k และ y โดยกำหนดค่า tracking เท่ากับลบ 156 Ps

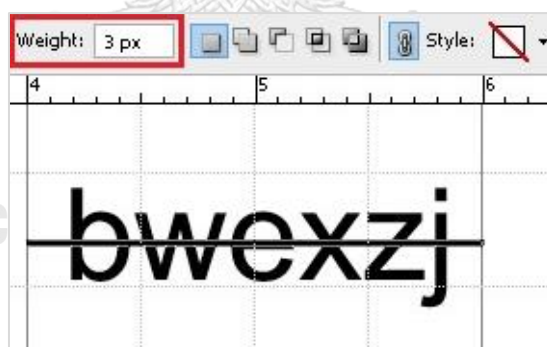
ความหนา (width) ของเส้นรบกวน เนื่องจากหน่วยพอยต์ในการวัดขนาดของตัวอักษร (font-size) กับหน่วยพอยต์ในการวัดความหนาของเส้นรบกวนนั้นไม่สามารถนำมาเปรียบเทียบกันได้ ดังตัวอย่างในภาพที่ 1.6 โดยในกรอบสี่เหลี่ยมแสดงการกำหนดค่าตัวอักษรขนาด 6 พอยต์ และความหนาของเส้น 6 พอยต์



ภาพที่ 1. 6 แสดงตัวอย่างตัวอักษร a และเส้นรบกวนที่มีขนาด 6 พอยต์เท่ากัน

ในการวิจัยนี้ได้กำหนดขนาด (weight) ด้วยโปรแกรม Adobe Photoshop CS3 ดังตัวอย่างภาพที่ 1.7 ในกรอบสี่เหลี่ยม คือการกำหนดขนาดของเส้นเท่ากับ 3 พิกเซล และงานวิจัยนี้เลือกกำหนดค่าความหนาของเส้นสามขนาด ได้แก่ 1, 3, และ 5 พิกเซล เนื่องจากเมื่อเพิ่มความหนาของ

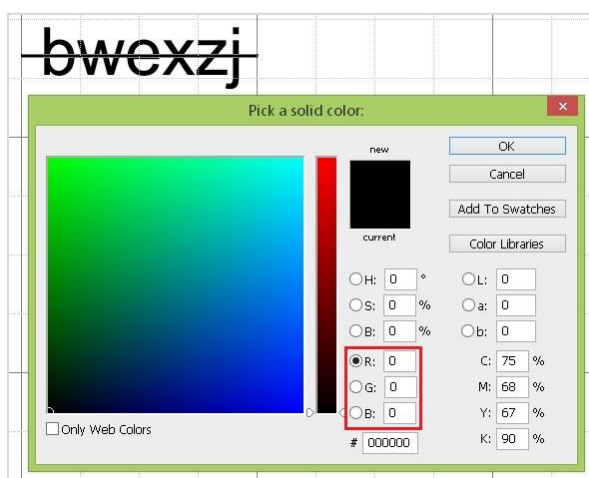
เส้นรบกวน 1-2 พิกเซล เข้าไปในชุดอักษรแคปซูลข้อความแล้วมนุษย์ยังสามารถอ่านและตอบแคปซูลข้อความได้ถูกต้องประมาณ 90% หรือเทียบเท่ากับแคปซูลข้อความที่ไม่มีเส้นรบกวนพาดข้าม แต่ความหนาที่เพิ่มขึ้นของเส้นรบกวนจะลดความสามารถในการตอบแคปซูลข้อความของมนุษย์ โดยเมื่อเพิ่มความหนาของเส้นรบกวน 3 พิกเซล เข้าไปในชุดอักษรแคปซูลข้อความแล้วมนุษย์ยังสามารถอ่านและตอบแคปซูลข้อความได้ถูกต้องประมาณ 80% และเมื่อเพิ่มความหนาของเส้นรบกวน 5 พิกเซล เข้าไปในชุดอักษรแคปซูลข้อความแล้วมนุษย์ยังสามารถอ่านและตอบแคปซูลข้อความได้ถูกต้องประมาณ 55% แต่เมื่อเพิ่มความหนาของเส้นรบกวน 6 พิกเซล เข้าไปในชุดอักษรแคปซูลข้อความแล้วมนุษย์ยังสามารถอ่านและตอบแคปซูลข้อความได้ถูกต้องประมาณ 34% เท่านั้น (Bursztein E. et al., 2014) นอกจากนี้ Yan & El Ahmad (2008) ได้อธิบายกระบวนการแบ่งส่วนของโปรแกรมอัตโนมัติ โดยในขั้นตอนการวิเคราะห์เพื่อนำสิ่งรบกวนออกจากตัวอักษรแคปซูลข้อความ โปรแกรมมักมีข้อผิดพลาดในการนำเส้นรบกวนออกเมื่อเส้นรบกวนมีความหนามาก (thick) ทั้งนี้ผู้วิจัยต้องการวิเคราะห์คุณลักษณะของเส้นรบกวนเพื่อเพิ่มอัตราความทนทานของแคปซูลข้อความต่อโปรแกรมอัตโนมัติให้ได้มากที่สุดโดยที่มนุษย์ยังสามารถอ่านและตอบได้อย่างถูกต้องจึงเลือกศึกษาความหนาของเส้นรบกวนที่มีความหนา 1, 3, และ 5 พิกเซล ที่เพิ่มเข้าไปในลักษณะพาดข้ามกึ่งกลางของตัวอักษรแคปซูลข้อความ



ภาพที่ 1. 7 ตัวอย่างการกำหนดความหนาของเส้นรบกวนเป็น 3 พิกเซล ด้วยโปรแกรม Adobe Photoshop CS3

อย่างไรก็ตามผู้วิจัยยังไม่พบผลการวิเคราะห์เชิงประจักษ์ (empirical) เกี่ยวกับผลของสีของเส้นรบกวนที่มีต่อความทนทานของแคปซูลข้อความ ทั้งที่สีของเส้นรบกวนเป็นตัวแปรสำคัญในการออกแบบแคปซูลข้อความ การวิจัยนี้จึงเลือกศึกษาสีของเส้นรบกวนในสองลักษณะ ได้แก่ (1) สีเดียวกับอักษร (ดูตัวอย่างภาพที่ 1.15) โดยเมื่อเพิ่มเส้นรบกวนที่มีสีเดียวกับอักษรเข้าไปในแคปซูลข้อความแล้วอาจทำให้โปรแกรมอัตโนมัติไม่สามารถตัดสินใจว่าส่วนของเส้นรบกวนนั้นเป็นส่วนหนึ่งของ

ตัวอักษร และ (2) สีเดียวกับพื้นหลัง (ดูตัวอย่างภาพที่ 1.16) โดยเมื่อเพิ่มเส้นรบกวนที่มีสีเดียวกับพื้นหลังเข้าไปในแคปซูลข้อความแล้วจะทำให้เส้นของตัวอักษรของแคปซูลข้อความขาดหายเป็นบางส่วนตามแนวเส้นรบกวนนั้น โปรแกรมอัตโนมัติอาจไม่สามารถวิเคราะห์รูปร่างของตัวอักษรได้ ดังนั้นงานวิจัยนี้ได้กำหนดสีของเส้นรบกวนด้วยโปรแกรม Adobe Photoshop CS3 ดังกรอบสี่เหลี่ยมในตัวอย่างภาพที่ 1.8 ทั้งนี้รายละเอียดการกำหนดสีของอักษรและสีพื้นหลังได้แสดงในบทที่ 3



ภาพที่ 1. 8 ตัวอย่างการกำหนดสีของเส้นรบกวนเป็นสีเดียวกับอักษรด้วยโปรแกรม Adobe Photoshop CS3

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปซูลข้อความ ระหว่างแคปซูลข้อความที่กำหนดระยะห่างระหว่างอักษรสามแบบคือ (1) ปกติ (2) ซิด (0 พอยต์) และ (3) ทับซ้อน (ลบ 3.75 พอยต์)
2. เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปซูลข้อความ ระหว่างแคปซูลข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนาสามแบบคือ (1) 1 พิกเซล (2) 3 พิกเซล และ (3) 5 พิกเซล
3. เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปซูลข้อความ ระหว่างแคปซูลข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีสีสองแบบคือ (1) สีเดียวกับอักษร และ (2) สีเดียวกับพื้นหลัง

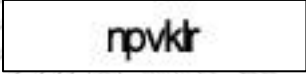
1.3 ตัวแปรสำคัญที่ศึกษา

1. ตัวแปรต้นหรือตัวแปรอิสระ (Independent Variables) มีสามตัวแปร ดังนี้

1. ระยะห่างระหว่างอักษร คือ ระยะห่างระหว่างอักษรที่อยู่ต่อเนื่องกัน (Bursztein E. et al., 2014) วัดจากระยะห่างด้านขวาของภาพตัวอักษรแรกรวมกับระยะห่างด้านซ้ายของภาพตัวอักษรหลัง (วงศกร รัตติรุจิเศวต, 2557) โดยจำแนกค่าของตัวแปรได้สามค่าคือ (1) ปกติ (2) ชิด (0 พอยต์) และ (3) ทับซ้อน (ลบ 3.75 พอยต์) ตามที่แสดงในภาพที่ 1.9, 1.10, และ 1.11 ตามลำดับ



ภาพที่ 1. 9 ตัวอย่างแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าปกติ



ภาพที่ 1. 10 ตัวอย่างแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าชิด

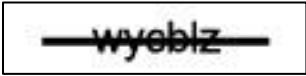


ภาพที่ 1. 11 ตัวอย่างแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าทับซ้อน

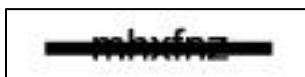
2. ความหนาของเส้นรบกวน คือ ความหนาของเส้นที่เพิ่มเข้าไปในลักษณะที่พาดข้ามชุดอักษรแคปต์ชาข้อความทั้งหมด โดยผู้วิจัยได้กำหนดความหนาของเส้นด้วยโปรแกรม Adobe Photoshop CS3 โดยจำแนกค่าของตัวแปรได้สามค่าคือ (1) 1 พิกเซล (2) 3 พิกเซล และ (3) 5 พิกเซล ตามที่แสดงในภาพที่ 1.12, 1.13, และ 1.14 ตามลำดับ



ภาพที่ 1. 12 ตัวอย่างแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล

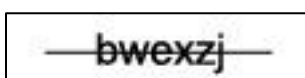


ภาพที่ 1. 13 ตัวอย่างแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล

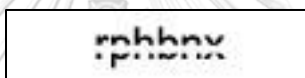


ภาพที่ 1. 14 ตัวอย่างแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

3. สีของเส้นรบกวน คือ สีของเส้นรบกวนที่พาดข้ามชุดอักษรแคปต์ชาข้อความ กำหนดสีของเส้นด้วยโปรแกรม Adobe Photoshop CS3 โดยจำแนกค่าของตัวแปรได้สองค่าคือ (1) สีเดียวกับอักษร และ (2) สีเดียวกับพื้นหลัง ตามที่แสดงในภาพที่ 1.15 และ 1.16 ตามลำดับ



ภาพที่ 1. 15 ตัวอย่างแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับอักษร และมีขนาด 1 พิกเซล



ภาพที่ 1. 16 ตัวอย่างแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับพื้นหลัง และมีขนาด 3 พิกเซล

2. ตัวแปรตาม (Dependent Variables) มีสองตัวแปร ดังนี้

1. อัตราการยืนยันความเป็นมนุษย์ (Human Affirmative Rate) หมายถึง จำนวนครั้งที่ผู้ใช้สามารถระบุอักษรที่ปรากฏในแคปต์ชาข้อความได้ถูกต้องจนสามารถเข้าใช้บริการของเว็บไซต์ได้ โดยวัดจากสัดส่วนของจำนวนผู้ตอบแคปต์ชาข้อความได้ถูกต้องต่อจำนวนผู้ตอบแคปต์ชาข้อความทั้งหมด (Yan J. & El Ahmad A. S., 2008; มณีรัตน์ ชาตธีรังสรรค์, 2555)

2. อัตราความทนทานของแคปต์ชาข้อความ (Robustness Rate of CAPTCHA) หมายถึง ความสามารถของแคปต์ชาข้อความในการป้องกันการโจมตีของโปรแกรมโอซีอาร์ได้ โดยวัดจากอัตราการตอบแคปต์ชาข้อความที่ไม่ถูกต้องของโปรแกรมโอซีอาร์ต่อจำนวนการตอบแคปต์ชาข้อความทั้งหมด (Yan J. & El Ahmad A. S., 2008; มณีรัตน์ ชาตธีรังสรรค์, 2555)

1.4 คำจำกัดความของการวิจัย

1. แคปTCHAข้อความ (Text-Based CAPTCHA) หมายถึง แคปTCHAซึ่งแสดงในรูปแบบของชุดอักษร โดยแคปTCHAเป็นโปรแกรมทดสอบเพื่อเข้าสู่ระบบคอมพิวเตอร์แบบโต้ตอบชนิดหนึ่ง เพื่อทดสอบว่าผู้ใช้งานเป็นมนุษย์จริงหรือไม่ (Luis von Ahn et al., 2004)
2. เส้นรบกวน (Noise) หมายถึง เส้นที่ถูกเพิ่มเข้าไปในลักษณะที่พาดข้ามชุดอักษรแคปTCHA ข้อความในลักษณะแฉกแฉก เพื่อเพิ่มความสามารถของแคปTCHAข้อความในการป้องกันการโจมตีด้านการแบ่งส่วนจากโปรแกรมอัตโนมัติ (Bursztein E. et al., 2011)
3. การป้องกันการโจมตีด้านการแบ่งส่วน (Anti-segmentation) หมายถึง การป้องกันแคปTCHAข้อความจากการโจมตีของโปรแกรมโอซีอาร์ด้านการจำแนกชุดอักษรของแคปTCHAข้อความ ออกเป็นแต่ละตัวอักษร (Bursztein E. et al., 2014)
4. การป้องกันการโจมตีด้านการรู้จำ (Anti-recognition) หมายถึง การป้องกันแคปTCHA ข้อความจากการโจมตีของโปรแกรมโอซีอาร์ด้านการรู้จำของอักษรแต่ละตัว (Bursztein E. et al., 2014)
5. โปรแกรมแคปTCHAโซลเวอร์ (CAPTCHA Solver) หมายถึง โปรแกรมที่ใช้สำหรับวิเคราะห์ แคปTCHAข้อความ โดยใช้คุณสมบัติการแปลงภาพเป็นข้อความของโปรแกรมโอซีอาร์ที่แปลงภาพของข้อความจากการเขียนหรือจากการพิมพ์เป็นข้อความที่สามารถแก้ไขและเก็บไว้ได้โดยเครื่องคอมพิวเตอร์ (วงศกร รัตติรุจิเศวต, 2557)

1.5 ขอบเขตของการวิจัย

1. ผู้วิจัยได้พัฒนาชุดอักษรแคปTCHAข้อความด้วยอักษรพิมพ์เล็ก แบบอักษร Arial ขนาด 24 พอยต์ จำนวน 6 ตัวอักษรต่อหนึ่งชุดอักษรแคปTCHAข้อความ และเลือกใช้อักษรสีดำบนพื้นหลังสีขาว โดยอ้างอิงจากงานวิจัยของ Bursztein et al. (2014) และวงศกร รัตติรุจิเศวต (2557) เพื่อนำมาวิเคราะห์ต่อยอดในงานวิจัยนี้ และกำหนดสามส่วนสำคัญในแคปTCHAที่ออกแบบคือ (1) ระยะห่างระหว่างอักษรสามแบบคือ ปกติ, ซิด (0 พอยต์), และ ทับซ้อน (ลบ 3.75 พอยต์) (2) ความหนาของเส้นรบกวนสามแบบคือ 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล และ (3) สีของเส้นรบกวนสองแบบคือ สีเดียวกับอักษร และสีเดียวกับพื้นหลัง
2. ผู้วิจัยได้พัฒนาเว็บไซต์เพื่อตรวจสอบการเข้าสู่ระบบของผู้ใช้ เมื่อผู้ใช้ต้องการเข้าสู่ระบบ จะต้องกรอกแคปTCHAข้อความที่ผู้ใช้ได้เห็น เพื่อเป็นเครื่องมือช่วยในการทดสอบความเป็นมนุษย์ของผู้ใช้

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อเสนอแนวคิดในการพัฒนาแคปซูลข้อความให้มีความสามารถป้องกันการโจมตีด้านการแบ่งส่วนจากโปรแกรมไอซีอาร์
2. เพื่อเสนอข้อมูลในการเลือกระดับระยะห่างระหว่างอักขร ความหนาของเส้นรบกวน และสีของเส้นรบกวน ในการพัฒนาแคปซูลข้อความให้ทนทานต่อการโจมตีด้านการแบ่งส่วนจากโปรแกรมไอซีอาร์ โดยที่มนุษย์ยังสามารถเข้าใจได้
3. เพื่อเสนอแนวทางการเลือกใช้แคปซูลข้อความให้กับเจ้าของเว็บไซต์พาณิชย์อิเล็กทรอนิกส์ที่ต้องการป้องกันการรบกวนจากผู้ไม่ประสงค์ดี



บทที่ 2

การทบทวนวรรณกรรม

2.1 ความนำ

ในบทนี้เป็นการนำเสนอผลของการทบทวนวรรณกรรม (Literature review) ที่เกี่ยวข้อง เพื่อชี้ให้เห็นถึงประเด็นการออกแบบแคปTCHAข้อความ การนำเสนอประกอบด้วยหัวข้อย่อยคือ (1) รูปแบบของแคปTCHAข้อความ (2) การยืนยันความเป็นมนุษย์ของแคปTCHAข้อความ (3) โปรแกรมแคปTCHAโซลเวอร์ (CAPTCHA Solver) (4) การป้องกันโปรแกรมอัตโนมัติโจมตีแคปTCHAข้อความ (5) ระยะห่างระหว่างอักษร (6) สิ่งรบกวนของแคปTCHAข้อความ (noise) และประเด็นสุดท้ายของบทจะสรุปวรรณกรรมที่เกี่ยวข้องกับงานวิจัยนี้ เพื่อแสดงให้เห็นถึงความสำคัญในการออกแบบแคปTCHAข้อความ

2.2 รูปแบบของแคปTCHAข้อความ

แคปTCHA (CAPTCHA; Completely Automated Public Turing Test to Tell Computers and Humans Apart) คือ โปรแกรมทดสอบแบบโต้ตอบชนิดหนึ่งเพื่อเข้าสู่ระบบคอมพิวเตอร์ ใช้สำหรับทดสอบว่าผู้ใช้งานที่กำลังขอรับบริการเป็นมนุษย์จริงหรือไม่ (Luis von Ahn et al., 2004) การทดสอบนั้นสามารถจำแนกความเป็นมนุษย์ได้จากการตอบแบบทดสอบแคปTCHAได้อย่างถูกต้อง โดยที่โปรแกรมคอมพิวเตอร์อัตโนมัติไม่สามารถตอบแบบทดสอบได้ถูกต้อง แคปTCHAเป็นเทคโนโลยีด้านความปลอดภัยที่ได้รับความนิยมมากที่สุด แคปTCHAข้อความถูกพัฒนาขึ้นและนำมาใช้อย่างแพร่หลายสำหรับป้องกันการโจมตีจากโปรแกรมคอมพิวเตอร์ที่ไม่ประสงค์ดี (Yan J. & El Ahmad A. S., 2008)

แคปTCHAสามารถจำแนกได้เป็นสองกลุ่มหลักๆคือ (1) กลุ่มแคปTCHAแบบรู้จำอักษรทางภาพ (OCR-based CAPTCHA) แคปTCHAกลุ่มนี้ได้แก่ แคปTCHAข้อความ (Text-based CAPTCHA) โดยจะแสดงชุดอักษรที่มีลักษณะบิดเบือน และกำหนดให้ผู้ผู้ใช้งานต้องพิมพ์ชุดอักษรที่แสดงอยู่ให้ถูกต้องเพื่อยืนยันความเป็นมนุษย์ และ (2) กลุ่มแคปTCHAแบบไม่รู้จำอักษรทางภาพ (Non-OCR based CAPTCHA) แคปTCHAกลุ่มนี้ได้แก่ แคปTCHAรูปภาพ (Image-based CAPTCHA) โดยจะแสดงรูปภาพและกำหนดให้ผู้ผู้ใช้งานทำตามคำสั่งให้ถูกต้องเพื่อยืนยันความเป็นมนุษย์ แคปTCHAเสียง (Audio-based CAPTCHA) โดยจะให้ผู้ผู้ใช้งานฟังเสียง และกำหนดให้ผู้ผู้ใช้งานต้องพิมพ์ชุดอักษรตามเสียงที่ได้ยินให้ถูกต้อง

เพื่อยืนยันความเป็นมนุษย์ หรือแคปต์ชาวิดีโอ (Video-based CAPTCHA) โดยจะแสดงวิดีโอเป็นแบบทดสอบ และกำหนดให้ผู้ใช้กรอกคำที่สื่อด้วยภาพในวิดีโอให้ถูกต้องเพื่อยืนยันความเป็นมนุษย์ (Montree Imsamai & Suphakant Phimoltares, 2010; มณีรัตน์ ชาตธีรังสรรค์, 2555)

การวิจัยนี้จึงเลือกศึกษาแคปต์ชาข้อความ เนื่องจากแคปต์ชาข้อความนี้เป็นประเภทของแคปต์ชาที่ถูกนำมาใช้งานมากที่สุดและสามารถสร้างแบบทดสอบได้หลายรูปแบบ (Kolupave A. & Ogijenko J., 2008) โดยรูปแบบของแคปต์ชาข้อความที่ Bursztein et al. (2011) ได้เลือกตัวอย่างแคปต์ชาข้อความเพื่อวิจัยความสามารถในการตอบแคปต์ชาของมนุษย์ ได้แก่ (1) Authorize (2) Baidu (3) captchas.net (4) digg.com (5) eBay (6) Google (7) Mail.ru (8) Microsoft (9) Recaptcha (10) Skyrock (11) Slashdot (12) Blizzard และ (13) Yahoo (ตัวอย่างในภาพที่ 2.1 ถึง 2.13 ตามลำดับ) และพบว่าแคปต์ชาข้อความส่วนใหญ่ที่ถูกนำมาใช้ในการทดสอบอย่างแพร่หลายนั้นยากเกินกว่าที่ควรจะเป็นสำหรับมนุษย์ที่จะสามารถตอบได้อย่างถูกต้อง โดยรายละเอียดรูปแบบของแคปต์ชาข้อความแต่ละแบบมีดังนี้

แคปต์ชาข้อความจาก authorize.net ในภาพที่ 2.1 ประกอบด้วยตัวเลขและตัวอักษรพิมพ์เล็กสีดำจำนวนห้าตัวได้แก่ 3 n c 9 และ z บนพื้นหลังสีขาว การบิดเบือนแคปต์ชาข้อความกระทำโดยการบิดเบือนตัวเลขและตัวอักษรด้วยองศาที่หลากหลาย และเพิ่มจุดรบกวน (Noise) สีเทาขนาดเล็กทั้งพื้นหลังและบนตัวเลขและตัวอักษร



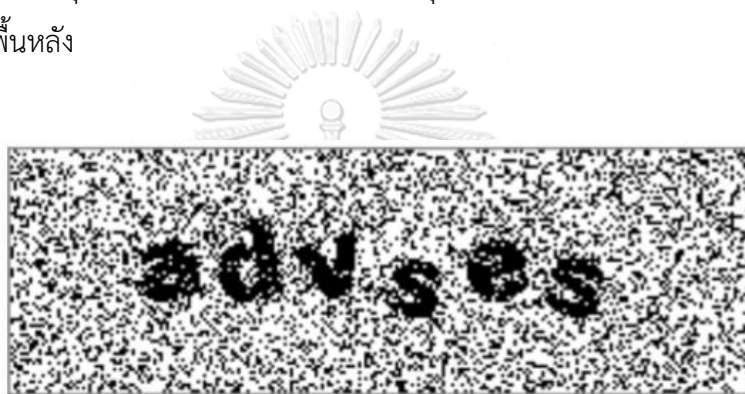
ภาพที่ 2.1 ตัวอย่างแคปต์ชาข้อความจาก authorize.net

แคปต์ชาข้อความจาก baidu.com ในภาพที่ 2.2 ประกอบด้วยตัวเลขและตัวอักษรพิมพ์ใหญ่สีดำจำนวนสี่ตัวได้แก่ R A E และ 3 บนพื้นหลังสีขาว การบิดเบือนแคปต์ชาข้อความกระทำโดยการเพิ่มเส้นรบกวนพาดข้ามชุดข้อความในตอนกลางของข้อความ ตัวเลขและตัวอักษรแต่ละตัวจะถูกทำให้เอียง และมีตำแหน่งที่ทับซ้อนกัน



ภาพที่ 2. 2 ตัวอย่างแคปช่าข้อความจาก baidu.com

แคปช่าข้อความจาก captcha.net ในภาพที่ 2.3 ประกอบด้วยตัวอักษรพิมพ์เล็กสีดำ จำนวนหกตัว ได้แก่ a d v s e และ s บนพื้นหลังสีขาว การบิดเบือนแคปช่าข้อความกระทำโดยการสุมตัวอักษรเพื่อหมุน ยกขึ้น หรือเลื่อนลง และเพิ่มจุดรบกวนสีขาวขนาดเล็กบนตัวอักษร และจุดรบกวนสีดำบนพื้นหลัง



ภาพที่ 2. 3 ตัวอย่างแคปช่าข้อความจาก captcha.net

แคปช่าข้อความจาก digg.com ในภาพที่ 2.4 ประกอบด้วยตัวเลขและตัวอักษรพิมพ์ (ทั้งพิมพ์เล็กและพิมพ์ใหญ่) สีดำจำนวนห้าตัว ได้แก่ 2 p v c p บนพื้นหลังสีขาวที่มีลายเส้นสีเทา การบิดเบือนแคปช่าข้อความกระทำโดยการสุมตัวอักษรเพื่อหมุน ยกขึ้น หรือเลื่อนลง และเพิ่มเส้นรบกวนสีดำและสีเทาพาดข้ามแคปช่าข้อความทแยงไปมา



ภาพที่ 2. 4 ตัวอย่างแคปช่าข้อความจาก digg.com

แคปTCHAข้อความจาก ebay.com ในภาพที่ 2.5 ประกอบด้วยตัวอักษรสีเข้มจำนวนหกตัว ได้แก่ 9 8 4 5 0 และ 5 บนพื้นหลังสีขาว การบิดเบือนแคปTCHAข้อความกระทำด้วยการสุมตัวอักษร เพื่อหมุนหรือบิดตัวอักษร ยกขึ้น หรือเลื่อนลง และจัดให้มีตำแหน่งที่ทับซ้อนเหลื่อมกัน



ภาพที่ 2. 5 ตัวอย่างแคปTCHAข้อความจาก ebay.com

แคปTCHAข้อความจาก google.com ในภาพที่ 2.6 ประกอบด้วยตัวอักษรสีแดง สีเขียว หรือสีฟ้าจำนวนสิบตัว ได้แก่ d e a c t i e s s และ e บนพื้นหลังสีขาว โดยรูปแบบของชุดอักษรจะมีลักษณะคล้ายคำที่มีความหมายในภาษาอังกฤษ การบิดเบือนแคปTCHAข้อความกระทำโดยการบิดหมุนเอียงตัวอักษร จัดให้มีตำแหน่งที่ทับซ้อนกัน และชุดอักษรจะถูกจัดเรียงในลักษณะโค้งเป็นคลื่น



ภาพที่ 2. 6 ตัวอย่างแคปTCHAข้อความจาก google.com

แคปTCHAข้อความจาก mail.ru ในภาพที่ 2.7 ประกอบด้วยอักษรจำนวนหกตัว ได้แก่ F 5 A C M และ X ที่แสดงเฉพาะโครงเป็นกรอบสีฟ้าบนพื้นหลังสีขาว การบิดเบือนแคปTCHAข้อความกระทำโดยการหมุนเอียงตัวอักษร ทำให้โค้งงอ ยกขึ้น หรือเลื่อนลง และเพิ่มเส้นกรอบของตัวเลขและตัวอักษรอื่น ๆ บนพื้นหลัง



ภาพที่ 2. 7 ตัวอย่างแคปช่าข้อความจาก mail.ru

แคปช่าข้อความจาก live.com ในภาพที่ 2.8 ประกอบด้วยตัวเลขและตัวอักษรพิมพ์ใหญ่สีน้ำเงินจำนวนแปดตัว ได้แก่ M Y 6 K C 8 6 และ C บนพื้นหลังสีเทา การบิดเบือนแคปช่าข้อความกระทำโดยการบีบ หมุนเอียงตัวอักษร จัดให้เส้นของตัวอักษรมีขนาดใหญ่มากต่างกัน และจัดให้มีตำแหน่งที่ทับซ้อนกัน ชุดอักษรจะถูกจัดเรียงในลักษณะโค้งเป็นคลื่น



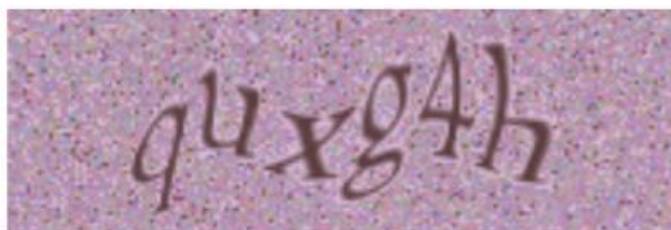
ภาพที่ 2. 8 ตัวอย่างแคปช่าข้อความจาก live.com

แคปช่าข้อความจาก recaptcha.net ในภาพที่ 2.9 ประกอบด้วยชุดอักษรสองชุดคือ 3-2 และ parks ด้วยอักษรสีดำบนพื้นหลังสีขาว การบิดเบือนแคปช่าข้อความกระทำโดยการใช้ภาพของคำที่ได้จากการตรวจจรวด (scan) แล้วโปรแกรมอัตโนมัติไม่สามารถรู้จำ (Recognize) คำนั้นได้ นอกจากนี้ชุดอักษรจะถูกบีบและจัดเรียงในลักษณะโค้งเป็นคลื่น

3-2 parks

ภาพที่ 2. 9 ตัวอย่างแคปช่าข้อความจาก recaptcha.net

แคปต์ชาข้อความจาก skyrock.com ในภาพที่ 2.10 ประกอบด้วยตัวเลขและตัวอักษรพิมพ์เล็กสี่เหลี่ยมจำนวนหกตัว ได้แก่ q u x g 4 h บนพื้นหลังที่ใช้โทนสีสว่างกว่า การบิดเบือนแคปต์ชาข้อความกระทำโดยการบีบและจัดเรียงในลักษณะโค้งงอเป็นคลื่น



ภาพที่ 2. 10 ตัวอย่างแคปต์ชาข้อความจาก skyrock.com

แคปต์ชาข้อความจาก slashdot.org ในภาพที่ 2.11 ประกอบด้วยคำที่มีความหมายในภาษาอังกฤษหนึ่งคำ ได้แก่ apricot ซึ่งประกอบด้วยตัวอักษรเจ็ดตัว โดยใช้ตัวอักษรพิมพ์เล็กสี่เหลี่ยมพื้นหลังสีขาว การบิดเบือนแคปต์ชาข้อความกระทำโดยตัวอักษรที่บางตัวจะแสดงแค่เส้นกรอบของตัวอักษร บางตัวเป็นเส้นทึบ และมีการเพิ่มเส้นรบกวนในลักษณะพื้นปลาพาดข้ามชุดอักษร และเพิ่มจุดรบกวนสี่ด้านบนพื้นหลัง ตามที่แสดงในภาพที่ 2.11



ภาพที่ 2. 11 ตัวอย่างแคปต์ชาข้อความจาก slashdot.org

แคปต์ชาข้อความจาก worldofwarcraft.com ในภาพที่ 2.12 ประกอบด้วยตัวอักษรสี่เหลี่ยมจำนวนหกตัว ได้แก่ p m y m k และ u บนพื้นหลังสี่เหลี่ยมที่เป็นแบบรูป (pattern) การบิดเบือนแคปต์ชาข้อความกระทำโดยหมุนตัวอักษรเล็ก บีบ ยกขึ้น และเลื่อนลง



ภาพที่ 2. 12 ตัวอย่างแคปท์ซาข้อความจาก worldofwarcraft.com

แคปท์ซาข้อความจาก yahoo.com ในภาพที่ 2.13 ประกอบด้วยตัวเลขและตัวอักษรพิมพ์ใหญ่สีดำจำนวนเจ็ดตัว ได้แก่ W 6 T 4 V 4 และ M บนพื้นหลังสีขาว การบิดเบือนแคปท์ซาข้อความกระทำโดยการบีบตัวอักษร ทำให้โค้งงอในลักษณะขึ้นลง และจัดให้มีตำแหน่งที่ทับซ้อนกัน โดยชุดอักษรจะถูกจัดเรียงในลักษณะโค้งเป็นคลื่น



ภาพที่ 2. 13 ตัวอย่างแคปท์ซาข้อความจาก yahoo.com

จากตัวอย่างแคปท์ซาข้อความข้างต้นในภาพที่ 2.1 ถึง 2.13 จะเห็นได้ว่า ผู้พัฒนาแคปท์ซาข้อความได้พยายามป้องกันการโจมตีจากโปรแกรมอัตโนมัติไม่ให้อ่านได้ด้วยการบิดเบือนตัวอักษรในหลายลักษณะ เช่น การหมุน บีบ ทำให้โค้ง หรือวางให้ตำแหน่งของอักษรทับซ้อนกัน ตลอดจนเพิ่มสิ่งรบกวนเข้าไปในแคปท์ซาข้อความ เช่น เส้นรบกวน จุดรบกวน หรือภาพพื้นหลังรบกวน โดยยังคงหวังว่ามนุษย์หรือผู้ใช้จะยังสามารถอ่านและตอบแคปท์ซาข้อความได้อย่างถูกต้อง

2.3 การยืนยันความเป็นมนุษย์ของแคปท์ซาข้อความ

แคปท์ซาข้อความต้องทนทาน (Robust) ต่อการวิเคราะห์ของโปรแกรมอัตโนมัติ (Bot) อย่างไรก็ตามแคปท์ซาข้อความควรง่ายต่อผู้ใช้เพื่อให้ผู้ใช้สามารถตอบได้ถูกต้อง อัตราการยืนยันความเป็นมนุษย์สามารถวัดจากสัดส่วนของจำนวนผู้ตอบแคปท์ซาข้อความได้ถูกต้องต่อจำนวนผู้ตอบแคปท์ซาข้อความทั้งหมด (Yan J. & El Ahmad A. S., 2008; มณีรัตน์ ชาตธีรังสรรค์, 2555)

Bursztein et al. (2010) ได้วิจัยความสามารถในการตอบแคปต์ชาของมนุษย์ ทั้งแคปต์ชาในรูปแบบข้อความและแบบเสียง โดยเลือกตัวอย่างแคปต์ชาจากยี่สิบเอ็ดรูปแบบที่ได้รับความนิยมมากที่สุด (แคปต์ชาข้อความสิบสามรูปแบบ และแคปต์ชาเสียงแปดรูปแบบ) สำหรับรูปแบบของแคปต์ชาข้อความที่นำมาวิเคราะห์ความสามารถในการตอบแคปต์ชาข้อความประกอบด้วย (1) Authorize (2) Baidu (3) captchas.net (4) digg.com (5) eBay (6) Google (7) Mail.ru (8) Microsoft (9) Recaptcha (10) Skyrock (11) Slashdot (12) Blizzard และ (13) Yahoo โดยกำหนดให้หน่วยทดลองตอบแคปต์ชาข้อความ

ผลการวิจัยพบว่าแคปต์ชาข้อความส่วนใหญ่ยากเกินกว่าที่ควร โดยค่าเฉลี่ยระยะเวลาที่ใช้ในการตอบแคปต์ชาข้อความ (solving time) คือ 9.8 วินาทีต่อหนึ่งแคปต์ชาข้อความ อย่างไรก็ตาม เนื่องจากแคปต์ชาข้อความที่นำมาวิเคราะห์ Bursztein และคณะ (2010) ไม่ได้สร้างขึ้นเองจึงไม่สามารถรู้คำตอบที่ถูกต้อง การวัดค่าความถูกต้องในการตอบแคปต์ชาข้อความจึงใช้วิธีพิจารณาความแตกต่างของคำตอบของหน่วยทดลอง ทั้งนี้ผลการวิจัยพบว่าหน่วยทดลองตอบแคปต์ชาได้ถูกต้องเพียง 71% เท่านั้น โดยงานวิจัยนี้ได้วัดค่าความถูกต้องจากการตรวจดูความเห็นที่ตรงกันตรงกัน (solving agreement) ของหน่วยทดลองคือ คำตอบเหมือนกัน จะเห็นได้ว่าการออกแบบแคปต์ชาต้องคำนึงถึงความทนทานต่อโปรแกรมอัตโนมัติและความง่ายในการตอบของมนุษย์ด้วยเพื่อไม่ให้เกิดความยุ่งยากในการใช้งานของผู้ใช้ปกติ อย่างไรก็ตามงานวิจัยนี้จะมีข้อบกพร่องในการวัดอัตราการยืนยันยืนยันความเป็นมนุษย์ เนื่องจากไม่สามารถรู้คำตอบที่ถูกต้องที่แท้จริงได้

2.4 โปรแกรมแคปต์ชาโซลเวอร์ (CAPTCHA Solver) วิทยาลัย

เนื่องจากแคปต์ชาข้อความคือ การทดสอบแบบโต้ตอบบนระบบคอมพิวเตอร์ชนิดหนึ่ง เพื่อทดสอบว่าผู้ใช้งานที่กำลังขอรับบริการเป็นมนุษย์จริงหรือไม่ (Luis von Ahn et al., 2004) แคปต์ชาข้อความที่ดีต้องทนทานต่อการโจมตีของโปรแกรมอัตโนมัติ โดยอัตราการความทนทานของแคปต์ชาข้อความสามารถวัดจากจำนวนครั้งการโจมตีแคปต์ชาข้อความของโปรแกรมอัตโนมัติที่ไม่ถูกต้องต่อจำนวนการโจมตีทั้งหมด (El Ahmand A. S. et al., 2011) ดังนั้นผู้วิจัยจึงได้ทบทวนวรรณกรรมเกี่ยวกับกระบวนการวิเคราะห์แคปต์ชาข้อความของโปรแกรมอัตโนมัติเพื่อพิจารณาเลือกโปรแกรมอัตโนมัติเพื่อใช้ในการวัดอัตราการความทนทานของแคปต์ชาข้อความ

โปรแกรมแคปต์ชาโซลเวอร์ (CAPTCHA Solver) คือ โปรแกรมอัตโนมัติที่สามารถแปลงภาพข้อความด้วยเทคนิคการแบ่งส่วน (Segment) หรือเทคนิคการรู้จำ (Recognize) ออกเป็นอักขระอย่างถูกต้อง ทั้งนี้การแปลงข้อความมาจากหลักของโปรแกรมโอซีอาร์ (Optical Character Recognition;

OCR) โดยโปรแกรมโอซีอาร์คือ โปรแกรมที่สามารถแปลงภาพใด ๆ ออกเป็นข้อมูลสามารถแก้ไขและเก็บไว้ได้โดยเครื่องคอมพิวเตอร์ (Text) (Kazem Taghva, Julie Borsack, & Allen Condit, 1996; Montree Imsamai & Suphakant Phimoltares, 2010; มณีรัตน์ ชาตธีรธรรม, 2555) กล่าวคือ การทำงานหลักของโปรแกรมโอซีอาร์เป็นการ (1) แบ่งส่วน (Segment) หรือ (2) รู้จำ (Recognize) นั้นหมายความว่า แคปต์ชาข้อความที่ดีจะต้องถูกออกแบบให้ทนทาน (Robust) ต่อการแบ่งส่วน หรือ การรู้จำจากการโจมตีของโปรแกรมโอซีอาร์ โดยโปรแกรมแคปต์ชาโซลเวอร์เป็นโปรแกรมที่รวบรวม หรือบูรณาการความสามารถของโปรแกรมโอซีอาร์หลายโปรแกรมเข้าด้วยกัน (Marti Motoyama et al., 2010)

ดังนั้นโปรแกรมแคปต์ชาโซลเวอร์จึงประกอบด้วยกระบวนการหลักในการวิเคราะห์แคปต์ชาห้า ขั้นตอนคือ (1) ขั้นตอนการประมวลผล (Pre-processing) ขั้นตอนนี้จะนำภาพพื้นหลัง หรือรูปภาพที่อื่น ๆ ที่ไม่ใช่ชุดอักษรและสามารถลบการแบ่งส่วนออกไป (2) ขั้นตอนการแบ่งส่วน (Segmentation) ขั้นตอนนี้จะจำแนกชุดอักษรออกเป็นแต่ละตัวอักษร เป็นการประมวลผลว่าในแคปต์ชาข้อความนี้มีกี่ตัวอักษร (3) ขั้นตอนหลังจากการแบ่งส่วน (Post-segmentation) ขั้นตอนนี้จะปรับขนาดของรูปภาพอักษรที่ได้จากขั้นตอนการแบ่งส่วนให้มีขนาดปกติ (4) ขั้นการรู้จำ (Recognition) ขั้นตอนนี้จะระบุว่ารูปภาพอักษรนั้นเป็นตัวอักษรใด และ (5) ขั้นตอนหลังจากการประมวลผล (Post-processing) ขั้นตอนนี้จะตรวจสอบความถูกต้องเพิ่มเติมหลังจากการระบุตัวอักษรในขั้นตอนนี้ เช่น การตรวจสอบการสะกดคำสำหรับแคปต์ชาที่ใช้คำที่มีความหมาย เป็นต้น (Bursztein E. et al., 2011)

โปรแกรมแคปต์ชาโซลเวอร์เป็นโปรแกรมอัตโนมัติที่มีคุณสมบัติเฉพาะเพื่อโจมตีแคปต์ชาข้อความ กระบวนการทำงานหรือการโจมตีของโปรแกรมแคปต์ชาโซลเวอร์จึงมีขั้นตอนการกำจัดสิ่งรบกวน (Noise) จากชุดอักษร เพื่อให้ในขั้นตอนการจำแนกตัวอักษร การปรับขนาดตัวอักษร และการระบุตัวอักษรส่งผลให้การโจมตีทำได้มากขึ้น ดังนั้นเพื่อให้ค่าที่วัดได้ใกล้เคียงกับสถานการณ์จริงมากที่สุด การวัดอัตราความทนทานของแคปต์ชาข้อความ (Robusness Rate of CAPTCHA) ผู้วิจัยเลือกใช้ GSA Captcha Breaker ในการวัดความสามารถของแคปต์ชาข้อความในการศึกษาครั้งนี้ เพราะเป็นโปรแกรมแคปต์ชาโซลเวอร์ที่ใช้กระบวนการโอซีอาร์ขั้นสูงในการตอบแคปต์ชาข้อความ และมีอัตราการตอบได้ถูกต้องมากที่สุดจากแคปต์ชาข้อความหลายรูปแบบ (Beheshti S. M. R. S. & Liatsis P, 2015)

2.5 การป้องกันโปรแกรมอัตโนมัติโจมตีแคปช่าข้อความ

เนื่องจากแคปช่าข้อความที่ติดตั้งบนหน้าจอคอมพิวเตอร์ของโปรแกรมอัตโนมัติหรือในที่นี้คือโปรแกรมแคปช่าโซลเวอ์ ผู้วิจัยได้ทบทวนวรรณกรรมเกี่ยวกับการป้องกันโปรแกรมอัตโนมัติโจมตีแคปช่าข้อความ เพื่อเป็นหลักในการวัดความสามารถของแคปช่าข้อความที่ได้ออกแบบภายใต้สถานการณ์ต่าง ๆ กัน

การป้องกันโปรแกรมอัตโนมัติจากการโจมตีแคปช่าข้อความคือ การออกแบบแคปช่าข้อความให้ทนทานต่อการโจมตีจากโปรแกรมอัตโนมัติ โดยแบ่งออกเป็นสองประเด็นหลักๆ ได้แก่ (1) ป้องกันการโจมตีด้านการรู้จำ (anti-recognition) คือการป้องกันไม่ให้โปรแกรมอัตโนมัติสามารถระบุได้ว่ารูปภาพอักษรเป็นตัวอักษรใด และ (2) ป้องกันการโจมตีด้านการแบ่งส่วน (anti-segmentation) คือการป้องกันไม่ให้แคปช่าขานั้นถูกจำแนกออกเป็นอักษรแต่ละตัว ซึ่งทำให้โปรแกรมอัตโนมัติจะทราบถึงจำนวนตัวอักษรในชุดอักษรนั้น (Bursztein E. et al., 2011)

(1) การป้องกันการโจมตีด้านการรู้จำ

Bursztein et al. (2014) ได้วิจัยคุณลักษณะของแคปช่าข้อความเพื่อให้ทนทานต่อการโจมตีด้านการรู้จำ (anti-recognition features) คุณลักษณะที่ว่ามีห้าคุณลักษณะคือ (1) จำนวนตัวอักษรของชุดอักษรแคปช่าข้อความที่หมุน (rotates character counts) (2) องศาการหมุนของตัวอักษร (rotates character degrees) (3) การเปลี่ยนขนาดของตัวอักษร (vertical shifting sizes) (4) ความหลากหลายของขนาดตัวอักษร (character size variations) และ (5) การบิดเบือนตัวอักษร (character distortions) ผลการวิจัยพบว่าแคปช่าที่หมุนอักษรประมาณสามสิบองศาจากเส้นแนวตั้ง มนุษย์จะยังสามารถตอบคำถามได้ถูกต้องโดยเร็ว แต่องศาการหมุนที่มีค่าเพิ่มมากขึ้นจะทำให้ความสามารถในการยืนยันความเป็นมนุษย์ลดลง การเปลี่ยนขนาดของตัวอักษรและความหลากหลายของขนาดตัวอักษรนั้นไม่มีผลต่อการอ่านแคปช่าข้อความ ทั้งนี้ณิรัตน์ ชาตริงสรรค์ (2555) ได้วิจัยคุณลักษณะที่ป้องกันการโจมตีด้านการรู้จำ โดยคุณลักษณะด้านแบบอักษรคือ (1) อักษรตัวพิมพ์ และ (2) อักษรตัวเขียน คุณลักษณะด้านจำนวนอักษร ได้แก่ (1) จำนวนอักษรสี่ตัว และ (2) จำนวนอักษรสิบตัว และคุณลักษณะด้านชุดอักขระของแคปช่าข้อความ ได้แก่ (1) อักษรตัวใหญ่ทั้งหมด (2) อักษรตัวเล็กทั้งหมด และ (3) อักษรตัวใหญ่ผสมตัวเล็กในสัดส่วนครึ่งต่อครึ่ง ผลการวิจัยพบว่าแบบอักษรและชุดอักขระของแคปช่าข้อความมีผลต่อความสามารถในการยืนยันความเป็นมนุษย์ของแคปช่าข้อความ ส่วนจำนวนอักษรไม่มีผลต่อความสามารถในการยืนยันความเป็นมนุษย์ หรือความทนทานของแคปช่าข้อความ ส่วนวงศกร รัตริฐจิเสวต (2557) นั้นได้วิจัยคุณลักษณะที่ป้องกันการโจมตีด้านการรู้จำ ดังนี้ คุณลักษณะด้านจำนวนอักษร ได้แก่ (1) จำนวนห้าตัว และ (2) จำนวนอักษร

สิบห้าตัว และคุณสมบัติด้านคู่สีอักษรและสีพื้นหลัง ได้แก่ (1) อักษรสีดำพื้นหลังสีขาว และ (2) อักษรสีขาวพื้นหลังสีน้ำเงิน ผลการวิจัยพบว่าจำนวนอักษรของแคปต์ชาข้อความมีผลต่อความสามารถในการยืนยันความเป็นมนุษย์ แต่ไม่มีผลต่อความทนทานของแคปต์ชาข้อความ ส่วนคู่สีอักษรและสีพื้นหลังไม่มีผลต่อความสามารถในการยืนยันความเป็นมนุษย์และความทนทานของแคปต์ชาข้อความ

(2) การป้องกันการโจมตีด้านการแบ่งส่วน

Bursztein et al. (2014) ได้วิจัยคุณลักษณะของแคปต์ชาข้อความเพื่อให้ทนทานต่อการโจมตีด้านการแบ่งส่วน (anti-segmentation features) ของโปรแกรมแคปต์ชาโซลเวอร์ ดังนี้ (1) การทับซ้อนของตัวอักษร (character overlap) (2) การสุ่มขนาดของจุดรบกวน (random dot size) (3) การสุ่มจำนวนของจุดรบกวน (random dot counts) (4) ประเภทของเส้นรบกวน (line types) (5) จำนวนเส้นรบกวน (line counts) (6) ความหนาของเส้นรบกวน (line widths) (7) ตำแหน่งของเส้นรบกวน (line positions) และ (8) สีของพื้นหน้าและพื้นหลังที่คล้ายกัน (similar foreground/background colors)

ผลการวิจัยพบว่า การทับซ้อนของตัวอักษรโดยให้อักษรอยู่ชิดกันมากเกินไปทำให้ความสามารถในการยืนยันความเป็นมนุษย์ลดลง การใช้จุดรบกวนขนาด 1 พิกเซล จำนวน 2,000 จุดขึ้นไปจะทำให้ความสามารถในการยืนยันความเป็นมนุษย์ลดลง เส้นรบกวนที่มีความหนามากกว่า 2 พิกเซล ขึ้นไปจะทำให้ความสามารถในการยืนยันความเป็นมนุษย์ลดลง อย่างไรก็ตามสีของพื้นหน้าและพื้นหลังไม่มีผลต่อความสามารถในการยืนยันความเป็นมนุษย์ ทั้งนี้ มณีรัตน์ ชาตธีรังสรรค์ (2555) ได้วิจัยคุณลักษณะที่ป้องกันการโจมตีด้านการแบ่งส่วนคือ ด้านการหมุนเอียง ได้แก่ (1) เอียงลบ 45 องศา และ (2) เอียงบวก 45 องศา ผลการวิจัยพบว่า การหมุนเอียงไม่มีผลต่อความสามารถในการยืนยันความเป็นมนุษย์ หรือความทนทานของแคปต์ชาข้อความ ส่วน วงศกร รัตติรุจิเสวต (2557) ได้วิจัยคุณลักษณะที่ป้องกันการโจมตีด้านการแบ่งส่วน ดังนี้ ด้านระยะห่างระหว่างอักษร ได้แก่ (1) ปกติ (2) ขยาย 8 พอยต์ และ (3) บีบ 3.4 พอยต์ และคุณลักษณะด้านตำแหน่งอักษร ได้แก่ (1) ปกติ (2) ยกขึ้น 5 พอยต์ และ (3) ต่ำลง 5 พอยต์ ผลการวิจัยพบว่า ระยะห่างระหว่างอักษรมีผลต่อความสามารถในการยืนยันความเป็นมนุษย์ หรือความทนทานของแคปต์ชาข้อความ และตำแหน่งอักษรมีผลต่อความทนทานของแคปต์ชาข้อความ

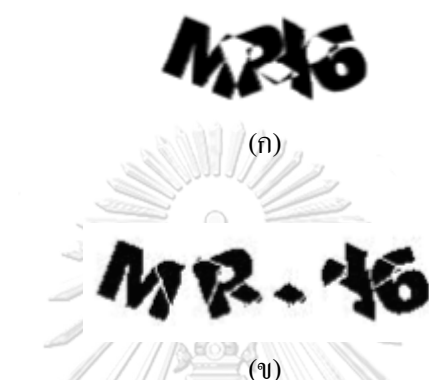
เนื่องจาก มณีรัตน์ ชาตธีรังสรรค์ (2555) และ วงศกร รัตติรุจิเสวต (2557) ได้วิเคราะห์คุณลักษณะของแคปต์ชาข้อความเพื่อป้องกันการรู้จำ โดย มณีรัตน์ ชาตธีรังสรรค์ (2555) ได้วิเคราะห์คุณลักษณะด้านการหมุนเอียง คุณลักษณะด้านแบบอักษร และคุณลักษณะด้านจำนวนอักษร โดยพบว่าคุณลักษณะด้านการหมุนเอียงเท่านั้นที่สามารถป้องกันการโจมตีด้านการแบ่งส่วน ในขณะที่

วงศ์กร รัตติรุจิเสวต (2557) ได้ศึกษาคุณลักษณะด้านระยะห่างระหว่างอักษร คุณลักษณะด้านตำแหน่งอักษร และ คุณลักษณะด้านจำนวนอักษร โดยคุณลักษณะที่ป้องกันการโจมตีด้านการแบ่งส่วน ได้แก่ คุณลักษณะด้านระยะห่างระหว่างอักษรและตำแหน่งอักษร เพื่อเพิ่มอัตราความทนทานของแคปช่าข้อความด้านการแบ่งส่วน อีกทั้ง Simard et al. (2003) สรุปร่วมมนุษย์มีความสามารถในการแบ่งส่วนข้อความ (segmentation) หรือการจำแนกข้อความออกเป็นตัวอักษรแต่ละตัวได้ดีกว่าโปรแกรมอัตโนมัติ โดยทั้ง Yan และ El Ahmad (2008) และ El Ahmad et al. (2010) ได้พัฒนาโปรแกรมอัตโนมัติเพื่อทดสอบโจมตีด้านการแบ่งส่วนกับแคปช่าข้อความ และพบว่าแคปช่าข้อความทนทานต่อการโจมตีด้านการแบ่งส่วน โดยผู้วิจัยเห็นว่าสามารถนำมาเพิ่มความทนทานให้กับแคปช่าข้อความได้

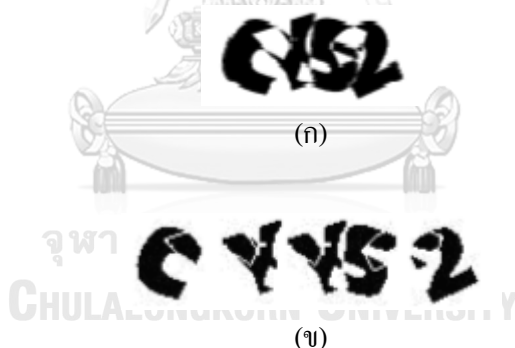
Yan และ El Ahmad (2008) พัฒนาโปรแกรมอัตโนมัติเพื่อโจมตีด้านการแบ่งส่วนของแคปช่าข้อความในไมโครซอฟต์ (Microsoft CAPTCHA) และสร้างขึ้นจาก MSN CAPTCHA scheme การโจมตีด้านการแบ่งส่วนนี้มีจุดประสงค์เพื่อ (1) ระบุและนำส่วนสิ่งรบกวนออกไป และ (2) ระบุตำแหน่งของตัวอักษรแต่ละตัวทั้งหมดตามลำดับจากทางด้านขวาไปซ้าย ผลของงานวิจัยนี้สรุปได้ว่าแคปช่าข้อความที่ทนทานต่อการโจมตีของโปรแกรมอัตโนมัติด้านการแบ่งส่วนสามประเภท ได้แก่ (1) โปรแกรมอัตโนมัติไม่สามารถนำเส้นรบกวนที่หนาออกไปได้ (2) โปรแกรมอัตโนมัติไม่สามารถระบุตำแหน่งของตัวอักษรที่มีลักษณะขีดติดกันได้ และ (3) โปรแกรมอัตโนมัติไม่สามารถแบ่งส่วนตัวอักษรที่ขีดติดกันได้ (even cut)

El Ahmad et al. (2010) ได้พัฒนาโปรแกรมอัตโนมัติ (Bot) เพื่อโจมตีด้านการแบ่งส่วนแคปช่าข้อความ ที่สร้างขึ้นโดย Megaupload CAPTCHA ระดับการโจมตีด้านการแบ่งส่วนของโปรแกรมนี้ ได้แก่ (1) ระบุตำแหน่งของเส้นของตัวอักษรที่เป็นสีดำ (2) ระบุตำแหน่งของระยะห่างระหว่างอักษรที่เป็นสีขาว และ (3) นำระยะห่างระหว่างอักษรมารวมอยู่ในตำแหน่งที่เหมาะสมกับเส้นของตัวอักษรที่เป็นสีดำเพื่อประกอบเป็นตัวอักษรที่สมบูรณ์ ผลจากงานวิจัยพบว่าแคปช่าข้อความที่ทนทานต่อการโจมตีของโปรแกรมอัตโนมัติด้านการแบ่งส่วนได้แก่ (1) โปรแกรมอัตโนมัติไม่สามารถแยกส่วนสีดำที่เป็นเส้นของตัวอักษรได้ เมื่อตัวอักษรมีลักษณะทับซ้อนกัน เนื่องจากเส้นของตัวอักษรคนละตัวอักษรกันทับซ้อนกลายเป็นเส้นเดียวกัน (ตัวอย่างในภาพที่ 2.14 (ก) และ (ข)) (2) โปรแกรมอัตโนมัติไม่สามารถวิเคราะห์ระยะห่างระหว่างอักษรได้ โดยเมื่อตัวอักษรทับซ้อนกัน โปรแกรมอัตโนมัติตรวจหาได้ว่าส่วนสีขาวที่ควรเป็นระยะห่างระหว่างอักษรเป็นส่วนหนึ่งของตัวอักษรที่จำเป็นต้องนำออกไป ทำให้เส้นของตัวอักษรบางส่วนขาดหายไป (ตัวอย่างในภาพที่ 2.15 (ก) และ (ข)) (3) โปรแกรมอัตโนมัติไม่สามารถระบุขอบเขตของระยะห่างระหว่างอักษรได้ เนื่องจากอักษรมีระยะห่างระหว่างอักษรแตกต่างกันแต่โปรแกรมอัตโนมัติจะระบุค่าระยะห่างระหว่างอักษรของทุกตัวอักษรใน

ข้อความเป็นค่าเดียวกัน (ตัวอย่างในภาพที่ 2.16 (ก) และ (ข)) และ (4) โปรแกรมอัตโนมัติไม่สามารถเติมเต็มเส้นของตัวอักษรได้อย่างถูกต้อง เนื่องจากตัวอักษรทับซ้อนกัน เมื่อระบุขอบเขตของระยะห่างระหว่างอักษรแล้วกำหนดให้เป็นสีขาวอาจทำให้เส้นของตัวอักษรบางส่วนขาดหายไป โปรแกรมอัตโนมัติ ต้องเพิ่มเส้นสีดำเพื่อความสมบูรณ์ของตัวอักษรบางตัว แต่โปรแกรมได้เพิ่มเส้นสีดำให้กับตัวอักษรตัวอื่นโดยไม่จำเป็น (ตัวอย่างในภาพที่ 2.17 (ก) และ (ข))



ภาพที่ 2. 14 (ก) แคปหน้าจอความต้นฉบับ (ข) แคปหน้าจอความที่โปรแกรมอัตโนมัติไม่สามารถแบ่งส่วน (Segment) เส้นของตัวอักษรได้

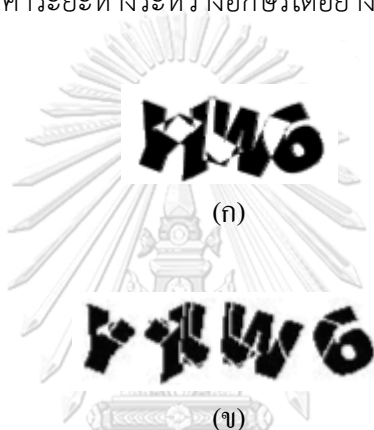


ภาพที่ 2. 15 (ก) แคปหน้าจอความต้นฉบับ (ข) แคปหน้าจอความที่โปรแกรมอัตโนมัติวิเคราะห์แล้วทำให้เส้นของตัวอักษรบางส่วนขาดหายไป

(ก)

(ข)

ภาพที่ 2. 16 (ก) แคปท์ซาข้อความต้นฉบับ (ข) แคปท์ซาข้อความที่โปรแกรมอัตโนมัติไม่สามารถระบุ
ค่าระยะห่างระหว่างอักษรได้อย่างถูกต้อง



(ก)

(ข)

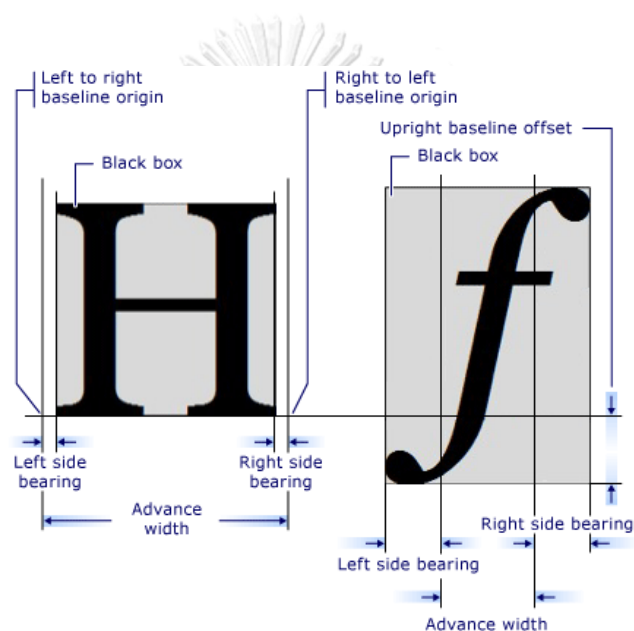
ภาพที่ 2. 17 (ก) แคปท์ซาข้อความต้นฉบับ (ข) แคปท์ซาข้อความที่โปรแกรมอัตโนมัติไม่สามารถเพิ่ม
เส้นของตัวอักษรได้อย่างถูกต้อง

การสำรวจงานวิจัยในอดีตพบว่าโปรแกรมอัตโนมัติจะไม่สามารถโจมตีแคปท์ซาข้อความด้าน
การแบ่งส่วนได้ เมื่อตัวอักษรในแคปท์ซาข้อความมีลักษณะทับซ้อนกัน หรือเมื่อเส้นรบกวนมีขนาด
หนาเข้าไปในแคปท์ซาข้อความ

2.6 ระยะห่างระหว่างอักษร (Spacing)

เนื่องจากแคปท์ซาข้อความที่ประกอบด้วยตัวอักษรที่ทับซ้อนกัน (หรือระยะห่างระหว่าง
อักษรถูกกำหนดเป็นศูนย์) จะยากต่อการโจมตีจากโปรแกรมอัตโนมัติ ผู้วิจัยจึงสนใจวิเคราะห์
ระยะห่างระหว่างอักษรเพื่อใช้ในการออกแบบแคปท์ซาข้อความให้มีความทนทานต่อการโจมตีด้าน
การแบ่งส่วนของโปรแกรมอัตโนมัติ

ระยะห่างระหว่างอักษร (Spacing) หมายถึง ระยะห่างด้านขวาของภาพอักษรแรกพร้อมกับ ระยะห่างด้านซ้ายของภาพตัวอักษรหลัง (วงศกร รัตติรุจิเสวต, 2557) โดยภาพอักษร (Glyph) หมายถึงภาพของตัวพิมพ์ (Type) ที่ใช้แสดงผลบนหน้าจอ ภาพที่ 2.18 คือตัวอย่างภาพอักษรของ ตัวอักษร H และตัวอักษร f จะเห็นจากภาพที่ 2.18 ได้ว่าตัวอักษร H มีความกว้างของภาพอักษร มากกว่าความกว้างของตัวอักษร ดังนั้นตัวอักษร H จะมีค่าระยะห่างด้านซ้ายและระยะห่างขวาเป็นค่า บวก ส่วนตัวอักษร f มีความกว้างของภาพอักษรมีค่าน้อยกว่าความกว้างของตัวอักษร ดังนั้นตัวอักษร f จะมีค่าระยะห่างด้านซ้ายและระยะห่างขวาเป็นค่าลบ (Microsoft Corporation, 2014; วงศกร รัตติรุจิเสวต, 2557)



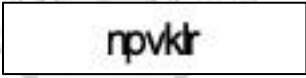
ภาพที่ 2. 18 ตัวอย่างภาพอักษร (Microsoft Corporation, 2014; วงศกร รัตติรุจิเสวต, 2557)

ตารางที่ 2. 1 ความหมายของส่วนประกอบต่าง ๆ ของภาพอักษร (โปรดดูภาพที่ 2.18 ประกอบ)

ส่วนประกอบของภาพอักษร	ความหมาย
Black box	ความกว้างของอักษร
Advance width	ความกว้างของภาพอักษร
Left side bearing	ระยะห่างด้านซ้ายของตัวอักษร
Right side bearing	ระยะห่างด้านขวาของตัวอักษร

การเพิ่มหรือลดระยะห่างระหว่างอักษรสามารถทำได้ด้วยการกำหนดค่า Tracking หรือค่า Kerning ของอักษร โดยที่ Tracking คือการกำหนดระยะห่างของอักษรทั้งหมดเมื่อจัดเป็นคำหรือรูปประโยค การกำหนด Tracking มีสองระดับได้แก่ (1) ขยาย (Expanded) คือการเพิ่มระยะห่างด้านขวาของภาพอักษร และ (2) บีบ (Condensed) คือการลดระยะห่างด้านขวาของภาพอักษร ส่วน Kerning คือการกำหนดระยะห่างระหว่างอักษรเฉพาะคู่ที่อยู่ติดกัน ดังนั้นค่า Kerning จึงขึ้นกับคู่อักษร เนื่องจากอักษรแต่ละตัวมีระยะห่างด้านซ้าย ด้านขวาไม่เท่ากัน ทั้งนี้ระยะห่างระหว่างอักษรมาจากระยะห่างด้านขวาของภาพอักษรตัวแรกรวมกับระยะห่างด้านซ้ายของภาพอักษรตัวหลัง

งานวิจัยนี้เลือกวิเคราะห์การลดระยะห่างระหว่างอักษรสองลักษณะ ได้แก่ (1) การลดระยะห่างระหว่างอักษรจนชิดกัน คือให้ระยะห่างด้านขวาของภาพอักษรแรกและระยะห่างด้านซ้ายของภาพอักษรหลังมีค่า 0 พอยต์ (ตัวอย่างแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าชิดในภาพที่ 2.19) และ (2) การลดระยะห่างระหว่างอักษรจนทับซ้อนกัน คือให้ระยะห่างด้านขวาของภาพอักษรแรกและระยะห่างด้านซ้ายของภาพอักษรหลังมีค่าลบ 3.75 พอยต์ (ตัวอย่างแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าทับซ้อนในภาพที่ 2.20)



ภาพที่ 2. 19 ตัวอย่างแคปท์ซาข้อความที่ได้ปรับระยะห่างระหว่างอักษรจนชิดกัน



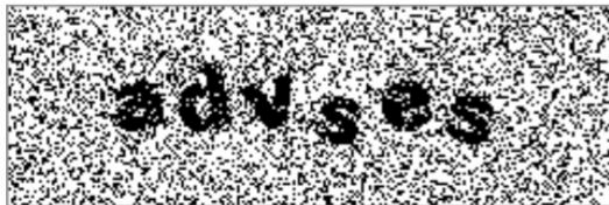
ภาพที่ 2. 20 ตัวอย่างแคปท์ซาข้อความที่ได้ปรับระยะห่างระหว่างอักษรจนทับซ้อน

เนื่องจากระยะห่างระหว่างอักษรของตัวอักษรแต่ละตัวมีค่าไม่เท่ากัน การลดระยะห่างระหว่างอักษรจนชิดติดกันจึงต้องกำหนดจากค่า Kerning ของแต่ละคู่อักษรที่อยู่ติดกัน ส่วนการลดระยะห่างระหว่างอักษรจนทับซ้อนกันผู้วิจัยได้กำหนดให้ระยะห่างด้านขวาของภาพอักษรแรกและระยะห่างด้านซ้ายของภาพอักษรหลังมีค่าคงที่เท่ากันทุกคู่อักษรคือ ลบ 3.75 พอยต์ จึงต้องกำหนดจากค่า Tracking โดยงานวิจัยนี้เลือกกำหนดระยะห่างระหว่างอักษรด้วยโปรแกรม Adobe Photoshop CS3 ได้แก่ (1) การลดระยะห่างระหว่างอักษรจนชิดติดกันสามารถกำหนดจากค่า Kerning (2) การลดระยะห่างระหว่างอักษรจนทับซ้อนสามารถกำหนดจากค่า Tracking

2.7 สิ่งรบกวนของแคปต์ชาข้อความ (Noise)

สิ่งรบกวนของแคปต์ชาข้อความหมายถึง ส่วนประกอบอื่นที่ไม่ใช่ส่วนสำคัญของอักขรแต่ถูกเพิ่มเข้าไปในแคปต์ชาข้อความ เพื่อให้แคปต์ชาข้อความทนทานต่อการโจมตีของโปรแกรมอัตโนมัติ Yan และ El Ahmad (2008) และ El Ahmad et al. (2010) พบว่าโปรแกรมอัตโนมัติไม่สามารถโจมตีด้านการแบ่งส่วนของแคปต์ชาข้อความได้ เนื่องจากการทับซ้อนกันของตัวอักษรและการนำส่วนประกอบอื่นที่ไม่ใช่ตัวอักษรออกจากแคปต์ชาข้อความ อีกทั้ง Bursztein (2011, 2014) ได้วิเคราะห์เทคนิคการป้องกันการโจมตีด้านการแบ่งส่วนดังนี้

Bursztein et al. (2011) ได้วิจัยคุณลักษณะสำหรับการออกแบบแคปต์ชาข้อความเพื่อป้องกันการโจมตีด้านการแบ่งส่วน ได้แก่ (1) ใช้ภาพพื้นหลังที่ซับซ้อน (complex background) (ตัวอย่างในภาพที่ 2.21) เพื่อเพิ่มความยุ่งยากให้การอ่านแคปต์ชาข้อความของโปรแกรมอัตโนมัติ เนื่องจากภาพพื้นหลังที่ซับซ้อนจะบดบังตัวอักษรในแคปต์ชาข้อความ (2) เพิ่มเส้นรบกวน (ตัวอย่างในภาพที่ 2.22) เพื่อป้องกันไม่ให้โปรแกรมจำแนกได้ว่าส่วนใดเป็นส่วนใดเป็นเส้นของตัวอักษร และ (3) การทับซ้อน (ตัวอย่างในภาพที่ 2.23) โดยการลดระยะห่างระหว่างอักขรเพื่อป้องกันการจำแนกตัวอักษรแต่ละตัว ผลของงานวิจัยนี้พบว่าการใช้ภาพพื้นหลังที่ซับซ้อนถึงแม้ว่าจะบดบังตัวอักษรได้ดี แต่โปรแกรมอัตโนมัติสามารถวิเคราะห์และนำภาพพื้นหลังออกได้ในขั้นก่อนการประมวลผล



ภาพที่ 2. 21 ตัวอย่างแคปต์ชาข้อความที่ใช้ภาพพื้นหลังที่ซับซ้อน



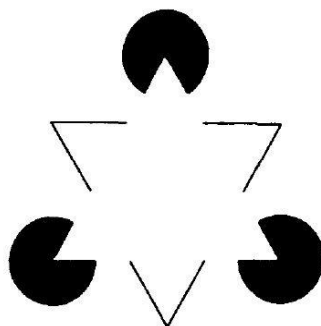
ภาพที่ 2. 22 ตัวอย่างแคปต์ชาข้อความที่เพิ่มเส้นรบกวน



ภาพที่ 2. 23 ตัวอย่างแคปท์ชาข้อความที่มีตัวอักษรทับซ้อนกัน

ทั้งนี้ Bursztein et al. (2011) ได้แนะนำว่าการเพิ่มเส้นรบกวน (noise) ควรมีลักษณะดังนี้ (1) ใช้เส้นที่มีขนาดใหญ่ และความหนาของเส้นรบกวนไม่ควรมีความหนาเท่ากับความหนาของเส้นที่เป็นส่วนประกอบของตัวอักษร (2) ตำแหน่งของเส้นควรอยู่ภายในแคปท์ชาข้อความในลักษณะพาดข้ามตัวอักษร (3) ไม่ควรใช้เส้นที่มีความชัน (slope) แตกต่างจากความชันเส้นของตัวอักษรในแคปท์ชาข้อความ เนื่องจากการใช้เส้นรบกวนที่มีความชันแตกต่างกันไปทำให้แคปท์ชาถูกโจมตีได้ง่าย โดยโปรแกรมอัตโนมัติจะทราบได้ว่าเส้นรบกวนนั้นไม่ใช่เส้นที่เป็นส่วนประกอบของตัวอักษร (4) กำหนดความชันของเส้นรบกวนให้ใกล้เคียงกับความชันของชุดอักษรของแคปท์ชาข้อความ เพื่อให้ยากต่อการวิเคราะห์ของโปรแกรมอัตโนมัติว่าเป็นเส้นรบกวนหรือเป็นเส้นที่เป็นส่วนประกอบของตัวอักษร (5) ใช้เส้นรบกวนสีเดียวกับสีของตัวอักษรในแคปท์ชาข้อความ เนื่องจากหากใช้เส้นรบกวนที่มีสีแตกต่างกับสีของตัวอักษร โปรแกรมอัตโนมัติจะไม่สามารถจำกัดเส้นรบกวนออกได้โดยง่าย และ (6) สุ่มใช้ขนาดความยาวของเส้นรบกวน เส้นรบกวนที่ถูกเพิ่มเข้าไปในแคปท์ชาข้อความมากกว่าหนึ่งเส้น แต่ละเส้นควรมีขนาดความยาวของเส้นแตกต่างกัน เพื่อไม่ให้โปรแกรมอัตโนมัติจดจำขนาดของความยาวเส้นรบกวนได้ อีกทั้งผลการวิจัยแนะนำว่าควรใช้การทับซ้อนของอักษรเป็นคุณสมบัติหลักช่วยป้องกันการโจมตีด้านการแบ่งส่วน โดยนักวิจัยไม่ควรลดระยะห่างระหว่างอักษรจนทับซ้อนกันมากเกินไป เนื่องจากระยะห่างระหว่างอักษรที่มีค่าตลลงน้อยกว่าลบ 5 พิกเซล หรือเท่ากับลบ 3.75 พอยต์ (1 พิกเซลเท่ากับ 0.75 พอยต์) จะทำให้ความสามารถในการยืนยันความเป็นมนุษย์ลดลง

Bursztein et al. (2014) ได้วิจัยเพิ่มเติมเกี่ยวกับคุณลักษณะของเส้นรบกวน โดยพบว่าเส้นรบกวนในลักษณะเส้นตรงเมื่อความหนาของเส้นรบกวนเป็น 1-2 พิกเซล และพาดข้ามชุดอักษรแคปท์ชาข้อความแล้ว มนุษย์ยังสามารถอ่านและตอบแคปท์ชาข้อความได้ถูกต้องประมาณ 90% แต่ความหนาของเส้นรบกวนที่เพิ่มขึ้นจะลดความสามารถในการตอบแคปท์ชาข้อความของมนุษย์ โดยเมื่อเพิ่มความหนาของเส้นรบกวนเป็น 3 พิกเซล เข้าไปในชุดอักษรแคปท์ชาข้อความแล้วมนุษย์ยังสามารถอ่านและตอบแคปท์ชาข้อความได้ถูกต้องประมาณ 80% และเมื่อเพิ่มความหนาของเส้นรบกวน 5 พิกเซล แล้วมนุษย์ยังสามารถตอบแคปท์ชาข้อความได้ถูกต้องประมาณ 55% แต่เมื่อเพิ่มความหนาของเส้นรบกวน 6 พิกเซล แล้วมนุษย์สามารถตอบแคปท์ชาข้อความได้ถูกต้องเพียง 34% เท่านั้น ดังนั้น การเพิ่มความหนาของเส้นรบกวนจะทำให้ความถูกต้องในการตอบแคปท์ชาของมนุษย์ลดลงและต้อง



ภาพที่ 2. 24 ตัวอย่างรูปสามเหลี่ยมที่ไม่มีเส้นเชื่อมต่อจนสมบูรณ์ แต่มนุษย์สามารถรับรู้ได้จากรูปร่างที่คล้ายกับรูปสามเหลี่ยม (Nuttanont Hongwarittorn & Suttikiat Meelap, 2015)

ด้วยทฤษฎีความสมบูรณ์ของเกสตัลท์ ผู้วิจัยจึงเลือกวิเคราะห์สีของเส้นรบกวนเป็นสีเดียวกับพื้นหลังของแคปuchaข้อความเปรียบเทียบกับสีของเส้นรบกวนที่มีสีเดียวกับสีของตัวอักษร ถึงแม้ว่าเส้นรบกวนที่มีสีเดียวกับพื้นหลังอาจทำให้เส้นของตัวอักษรบางส่วนขาดไป แต่การรับรู้ของมนุษย์จะสามารถรับรู้เป็นตัวอักษรสมบูรณ์ได้โดยอาศัยประสบการณ์เดิม

2.8 วรรณกรรมที่เกี่ยวข้อง: รูป

การทบทวนวรรณกรรมในบทนี้ แสดงให้เห็นถึงความสำคัญในการออกแบบแคปuchaข้อความที่ยากต่อการโจมตีจากโปรแกรมอัตโนมัติ และง่ายต่อผู้ใช้งาน อีกทั้งยังได้นำเสนอกระบวนการวิเคราะห์แคปuchaข้อความของโปรแกรมอัตโนมัติ การกำหนดค่าระยะห่างระหว่างอักษร สิ่งรบกวนที่สามารถเพิ่มเข้าไปได้ในแคปuchaข้อความ และการเลือกใช้เส้นรบกวนสีเดียวกับพื้นหลังโดยอ้างอิงจากกฎความสมบูรณ์ของเกสตัลท์ อันเป็นส่วนสำคัญในการออกแบบแคปuchaข้อความให้ยากต่อการโจมตีด้านการแบ่งส่วน

การทบทวนวรรณกรรมในอดีต พบว่าการวิจัยตัวแปรที่มีผลต่อการโจมตีแบ่งส่วนของโปรแกรมอัตโนมัติค่อนข้างน้อย เมื่อเทียบกับตัวแปรที่มีผลต่อการโจมตีด้านการรู้จำ ดังนั้นจึงเป็นโอกาสให้ผู้วิจัยสนใจวิเคราะห์ตัวแปรที่ส่งผลต่อการโจมตีการแบ่งส่วน โดยตัวแปรที่อยู่ในความสนใจของวิทยานิพนธ์ฉบับนี้ ประกอบด้วย (1) ระยะห่างระหว่างอักษร (2) ความหนาของเส้นรบกวน และ (3) สีของเส้นรบกวน

บทที่ 3 ระเบียบวิธีวิจัย

3.1 ความนำ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปช่าข้อความที่ออกแบบแตกต่างกันตามตัวแปรที่ศึกษา รายละเอียดในบทระเบียบวิธีวิจัยนี้นำเสนอแนวทางในการทำวิจัยเพื่อตอบวัตถุประสงค์ของการวิจัย การนำเสนอประกอบด้วย ประชากรและหน่วยทดลอง การทดสอบสมมติฐาน เครื่องมือเก็บข้อมูล ขั้นตอนการเก็บข้อมูล และเทคนิคที่ใช้วิเคราะห์ข้อมูลที่เก็บได้

3.2 ประชากรและหน่วยทดลอง

3.2.1 ประชากร

งานวิจัยนี้ต้องการวิเคราะห์การใช้แคปช่าข้อความเพื่อยืนยันความเป็นมนุษย์ของผู้ใช้อินเทอร์เน็ต และตรวจสอบอัตราความทนทานของแคปช่าข้อความเมื่อถูกโจมตีด้วยโปรแกรมแคปช่าโซลเวอร์ (CAPTCHA Solver) แต่เนื่องจากผู้ใช้อินเทอร์เน็ตเป็นประชากรขนาดใหญ่ ผู้วิจัยไม่ทราบขนาดของประชากรทั้งหมดทำให้ไม่สามารถเก็บข้อมูลจากประชากรได้ทุกหน่วย จึงต้องเก็บข้อมูลจากตัวอย่าง หรือหน่วยทดลองที่สามารถเป็นตัวแทนของประชากร (กัลยา วาณิชย์บัญชา, 2553)

3.2.2 หน่วยทดลอง

หน่วยทดลองที่จะเป็นตัวแทนที่ดีของประชากรควรมาจากการเลือกหน่วยทดลองตามความน่าจะเป็น (Probability Sampling) เพื่อให้การอ้างอิงไปยังประชากรสามารถเชื่อถือได้ แต่การเลือกหน่วยทดลองตามความน่าจะเป็นนั้นจะต้องทราบขนาดของประชากรทั้งหมด และรายชื่อของแต่ละหน่วยประชากร ทว่าผู้วิจัยไม่ทราบขนาดของประชากร ดังนั้นการเลือกหน่วยทดลองในงานวิจัยนี้จึงเป็นไปในเชิงไม่อิงกับความน่าจะเป็น (Non Probability Sampling) อีกทั้งข้อจำกัดของการใช้ห้องปฏิบัติการ ผู้วิจัยจึงจำเป็นต้องกำหนดให้หน่วยทดลองเป็นนิสิตปริญญาตรีของคณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย เพื่อเป็นตัวอย่างของผู้ใช้อินเทอร์เน็ตส่วนใหญ่ และมีพื้นฐานทางด้านการศึกษาใกล้เคียงกัน

3.3.3 จำนวนหน่วยทดลอง

วัตถุประสงค์ของงานนี้เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปซูลข้อความ เมื่อตัวแปรต้นทั้งสามตัวต่อไปนี้มีค่าแตกต่างกัน ได้แก่ (1) ระยะห่างระหว่างอักษรมีสามค่า ได้แก่ ปกติ, 0 พอยต์, และลบ 3.75 พอยต์ (2) ความหนาของเส้นรบกวนมีสามค่า ได้แก่ 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล และ (3) สีของเส้นรบกวนมีสองค่า ได้แก่ สีเดียวกับอักษร และสีเดียวกับพื้นหลัง ดังนั้นการทดลองนี้จึงประกอบด้วย $3*3*2 = 18$ กลุ่ม

งานวิจัยนี้กระทำในลักษณะของการทดลอง (Laboratory Experiment) ในห้องปฏิบัติการของศูนย์คอมพิวเตอร์ของคณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย เนื่องจากผู้วิจัยต้องการควบคุมตัวแปรอื่นๆที่ไม่เกี่ยวข้องกับการวิจัยให้คงที่ เพื่อให้ความแตกต่างที่อาจเกิดขึ้นเป็นผลจากตัวแปรต้น และเพื่อไม่ให้เกิดการเปรียบเทียบของหน่วยทดลองที่ได้รับทริตเมนต์ที่แตกต่างกันของการศึกษาตัวแปรในลักษณะเชิงปริมาณ (Quantitative) จำนวนหน่วยทดลองควรเป็นอย่างน้อยเจ็ดสิบสอง หน่วย (Nielsen J. & Pernice K., 2009) ดังนั้นจึงต้องมีหน่วยทดลองทั้งหมดอย่างน้อยเป็นจำนวน $18*20 = 360$ คน

3.3 การทดสอบสมมติฐาน

งานวิจัยนี้เป็นการวิจัยเชิงทดลอง (Experiment Research) โดยการทดลองอาจจะต้องสร้างสิ่งทดลองควบคุม เพื่อใช้เปรียบเทียบให้เห็นผลของสิ่งทดลองอื่นๆที่ต้องการวัดได้ชัดเจนขึ้น (กัลยา วาณิชยบัญชา, 2553) ผู้วิจัยได้กำหนดสมมติฐาน เพื่อสรุปลักษณะของประชากรโดยอ้างอิงผลจากหน่วยทดลอง จากงานวิจัยในอดีต ผู้วิจัยสามารถตั้งสมมติฐานจำนวนหกสมมติฐาน ดังต่อไปนี้

1. การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นปกติ, ซิด (0 พอยต์), และทับซ้อน (ลบ 3.75 พอยต์)

การลดระยะห่างระหว่างอักษรจนตัวอักษรซ้อนทับกัน คือค่าระยะห่างระหว่างอักษรเป็นค่าติดลบเป็นวิธีป้องกันการโจมตีด้านการแบ่งส่วนที่ดีที่สุด แต่การลดระยะห่างระหว่างอักษรจนมีระยะห่างติดลบมากกว่า 3.75 พอยต์ จะทำให้ความสามารถในการอ่านชุดอักษรของมนุษย์ลดลงอย่างมาก (Bursztein E. et al., 2011) ดังนั้น ผู้วิจัยจึงตั้งสมมติฐานได้ดังนี้

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นปกติ, ซิด, และทับซ้อน ไม่แตกต่างกัน

H_1 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ, ซิด, และทับซ้อน แตกต่างกัน โดยมีสมมติฐานย่อยเป็นรายคู่ดังนี้ (1) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ สูงกว่าชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักขรแบบซิด (2) เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ สูงกว่าชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักขรแบบทับซ้อน แต่ (3) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรแบบซิด ไม่เท่ากับชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักขรแบบทับซ้อน

2. การเปรียบเทียบอัตราความทนทานของแคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ, ซิด (0 พอยต์), และทับซ้อน (ลบ 3.75 พอยต์)

Bursztein et al. (2011) ได้แนะนำว่าเทคนิคหลักที่ควรนำมาใช้เพื่อป้องกันแคปต์ชาข้อความจากการโจมตีด้านการแบ่งส่วน คือการลดระยะห่างระหว่างอักขร อีกทั้ง El Ahmad et al. (2011) ได้วิเคราะห์กระบวนการแบ่งส่วนของโปรแกรมอัตโนมัติ พบว่าเมื่อชุดอักษรแคปต์ชาข้อความที่ตัวอักษรซ้อนทับกันจะทำให้เกิดข้อผิดพลาดในกระบวนการแบ่งส่วนของโปรแกรมอัตโนมัติ ดังนั้นผู้วิจัยจึงตั้งสมมติฐานได้ดังนี้

H_0 : อัตราความทนทานของแคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ, ซิด, และทับซ้อน ไม่แตกต่างกัน

H_1 : อัตราความทนทานของแคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ, ซิด, และทับซ้อน แตกต่างกัน โดยมีสมมติฐานย่อยเป็นรายคู่ดังนี้ (1) อัตราความทนทานของแคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ ต่ำกว่าชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักขรแบบซิด (2) เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรเป็นปกติ ต่ำกว่าชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักขรแบบทับซ้อน แต่ (3) อัตราความทนทานของแคปต์ชาข้อความ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักขรแบบซิด ไม่เท่ากับชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักขรแบบทับซ้อน

3. การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล

Bursztein et al. (2014) พบว่าเมื่อเพิ่มเส้นรบกวนหนา 1 ถึง 2 พิกเซล เข้าไปในลักษณะเส้นตรงที่พาดข้ามชุดอักษรแคปซูลข้อความแล้ว มนุษย์ยังสามารถอ่านและตอบแคปซูลข้อความได้ถูกต้อง โดยประสิทธิภาพในการอ่านและตอบแคปซูลข้อความของมนุษย์เทียบเท่ากับแคปซูลข้อความที่ไม่มีเส้นรบกวนพาดข้าม และการเพิ่มเส้นรบกวนเป็นคุณสมบัติการป้องกันการโจมตีด้านการแบ่งส่วนของแคปซูลข้อความจากโปรแกรมอัตโนมัติ นอกจากนี้ Yan & El Ahmad (2008) ได้อธิบายกระบวนการแบ่งส่วนของโปรแกรมอัตโนมัติ โดยในขั้นตอนการวิเคราะห์เพื่อนำสิ่งรบกวนออกจากตัวอักษรแคปซูลข้อความ โปรแกรมจะเกิดข้อผิดพลาดในการนำเส้นรบกวนออกเมื่อเส้นรบกวนมีความหนามาก (thick) ผู้วิจัยจึงตั้งสมมติฐานได้ดังนี้

H₀: อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล ไม่แตกต่างกัน

H₁: อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล แตกต่างกัน โดยมีสมมติฐานย่อยเป็นรายคู่ดังนี้ (1) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล (2) เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล และ (3) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล สูงกว่าชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

4. การเปรียบเทียบอัตราความทนทานของแคปซูลข้อความ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล

Bursztein et al. (2014) พบว่าการเพิ่มเส้นรบกวนเป็นคุณสมบัติการป้องกันการโจมตีด้านการแบ่งส่วนของแคปซูลข้อความจากโปรแกรมอัตโนมัติ และ Yan & El Ahmad (2008) ได้อธิบายกระบวนการแบ่งส่วนของโปรแกรมอัตโนมัติ โดยในขั้นตอนการวิเคราะห์เพื่อนำสิ่งรบกวนออกจากตัวอักษรแคปซูลข้อความ โปรแกรมจะเกิดข้อผิดพลาดในการนำเส้นรบกวนออกเมื่อเส้นรบกวนมีความหนามาก (thick) ผู้วิจัยจึงตั้งสมมติฐานได้ดังนี้

H₀: อัตราความทนทานของแคปซูลข้อความ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล ไม่แตกต่างกัน

H_1 : อัตราความทนทานของแคปต์ซาข้อความ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล แตกต่างกัน โดยมีสมมติฐานย่อยเป็นรายคู่ ดังนี้ (1) อัตราความทนทานของแคปต์ซาข้อความ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล ต่ำกว่าชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล (2) เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล ต่ำกว่าชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล แต่ (3) อัตราความทนทานของแคปต์ซาข้อความ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล ไม่เท่ากับชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

5. การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และสีเดียวกับพื้นหลังของแคปต์ซาข้อความ ผู้วิจัยยังไม่พบผลการวิเคราะห์เชิงประจักษ์ (empirical) เกี่ยวกับสีของเส้นรบกวนต่อความสามารถในการตอบแคปต์ซาของมนุษย์ เพื่อวิเคราะห์ความสามารถในการตอบแคปต์ซาของมนุษย์การวิจัยนี้จึงเลือกศึกษาสีของเส้นรบกวนสองลักษณะคือ เมื่อเส้นรบกวนมีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และสีเดียวกับพื้นหลังของแคปต์ซาข้อความ ถึงแม้ว่าเส้นรบกวนที่มีสีเดียวกับตัวอักษรอาจทำให้บังเส้นของตัวอักษรบางส่วน และเส้นรบกวนที่มีสีเดียวกับพื้นหลังอาจทำให้เส้นของตัวอักษรบางส่วนขาดไป แต่เนื่องจากทฤษฎีความสมบูรณ์ของเกสตัลที่ได้กล่าวไว้ว่าข้อมูลที่ขาดหายไปมนุษย์สามารถรับรู้ให้เป็นภาพสมบูรณ์ได้โดยอาศัยประสบการณ์เดิม (Wasserstein et al., 1987) ผู้วิจัยจึงตั้งสมมติฐานได้ดังนี้

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และชุดอักษรแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับพื้นหลังของแคปต์ซาข้อความ ไม่แตกต่างกัน

H_1 : อัตราการยืนยันความเป็นมนุษย์แตกต่างกัน เมื่อชุดอักษรแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และชุดอักษรแคปต์ซาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับพื้นหลังของแคปต์ซาข้อความ แตกต่างกัน

6. การเปรียบเทียบอัตราความทนทานของแคปต์ซาข้อความ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และสีเดียวกับพื้นหลังของแคปต์ซาข้อความ

ผู้วิจัยยังไม่พบผลการวิเคราะห์เชิงประจักษ์ (empirical) เกี่ยวกับสีของเส้นรบกวนต่อความทนทานของแคปซูลข้อความต่อโปรแกรมอัตโนมัติ เพื่อวิเคราะห์ความทนทานของแคปซูลข้อความเมื่อเส้นรบกวนมีสีเดียวกับตัวอักษรแคปซูลข้อความ และสีเดียวกับพื้นหลังของแคปซูลข้อความ เนื่องจากเมื่อเส้นรบกวนมีสีเดียวกับตัวอักษรก็จะสร้างความสับสนในการวิเคราะห์ว่าเส้นใดเป็นเส้นของตัวอักษร เส้นใดเป็นเส้นรบกวน และเมื่อเส้นรบกวนมีสีเดียวกับพื้นหลังจะทำให้เส้นที่ประกอบเป็นตัวอักษรบางช่วงขาดหายไป ไม่ได้ประกอบเป็นลักษณะปกติของอักษร ผู้วิจัยจึงตั้งสมมติฐานได้ดังนี้

H_0 : อัตราความทนทานของแคปซูลข้อความ เมื่อชุดอักษรแคปซูลข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปซูลข้อความ และชุดอักษรแคปซูลข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับพื้นหลังของแคปซูลข้อความ ไม่แตกต่างกัน

H_1 : อัตราความทนทานของแคปซูลข้อความแตกต่างกัน เมื่อชุดอักษรแคปซูลข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปซูลข้อความ และชุดอักษรแคปซูลข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับพื้นหลังของแคปซูลข้อความ แตกต่างกัน

3.4 เครื่องมือเก็บข้อมูล

3.4.1 ตัวแปรที่ต้องการเก็บ

งานวิจัยนี้จัดทำขึ้นโดยมีวัตถุประสงค์เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปซูลข้อความที่ออกแบบแตกต่างกันตามตัวแปรอิสระสามตัวแปร ซึ่งมีรูปแบบแคปซูลข้อความที่ต้องพัฒนาจำนวน 18 เงื่อนไข (แสดงดังตารางที่ 3.2) โดยงานวิจัยนี้เลือกใช้อักษรพิมพ์เล็ก แบบอักษร Arial ขนาด 24 พอยต์ จำนวนหกตัวอักษรแบบไม่ซ้ำกัน ทั้งนี้ผู้วิจัยกำหนดค่าของตัวแปรอิสระเป็นดังนี้

1. ระยะห่างระหว่างอักษร คือ ระยะห่างระหว่างอักษรที่อยู่ต่อเนื่องกัน (Bursztein E. et al., 2014) วัดจากระยะห่างด้านขวาของภาพตัวอักษรแรกรวมกับระยะห่างด้านซ้ายของภาพตัวอักษรหลัง (วงศกร รัตติรุจิเสวต, 2557) โดยจำแนกค่าของตัวแปรได้สามค่าคือ (1) ปกติ (2) 0 พอยต์ หรือให้อักษรอยู่ชิดกันพอดี และ (3) ลบ 3.75 พอยต์ หรือให้อักษรอยู่เกยกัน โดยรายละเอียดของแต่ละค่าเป็นดังนี้

- ก. ระยะห่างปกติ คือระยะห่างระหว่างอักษรที่เป็นค่าปกติของตัวอักษรแต่ละตัว ภาพที่ 3.1 แสดงตัวอย่างระยะห่างระหว่างอักษรของแคปซูลข้อความ โดยระยะห่างระหว่าง bwexzj มีค่าปกติ

และตารางที่ 3.1 แสดงค่าระยะห่างด้านซ้ายและระยะห่างด้านขวาของตัวอักษรที่เป็นค่าปกติแต่ละตัวของแบบอักษร Arial

bwexzj

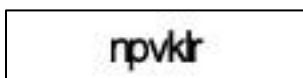
ภาพที่ 3. 1 ตัวอย่างแคปซันข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าปกติ

ตารางที่ 3. 1 ค่าระยะห่างด้านซ้าย และระยะห่างด้านขวาของตัวอักษรพิมพ์เล็ก แบบอักษร Arial (มีหน่วยเป็นพอยต์ ทั้งนี้การคำนวณค่าระยะห่างจากหน่วย U เป็นหน่วย พอยต์ อยู่ในสูตรที่ 1.2 ของบทที่ 1)

อักษร	ระยะห่างด้านซ้าย	ระยะห่างด้านขวา	อักษร	ระยะห่างด้านซ้าย	ระยะห่างด้านขวา	อักษร	ระยะห่างด้านซ้าย	ระยะห่างด้านขวา
a	0.58	0.68	j	-0.73	1.10	s	0.49	0.62
b	1.05	0.66	k	1.06	0.06	t	0.28	0.12
c	0.63	0.15	l	1.02	1.13	u	1.02	1.15
d	0.55	1.16	m	1.05	1.03	v	0.20	0.19
e	0.59	0.66	n	1.05	1.10	w	0.05	0.13
f	0.15	-0.55	o	0.53	0.59	x	0.12	0.12
g	0.52	1.07	p	1.05	0.64	y	0.26	0.14
h	1.05	1.09	q	0.56	1.15	z	0.31	0.34
i	1.06	1.09	r	1.04	-0.22			

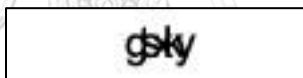
ข. ระยะห่าง 0 พอยต์ คือการลดระยะห่างระหว่างอักษรสองตัวที่อยู่ติดกันให้ชิดกัน โดยให้ระยะห่างด้านขวาของภาพอักษรแรกและระยะห่างด้านซ้ายของภาพอักษรหลังมีค่าเป็น 0 พอยต์ จากตารางที่ 1.1 หากต้องการลดระยะห่างระหว่างอักษร n กับ p ให้เหลือ 0 พอยต์ จะต้องลดค่าระยะห่างทางด้านขวาของอักษร n จำนวน 1.10 พอยต์ และลดค่าระยะห่างทางด้านซ้ายของอักษร p จำนวน 1.05 พอยต์ ภาพที่ 3.2 แสดงตัวอย่างระยะห่างระหว่างอักษรของแคปซันข้อความ โดยระยะห่างระหว่าง npvklr มีค่า 0 พอยต์

สำหรับบางตัวอักษรที่มีระยะห่างด้านขวา หรือระยะห่างด้านซ้ายติดลบ เช่น ตัวอักษร f มีค่าระยะห่างด้านขวาเท่ากับลบ 0.55 พอยต์ หากตัวอักษร f อยู่ทางด้านซ้ายของตัวอักษรใด การลดระยะห่างจนชิดติดกันจะหมายถึง การลดระยะห่างด้านซ้ายของตัวอักษรที่อยู่ทางด้านขวาของตัวอักษร f เท่านั้น



ภาพที่ 3. 2 ตัวอย่างแคปซันข้อความที่มีระยะห่างระหว่างอักษรเป็นค่า 0 พอยต์

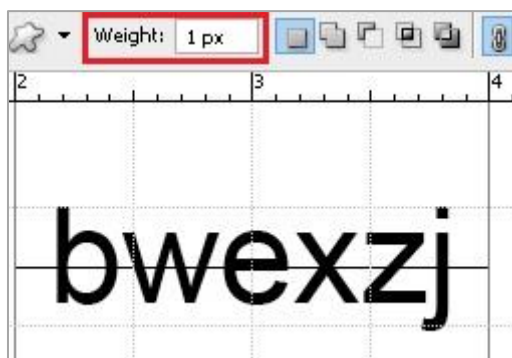
ค. ระยะห่างลบ 3.75 พอยต์ คือการลดระยะห่างระหว่างอักษรสองตัวที่อยู่ติดกันจนมีค่าลบ 3.75 พอยต์ เนื่องจากระยะห่างระหว่างอักษรที่มีค่าลดลงน้อยกว่าลบ 3.75 พอยต์ จะทำให้ความสามารถในการยืนยันความเป็นมนุษย์ลดลง (Bursztein E. et al., 2011) ภาพที่ 3.3 แสดงตัวอย่างระยะห่างระหว่างอักษรของแคปซันข้อความ โดยระยะห่างระหว่าง gtsxky มีค่าลบ 3.75 พอยต์



ภาพที่ 3. 3 ตัวอย่างแคปซันข้อความที่มีระยะห่างระหว่างอักษรเป็นค่าลบ 3.75 พอยต์

2. ความหนาของเส้นรบกวน คือ ความหนาของเส้นที่เพิ่มเข้าไปในลักษณะของเส้นตรงที่พาดข้ามชุดอักษรแคปซันข้อความทั้งหมด ในงานวิจัยนี้ผู้วิจัยกำหนดความยาวของเส้นรบกวนเป็น 3 เซนติเมตร เนื่องจากสามารถพาดข้ามตัวอักษร m จำนวนหกตัว ซึ่งตัวอักษร m เป็นตัวอักษรที่มีค่าความกว้างของอักษรมากที่สุด โดยผู้วิจัยได้กำหนดความหนาของเส้นด้วยโปรแกรม Adobe Photoshop CS3 โดยจำแนกค่าของตัวแปรได้สามค่าคือ (1) 1 พิกเซล (2) 3 พิกเซล และ (3) 5 พิกเซล ดังรายละเอียดต่อไปนี้

ก. เส้นรบกวนหนา 1 พิกเซล คือ การกำหนดค่า Weight ของเส้นตรงด้วยโปรแกรม Adobe Photoshop CS3 ให้เท่ากับ 1 พิกเซล แสดงตัวอย่างการกำหนดความหนาของเส้นรบกวน ดังภาพที่ 3.4 และภาพที่ 3.5 แสดงตัวอย่างเส้นรบกวนที่พาดข้ามชุดอักษรแคปซันข้อความทั้งหมด โดยเส้นรบกวนความหนา 1 พิกเซล พาดข้ามอักษร bwexzj

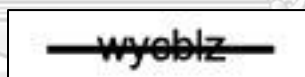


ภาพที่ 3. 4 ตัวอย่างการกำหนดความหนาของเส้นรบกวนเป็น 1 พิกเซล ด้วยโปรแกรม Adobe Photoshop CS3



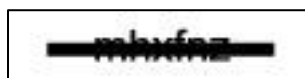
ภาพที่ 3. 5 ตัวอย่างเส้นรบกวนที่มีความหนา 1 พิกเซล พาดข้ามอักษร bwexzj

ข. เส้นรบกวนหนา 3 พิกเซล คือ การกำหนดค่า Weight ของเส้นตรงด้วยโปรแกรม Adobe Photoshop CS3 ให้เท่ากับ 3 พิกเซล ภาพที่ 3.6 แสดงตัวอย่างเส้นรบกวนที่พาดข้ามชุดอักษร แคปต์ซาข้อความทั้งหมด โดยเส้นรบกวนความหนา 3 พิกเซล พาดข้ามอักษร wycblz



ภาพที่ 3. 6 ตัวอย่างเส้นรบกวนที่มีความหนา 3 พิกเซล พาดข้ามอักษร wycblz

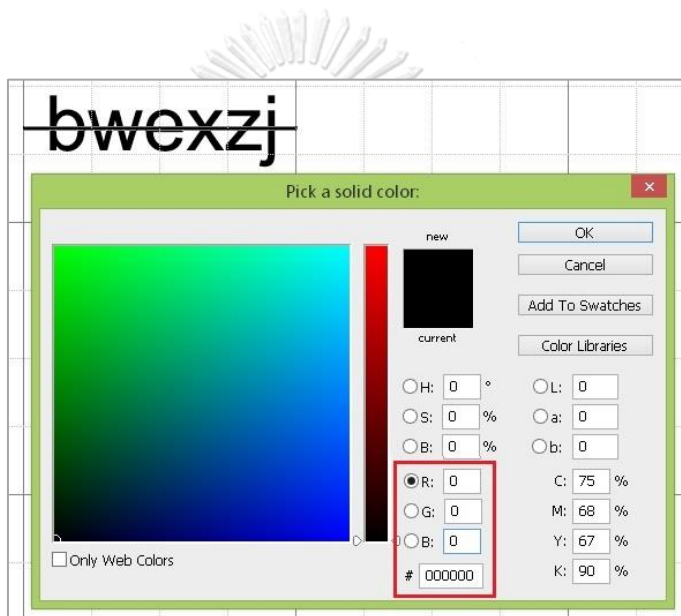
ค. เส้นรบกวนหนา 5 พิกเซล คือ การกำหนดค่า Weight ของเส้นตรงด้วยโปรแกรม Adobe Photoshop CS3 ให้เท่ากับ 5 พิกเซล ภาพที่ 3.7 แสดงตัวอย่างเส้นรบกวนที่พาดข้ามชุดอักษร แคปต์ซาข้อความทั้งหมด โดยเส้นรบกวนความหนา 5 พิกเซล พาดข้ามอักษร mhxfnz



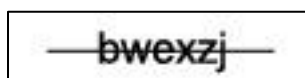
ภาพที่ 3. 7 ตัวอย่างเส้นรบกวนที่มีความหนา 5 พิกเซล พาดข้ามอักษร mhxfnz

3. สีของเส้นรบกวน คือ สีของเส้นรบกวนที่พาดข้ามชุดอักษรแคปทชาข้อความ ผู้วิจัย กำหนดสีของเส้นด้วยโปรแกรม Adobe Photoshop CS3 โดยจำแนกค่าของตัวแปรได้สองค่าคือ (1) สีเดียวกับอักษร และ (2) สีเดียวกับพื้นหลัง ดังรายละเอียดต่อไปนี้

ก. เส้นรบกวนสีเดียวกับอักษร คือ การกำหนดสีของเส้นตรงให้มีค่าเดียวกับสีของตัวอักษรของแคปทชาข้อความ โดยงานวิจัยนี้ใช้ตัวอักษรสีดำ มีรหัสสี #000000 แสดงตัวอย่างการกำหนดสีของเส้นรบกวนให้เป็นสีดำด้วยโปรแกรม Adobe Photoshop CS3 ดังภาพที่ 3.8 และภาพที่ 3.9 แสดงตัวอย่างเส้นรบกวนที่พาดข้ามชุดอักษรแคปทชาข้อความทั้งหมด โดยเส้นรบกวนขนาด 1 พิกเซล สีดำพาดข้ามอักษร bwexzj สีดำเหมือนกัน

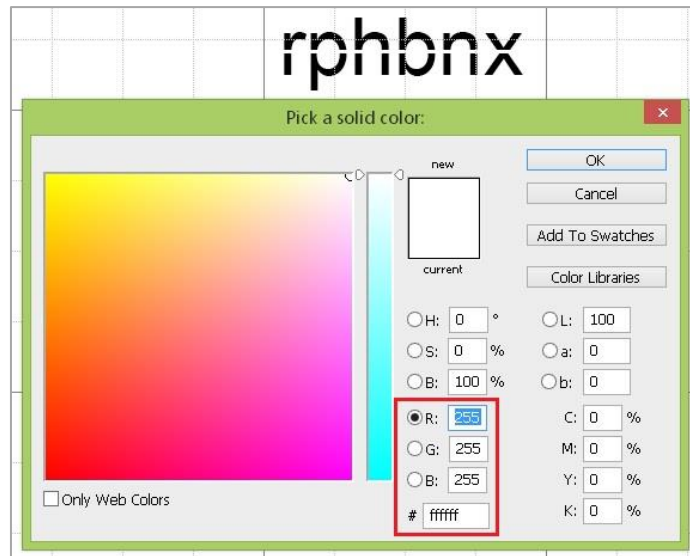


ภาพที่ 3. 8 ตัวอย่างการกำหนดสีของเส้นรบกวนให้เป็นสีดำด้วยโปรแกรม Adobe Photoshop CS3

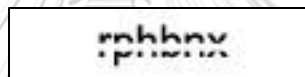


ภาพที่ 3. 9 ตัวอย่างเส้นรบกวนขนาด 1 พิกเซล สีดำพาดข้ามอักษร bwexzj

ข. เส้นรบกวนสีเดียวกับพื้นหลัง คือ การกำหนดสีของเส้นตรงให้มีค่าเดียวกับสีของพื้นหลังของแคปทชาข้อความ โดยงานวิจัยนี้ใช้พื้นหลังสีขาว มีรหัสสี #FFFFFF แสดงตัวอย่างการกำหนดสีของเส้นรบกวนให้เป็นสีขาวด้วยโปรแกรม Adobe Photoshop CS3 ดังภาพที่ 3.10 และภาพที่ 3.11 แสดงตัวอย่างเส้นรบกวนที่พาดข้ามชุดอักษรแคปทชาข้อความทั้งหมด โดยเส้นรบกวนขนาด 1 พิกเซล สีขาวพาดข้ามอักษร rphbnx สีดำบนพื้นหลังสีขาว



ภาพที่ 3. 10 ตัวอย่างการกำหนดสีของเส้นรบกวนให้เป็นสีขาวด้วยโปรแกรม Adobe Photoshop CS3



ภาพที่ 3. 11 ตัวอย่างเส้นรบกวนขนาด 3 พิกเซล สีขาวพาดข้ามอักษร rphbnx

ตารางที่ 3. 2 รูปแบบแคปต์ซาข้อความที่พัฒนามีจำนวน 18 เงื่อนไข

แบบที่	ระยะห่างระหว่างอักษร			ความหนาของเส้นรบกวน			สีของเส้นรบกวน	
	ปกติ	0 พอยต์	ลบ 3.75 พอยต์	1 พิกเซล	2 พิกเซล	3 พิกเซล	ดำ	ขาว
1	✓	-	-	✓	-	-	✓	-
2	✓	-	-	✓	-	-	-	✓
3	✓	-	-	-	✓	-	✓	-
4	✓	-	-	-	✓	-	-	✓
5	✓	-	-	-	-	✓	✓	-
6	✓	-	-	-	-	✓	-	✓
7	-	✓	-	✓	-	-	✓	-

ตารางที่ 3.2 รูปแบบแคปต์ซาข้อความที่พัฒนามีจำนวน 18 เงื่อนไข (ต่อ)

แบบที่	ระยะห่างระหว่างอักษร			ความหนาของเส้นรบกวน			สีของเส้นรบกวน	
	ปกติ	0 พอยต์	ลบ 3.75 พอยต์	1 พิกเซล	2 พิกเซล	3 พิกเซล	ดำ	ขาว
8	-	✓	-	✓	-	-	-	✓
9	-	✓	-	-	✓	-	✓	-
10	-	✓	-	-	✓	-	-	✓
11	-	✓	-	-	-	✓	✓	-
12	-	✓	-	-	-	✓	-	✓
13	-	-	✓	✓	-	-	✓	-
14	-	-	✓	✓	-	-	-	✓
15	-	-	✓	-	✓	-	✓	-
16	-	-	✓	-	✓	-	-	✓
17	-	-	✓	-	-	✓	✓	-
18	-	-	✓	-	-	✓	-	✓

การออกแบบแคปต์ซาข้อความมีจุดประสงค์หลัก เพื่อยืนยันความเป็นมนุษย์ของผู้ใช้ ดังนั้น การออกแบบแคปต์ซาข้อความควรคำนึงถึงความเหมาะสมเพื่อให้มนุษย์สามารถเข้าใจได้ และใน ขณะเดียวกันต้องทนทานต่อการวิเคราะห์ของโปรแกรมอัตโนมัติเพื่อป้องกันการสร้างความเสียหาย จากการประสงค์ร้าย เพื่อวัดความสามารถของแคปต์ซาข้อความที่ออกแบบแตกต่างกันตามตัวแปร อิสระสามตัวแปรสามารถทำได้ โดยการวัดค่าตัวแปรตามดังนี้

1. อัตราการยืนยันความเป็นมนุษย์ (Human Affirmative Rate) หมายถึง จำนวนครั้งที่ผู้ใช้ สามารถระบุอักษรที่ปรากฏในแคปต์ซาข้อความได้ถูกต้องจนสามารถเข้าใช้บริการของเว็บไซต์ได้ โดย วัดจากสัดส่วนของจำนวนผู้ตอบแคปต์ซาข้อความได้ถูกต้องต่อจำนวนผู้ตอบแคปต์ซาข้อความทั้งหมด (Yan J. & El Ahmad A. S., 2008; มณีรัตน์ ชาติรังสรรค์, 2555)

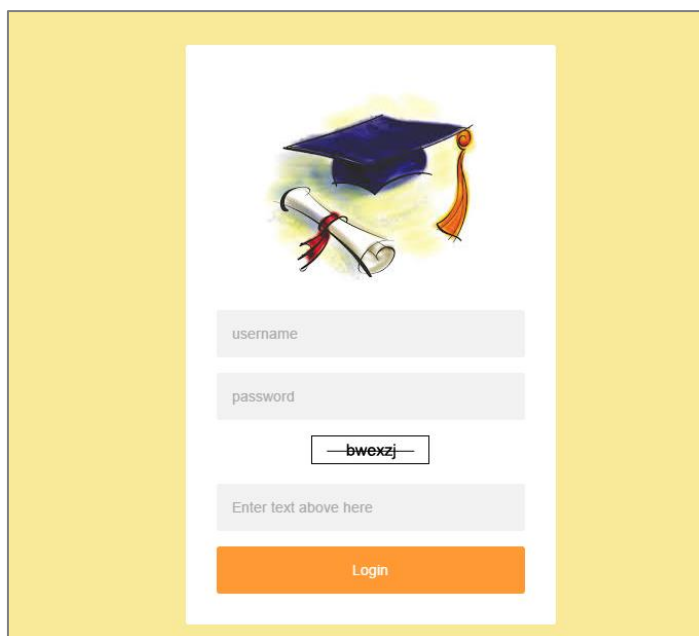
ผู้วิจัยได้พัฒนาเว็บไซต์เพื่อทดสอบตาบอดสี (รายละเอียดของเว็บไซต์นำเสนอในหัวข้อ 3.4.2) เมื่อผู้ใช้ต้องการเข้าสู่ระบบจะต้องกรอกแคปช่าข้อความที่ผู้ใช้ได้เห็น เพื่อเป็นเครื่องมือช่วยในการทดสอบความเป็นมนุษย์ของผู้ใช้ โดยหน่วยทดลองจะถูกกำหนดให้เข้าใช้งานเว็บไซต์ จากนั้นจะต้องเข้าสู่ระบบด้วยการกรอก username กับ password ที่ผู้วิจัยเตรียมไว้ให้ และหน่วยทดลองจะต้องตอบแคปช่าข้อความที่แสดงอยู่บนหน้าเว็บไซต์นั้น เพื่อให้สามารถกดปุ่ม Login เพื่อยืนยันการเข้าสู่ระบบ เมื่อหน่วยทดลองเข้าสู่ระบบได้เรียบร้อยแล้วจะต้องตอบคำถามเพื่อเก็บข้อมูลส่วนบุคคลสำหรับงานวิจัยครั้งนี้ และตอบแบบทดสอบตาบอดสี จากนั้นระบบจะแสดงผลการทดสอบแจ้งให้หน่วยทดลองทราบว่าตาบอดสีหรือไม่

2. อัตราความทนทานของแคปช่าข้อความ (Robustness Rate of CAPTCHA) หมายถึงความสามารถของแคปช่าข้อความในการป้องกันการโจมตีของโปรแกรมอัตโนมัติได้ โดยวัดจากอัตราการตอบแคปช่าข้อความที่ไม่ถูกต้องของโปรแกรมอัตโนมัติต่อจำนวนการตอบแคปช่าข้อความทั้งหมด (Yan J. & El Ahmad A. S., 2008; มณีรัตน์ ชาตธีรธรรม, 2555)

มณีรัตน์ ชาตธีรธรรม (2555) และ วงศกร รัตติรุจิเสวต (2557) ได้เลือกใช้โปรแกรม Omnipage Professional 18 เป็นการเลือกใช้คุณสมบัติของโปรแกรมโอซีอาร์มาช่วยในการวิเคราะห์ชุดอักขระของแคปช่าข้อความ ผู้วิจัยเห็นว่าในปัจจุบันมีโปรแกรมอัตโนมัติที่มีคุณสมบัติเพื่อวิเคราะห์แคปช่าข้อความโดยเฉพาะ โปรแกรมแคปช่าโซลเวอร์เป็นโปรแกรมอัตโนมัติที่มีคุณสมบัติการแปลภาพเป็นข้อความของโปรแกรมโอซีอาร์ และมีขั้นตอนการกำจัดสิ่งรบกวนจากชุดอักขระแคปช่าข้อความ เพื่อให้ในขั้นตอนการจำแนกตัวอักษร การปรับขนาดตัวอักษร และการระบุว่าเป็นตัวอักษรใดนั้นมีข้อผิดพลาดน้อยลง ผู้วิจัยจึงเลือกโปรแกรม GSA Captcha Breaker วัดอัตราความทนทานของแคปช่าข้อความ เพื่อให้ค่าที่วัดได้นั้นใกล้เคียงกับการสถานการณ์จริงที่สุด

3.4.2 เครื่องมือของการทดลอง

เมื่อเริ่มต้นการทดลองเว็บไซต์จะแสดงกล่องข้อความเพื่อตรวจสอบการเข้าสู่ระบบของหน่วยทดลอง โดยหน่วยทดลองจะต้องกรอก username และ password ที่ผู้วิจัยเตรียมไว้ให้ และถัดจากส่วนของข้อมูล username และ password จะแสดงรูปภาพของแคปช่าข้อความ หน่วยทดลองจะต้องกรอกตัวอักษรที่เห็นในภาพจึงจะสามารถกดปุ่ม Login เพื่อเข้าสู่ระบบ ดังภาพที่ 3.12 โดยแคปช่าข้อความที่แสดงในรูปนี้เป็นแคปช่าข้อความได้จากการสุ่มตัวอักษร bwexzj และกำหนดให้มีลักษณะเป็น (1) ระยะห่างระหว่างอักษรปกติ (2) ความหนาเส้นรบกวน 1 พิกเซล และ (3) เส้นรบกวนสีเดียวกับตัวอักษร



ภาพที่ 3. 12 ส่วนของเว็บไซต์สำหรับตรวจสอบการเข้าสู่ระบบ

หากหน่วยทดลองกรอก username และ password ได้ถูกต้อง และกรอกตัวอักษรแคปทชาข้อความแล้ว (โดยไม่จำเป็นจะต้องกรอกตัวอักษรแคปทชาได้ถูกต้อง) หน่วยทดลองจะสามารถเข้าสู่ระบบได้ ทั้งนี้ระบบจะนับและบันทึกว่าที่หน่วยทดลองสามารถตอบแคปทชาได้ถูกต้องหรือไม่สำหรับวัตถุประสงค์การยืนยันความเป็นมนุษย์ จากนั้นเว็บไซต์จะแสดงแบบสอบถามให้หน่วยทดลองตอบคำถาม ข้อมูลส่วนบุคคล ซึ่งจะถูกนำไปใช้ในการวิจัยครั้งนี้ มี 6 คำถาม ได้แก่ (1) เพศ (2) อายุ (3) หลักสูตร (4) ปัญหาตาบอดสี (5) ปัญหาเกี่ยวกับสายตา และ (6) การแก้ปัญหาเกี่ยวกับสายตาขณะทำการทดลอง ดังภาพที่ 3.13

ข้อมูลส่วนบุคคล

1. เพศ
 ชาย หญิง

2. อายุ
 ต่ำกว่า 20 ปี 20 ปีขึ้นไป

3. หลักสูตร
 บัญชี บริหารธุรกิจ สถิติ นานาชาติ

4. ปัญหาตาบอดสี
 มีปัญหาตาบอดสี ไม่มีปัญหาตาบอดสี
 ไม่ทราบว่ามีปัญหาหรือไม่

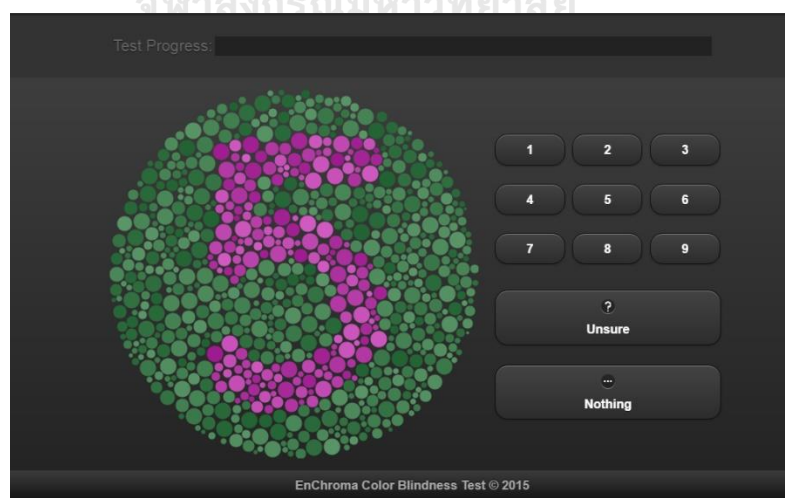
5. ปัญหาเกี่ยวกับสายตา
 สั้น ยาว เอียง ไม่มีปัญหาเกี่ยวกับสายตา

6. การแก้ปัญหาเกี่ยวกับสายตาขณะทำการทดลอง
 ใส่แว่น ใส่คอนแทคเลนส์ ยังไม่ต้องแก้ปัญหา

Next

ภาพที่ 3. 13 ส่วนของเว็บไซต์สำหรับตอบคำถามข้อมูลส่วนบุคคล

เมื่อหน่วยทดลองตอบคำถามข้อมูลส่วนบุคคลเรียบร้อยแล้วและกดปุ่ม Next เว็บไซต์จะแสดงแบบทดสอบตาบอดสี อันเป็นส่วนสุดท้ายของการทดลอง ผู้วิจัยเลือกใช้แบบทดสอบจากเว็บไซต์ enchroma.com ดังภาพที่ 3.14 โดยเมื่อหน่วยทดลองทำแบบทดสอบตาบอดสีเสร็จแล้ว ระบบจะแสดงผลการทดสอบแจ้งให้หน่วยทดลองทราบว่าตาบอดสีหรือไม่



ภาพที่ 3. 14 แบบทดสอบตาบอดสีจากเว็บไซต์ enchroma.com

3.5 ประเด็นความเชื่อถือได้ (Reliability) และความถูกต้อง (Validity)

งานวิจัยนี้เป็นการทดลองในห้องปฏิบัติการเพื่อตอบวัตถุประสงค์ทั้งหมดของงานวิจัย ดังนั้น สิ่งที่สำคัญที่สุดคือ การควบคุมการทดลองให้ได้ข้อมูลที่ดีสำหรับนำไปสู่ผลสรุปที่ถูกต้องและน่าเชื่อถือได้ ผู้วิจัยได้ควบคุมปัจจัยต่างๆ ดังต่อไปนี้

1. การกำหนดหน่วยทดลอง ผู้วิจัยเลือกนิสิตปริญญาตรีของคณะพาณิชยศาสตร์และการบัญชี หลักสูตรภาษาไทยและหลักสูตรนานาชาติเป็นหน่วยทดลอง เนื่องจากมีลักษณะความสามารถที่ทัดเทียมกัน กล่าวคือหน่วยทดลองจะเรียนในรายวิชาที่มีลักษณะเหมือนหรือคล้ายคลึงกัน มีอายุและประสบการณ์ใกล้เคียงกัน (ขวัญหทัย สันติบุตร, 2549; มณีรัตน์ ชาตธีรังสรรค์, 2555) ทำให้ได้หน่วยทดลองที่มีความใกล้เคียงกันมากที่สุด โดยหน่วยทดลองแต่ละคนจะถูกเลือกมาเพื่อตอบแคปซูลเพียงครั้งเดียว และขอความร่วมมือจากหน่วยทดลองแต่ละคนไม่ให้พูดคุยกัน หรือแลกเปลี่ยนข้อมูลระหว่างการทดลอง อีกทั้งไม่ให้กลุ่มตัวอย่างเดิมเข้ามาร่วมให้ข้อมูลอีก นอกจากนี้กลุ่มตัวอย่างกลุ่มเดียวกันจะถูกเก็บข้อมูลพร้อมกันในห้องเดียวกัน การที่ผู้วิจัยควบคุมการเก็บข้อมูลตามที่กล่าวมานี้ จะทำให้ผลการทดลองมีข้อมูลที่ถูกต้องและเชื่อถือได้

2. ความเข้าใจในขั้นตอนการทดลอง ผู้วิจัยจะอธิบายการใช้งานเว็บไซต์โดยสรุป และแจกใบงาน (ดูตัวอย่างในภาคผนวก ก) เพื่อให้หน่วยทดลองสามารถปฏิบัติตามขั้นตอนตามที่ผู้วิจัยกำหนดได้อย่างถูกต้อง แต่ผู้วิจัยจะไม่อธิบายวัตถุประสงค์ที่แท้จริงของการทดลอง เพื่อไม่ให้เกิดความคลาดเคลื่อนของผลการทดลอง หากหน่วยทดลองทราบวัตถุประสงค์ของการทดลอง (Nielsen J. & Pernice K., 2009; วงศกร รัตติรุจิเสวต, 2557)

3. เทคโนโลยีการสื่อสาร อันเป็นตัวแปรสำคัญที่มีผลต่อการทดลอง โดยเฉพาะการส่งข้อมูลระหว่างกัน ได้แก่ ความเร็วในการประมวลผลเซิร์ฟเวอร์ ความเร็วในการประมวลผลของเครื่องผู้ใช้ ความเร็วของสายส่งข้อมูลระหว่างเครื่องเซิร์ฟเวอร์กับเครื่องผู้ใช้ และความเร็วในการส่งผ่านข้อมูลในเครือข่าย ปัจจัยทุกข้อดังกล่าวนี้ ผู้วิจัยได้พยายามลดความคลาดเคลื่อนในเรื่องความเร็วที่แตกต่างกันที่อาจเกิดขึ้นขณะที่หน่วยทดลองกำลังทำงานที่ได้รับมอบหมายให้น้อยที่สุด ดังนั้น จึงจำเป็นต้องทดลองในห้องปฏิบัติการ และสถานที่ทดลองคือห้องปฏิบัติการคอมพิวเตอร์ของคณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อันประกอบไปด้วย เครื่องคอมพิวเตอร์รุ่นเดียวกัน และความเร็วของสายส่งข้อมูลระหว่างเครื่องเซิร์ฟเวอร์กับเครื่องผู้ใช้ใกล้เคียงกันมากที่สุด เพื่อผลการทดลองที่ออกมาจะอยู่บนพื้นฐานของระบบที่เหมือนกันและน่าเชื่อถือ ตลอดจนถูกต้องในระดับที่ยอมรับได้ (ขวัญหทัย สันติบุตร, 2549; มณีรัตน์ ชาตธีรังสรรค์, 2555)

4. ลักษณะของแคปช่าข้อความ ผู้วิจัยได้กำหนดลักษณะคือ ใช้แบบอักษร Arial แบบพิมพ์ เล็กขนาด 15 พอยต์ และใช้อักษรสีดำบนพื้นหลังสีขาว โดยอ้างอิงจากงานของ Bursztein et al. (2014) ได้ศึกษาการใช้เส้นรบกวน เพื่อเพิ่มความทนทานให้กับแคปช่าข้อความ และเลือกใช้ลักษณะ ขนาดดังกล่าว

3.6 ขั้นตอนการเก็บข้อมูล

การเก็บข้อมูลของงานวิจัยนี้ทำในห้องปฏิบัติการของศูนย์คอมพิวเตอร์ โดยเลือกหน่วย ทดลองจากนิสิตปริญญาตรี คณะพาณิชยศาสตร์และการบัญชี จำนวนอย่างน้อย 360 คน สำหรับเป็น หน่วยทดลอง 18 เงื่อนไข จำนวนรูปแบบของแคปช่าที่พัฒนาในงานวิจัยนี้ เงื่อนไขอย่างน้อย 20 คน

ผู้วิจัยได้พัฒนาเว็บไซต์ตรวจสอบการเข้าสู่ระบบของผู้ใช้ (รายละเอียดของเว็บไซต์นำเสนอใน หัวข้อ 3.4.2) โดยหน่วยทดลองจะต้องกรอก username กับ password และตอบแคปช่าข้อความ ที่ผู้ใช้ได้เห็น เพื่อเป็นเครื่องมือช่วยในการทดสอบความเป็นมนุษย์ของผู้ใช้ จากนั้นกดปุ่ม Login เพื่อยืนยันการเข้าสู่ระบบ เมื่อหน่วยทดลองเข้าสู่ระบบได้เรียบร้อยแล้วจะต้องตอบคำถามเพื่อเก็บข้อมูล ส่วนบุคคลสำหรับงานวิจัยครั้งนี้ และตอบแบบทดสอบตาบอดสีด้วย

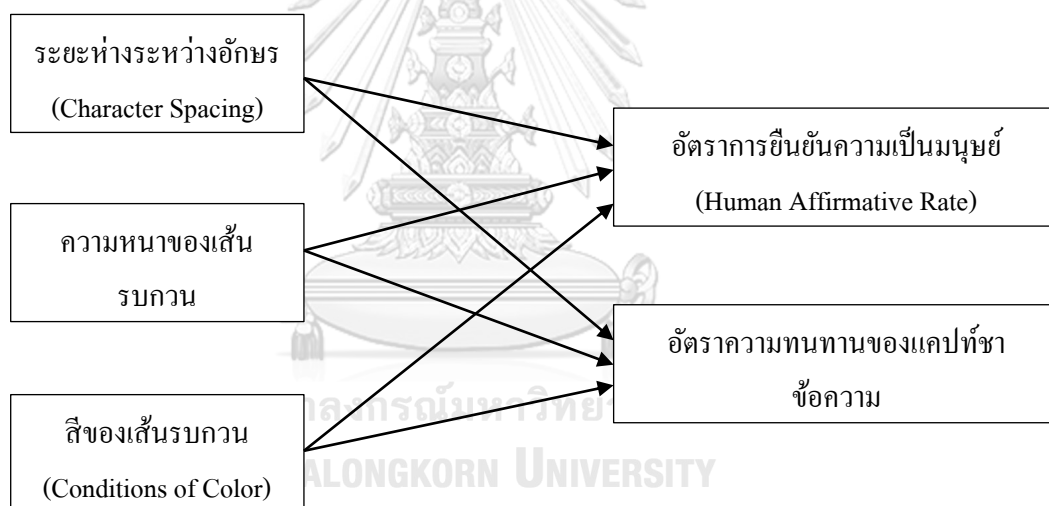
ก่อนเริ่มการทดลอง ผู้วิจัยจะแจกใบงานเพื่อแสดงขั้นตอนให้หน่วยทดลองปฏิบัติตาม (ดู ตัวอย่างในภาคผนวก ก) จากนั้นผู้วิจัยจะอธิบายการใช้งานเว็บไซต์โดยสรุป แต่จะไม่อธิบาย วัตถุประสงค์ที่แท้จริงของการทดลอง เพื่อไม่ให้เกิดความคลาดเคลื่อนของผลการทดลอง หากหน่วย ทดลองทราบวัตถุประสงค์ผลของการทดลอง (Nielsen J. & Pernice K., 2009; วงศกร รัตติรุจิเสวต, 2557)

การเก็บอัตราความทนทานของแคปช่าข้อความ จะกระทำเมื่อการทดลองการเก็บข้อมูล อัตราการยืนยันความเป็นมนุษย์ในแต่ละกลุ่มการทดลองเสร็จสิ้นลง เนื่องจากผู้วิจัยต้องการนำแคปช่า ข้อความชุดเดียวกันกับที่หน่วยทดลองได้ตอบในการทดลอง การควบคุมการเก็บข้อมูลอัตราความ ทนทานของแคปช่าข้อความนี้ เพื่อให้ได้ผลสรุปที่กระจ่างถูกต้องมากขึ้น เมื่อการทดลองแต่ละกลุ่ม เสร็จสิ้นลง ผู้วิจัยจะนำแคปช่าข้อความเดียวกันกับที่หน่วยทดลองได้ตอบในระหว่างการทดลอง มา ทดสอบอัตราความทนทานของแคปช่าข้อความด้วยโปรแกรม GSA Captcha Breaker

3.7 เทคนิคที่ใช้วิเคราะห์ข้อมูลที่เก็บได้

การวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์สามข้อของงานวิจัย และการเก็บรวบรวมข้อมูลของงานวิจัยเป็นการเก็บข้อมูลจากหน่วยทดลอง โดยหน่วยทดลองเข้าใช้เว็บไซต์ที่มีการแสดงแคปucha ข้อความ เพื่อวัดอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปucha ข้อความในรูปแบบต่างๆ จะกระทำดังรายละเอียดต่อไปนี้

สำหรับตัวแปรที่ศึกษาในงานวิจัยนี้แบ่งเป็นสองกลุ่ม คือ กลุ่มที่ 1 ตัวแปรต้นหรือตัวแปรอิสระ (Independent Variable) ได้แก่ (1) ระยะห่างระหว่างตัวอักษร (2) ความหนาของเส้นรบกวน และ (3) สีของเส้นรบกวน กลุ่มที่ 2 ตัวแปรตาม (Dependent Variable) ได้แก่ (1) อัตราการยืนยันความเป็นมนุษย์ และ (2) อัตราความทนทานของแคปucha ข้อความ โดยผู้วิจัยสนใจว่าตัวแปรอิสระมีอิทธิพลอย่างไรต่อตัวแปรตาม ภาพที่ 3.15 แสดงกรอบความคิดของงานวิจัยนี้



ภาพที่ 3. 15 กรอบความคิดของงานวิจัยนี้

ข้อมูลการยืนยันความเป็นมนุษย์ จะเก็บเมื่อหน่วยทดลองกรอกอักษรตามแคปucha ข้อความที่ปรากฏ หรือเรียกว่าการตอบแคปucha ข้อความ และกดปุ่มตกลง จากนั้นข้อมูลการตอบแคปucha ข้อความจะถูกบันทึกลงฐานข้อมูล งานวิจัยนี้จะเก็บข้อมูลการตอบแคปucha ข้อความของหน่วยทดลองทั้งหมด ไม่ว่าจะหน่วยทดลองจะตอบแคปucha ข้อความผิดหรือถูก เพื่อนำมาวิเคราะห์และตอบวัตถุประสงค์ โดยข้อมูลอัตราการยืนยันความเป็นมนุษย์จะคำนวณจากร้อยละของจำนวนการตอบ

แคปช่าข้อความได้ถูกต้องของหน่วยทดลองต่อจำนวนการตอบแคปช่าข้อความทั้งหมดของหน่วยทดลอง (มณีรัตน์ ชาตริงสรรค์, 2555) ตามสูตรที่ 3.1 ดังนี้

$$P = \frac{n}{N} \times 100 \quad \dots\dots\dots (3.1)$$

เมื่อ	P	คือ	อัตราการยืนยันความเป็นมนุษย์ (ร้อยละ)
	n	คือ	จำนวนการตอบแคปช่าข้อความได้ถูกต้องของหน่วยทดลอง
	N	คือ	จำนวนการตอบแคปช่าข้อความทั้งหมดของหน่วยทดลอง

ข้อมูลความทนทานของแคปช่าข้อความ จะเก็บจากการโจมตีแคปช่าข้อความของโปรแกรม GSA Captcha Breaker โดยหลังจากการทดลองเสร็จสิ้นลง ผู้วิจัยจะนำแคปช่าเดียวกันกับที่หน่วยทดลองได้โต้ตอบด้วยให้โปรแกรม GSA Captcha Breaker โจมตี ข้อมูลอัตราความทนทานของแคปช่าข้อความคำนวณจากร้อยละของจำนวนครั้งการโจมตีแคปช่าข้อความที่ไม่ถูกต้องต่อจำนวนครั้งการโจมตีแคปช่าข้อความทั้งหมด ตามสูตรที่ 3.2 ดังนี้

$$r = \frac{a}{A} \times 100 \quad \dots\dots\dots (3.2)$$

เมื่อ	r	คือ	อัตราความทนทานของแคปช่าข้อความ (ร้อยละ)
	a	คือ	จำนวนครั้งการโจมตีแคปช่าข้อความที่ไม่ถูกต้องของโปรแกรม GSA Captcha Breaker
	A	คือ	จำนวนครั้งการโจมตีแคปช่าข้อความทั้งหมดของโปรแกรม GSA Captcha Breaker (โดย A เป็นจำนวนเดียวกับ N)

ข้อมูลส่วนบุคคลของหน่วยทดลอง ได้แก่ (1) เพศ (2) อายุ (3) หลักสูตร (4) ลักษณะสายตา และ (5) การใช้อุปกรณ์ช่วยในการมองเห็น จะวิเคราะห์คำตอบของหน่วยทดลองด้วยสถิติเชิงพรรณนา (Descriptive Statistic)

ผู้วิจัยเลือกใช้วิธีการทางสถิติอิงพารามิเตอร์ (Parametric Statistical Technique) คือการใช้สถิติทดสอบ Z (Z-test) แบบทางเดียว เพื่อทดสอบสมมติฐานของผลต่างระหว่างสัดส่วนของประชากรสองประชากร และการใช้สถิติทดสอบ Z แบบสองทาง เพื่อทดสอบสมมติฐานของผลต่างระหว่างสัดส่วนของประชากรสองประชากรว่าแตกต่างกันหรือไม่ (กัลยา วาณิชย์บัญชา, 2553)



บทที่ 4

ผลการวิเคราะห์ข้อมูล

4.1 ความน่า

การนำเสนอผลการวิเคราะห์ข้อมูลที่ได้จากการทดลองนี้เพื่อใช้ตอบวัตถุประสงค์ที่ต้องการตรวจสอบผลของ (1) ระยะห่างระหว่างอักษร (2) ความหนาของเส้นรบกวน และ (3) สีของเส้นรบกวน ที่มีต่อ (1) อัตราการยืนยันความเป็นมนุษย์ และ (2) อัตราความทนทานของแคปTCHA ข้อความ

การนำเสนอประกอบด้วยการวิเคราะห์ข้อมูลลักษณะสถิติเชิงพรรณนา (Descriptive Statistic) และสถิติเชิงอนุมาน (Inferential Statistic)

4.2 ลักษณะทางประชากรของหน่วยทดลอง

งานวิจัยนี้เก็บข้อมูลในห้องปฏิบัติการของศูนย์คอมพิวเตอร์ คณะพาณิชยศาสตร์และการบัญชี โดยเก็บข้อมูลจากหน่วยทดลองจำนวน 360 คน ทั้งนี้ตารางที่ 4.1 ได้แสดงข้อมูลส่วนบุคคลของหน่วยทดลองจำแนกตามเพศ อายุ หลักสูตรที่ศึกษา ปัญหาตาบอดสี ปัญหาเกี่ยวกับสายตา และการแก้ไขปัญหาเกี่ยวกับสายตาขณะทำการทดลอง การจำแนกตามเพศ พบว่าหน่วยทดลองเป็นเพศชาย และเพศหญิงคิดเป็นร้อยละ 29.7 และ 70.3 ตามลำดับ การจำแนกตามอายุ พบว่าหน่วยทดลองอายุต่ำกว่า 20 ปี และอายุ 20 ปีขึ้นไปคิดเป็นร้อยละ 36.7 และ 63.3 ตามลำดับ การจำแนกตามหลักสูตรที่ศึกษา พบว่าหน่วยทดลองกำลังศึกษาอยู่ในหลักสูตรบัญชีบัณฑิตมากที่สุด คิดเป็นร้อยละ 46.7 รองลงมาคือ หลักสูตรบริหารธุรกิจบัณฑิต หลักสูตรสถิติศาสตร์บัณฑิต และหลักสูตรนานาชาติ คิดเป็นร้อยละ 40.8, 11.4, และ 1.1 ตามลำดับ การจำแนกตามปัญหาตาบอดสี พบว่าหน่วยทดลองอยู่ในกลุ่มไม่มีปัญหาด้านตาบอดสีมากที่สุด คิดเป็นร้อยละ 89.5 รองลงมาคือ ไม่ทราบว่าไม่มีปัญหา และมีปัญหาตาบอดสี คิดเป็นร้อยละ 7.2 และ 3.3 ตามลำดับ การจำแนกตามปัญหาเกี่ยวกับสายตา พบว่าหน่วยทดลองอยู่ในกลุ่มสายตาสั้นและเอียงมากที่สุด คิดเป็นร้อยละ 37.8 รองลงมาคือ สายตาปกติ สายตาสั้นเพียงอย่างเดียว สายตาเอียงเพียงอย่างเดียว และสายตาวาวและเอียง คิดเป็นร้อยละ 29.4, 29.2, 2.8, และ 0.8 ตามลำดับ และการจำแนกตามการแก้ไขปัญหาเกี่ยวกับสายตาขณะทำการทดลอง พบว่ามีหน่วยทดลองใส่แว่นสายตามากที่สุด คิดเป็นร้อยละ 48.6 รองลงมาคือ ไม่ต้องแก้ปัญหา และใส่คอนแทคเลนส์ คิดเป็นร้อยละ 37.5 และ 13.9 ตามลำดับ

ตารางที่ 4. 1 ข้อมูลส่วนบุคคลของหน่วยทดลองจำนวน 360 คน

คุณสมบัติ		จำนวน (คน)	ร้อยละ
เพศ	ชาย	107	29.7
	หญิง	253	70.3
อายุ	ต่ำกว่า 20 ปี	132	36.7
	อายุ 20 ปีขึ้นไป	228	63.3
หลักสูตร	บัญชี	168	46.7
	บริหารธุรกิจ	147	40.8
	สถิติศาสตร์	41	11.4
	นานาชาติ	4	1.1
ปัญหาตาบอดสี	มีปัญหาตาบอดสี	12	3.3
	ไม่มีปัญหาตาบอดสี	322	89.5
	ไม่ทราบว่ามีปัญหาหรือไม่	26	7.2
ปัญหาเกี่ยวกับสายตา	สายตาปกติ	106	29.4
	สายตาสั้น	105	29.2
	สายตาสั้นและเอียง	136	37.8
	สายตายาวและเอียง	3	0.8
	สายตาเอียง	10	2.8
การแก้ปัญหาเกี่ยวกับสายตาขณะทำการทดลอง	ใส่แว่น	175	48.6
	ใส่คอนแทคเลนส์	50	13.9
	ไม่ต้องแก้ปัญหา	135	37.5

ตารางที่ 4.2 แสดงข้อมูลปัญหาตาบอดสีของหน่วยทดลอง จำแนกตามผลการทดสอบตาบอดสี พบว่าหน่วยทดลองที่ตอบคำถามส่วนบุคคลว่ามีปัญหาตาบอดสีจำนวน 12 คน เมื่อตอบแบบทดสอบตาบอดสีแล้วได้ผลการทดสอบว่าตาบอดสีจริงคิดเป็นร้อยละ 3.3 ส่วนหน่วยทดลองที่ตอบคำถามส่วนบุคคลว่าไม่มีปัญหาตาบอดสีจำนวน 322 คน เมื่อตอบแบบทดสอบตาบอดสีแล้วก็ได้ผลการทดสอบว่าปกติคิดเป็นร้อยละ 89.5 และหน่วยทดลองที่ตอบคำถามส่วนบุคคลว่าไม่ทราบว่าตาบอดสีจำนวน 26 คน เมื่อตอบแบบทดสอบตาบอดสีแล้วได้ผลการทดสอบว่าปกติ และตาบอดสีคิดเป็นร้อยละ 7.0 และ 0.3 ตามลำดับ

ตารางที่ 4. 2 ข้อมูลปัญหาตาบอดสี จำแนกตามผลการทดสอบตาบอดสีจากแบบทดสอบ

ปัญหาตาบอดสี	ผลจากการทดสอบ			
	จำนวน (คน)		ร้อยละ (เมื่อเทียบกับจำนวนทั้งหมด 360 คน)	
	ปกติ	ตาบอดสีจริง	ปกติ	ตาบอดสีจริง
มีปัญหาตาบอดสี	-	12	-	3.3
ไม่มีปัญหาตาบอดสี	322	-	89.5	-
ไม่ทราบว่าปัญหาหรือไม่	25	1	7.0	0.3

4.3 อัตราการยืนยันความเป็นมนุษย์

ผู้วิจัยได้นำเสนออัตราการยืนยันความเป็นมนุษย์จำแนกตามรูปแบบแคปต์ชาข้อความที่ได้ นำเสนอไปแล้วในตารางที่ 3.2 และจำแนกตามตัวแปรที่ศึกษา ได้แก่ (1) ระยะห่างระหว่างอักษร (2) ความหนาของเส้นรบกวน และ (3) สีของเส้นรบกวน ดังตารางที่ 4.3 ถึง ตารางที่ 4.6 โดยในตารางที่ 4.3 เป็นการจำแนกตามรูปแบบแคปต์ชาข้อความ ตารางที่ 4.4 เป็นการจำแนกตามระยะห่างระหว่างอักษร ตารางที่ 4.5 เป็นการจำแนกตามความหนาของเส้นรบกวน และตารางที่ 4.6 เป็นการจำแนกตามสีของเส้นรบกวน

ตารางที่ 4. 3 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามรูปแบบแคปท์ชาข้อความ โดยมีจำนวนผู้ตอบแคปท์ชาข้อความรูปแบบละ 20 คน

รูปแบบ แคปท์ชา ข้อความ	ระยะห่าง ระหว่างอักษร (พอยต์)	ความหนาของ เส้นรบกวน (พิกเซล)	สีของ เส้นรบกวน	จำนวนผู้ตอบ ถูกต้อง (คน)	อัตราการยืนยัน ความเป็นมนุษย์ (%)
1	ปกติ	1	สีเดียวกับ ตัวอักษร	20	100
2	ปกติ	1	สีเดียวกับ พื้นหลัง	20	100
3	ปกติ	3	สีเดียวกับ ตัวอักษร	17	85
4	ปกติ	3	สีเดียวกับ พื้นหลัง	9	45
5	ปกติ	5	สีเดียวกับ ตัวอักษร	9	45
6	ปกติ	5	สีเดียวกับ พื้นหลัง	9	45
7	ขีด	1	สีเดียวกับ ตัวอักษร	19	95
8	ขีด	1	สีเดียวกับ พื้นหลัง	19	95
9	ขีด	3	สีเดียวกับ ตัวอักษร	20	100
10	ขีด	3	สีเดียวกับ พื้นหลัง	11	55
11	ขีด	5	สีเดียวกับ ตัวอักษร	8	40

ตารางที่ 4.3 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามรูปแบบแคปTCHAข้อความ โดยมีจำนวนผู้ตอบแคปTCHAข้อความรูปแบบละ 20 คน (ต่อ)

รูปแบบแคปTCHAข้อความ	ระยะห่างระหว่างอักษร (พอยต์)	ความหนาของเส้นรบกวน (พิกเซล)	สีของเส้นรบกวน	จำนวนผู้ตอบถูกต้อง (คน)	อัตราการยืนยันความเป็นมนุษย์ (%)
12	ชิด	5	สีเดียวกับพื้นหลัง	5	25
13	ทับซ้อน	1	สีเดียวกับตัวอักษร	8	40
14	ทับซ้อน	1	สีเดียวกับพื้นหลัง	12	60
15	ทับซ้อน	3	สีเดียวกับตัวอักษร	8	40
16	ทับซ้อน	3	สีเดียวกับพื้นหลัง	8	40
17	ทับซ้อน	5	สีเดียวกับตัวอักษร	5	25
18	ทับซ้อน	5	สีเดียวกับพื้นหลัง	5	25

ตารางที่ 4.3 พบว่าแคปTCHAข้อความในสามลักษณะที่มีอัตราการยืนยันความเป็นมนุษย์ที่สูงที่สุด ได้แก่ (1) แคปTCHAข้อความรูปแบบที่ 1 อันเป็นแคปTCHAข้อความที่มีระยะห่างระหว่างอักษรปกติ ถูกพาดข้ามด้วยเส้นรบกวนหนา 1 พิกเซล โดยเส้นรบกวนมีสีเดียวกับตัวอักษร (สีดำ) (2) แคปTCHAข้อความรูปแบบที่ 2 อันเป็นแคปTCHAข้อความที่มีระยะห่างระหว่างอักษรปกติ ถูกพาดข้ามด้วยเส้นรบกวนหนา 1 พิกเซล โดยเส้นรบกวนมีสีเดียวกับพื้นหลัง (สีขาว) และ (3) แคปTCHAข้อความรูปแบบที่ 9 อันเป็นแคปTCHAข้อความที่มีระยะห่างระหว่างอักษรชิดติดกัน (0 พอยต์) ถูกพาดข้ามด้วยเส้น

รบกวนหนา 3 พิกเซล โดยเส้นรบกวนมีสีเดียวกับตัวอักษร (สีดำ) โดยที่อัตราการยืนยันความเป็นมนุษย์ที่สูงที่สุดคิดเป็นร้อยละ 100 เท่ากันทั้งสามลักษณะ

ส่วนแคปทชาข้อความที่มีอัตราการยืนยันความเป็นมนุษย์ที่ต่ำที่สุดมีสามลักษณะ ได้แก่ (1) แคปทชาข้อความรูปแบบที่ 12 อันเป็นแคปทชาข้อความที่มีระยะห่างระหว่างอักษรชิดติดกัน (0 พอยต์) ถูกพาดข้ามด้วยเส้นรบกวนหนา 5 พิกเซล โดยเส้นรบกวนมีสีเดียวกับพื้นหลัง (สีขาว) (2) แคปทชาข้อความรูปแบบที่ 17 อันเป็นแคปทชาข้อความที่มีระยะห่างระหว่างอักษรทับซ้อนกัน (ลบ 3.75 พอยต์) ถูกพาดข้ามด้วยเส้นรบกวนหนา 5 พิกเซล โดยเส้นรบกวนมีสีเดียวกับตัวอักษร (สีดำ) และ (3) แคปทชาข้อความรูปแบบที่ 18 อันเป็นแคปทชาข้อความที่มีระยะห่างระหว่างอักษรทับซ้อนกัน (ลบ 3.75 พอยต์) ถูกพาดข้ามด้วยเส้นรบกวนหนา 5 พิกเซล โดยเส้นรบกวนมีสีเดียวกับพื้นหลัง (สีขาว) โดยมีอัตราการยืนยันความเป็นมนุษย์ที่ต่ำที่สุดคิดเป็นร้อยละ 25 เท่ากันทั้งสามลักษณะ

ตารางที่ 4. 4 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามระยะห่างระหว่างอักษร

ระยะห่างระหว่างอักษร	จำนวนผู้ตอบถูกต้อง (คน)	จำนวนผู้ตอบทั้งหมด (คน)	อัตราการยืนยันความเป็นมนุษย์ (%)
ปกติ	84	120	70.0
ชิด	82	120	68.3
ทับซ้อน	46	120	38.3

ตารางที่ 4.4 พบว่าแคปทชาข้อความที่มีระยะห่างระหว่างอักษรปกติมีอัตราการยืนยันความเป็นมนุษย์สูงกว่าแคปทชาข้อความที่มีระยะห่างระหว่างอักษรแบบชิดและทับซ้อน ส่วนแคปทชาข้อความที่มีระยะห่างระหว่างอักษรแบบชิดก็มีอัตราการยืนยันความเป็นมนุษย์สูงกว่าแคปทชาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

ตารางที่ 4. 5 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามความหนาของเส้นรบกวน

ความหนาของเส้นรบกวน	จำนวนผู้ตอบถูกต้อง (คน)	จำนวนผู้ตอบทั้งหมด (คน)	อัตราการยืนยันความเป็นมนุษย์ (%)
1 พิกเซล	98	120	81.7
3 พิกเซล	73	120	60.8
5 พิกเซล	41	120	34.2

ตารางที่ 4.5 พบว่าแคปต์ชาข้อความที่ถูกพาดข้ามด้วยเส้นรบกวนที่มีความหนา 1 พิกเซล มีอัตราการยืนยันความเป็นมนุษย์สูงกว่าแคปต์ชาข้อความที่ถูกพาดข้ามด้วยเส้นรบกวนที่มีความหนา 3 พิกเซล และ 5 พิกเซล ส่วนแคปต์ชาที่ถูกพาดข้ามด้วยเส้นรบกวนที่มีความหนา 3 พิกเซล ก็มีอัตราการยืนยันความเป็นมนุษย์สูงกว่าแคปต์ชาข้อความที่ถูกพาดข้ามด้วยเส้นรบกวนที่มีความหนา 5 พิกเซล

ตารางที่ 4. 6 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามสีของเส้นรบกวน

สีของเส้นรบกวน	จำนวนผู้ตอบถูกต้อง (คน)	จำนวนผู้ตอบทั้งหมด (คน)	อัตราการยืนยันความเป็นมนุษย์ (%)
สีเดียวกับตัวอักษร	114	180	63.3
สีเดียวกับพื้นหลัง	98	180	54.3

ตารางที่ 4.6 พบว่าแคปต์ชาที่ถูกพาดข้ามด้วยเส้นรบกวนที่มีสีเดียวกับตัวอักษร มีอัตราการยืนยันความเป็นมนุษย์สูงกว่าแคปต์ชาข้อความที่ถูกพาดข้ามด้วยเส้นรบกวนที่มีสีเดียวกับพื้นหลัง

4.4 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อแคปต์ชาข้อความแสดงด้วยระยะห่างระหว่างอักษรที่แตกต่างกัน

ผู้วิจัยต้องการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ, ซิด (0 พอยต์), และทับซ้อน (ลบ 3.75 พอยต์) เพื่อทดสอบสมมติฐานต่อไปนี้

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นปกติ ต่ำกว่าหรือเท่ากับชุดอักษรแคปซูลข้อความที่มีระยะห่างระหว่างอักษรแบบขีด

H_1 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นปกติ สูงกว่าชุดอักษรแคปซูลข้อความที่มีระยะห่างระหว่างอักษรแบบขีด

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านขวา (Right-tailed test) พบว่าไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ เพราะค่านวนค่า z ได้ 0.2851 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า 1.6577 จึงตกอยู่ในบริเวณยอมรับ ทำให้ไม่สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยืนยันความเป็นมนุษย์ เมื่อแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นปกติ ไม่แตกต่าง (หรือหมายความว่าเท่ากับหรือต่ำกว่า) แคปซูลข้อความที่มีระยะห่างระหว่างอักษรแบบขีด

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นปกติ ต่ำกว่าหรือเท่ากับชุดอักษรแคปซูลข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

H_1 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นปกติ สูงกว่าชุดอักษรแคปซูลข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านขวา (Right-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่านวนค่า z ได้ 4.9281 (ดูรายละเอียดในภาคผนวก ข) มีค่ามากกว่า 1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยืนยันความเป็นมนุษย์ เมื่อแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นปกติ สูงกว่าแคปซูลข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อนอย่างมีนัยสำคัญทางสถิติ

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นขีด และทับซ้อน ไม่แตกต่างกัน

H_1 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปซูลข้อความมีระยะห่างระหว่างอักษรเป็นขีด และทับซ้อน แตกต่างกัน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบสองทาง (Two-tailed test) พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่านวนค่า z ได้ 4.6579 (ดูรายละเอียดใน

ภาคผนวก ข) มีค่ามากกว่า -1.9799 และ 1.9799 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยืนยันความเป็นมนุษย์ เมื่อแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นขีด และแบบทับซ้อน แตกต่างกัน

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นขีด ต่ำกว่าหรือเท่ากับชุดอักษรแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

H_1 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นขีด สูงกว่าชุดอักษรแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านขวา (Right-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่า z ได้ 4.6579 (ดูรายละเอียดในภาคผนวก ข) มีค่ามากกว่า 1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยืนยันความเป็นมนุษย์ เมื่อแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นขีด สูงกว่าแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อนอย่างมีนัยสำคัญทางสถิติ

4.5 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อแคปท์ซาข้อความมีเส้นรบกวนที่มีความหนาแตกต่างกัน

ผู้วิจัยต้องการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล เพื่อทดสอบสมมติฐานต่อไปนี้

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล ต่ำกว่าหรือเท่ากับชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล

H_1 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านขวา (Right-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่า z ได้ 3.5655 (ดูรายละเอียดในภาคผนวก ข) มีค่ามากกว่า 1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธ

สมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยื่นยันความเป็นมนุษย์ เมื่อแคปท์ซาข้อความถูกรบกวน ด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล อย่างมีนัยสำคัญทางสถิติ

H_0 : อัตราการยื่นยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล ต่ำกว่าหรือเท่ากับชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

H_1 : อัตราการยื่นยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านขวา (Right-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่านวนค่า z ได้ 7.4572 (ดูรายละเอียดในภาคผนวก ข) มีค่ามากกว่า 1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยื่นยันความเป็นมนุษย์ เมื่อแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติ

H_0 : อัตราการยื่นยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล ไม่แตกต่างกัน

H_1 : อัตราการยื่นยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล แตกต่างกัน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบสองทาง (Two-tailed test) พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่านวนค่า z ได้ 4.1364 (ดูรายละเอียดในภาคผนวก ข) มีค่ามากกว่า -1.9799 และ 1.9799 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยื่นยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล แตกต่างกัน

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล ต่ำกว่าหรือเท่ากับชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

H_1 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล สูงกว่าชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านขวา (Right-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่าจำนวนค่า z ได้ 4.1364 (ดูรายละเอียดในภาคผนวก ข) มีค่ามากกว่า 1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยืนยันความเป็นมนุษย์ เมื่อแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล สูงกว่าแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

4.6 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อแคปต์ซาข้อความแสดงด้วยสีของเส้นรบกวนที่แตกต่างกัน

ผู้วิจัยต้องการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และสีเดียวกับพื้นหลังของแคปต์ซาข้อความ เพื่อทดสอบสมมติฐานต่อไปนี้

H_0 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และสีเดียวกับพื้นหลังของแคปต์ซาข้อความ ไม่แตกต่างกัน

H_1 : อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และสีเดียวกับพื้นหลังของแคปต์ซาข้อความ แตกต่างกัน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบสองทาง (Two-tailed test) พบว่าไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ เพราะค่าจำนวนค่า z ได้ 1.7138 (ดูรายละเอียดในภาคผนวก ข) มีค่ามากกว่า -1.9799 และน้อยกว่า 1.9799 จึงตกอยู่ในบริเวณวิกฤตยอมรับ ทำให้ไม่สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ซาข้อความ และสีเดียวกับพื้นหลังของแคปต์ซาข้อความ ไม่แตกต่างกัน

4.7 อัตราความทนทานของแคปต์ชาข้อความ

ผู้วิจัยได้นำแคปต์ชาข้อความชุดเดียวกันกับที่ใช้ในการทดลองกับหน่วยทดลอง จำนวนทั้งสิ้น 360 แคปต์ชาข้อความ มาทดสอบความทนทานของแคปต์ชาข้อความด้วยโปรแกรม GSA Captcha Breaker และนำผลการทดสอบมาเปรียบเทียบและนำเสนออัตราความทนทานของแคปต์ชาข้อความ จำแนกตามรูปแบบแคปต์ชาข้อความดังตารางที่ 3.2 และจำแนกตามตัวแปรที่ศึกษา ได้แก่ (1) ระยะห่างระหว่างอักษร (2) ความหนาของเส้นรบกวน และ (3) สีของเส้นรบกวน ดังตารางที่ 4.7 ถึง ตารางที่ 4.10 โดยในตารางที่ 4.7 เป็นการจำแนกตามรูปแบบแคปต์ชาข้อความ ตารางที่ 4.8 เป็นการจำแนกตามระยะห่างระหว่างอักษร ตารางที่ 4.9 เป็นการจำแนกตามความหนาของเส้นรบกวน และ ตารางที่ 4.10 เป็นการจำแนกตามสีของเส้นรบกวน

ตารางที่ 4. 7 อัตราความทนทานของแคปต์ชาข้อความ จำแนกตามรูปแบบแคปต์ชาข้อความ โดยจำนวนครั้งที่ GSA ตอบแคปต์ชาข้อความรูปแบบละ 20 ครั้ง

รูปแบบแคปต์ชาข้อความ	ระยะห่างระหว่างอักษร (พอยต์)	ความหนาของเส้นรบกวน (พิกเซล)	สีของเส้นรบกวน	จำนวนครั้งที่ GSA ตอบไม่ได้ (ครั้ง)	อัตราความทนทานของแคปต์ชาข้อความ (%)
1	ปกติ	1	สีเดียวกับตัวอักษร	0	0
2	ปกติ	1	สีเดียวกับพื้นหลัง	0	0
3	ปกติ	3	สีเดียวกับตัวอักษร	19	95
4	ปกติ	3	สีเดียวกับพื้นหลัง	15	75
5	ปกติ	5	สีเดียวกับตัวอักษร	19	95

ตารางที่ 4.7 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามรูปแบบแคปท์ซาข้อความ โดยจำนวนครั้งที่ GSA ตอบแคปท์ซาข้อความรูปแบบละ 20 ครั้ง (ต่อ)

รูปแบบแคปท์ซาข้อความ	ระยะห่างระหว่างอักษร (พอยต์)	ความหนาของเส้นรบกวน (พิกเซล)	สีของเส้นรบกวน	จำนวนครั้งที่ GSA ตอบไม่ได้ (ครั้ง)	อัตราความทนทานของแคปท์ซาข้อความ (%)
6	ปกติ	5	สีเดียวกับพื้นหลัง	19	95
7	ชิด	1	สีเดียวกับตัวอักษร	20	100
8	ชิด	1	สีเดียวกับพื้นหลัง	14	70
9	ชิด	3	สีเดียวกับตัวอักษร	19	95
10	ชิด	3	สีเดียวกับพื้นหลัง	18	90
11	ชิด	5	สีเดียวกับตัวอักษร	20	100
12	ชิด	5	สีเดียวกับพื้นหลัง	20	100
13	ทับซ้อน	1	สีเดียวกับตัวอักษร	20	100
14	ทับซ้อน	1	สีเดียวกับพื้นหลัง	20	100
15	ทับซ้อน	3	สีเดียวกับตัวอักษร	20	100

ตารางที่ 4.7 อัตราความทนทานของแคปทชาข้อความ จำแนกตามรูปแบบแคปทชาข้อความ โดยจำนวนครั้งที่ GSA ตอบแคปทชาข้อความรูปแบบละ 20 ครั้ง (ต่อ)

รูปแบบแคปทชาข้อความ	ระยะห่างระหว่างอักษร (พอยต์)	ความหนาของเส้นรบกวน (พิกเซล)	สีของเส้นรบกวน	จำนวนครั้งที่ GSA ตอบไม่ได้ (ครั้ง)	อัตราความทนทานของแคปทชาข้อความ (%)
16	ทับซ้อน	3	สีเดียวกับพื้นหลัง	20	100
17	ทับซ้อน	5	สีเดียวกับตัวอักษร	20	100
18	ทับซ้อน	5	สีเดียวกับพื้นหลัง	20	100

ตารางที่ 4.7 พบว่าแคปทชาข้อความในสามลักษณะที่มีอัตราความทนทานต่ำสุด อันหมายถึงโปรแกรม GSA Captcha Breaker สามารถตอบแคปทชาได้มากที่สุด ได้แก่ (1) แคปทชาข้อความรูปแบบที่ 1 อันเป็นชุดอักษรแคปทชาที่มีระยะห่างระหว่างอักษรแบบปกติ ถูกพาดข้ามด้วยเส้นรบกวนสีเดียวกับตัวอักษรของแคปทชาข้อความที่มีความหนา 1 พิกเซล (2) แคปทชาข้อความรูปแบบที่ 2 อันเป็นชุดอักษรแคปทชาที่มีระยะห่างระหว่างอักษรแบบปกติ ถูกพาดข้ามด้วยเส้นรบกวนสีเดียวกับพื้นหลังของแคปทชาข้อความที่มีความหนา 1 พิกเซล และ (3) แคปทชาข้อความรูปแบบที่ 8 อันเป็นชุดอักษรแคปทชาที่มีระยะห่างระหว่างอักษรแบบชิด ถูกพาดข้ามด้วยเส้นรบกวนสีเดียวกับพื้นหลังของแคปทชาข้อความที่มีความหนา 1 พิกเซล โดยแคปทชาข้อความสองลักษณะแรกมีอัตราความทนทานของแคปทชาข้อความคิดเป็นร้อยละ 0 ส่วนลักษณะที่สามคิดเป็นร้อยละ 70

ตารางที่ 4. 8 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามระยะห่างระหว่างอักษร

ระยะห่าง ระหว่าง อักษร	จำนวนครั้งที่ GSA ตอบ แคปท์ซาไม่ได้ (ครั้ง)	จำนวนครั้งที่ GSA ตอบ แคปท์ซาทั้งหมด (ครั้ง)	อัตราความทนทานของ แคปท์ซาข้อความ (%)
ปกติ	72	120	60.0
ชิด	111	120	92.5
ทับซ้อน	120	120	100.0

ตารางที่ 4.8 พบว่าแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบปกติมีอัตราความทนทานของแคปท์ซาข้อความต่ำกว่าแบบชิดและแบบทับซ้อน และแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบทับซ้อนมีอัตราความทนทานของแคปท์ซาข้อความสูงที่สุด

ตารางที่ 4. 9 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามความหนาของเส้นรบกวน

ความหนาของ เส้นรบกวน	จำนวนครั้งที่ GSA ตอบ แคปท์ซาไม่ได้ (ครั้ง)	จำนวนครั้งที่ GSA ตอบ แคปท์ซาทั้งหมด (ครั้ง)	อัตราความทนทานของ แคปท์ซาข้อความ (%)
1 พิกเซล	74	120	61.7
3 พิกเซล	111	120	92.5
5 พิกเซล	118	120	98.3

ตารางที่ 4.9 พบว่าแคปท์ซาข้อความที่ถูกรบกวนด้วยเส้นรบกวนขนาด 1 พิกเซล มีอัตราความทนทานของแคปท์ซาข้อความต่ำกว่าเส้นรบกวนขนาด 3 และ 5 พิกเซล และแคปท์ซาข้อความที่ถูกรบกวนด้วยเส้นรบกวนขนาด 5 พิกเซลมีอัตราความทนทานของแคปท์ซาข้อความสูงที่สุด

ตารางที่ 4. 10 อัตราความทนทานของแคปท์ชาข้อความ จำแนกตามสีของเส้นรบกวน

สีของเส้นรบกวน	จำนวนครั้งที่ GSA ตอบแคปท์ชาไม่ได้ (ครั้ง)	จำนวนครั้งที่ GSA ตอบแคปท์ชาทั้งหมด (ครั้ง)	อัตราความทนทานของแคปท์ชาข้อความ (%)
สีเดียวกับตัวอักษร	157	180	87.2
สีเดียวกับพื้นหลัง	146	180	81.1

ตารางที่ 4.10 พบว่าแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปท์ชาข้อความ มีอัตราความทนทานของแคปท์ชาข้อความสูงกว่าแคปท์ชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับพื้นหลังของแคปท์ชาข้อความ

4.8 การเปรียบเทียบอัตราความทนทานของแคปท์ชาข้อความ เมื่อแคปท์ชาข้อความแสดงด้วยระยะห่างระหว่างอักษรที่แตกต่างกัน

ผู้วิจัยต้องการเปรียบเทียบอัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ, ซิด (0 พอยต์), และทับซ้อน (ลบ 3.75 พอยต์) เพื่อทดสอบสมมติฐานต่อไปนี้

H_0 : อัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ สูงกว่าหรือเท่ากับชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบซิด

H_1 : อัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ ต่ำกว่าชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบซิด

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านซ้าย (Left-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่า z ได้ -5.9157 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า -1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความทนทานของแคปท์ชาข้อความ เมื่อแคปท์ชาข้อความมี

ระยะห่างระหว่างอักษรเป็นปกติ ต่ำกว่าแคปทซ์ข้อความที่มีระยะห่างระหว่างอักษรแบบชิดอย่างมีนัยสำคัญทางสถิติ

H_0 : อัตราความหนาแน่นของแคปทซ์ข้อความ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นปกติ สูงกว่าหรือเท่ากับชุดอักษรแคปทซ์ข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

H_1 : อัตราความหนาแน่นของแคปทซ์ข้อความ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นปกติ ต่ำกว่าชุดอักษรแคปทซ์ข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านซ้าย (Left-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่า z ได้ -7.7460 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า -1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความหนาแน่นของแคปทซ์ข้อความ เมื่อแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นปกติ ต่ำกว่าแคปทซ์ข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อนอย่างมีนัยสำคัญทางสถิติ

H_0 : อัตราความหนาแน่นของแคปทซ์ข้อความ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นชิด และทับซ้อน ไม่แตกต่างกัน

H_1 : อัตราความหนาแน่นของแคปทซ์ข้อความ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นชิด และทับซ้อน แตกต่างกัน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบสองทาง (Two-tailed test) พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่า z ได้ -3.0579 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า -1.9799 และ 1.9799 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความหนาแน่นของแคปทซ์ข้อความ เมื่อแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นชิด และแบบทับซ้อน แตกต่างกัน

H_0 : อัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นซิด สูงกว่าหรือเท่ากับชุดอักษรแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

H_1 : อัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นซิด ต่ำกว่าชุดอักษรแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านซ้าย (Left-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่า z ได้ -3.0579 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า -1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความทนทานของแคปท์ซาข้อความ เมื่อแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นซิด ต่ำกว่าแคปท์ซาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อนอย่างมีนัยสำคัญทางสถิติ

4.9 การเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อแคปท์ซาข้อความมีเส้นรบกวนที่มีความหนาแตกต่างกัน

ผู้วิจัยต้องการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล เพื่อทดสอบสมมติฐานต่อไปนี้

H_0 : อัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าหรือเท่ากับชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล

H_1 : อัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล ต่ำกว่าชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านซ้าย (Left-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่า z ได้ -5.6825 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า -1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความทนทานของแคปท์ซาข้อความ เมื่อแคปท์ซาข้อความถูก

รบกวนด้วยเส้นที่มีความหนา 1 พิกเซล ต่ำกว่าชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล อย่างมีนัยสำคัญทางสถิติ

H_0 : อัตราความทนทานของแคปซูลข้อความ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าหรือเท่ากับชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

H_1 : อัตราความทนทานของแคปซูลข้อความ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล ต่ำกว่าชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านซ้าย (Left-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่านวนค่า z ได้ -7.1005 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า -1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความทนทานของแคปซูลข้อความ เมื่อแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล ต่ำกว่าชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติ

H_0 : อัตราความทนทานของแคปซูลข้อความ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล ไม่แตกต่างกัน

H_1 : อัตราความทนทานของแคปซูลข้อความ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล แตกต่างกัน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบสองทาง (Two-tailed test) พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่านวนค่า z ได้ -2.1607 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า -1.9799 และ 1.9799 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความทนทานของแคปซูลข้อความ เมื่อชุดอักษรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล แตกต่างกัน

H_0 : อัตราความหนาแน่นของแคปซูลชาข้อความ เมื่อชุดอักษรแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล สูงกว่าหรือเท่ากับชุดอักษรแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

H_1 : อัตราความหนาแน่นของแคปซูลชาข้อความ เมื่อชุดอักษรแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล ต่ำกว่าชุดอักษรแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบทางเดียวด้านซ้าย (Left-tailed test) พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่าจำนวนค่า z ได้ -2.1607 (ดูรายละเอียดในภาคผนวก ข) มีค่าน้อยกว่า -1.6577 จึงตกอยู่ในบริเวณวิกฤต ทำให้สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความหนาแน่นของแคปซูลชาข้อความ เมื่อแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล ต่ำกว่าชุดอักษรแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติ

4.10 การเปรียบเทียบอัตราความหนาแน่นของแคปซูลชาข้อความ เมื่อแคปซูลชาข้อความแสดงสีของเส้นรบกวนที่แตกต่างกัน

ผู้วิจัยต้องการเปรียบเทียบอัตราความหนาแน่นของแคปซูลชาข้อความ เมื่อชุดอักษรแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปซูลชาข้อความ และสีเดียวกับพื้นหลังของแคปซูลชาข้อความ เพื่อทดสอบสมมติฐานต่อไปนี้

H_0 : อัตราความหนาแน่นของแคปซูลชาข้อความ เมื่อชุดอักษรแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปซูลชาข้อความ และสีเดียวกับพื้นหลังของแคปซูลชาข้อความ ไม่แตกต่างกัน

H_1 : อัตราความหนาแน่นของแคปซูลชาข้อความ เมื่อชุดอักษรแคปซูลชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปซูลชาข้อความ และสีเดียวกับพื้นหลังของแคปซูลชาข้อความ แตกต่างกัน

การทดสอบสมมติฐานด้วยสถิติทดสอบ Z แบบสองทาง (Two-tailed test) พบว่าไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่าจำนวนค่า z ได้ 1.5881 (ดูรายละเอียดในภาคผนวก ข) มีค่ามากกว่า -1.9799 และน้อยกว่า 1.9799 จึงตกอยู่ในบริเวณยอมรับ ทำให้ไม่สามารถปฏิเสธสมมติฐาน H_0 ได้ ดังนั้นสรุปได้ว่า อัตราความหนาแน่นของแคปซูลชาข้อความ เมื่อชุดอักษรแคปซูลชาข้อ

ความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ชาข้อความ และสีเดียวกับพื้นหลังของแคปต์ชาข้อความ ไม่แตกต่างกัน

4.11 สรุปผลการวิเคราะห์อัตราการยืนยันความเป็นมนุษย์ และอัตราความทนทานของแคปต์ชาข้อความ

ผู้วิจัยได้นำผลการทดสอบสมมติฐานตามที่ได้กำหนดไว้ในหัวข้อ 4.4 ถึง 4.10 มาสรุปเพื่อความกระจ่างมากขึ้น ดังตารางที่ 4.11 ถึง 4.13

ตารางที่ 4. 11 สรุปผลการวิเคราะห์ระยะห่างระหว่างอักษรของแคปต์ชาข้อความ

ระยะห่างระหว่างอักษร	
อัตราการยืนยันความเป็นมนุษย์	แคปต์ชาข้อความที่ระยะห่างระหว่างอักษรเป็นปกติมีอัตราการยืนยันความเป็นมนุษย์ ไม่แตกต่าง (หรือหมายความว่าเท่ากับหรือต่ำกว่า) แคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบชิด
	แคปต์ชาข้อความที่ระยะห่างระหว่างอักษรเป็นปกติมีอัตราการยืนยันความเป็นมนุษย์ สูงกว่าแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05
	แคปต์ชาข้อความที่ระยะห่างระหว่างอักษรเป็นชิดมีอัตราการยืนยันความเป็นมนุษย์ สูงกว่าแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

ตารางที่ 4.11 สรุปผลการวิเคราะห์ระยะห่างระหว่างอักษรของแคปต์ชาข้อความ (ต่อ)

ระยะห่างระหว่างอักษร	
อัตราความทนทาน ของแคปต์ชาข้อความ	แคปต์ชาข้อความที่ระยะห่างระหว่างอักษรเป็นปกติมี อัตราความทนทานของแคปต์ชาข้อความ ต่ำกว่าแคปต์ชา ข้อความที่มีระยะห่างระหว่างอักษรแบบชิด อย่างมี นัยสำคัญทางสถิติที่ระดับ 0.05
	แคปต์ชาข้อความที่ระยะห่างระหว่างอักษรเป็นปกติมี อัตราความทนทานของแคปต์ชาข้อความ ต่ำกว่าแคปต์ชา ข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน อย่างมี นัยสำคัญทางสถิติที่ระดับ 0.05
	แคปต์ชาข้อความที่ระยะห่างระหว่างอักษรเป็นชิดมีอัตรา ความทนทานของแคปต์ชาข้อความ ต่ำกว่าแคปต์ชา ข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน อย่างมี นัยสำคัญทางสถิติที่ระดับ 0.05

ตารางที่ 4. 12 สรุปผลการวิเคราะห์ความหนาของเส้นรบกวนที่พาดข้ามชุดอักษรแคปต์ชาข้อความ

ความหนาของเส้นรบกวน	
อัตราการยืนยันความเป็นมนุษย์	แคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 1 พิก เซล มีอัตราการยืนยันความเป็นมนุษย์ สูงกว่าแคปต์ชาข้อ ความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล อย่างมี นัยสำคัญทางสถิติที่ระดับ 0.05
	แคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 1 พิก เซล มีอัตราการยืนยันความเป็นมนุษย์ สูงกว่าแคปต์ชาข้อ ความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมี นัยสำคัญทางสถิติที่ระดับ 0.05

ตาราง 4.12 สรุปผลการวิเคราะห์ความหนาของเส้นรบกวนที่พาดข้ามชุดอักษรแคปทีชาข้อความ (ต่อ)

ความหนาของเส้นรบกวน	
อัตราการยืนยันความเป็นมนุษย์	แคปทีชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล มีอัตราการยืนยันความเป็นมนุษย์ สูงกว่าแคปทีชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05
อัตราความทนทานของ แคปทีชาข้อความ	แคปทีชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล มีอัตราความทนทานของแคปทีชาข้อความ ต่ำกว่าชุดอักษรแคปทีชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05
	แคปทีชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล มีอัตราความทนทานของแคปทีชาข้อความ ต่ำกว่าชุดอักษรแคปทีชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05
	แคปทีชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล มีอัตราความทนทานของแคปทีชาข้อความ ต่ำกว่าชุดอักษรแคปทีชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

ตารางที่ 4. 13 สรุปผลการวิเคราะห์สีของเส้นรบกวนที่พาดข้ามชุดอักษรแคปทซ์ข้อความ

สีของเส้นรบกวน	
อัตราการยืนยันความเป็นมนุษย์	แคปทซ์ข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปทซ์ข้อความ และสีเดียวกับพื้นหลังของแคปทซ์ข้อความมีอัตราการยืนยันความเป็นมนุษย์ ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติ
อัตราความทนทานของแคปทซ์ข้อความ	แคปทซ์ข้อความที่ถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปทซ์ข้อความ และสีเดียวกับพื้นหลังของแคปทซ์ข้อความมีอัตราความทนทานขอแคปทซ์ข้อความ ไม่แตกต่างกัน อย่างมีนัยสำคัญทางสถิติ

4.12 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ จำแนกตามข้อมูลส่วนบุคคลของหน่วยทดลอง

ผู้วิจัยได้วิเคราะห์ผลการทดลอง และสรุปผลของอัตราการยืนยันความเป็นมนุษย์จำแนกตามข้อมูลส่วนบุคคลของหน่วยทดลอง ดังตารางที่ 4.14

ตารางที่ 4. 14 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามข้อมูลส่วนบุคคลของหน่วยทดลอง

คุณสมบัติ		จำนวนผู้ตอบแคปทซ์ชาได้ถูกต้อง (คน)	จำนวนผู้ตอบแคปทซ์ชาทั้งหมด (คน)	อัตราการยืนยันความเป็นมนุษย์ (%)
เพศ	ชาย	68	107	63.5
	หญิง	144	253	56.9
อายุ	ต่ำกว่า 20 ปี	78	132	59.1
	อายุ 20 ปีขึ้นไป	134	228	58.8

ตารางที่ 4.14 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามข้อมูลส่วนบุคคลของหน่วยทดลอง (ต่อ)

คุณสมบัติ		จำนวนผู้ตอบ แคปซูลที่ได้ ถูกต้อง (คน)	จำนวน ผู้ตอบแคปซูล ชาทั้งหมด (คน)	อัตราการ ยืนยันความ เป็นมนุษย์ (%)
หลักสูตร	บัญชี	100	168	59.5
	บริหารธุรกิจ	90	147	61.2
	สถิติศาสตร์	22	41	53.7
	นานาชาติ	0	4	0.0
ปัญหาตาบอดสี	มีปัญหาตาบอดสี	7	12	58.3
	ไม่มีปัญหาตาบอดสี	187	322	58.1
	ไม่ทราบว่ามีปัญหา	18	26	69.2
ปัญหาเกี่ยวกับสายตา	สายตาปกติ	56	106	52.8
	สายตาสั้น	70	105	66.7
	สายตาสั้นและเอียง	80	136	58.8
	สายตายาว	0	0	-
	สายตายาวและเอียง	1	3	33.3
	สายตาเอียง	5	10	50.0
การแก้ปัญหาเกี่ยวกับ สายตาขณะทำการ ทดลอง	ใส่แว่น	107	175	61.1
	ใส่คอนแทคเลนส์	32	50	64.0
	ไม่ต้องแก้ปัญหา	73	135	54.1

ตารางที่ 4.14 เมื่อจำแนกตามเพศ พบว่าหน่วยทดลองเพศชายมีอัตราการยืนยันความเป็นมนุษย์สูงกว่าเพศหญิง เมื่อจำแนกตามอายุ พบว่าหน่วยทดลองที่มีอายุต่ำกว่า 20 ปี มีอัตราการยืนยัน

ความเป็นมนุษย์สูงกว่าหน่วยทดลองที่มีอายุ 20 ปีขึ้นไป เมื่อจำแนกตามหลักสูตร พบว่าหน่วยทดลองที่กำลังศึกษาอยู่ในหลักสูตรบริหารธุรกิจมีอัตราการยืนยันความเป็นมนุษย์สูงสุด คิดเป็นร้อยละ 61.2 เมื่อจำแนกตามปัญหาตาบอดสี พบว่าหน่วยทดลองที่ไม่ทราบว่ามีปัญหาอัตราการยืนยันความเป็นมนุษย์สูงสุด คิดเป็นร้อยละ 69.2 เมื่อจำแนกตามปัญหาเกี่ยวกับสายตา พบว่าหน่วยทดลองที่มีสายตาสั้นมีอัตราการยืนยันความเป็นมนุษย์สูงสุด คิดเป็นร้อยละ 66.7 และเมื่อจำแนกตามการแก้ปัญหาเกี่ยวกับสายตา พบว่าหน่วยทดลองที่ใส่คอนแทคเลนส์มีอัตราการยืนยันความเป็นมนุษย์สูงสุด คิดเป็นร้อยละ 64.0

4.13 การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ จำแนกตามตัวอักษร

ผู้วิจัยได้วิเคราะห์ผลการทดลอง และสรุปผลของอัตราการยืนยันความเป็นมนุษย์จำแนกตามตัวอักษร ดังตารางที่ 4.15

ตารางที่ 4. 15 อัตราการยืนยันความเป็นมนุษย์จำแนกตามตัวอักษร

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
a	84	81	96.4
b	97	90	92.8
c	96	52	54.2
d	78	76	97.4
e	81	43	53.1
f	89	83	93.3
g	82	80	97.6
h	70	68	97.1
i	81	75	92.6
j	81	76	93.8
k	95	95	100.0

ตารางที่ 4.15 อัตราการยืนยันความเป็นมนุษย์จำแนกตามตัวอักษร (ต่อ)

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
l	71	57	80.3
m	83	83	100.0
n	76	75	98.7
o	90	70	77.8
p	79	77	97.5
q	56	49	87.5
r	78	71	91.0
s	90	53	58.9
t	87	75	86.2
u	101	101	100.0
v	82	81	98.8
w	82	82	100.0
x	82	81	98.8
y	79	77	97.5
z	90	89	98.9

ตารางที่ 4.15 พบว่าตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์สูงสุด ได้แก่ ตัวอักษร k, m, u, และ w โดยมีอัตราการยืนยันความเป็นมนุษย์ร้อยละ 100.0 และตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์ต่ำสุด ได้แก่ ตัวอักษร e โดยมีอัตราการยืนยันความเป็นมนุษย์ร้อยละ 53.1

ตารางที่ 4. 16 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามระยะห่างระหว่าง
อักษร

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม			จำนวนครั้งที่ตอบถูก			อัตราการยืนยันความเป็นมนุษย์ (%)		
	ปกติ	ชิด	ทับซ้อน	ปกติ	ชิด	ทับซ้อน	ปกติ	ชิด	ทับซ้อน
a	28	29	27	28	29	24	100.0%	100.0%	88.9%
b	26	32	39	26	30	34	100.0%	93.8%	87.2%
c	30	40	26	17	28	7	56.7%	70.0%	26.9%
d	37	17	24	37	17	22	100.0%	100.0%	91.7%
e	29	31	21	15	18	10	51.7%	58.1%	47.6%
f	29	29	31	28	29	26	96.6%	100.0%	83.9%
g	28	23	31	28	22	30	100.0%	95.7%	96.8%
h	23	17	30	22	17	29	95.7%	100.0%	96.7%
i	27	30	24	27	29	19	100.0%	96.7%	79.2%
j	28	22	31	28	21	27	100.0%	95.5%	87.1%
k	33	30	32	33	30	32	100.0%	100.0%	100.0%
l	22	20	29	22	18	17	100.0%	90.0%	58.6%
m	23	31	29	23	31	29	100.0%	100.0%	100.0%
n	28	27	21	28	27	20	100.0%	100.0%	95.2%
o	29	27	34	27	21	22	93.1%	77.8%	64.7%
p	26	31	22	25	31	21	96.2%	100.0%	95.5%
q	16	21	19	16	20	13	100.0%	95.2%	68.4%
r	29	23	26	28	23	20	96.6%	100.0%	76.9%
s	22	34	34	14	20	19	63.6%	58.8%	55.9%
t	27	31	29	26	31	18	96.3%	100.0%	62.1%

ตารางที่ 4.16 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามระยะห่างระหว่างอักษร (ต่อ)

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม			จำนวนครั้งที่ตอบถูก			อัตราการยืนยันความเป็นมนุษย์ (%)		
	ปกติ	ชิด	ทับซ้อน	ปกติ	ชิด	ทับซ้อน	ปกติ	ชิด	ทับซ้อน
u	32	32	37	32	32	37	100.0%	100.0%	100.0%
v	31	27	24	31	27	23	100.0%	100.0%	95.8%
w	27	30	25	27	30	25	100.0%	100.0%	100.0%
x	24	30	28	24	30	27	100.0%	100.0%	96.4%
y	30	23	26	30	22	25	100.0%	95.7%	96.2%
z	36	33	21	36	33	20	100.0%	100.0%	95.2%

ตารางที่ 4.16 พบว่าตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์ต่ำสุดสามอันดับ ได้แก่ ตัวอักษร c ที่มีระยะห่างระหว่างอักษรแบบทับซ้อน ตัวอักษร e ที่มีระยะห่างระหว่างอักษรแบบปกติ และทับซ้อน โดยมีอัตราการยืนยันความเป็นมนุษย์ร้อยละ 26.9, 51.7, และ 47.6 ตามลำดับ

ตารางที่ 4. 17 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามขนาดความหนาของเส้นรบกวน (พิกเซล)

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม			จำนวนครั้งที่ตอบถูก			อัตราการยืนยันความเป็นมนุษย์ (%)		
	1	3	5	1	3	5	1	3	5
a	27	29	28	27	28	26	100.0%	96.6%	92.9%
b	31	35	31	29	33	28	93.5%	94.3%	90.3%
c	35	31	30	30	18	4	85.7%	58.1%	13.3%
d	22	30	26	22	29	25	100.0%	96.7%	96.2%
e	30	28	23	27	12	4	90.0%	42.9%	17.4%
f	37	23	29	34	20	29	91.9%	87.0%	100.0%

ตารางที่ 4.17 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามขนาดความหนาของเส้นรบกวน (พิกเซล) (ต่อ)

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม			จำนวนครั้งที่ตอบถูก			อัตราการยืนยันความเป็นมนุษย์ (%)		
	1	3	5	1	3	5	1	3	5
g	31	23	28	30	22	28	96.8%	95.7%	100.0%
h	20	27	23	20	26	22	100.0%	96.3%	95.7%
i	24	30	27	22	28	25	91.7%	93.3%	92.6%
j	31	28	22	31	25	20	100.0%	89.3%	90.9%
k	28	31	36	28	31	36	100.0%	100.0%	100.0%
l	23	24	24	16	21	20	69.6%	87.5%	83.3%
m	28	31	24	28	31	24	100.0%	100.0%	100.0%
n	22	24	30	22	24	29	100.0%	100.0%	96.7%
o	37	29	24	33	21	16	89.2%	72.4%	66.7%
p	26	32	21	25	31	21	96.2%	96.9%	100.0%
q	16	20	20	14	17	18	87.5%	85.0%	90.0%
r	29	26	23	27	23	21	93.1%	88.5%	91.3%
s	30	26	34	30	23	0	100.0%	88.5%	0.0%
t	29	28	30	25	23	27	86.2%	82.1%	90.0%
u	38	31	32	38	31	32	100.0%	100.0%	100.0%
v	22	32	28	22	31	28	100.0%	96.9%	100.0%
w	23	24	35	23	24	35	100.0%	100.0%	100.0%
x	29	25	28	29	24	28	100.0%	96.0%	100.0%
y	23	23	33	23	23	31	100.0%	100.0%	93.9%
z	29	30	31	29	29	31	100.0%	96.7%	100.0%

ตารางที่ 4.17 พบว่าตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์ต่ำสุดสามอันดับ ได้แก่ ตัวอักษร s, c, และ e ที่ถูกรบกวนด้วยเส้นหนา 5 พิกเซล โดยมีอัตราการยืนยันความเป็นมนุษย์ร้อยละ 0, 13.3, และ 17.4 ตามลำดับ

ตารางที่ 4. 18 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามสีของเส้นรบกวน

ตัวอักษร	จำนวนครั้งที่ถูกลွ่ม		จำนวนครั้งที่ตอบถูก		อัตราการยืนยันความเป็นมนุษย์ (%)	
	สีเดียวกับอักษร	สีเดียวกับพื้นหลัง	สีเดียวกับอักษร	สีเดียวกับพื้นหลัง	สีเดียวกับอักษร	สีเดียวกับพื้นหลัง
a	38	46	37	44	97.4%	95.7%
b	48	49	43	47	89.6%	95.9%
c	38	58	23	29	60.5%	50.0%
d	34	44	34	42	100.0%	95.5%
e	39	42	30	13	76.9%	31.0%
f	46	43	43	40	93.5%	93.0%
g	39	43	38	42	97.4%	97.7%
h	40	30	39	29	97.5%	96.7%
i	39	42	36	39	92.3%	92.9%
j	46	35	43	33	93.5%	94.3%
k	52	43	52	43	100.0%	100.0%
l	39	32	32	25	82.1%	78.1%
m	40	43	40	43	100.0%	100.0%
n	43	33	43	32	100.0%	97.0%
o	44	46	36	34	81.8%	73.9%
p	39	40	38	39	97.4%	97.5%

ตารางที่ 4.18 อัตราการยืนยันความเป็นมนุษย์ของแต่ละตัวอักษร จำแนกตามสีของเส้นรบกวน (ต่อ)

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม		จำนวนครั้งที่ตอบถูก		อัตราการยืนยันความเป็นมนุษย์ (%)	
	สีเดียวกับอักษร	สีเดียวกับพื้นหลัง	สีเดียวกับอักษร	สีเดียวกับพื้นหลัง	สีเดียวกับอักษร	สีเดียวกับพื้นหลัง
q	20	36	18	31	90.0%	86.1%
r	41	37	39	32	95.1%	86.5%
s	47	43	29	24	61.7%	55.8%
t	44	43	35	40	79.5%	93.0%
u	58	43	58	43	100.0%	100.0%
v	42	40	41	40	97.6%	100.0%
w	38	44	38	44	100.0%	100.0%
x	43	39	42	39	97.7%	100.0%
y	39	40	39	38	100.0%	95.0%
z	44	46	43	46	97.7%	100.0%

ตารางที่ 4.18 พบว่าตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์ต่ำสุดสามอันดับ ได้แก่ ตัวอักษร e, c, และ s ที่ถูกรบกวนด้วยเส้นสีเดียวกับพื้นหลังของแป้นพิมพ์ข้อความ โดยมีอัตราการยืนยันความเป็นมนุษย์ร้อยละ 31, 50, และ 55.8 ตามลำดับ

4.14 การเปรียบเทียบอัตราการความทนทานของแป้นพิมพ์ข้อความ จำแนกตามตัวอักษร

ผู้วิจัยได้วิเคราะห์ผลการทดลอง และสรุปผลของอัตราการความทนทานของแป้นพิมพ์ข้อความจำแนกตามตัวอักษร ดังตารางที่ 4.19

ตารางที่ 4. 19 อัตราความทนทานของแคปซูลข้อความจำแนกตามตัวอักษร

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ GSA ตอบ ไม่ถูกต้อง (ครั้ง)	อัตราความทนทานของ ตัวอักษร (%)
a	84	66	78.6
b	97	76	78.4
c	96	72	75.0
d	78	52	66.7
e	81	66	81.5
f	89	72	80.9
g	82	63	76.8
h	70	55	78.6
i	81	72	88.9
j	81	63	77.8
k	95	82	86.3
l	71	58	81.7
m	83	66	79.5
n	76	58	76.3
o	90	68	75.6
p	79	53	67.1
q	56	46	82.1
r	78	58	74.4
s	90	81	90.0
t	87	67	77.0

ตารางที่ 4.19 อัตราความทนทานของแคปซูลชาข้อความจำแนกตามตัวอักษร (ต่อ)

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ GSA ตอบ ไม่ถูกต้อง (ครั้ง)	อัตราความทนทานของ ตัวอักษร (%)
u	101	76	75.2
v	82	59	72.0
w	82	64	78.0
x	82	64	78.0
y	79	64	81.0
z	90	68	75.6

ตารางที่ 4.19 พบว่าตัวอักษรที่มีอัตราความทนทานสูงสุด ได้แก่ ตัวอักษร s โดยมีอัตราความทนทานร้อยละ 90.0 และตัวอักษรที่มีอัตราความทนทานต่ำสุด ได้แก่ ตัวอักษร d โดยมีอัตราความทนทานร้อยละ 66.7

ตารางที่ 4. 20 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามระยะห่างระหว่างอักษร

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม			จำนวนครั้งที่ GSA ตอบไม่ถูก			อัตราความทนทานของ ตัวอักษร (%)		
	ปกติ	ชิด	ทับซ้อน	ปกติ	ชิด	ทับซ้อน	ปกติ	ชิด	ทับซ้อน
a	28	29	27	15	24	27	53.6%	82.8%	100.0%
b	26	32	39	11	26	39	42.3%	81.3%	100.0%
c	30	40	26	14	32	26	46.7%	80.0%	100.0%
d	37	17	24	12	16	24	32.4%	94.1%	100.0%
e	29	31	21	16	29	21	55.2%	93.5%	100.0%
f	29	29	31	17	24	31	58.6%	82.8%	100.0%

ตารางที่ 4.20 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามระยะห่างระหว่างอักษร (ต่อ)

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม			จำนวนครั้งที่ GSA ตอบไม่ถูก			อัตราความทนทานของ ตัวอักษร (%)		
	ปกติ	ชิด	ทับซ้อน	ปกติ	ชิด	ทับซ้อน	ปกติ	ชิด	ทับซ้อน
g	28	23	31	14	18	31	50.0%	78.3%	100.0%
h	23	17	30	11	14	30	47.8%	82.4%	100.0%
i	27	30	24	19	29	24	70.4%	96.7%	100.0%
j	28	22	31	15	17	31	53.6%	77.3%	100.0%
k	33	30	32	22	28	32	66.7%	93.3%	100.0%
l	22	20	29	10	19	29	45.5%	95.0%	100.0%
m	23	31	29	9	28	29	39.1%	90.3%	100.0%
n	28	27	21	13	24	21	46.4%	88.9%	100.0%
o	29	27	34	11	23	34	37.9%	85.2%	100.0%
p	26	31	22	9	22	22	34.6%	71.0%	100.0%
q	16	21	19	10	17	19	62.5%	81.0%	100.0%
r	29	23	26	13	19	26	44.8%	82.6%	100.0%
s	22	34	34	16	31	34	72.7%	91.2%	100.0%
t	27	31	29	11	27	29	40.7%	87.1%	100.0%
u	32	32	37	11	28	37	34.4%	87.5%	100.0%
v	31	27	24	14	21	24	45.2%	77.8%	100.0%
w	27	30	25	14	25	25	51.9%	83.3%	100.0%
x	24	30	28	9	27	28	37.5%	90.0%	100.0%
y	30	23	26	17	21	26	56.7%	91.3%	100.0%
z	36	33	21	18	29	21	50.0%	87.9%	100.0%

ตารางที่ 4.20 พบว่าตัวอักษรที่มีอัตราความทนทานต่ำสุดสามอันดับ ได้แก่ ตัวอักษร d, u, และ p ที่มีระยะห่างระหว่างอักษรแบบปกติ โดยมีอัตราความทนทานร้อยละ 32.4, 34.4, และ 34.6 ตามลำดับ

ตารางที่ 4. 21 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามขนาดความหนาของเส้นรบกวน (พิกเซล)

ตัวอักษร	จำนวนครั้งที่ถูกสุม			จำนวนครั้งที่ GSA ตอบไม่ถูก			อัตราความทนทานของตัวอักษร (%)		
	1	3	5	1	3	5	1	3	5
a	27	29	28	18	23	25	66.7%	79.3%	89.3%
b	31	35	31	21	24	31	67.7%	68.6%	100.0%
c	35	31	30	19	24	29	54.3%	77.4%	96.7%
d	22	30	26	11	21	20	50.0%	70.0%	76.9%
e	30	28	23	18	25	23	60.0%	89.3%	100.0%
f	37	23	29	23	20	29	62.2%	87.0%	100.0%
g	31	23	28	17	19	27	54.8%	82.6%	96.4%
h	20	27	23	14	22	19	70.0%	81.5%	82.6%
i	24	30	27	18	27	27	75.0%	90.0%	100.0%
j	31	28	22	16	25	22	51.6%	89.3%	100.0%
k	28	31	36	17	29	36	60.7%	93.5%	100.0%
l	23	24	24	14	20	24	60.9%	83.3%	100.0%
m	28	31	24	18	27	21	64.3%	87.1%	87.5%
n	22	24	30	14	17	27	63.6%	70.8%	90.0%
o	37	29	24	20	27	21	54.1%	93.1%	87.5%

ตารางที่ 4.21 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามขนาดความหนาของเส้นรบกวน (พิกเซล) (ต่อ)

ตัวอักษร	จำนวนครั้งที่ถูกสุม			จำนวนครั้งที่ GSA ตอบไม่ถูก			อัตราความทนทานของตัวอักษร (%)		
	1	3	5	1	3	5	1	3	5
p	26	32	21	10	23	20	38.5%	71.9%	95.2%
q	16	20	20	10	17	19	62.5%	85.0%	95.0%
r	29	26	23	16	21	21	55.2%	80.8%	91.3%
s	30	26	34	24	23	34	80.0%	88.5%	100.0%
t	29	28	30	19	21	27	65.5%	75.0%	90.0%
u	38	31	32	21	25	30	55.3%	80.6%	93.8%
v	22	32	28	11	22	26	50.0%	68.8%	92.9%
w	23	24	35	15	16	33	65.2%	66.7%	94.3%
x	29	25	28	19	19	26	65.5%	76.0%	92.9%
y	23	23	33	14	17	33	60.9%	73.9%	100.0%
z	29	30	31	11	28	29	37.9%	93.3%	93.5%

ตารางที่ 4.21 พบว่าตัวอักษรที่มีอัตราความทนทานต่ำสุดสามอันดับ ได้แก่ ตัวอักษร z, p, d, และ v ที่ถูกรบกวนด้วยเส้นหนา 1 พิกเซล โดยตัวอักษร z และ p มีอัตราความทนทานร้อยละ 37.9 และ 38.5 ตามลำดับ ส่วนตัวอักษร d และ v มีอัตราความทนทานเท่ากันคือร้อยละ 50

ตารางที่ 4. 22 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามสีของเส้นรบกวน

ตัวอักษร	จำนวนครั้งที่ถูกสุ่ม		จำนวนครั้งที่ GSA ตอบไม่ถูก		อัตราความทนทานของตัวอักษร (%)	
	สีเดียวกับตัวอักษร	สีเดียวกับพื้นหลัง	สีเดียวกับตัวอักษร	สีเดียวกับพื้นหลัง	สีเดียวกับตัวอักษร	สีเดียวกับพื้นหลัง
a	38	46	33	33	86.8%	71.7%
b	48	49	42	34	87.5%	69.4%
c	38	58	33	39	86.8%	67.2%
d	34	44	28	24	82.4%	54.5%
e	39	42	29	37	74.4%	88.1%
f	46	43	40	32	87.0%	74.4%
g	39	43	29	34	74.4%	79.1%
h	40	30	36	19	90.0%	63.3%
i	39	42	37	35	94.9%	83.3%
j	46	35	40	23	87.0%	65.7%
k	52	43	47	35	90.4%	81.4%
l	39	32	34	24	87.2%	75.0%
m	40	43	32	34	80.0%	79.1%
n	43	33	38	20	88.4%	60.6%
o	44	46	35	33	79.5%	71.7%
p	39	40	31	22	79.5%	55.0%
q	20	36	17	29	85.0%	80.6%
r	41	37	34	24	82.9%	64.9%
s	47	43	42	39	89.4%	90.7%

ตารางที่ 4.22 อัตราความทนทานของแต่ละตัวอักษร จำแนกตามสีของเส้นรบกวน (ต่อ)

ตัวอักษร	จำนวนครั้งที่ถูกล้อม		จำนวนครั้งที่ GSA ตอบไม่ถูก		อัตราความทนทานของตัวอักษร (%)	
	สีเดียวกับตัวอักษร	สีเดียวกับพื้นหลัง	สีเดียวกับตัวอักษร	สีเดียวกับพื้นหลัง	สีเดียวกับตัวอักษร	สีเดียวกับพื้นหลัง
t	44	43	39	28	88.6%	65.1%
u	58	43	47	29	81.0%	67.4%
v	42	40	37	22	88.1%	55.0%
w	38	44	31	33	81.6%	75.0%
x	43	39	40	24	93.0%	61.5%
y	39	40	35	29	89.7%	72.5%
z	44	46	33	35	75.0%	76.1%

ตารางที่ 4.22 พบว่าตัวอักษรที่มีอัตราความทนทานต่ำสุดสามอันดับ ได้แก่ ตัวอักษร d, p, และ v ที่ถูกรบกวนด้วยเส้นสีเดียวกับพื้นหลังของแคปท์ชาข้อความ โดยตัวอักษร d มีอัตราความทนทานร้อยละ 54.5 ส่วนตัวอักษร p และ v มีอัตราความทนทานเท่ากันคือร้อยละ 55

บทที่ 5 สรุปและอภิปรายผล

5.1 ความนำ

บทสุดท้ายของวิทยานิพนธ์นี้ประกอบด้วยสรุปผลจากการทดลองในห้องปฏิบัติการ เพื่อตอบวัตถุประสงค์ของการวิจัย การอภิปรายเกี่ยวกับผลที่ได้และการนำงานวิจัยนี้ไปใช้ประโยชน์ทั้งเชิงทฤษฎี และเชิงประยุกต์ ตลอดจนข้อจำกัดของงานวิจัยและข้อเสนอแนะสำหรับการศึกษในอนาคต

5.2 การทดลองและลักษณะของหน่วยทดลอง

งานวิจัยนี้เป็นงานวิจัยเชิงทดลองในห้องปฏิบัติการ โดยใช้หน่วยทดลองเป็นนิสิตที่กำลังศึกษาในระดับปริญญาตรีของคณะพาณิชยศาสตร์และการบัญชี แห่งจุฬาลงกรณ์มหาวิทยาลัย จำนวนทั้งสิ้น 360 คน โดยการทดลองแบ่งออกเป็น 18 กลุ่ม แต่ละกลุ่มมีการสุ่มหน่วยทดลองจำนวน 20 คน เพื่อทำงานตามเงื่อนไขที่ผู้วิจัยมอบหมาย

ข้อมูลส่วนบุคคลของหน่วยทดลองจากตารางที่ 4.1 พบว่ามีหน่วยทดลองอายุต่ำกว่า 20 ปี และอายุ 20 ปีขึ้นไปคิดเป็นร้อยละ 36.7 และ 63.3 ตามลำดับ และหน่วยทดลองกำลังศึกษาอยู่ในระดับปริญญาตรีในหลักสูตรต่างๆ ได้แก่ หลักสูตรบัญชีบัณฑิตมากที่สุด คิดเป็นร้อยละ 46.7 รองลงมาคือ หลักสูตรบริหารธุรกิจบัณฑิต หลักสูตรสถิติศาสตร์บัณฑิต และหลักสูตรนานาชาติ คิดเป็นร้อยละ 40.8, 11.4, และ 1.1 ตามลำดับ แสดงให้เห็นว่าหน่วยทดลองมีคุณสมบัติใกล้เคียงกัน และคุณสมบัติด้านระดับการศึกษาของหน่วยทดลองสอดคล้องกับกลุ่มผู้ใช้อินเทอร์เน็ตในประเทศไทย โดยผลสำรวจจำแนกตามระดับการศึกษาพบว่า กลุ่มผู้ใช้งานอินเทอร์เน็ตในประเทศไทยส่วนใหญ่มีการศึกษาระดับปริญญาตรี (สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์, 2556) จึงกล่าวได้ว่าหน่วยทดลองของงานวิจัยนี้เป็นตัวแทนที่ดีของประชากร

5.3 อัตราการยืนยันความเป็นมนุษย์กับระยะห่างระหว่างอักษรของแคปทชาข้อความ

การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ, ชิด (0 พอยต์), และทับซ้อน (ลบ 3.75 พอยต์) พบว่าอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ ไม่แตกต่างกับชุดอักษร

แคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบขีด การพบว่าอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติอาจจะเท่ากับชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบขีดในครั้งนี้อาจเนื่องมาจากระยะห่างทางด้านซ้ายและระยะห่างทางด้านขวาของภาพอักษรบางตัวอักษรมีค่าเป็นติดลบ ดังข้อมูลแสดงในตารางที่ 3.1 โดยจะเห็นว่า ตัวอักษร f และ r มีระยะห่างทางด้านขวามีค่าติดลบ ได้แก่ ลบ 0.55 พอยต์ และ ลบ 0.22 พอยต์ ตามลำดับ และตัวอักษร j มีระยะห่างทางด้านซ้ายมีค่าติดลบ ได้แก่ ลบ 0.73 พอยต์ เมื่อต้องลดระยะห่างระหว่างอักษรของชุดอักษรแคปท์ชาข้อความให้เป็นแบบขีด ค่าระยะห่างระหว่างอักษรอาจเป็นค่าเดียวกันกับระยะห่างระหว่างอักษรของชุดอักษรแคปท์ชาข้อความแบบปกติ ทำให้ผลการเปรียบเทียบพบว่าอัตราการยืนยันความเป็นมนุษย์ไม่แตกต่างกัน

แต่อย่างไรก็ตาม เนื่องจาก (1) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ สูงกว่าชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อนอย่างมีนัยสำคัญทางสถิติ และ (2) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นขีด สูงกว่าชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อนอย่างมีนัยสำคัญทางสถิติ สอดคล้องกับงานวิจัยในอดีตที่กล่าวว่าระยะห่างระหว่างอักษรที่ลดลง (หรือการที่อักษรอยู่ชิดติดกัน) จะทำให้ความสามารถในการยืนยันความเป็นมนุษย์ลดลง และระยะห่างระหว่างอักษรไม่ควรต่ำกว่าลบ 3.75 พอยต์ (5 พิกเซล) หรือไม่ควรให้เกยกัน เพราะจะทำให้ความสามารถในการเข้าใจแคปท์ชาข้อความของมนุษย์ลดลงมากจนถึงไม่สามารถเข้าใจแคปท์ชาข้อความนั้นได้เลย (Bursztein E. et al., 2011)

จุฬาลงกรณ์มหาวิทยาลัย

5.4 อัตราความทนทานของแคปท์ชาข้อความกับระยะห่างระหว่างอักษรของแคปท์ชาข้อความ

การเปรียบเทียบอัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ, ขีด (0 พอยต์), และทับซ้อน (ลบ 3.75 พอยต์) พบว่า (1) อัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติต่ำกว่าชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบขีดอย่างมีนัยสำคัญทางสถิติ (2) อัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ ต่ำกว่าชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อนอย่างมีนัยสำคัญทางสถิติ และ (3) อัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความมีระยะห่างระหว่างอักษรเป็นขีด ต่ำกว่าชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อนอย่างมีนัยสำคัญทางสถิติ

อัตราความทนทานของแคปซูลข้อความกับระยะห่างระหว่างอักขรของแคปซูลข้อความจากผลการทดลองข้างต้นสอดคล้องกับงานวิจัยในอดีต (Bursztein E. et al., 2014) ที่กล่าวว่า การลดระยะห่างระหว่างอักขรจะเพิ่มความทนทานให้กับแคปซูลข้อความจากการโจมตีของโปรแกรมอัตโนมัติ โดยการลดระยะห่างระหว่างอักขรนั้นจะป้องกันการโจมตีด้านการแบ่งส่วน (anti-segmentation) ของโปรแกรมอัตโนมัติ ทำให้เมื่อลดระยะห่างระหว่างอักขรให้น้อยลงจะทำให้โปรแกรมอัตโนมัติไม่สามารถอ่านและตอบแคปซูลข้อความได้ แสดงถึงการเพิ่มความทนทานจากการโจมตีของโปรแกรมอัตโนมัติให้กับแคปซูลข้อความ และในงานวิจัยนี้สรุปได้ว่าชุดอักขรแคปซูลข้อความที่มีระยะห่างระหว่างอักขรแบบทับซ้อนมีอัตราความทนทานของแคปซูลข้อความสูงที่สุด โดยขอค้นพบแนะนำว่า ระยะห่างระหว่างอักขรควรเป็นแบบทับซ้อนเพื่อให้ทนทานที่สุด

5.5 อัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปซูลข้อความกับระยะห่างระหว่างอักขรของแคปซูลข้อความ

เมื่อพิจารณาร่วมกันระหว่างอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปซูลข้อความกับระยะห่างระหว่างอักขรแล้วพบว่า ระยะห่างระหว่างอักขรแบบชิดจะเหมาะสม เพราะเป็นลักษณะที่แคปซูลข้อความทั้งทนทานต่อโปรแกรมอัตโนมัติและมนุษย์ยังสามารถเข้าใจได้

5.6 อัตราการยืนยันความเป็นมนุษย์กับความหนาของเส้นรบกวนของแคปซูลข้อความ

การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักขรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล พบว่า (1) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักขรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าชุดอักขรแคปซูลข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล อย่างมีนัยสำคัญทางสถิติ (2) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักขรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าชุดอักขรแคปซูลข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติ และ (3) อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักขรแคปซูลข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล สูงกว่าชุดอักขรแคปซูลข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติ

อัตราการยืนยันความเป็นมนุษย์กับความหนาของเส้นรบกวนจากผลการทดลองข้างต้นสอดคล้องกับงานวิจัยในอดีต (Bursztein E. et al., 2014) ที่กล่าวว่าเส้นรบกวนที่มีความหนา มากกว่า 5 พิกเซล จะลดความสามารถในการยืนยันความเป็นมนุษย์ เนื่องจากมนุษย์สามารถอ่าน

และตอบแคปท์ชาข้อความได้ถูกต้องน้อยกว่า 50% โดยเส้นรบกวนในลักษณะเส้นตรงเมื่อความหนาของเส้นรบกวนเป็น 1-2 พิกเซล และพาดข้ามชุดอักษรแคปท์ชาข้อความแล้ว มนุษย์ยังสามารถอ่านและตอบแคปท์ชาข้อความได้ถูกต้องประมาณ 90% เมื่อเพิ่มความหนาของเส้นรบกวนเป็น 3 พิกเซล เข้าไปในชุดอักษรแคปท์ชาข้อความแล้วมนุษย์ยังสามารถอ่านและตอบแคปท์ชาข้อความได้ถูกต้องประมาณ 80% และเมื่อเพิ่มความหนาของเส้นรบกวน 5 พิกเซล แล้วมนุษย์ยังสามารถตอบแคปท์ชาข้อความได้ถูกต้องประมาณ 55% แต่เมื่อเพิ่มความหนาของเส้นรบกวน 6 พิกเซล แล้วมนุษย์สามารถตอบแคปท์ชาข้อความได้ถูกต้องเพียง 34% เท่านั้น

ดังนั้นในงานวิจัยนี้เมื่อชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 1 พิกเซล มนุษย์สามารถอ่านและตอบแคปท์ชาข้อความได้ 81.7% เมื่อเพิ่มความหนาของเส้นรบกวนเป็น 3 พิกเซล มนุษย์สามารถอ่านและตอบแคปท์ชาข้อความได้ 60.8% และเมื่อเพิ่มความหนาของเส้นรบกวนเป็น 5 พิกเซล มนุษย์สามารถอ่านและตอบแคปท์ชาข้อความได้ 34.2% ดังนั้นเส้นรบกวนที่ถูกเพิ่มเข้าไปในแคปท์ชาข้อความนั้นควรมีความหนาของเส้นรบกวนไม่เกิน 5 พิกเซล เพราะเส้นรบกวนที่มีความหนามากกว่า 5 พิกเซล จะลดความสามารถในการยืนยันความเป็นมนุษย์ คือมนุษย์จะอ่านและตอบแคปท์ชาข้อความได้ถูกต้องน้อยมากจนอาจจะไม่สามารถอ่านและตอบแคปท์ชาข้อความได้เลย

5.7 อัตราความทนทานของแคปท์ชาข้อความกับความหนาของเส้นรบกวนของแคปท์ชาข้อความ

การเปรียบเทียบอัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล, 3 พิกเซล, และ 5 พิกเซล พบว่า (1) อัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าชุดอักษรแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล อย่างมีนัยสำคัญทางสถิติ (2) อัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล สูงกว่าชุดอักษรแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติ และ (3) อัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล สูงกว่าชุดอักษรแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล อย่างมีนัยสำคัญทางสถิติ

อัตราการความทนทานของแคปท์ชาข้อความกับความหนาของเส้นรบกวนจากผลการทดลองข้างต้นสอดคล้องกับงานวิจัยในอดีตที่กล่าวว่าโปรแกรมอัตโนมัติจะไม่สามารถตอบแคปท์ชาข้อความได้ถูกต้อง เมื่อแคปท์ชาข้อความถูกรบกวนด้วยเส้นรบกวนที่มีความหนามาก (thick) (Yan J. & El

Ahmad A. S., 2008) โดยขั้นตอนการแบ่งส่วนของโปรแกรมอัตโนมัติ นั้น โปรแกรมจะไม่สามารถวิเคราะห์และนำเส้นรบกวนออกไปได้ ทำให้ในขั้นตอนการรู้จำ โปรแกรมอัตโนมัติก็จะไม่สามารถวิเคราะห์ตัวอักษรได้ ส่งผลให้แคปต์ชาข้อความมีความทนทานต่อการโจมตีของโปรแกรมอัตโนมัติ และในงานวิจัยนี้สรุปได้ว่า เมื่อชุดอักขรแคปต์ชาข้อความถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 5 พิกเซล จะมีอัตราความทนทานของแคปต์ชาข้อความสูงที่สุด

5.8 อัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปต์ชาข้อความกับความหนาของเส้นรบกวนของแคปต์ชาข้อความ

เมื่อพิจารณาร่วมกันระหว่างอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปต์ชาข้อความกับความหนาของเส้นรบกวนแล้วพบว่า เส้นรบกวนที่มีความหนา 3 พิกเซล จะเหมาะสมเพราะเป็นลักษณะที่แคปต์ชาข้อความทั้งทนทานต่อโปรแกรมอัตโนมัติและมนุษย์ยังสามารถเข้าใจได้

5.9 อัตราการยืนยันความเป็นมนุษย์กับสีของเส้นรบกวนของแคปต์ชาข้อความ

การเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักขรแคปต์ชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ชาข้อความ และสีเดียวกับพื้นหลังของแคปต์ชาข้อความ พบว่า อัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักขรแคปต์ชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปต์ชาข้อความ และสีเดียวกับพื้นหลังของแคปต์ชาข้อความ ไม่แตกต่างกัน

การไม่พบความแตกต่างนี้ อาจเนื่องมาจากการรับรู้ของมนุษย์ อ้างอิงจากทฤษฎีความสมบูรณ์ของเกสตัลท์ โดยกล่าวไว้ว่าแม้มีข้อมูลที่ขาดหายไป แต่มนุษย์สามารถรับรู้ให้เป็นภาพสมบูรณ์ได้โดยอาศัยประสบการณ์เดิม (Wasserstein et al., 1987) เพราะชุดอักขรแคปต์ชาข้อความไม่ว่าจะถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรของแคปต์ชาข้อความ หรือเส้นที่มีสีเดียวกับพื้นหลังของแคปต์ชาข้อความก็ทำให้ตัวอักษรในแคปต์ชาข้อความมีลักษณะไม่สมบูรณ์เช่นเดียวกันคือเส้นอักษรบางส่วนถูกบดบังหรือขาดหายไป แต่มนุษย์นั้นอาศัยประสบการณ์เดิมเพื่อใช้ในการตอบแคปต์ชาข้อความได้เช่นเดียวกัน

5.10 อัตราความทนทานของแคปท์ชาข้อความกับสีของเส้นรบกวนของแคปท์ชาข้อความ

การเปรียบเทียบอัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปท์ชาข้อความ และสีเดียวกับพื้นหลังของแคปท์ชาข้อความ พบว่าอัตราความทนทานของแคปท์ชาข้อความ เมื่อชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปท์ชาข้อความ และสีเดียวกับพื้นหลังของแคปท์ชาข้อความ ไม่แตกต่างกัน

การไม่พบความแตกต่างนี้ อาจเนื่องมาจากการเพิ่มเส้นรบกวนที่มีความหนา (thick) จะลดความสามารถในการโจมตีด้านการแบ่งส่วนของโปรแกรมอัตโนมัติ (Yan J. & El Ahmad A. S., 2008) เมื่อแคปท์ชาข้อความมีเส้นรบกวนที่ใช้สีเดียวกับตัวอักษรของแคปท์ชาข้อความนั้น โปรแกรมอัตโนมัติไม่สามารถนำเส้นรบกวนออกไปได้ในขั้นตอนของการแบ่งส่วน ทำให้ไม่สามารถวิเคราะห์และตอบตัวอักษรได้ในขั้นตอนของการรู้จำ และเมื่อแคปท์ชาข้อความถูกรบกวนด้วยเส้นรบกวนที่มีสีเดียวกับพื้นหลังของแคปท์ชาข้อความ โปรแกรมอัตโนมัติอาจไม่พบสิ่งรบกวน แต่ในขั้นตอนการรู้จำก็ไม่สามารถวิเคราะห์และตอบตัวอักษรได้เช่นกัน เนื่องจากความไม่สมบูรณ์ของเส้นของตัวอักษรในชุดอักษรแคปท์ชาข้อความ

จากข้อมูลในตารางที่ 4.9 พบว่า เมื่อชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 1 พิกเซล แคปท์ชาข้อความจะมีอัตราความทนทาน 61.7% ชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 3 พิกเซล แคปท์ชาข้อความจะมีอัตราความทนทาน 92.5% และชุดอักษรแคปท์ชาข้อความถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 5 พิกเซล แคปท์ชาข้อความจะมีอัตราความทนทาน 98.3% แสดงให้เห็นว่าเมื่อความหนาของเส้นรบกวนเท่ากับ 5 พิกเซล ชุดอักษรแคปท์ชาข้อความมีอัตราความทนทานของแคปท์ชาที่สูงมาก คือโปรแกรมอัตโนมัติสามารถวิเคราะห์และตอบตัวอักษรได้อย่างถูกต้องได้น้อยมาก และจากข้อมูลการวิเคราะห์รายอักษรในตารางที่ 4.16 พบว่าตัวอักษรที่มีอัตราความทนทานต่ำที่สุดได้แก่ ตัวอักษร d โดยมีอัตราความทนทาน 66.7%

5.11 อัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปท์ชาข้อความกับสีของเส้นรบกวนของแคปท์ชาข้อความ

เมื่อพิจารณาร่วมกันระหว่างอัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปท์ชาข้อความกับสีของเส้นรบกวนแล้วยังไม่พบความแตกต่างระหว่างเส้นรบกวนที่มีสีเดียวกับตัวอักษรแคปท์ชาข้อความและเส้นรบกวนที่มีสีเดียวกับพื้นหลังของแคปท์ชาข้อความ ดังนั้นสีของเส้นรบกวนอาจจะใช้สีเดียวกับตัวอักษรแคปท์ชาข้อความหรือสีเดียวกับพื้นหลังของแคปท์ชาข้อความก็ได้

5.12 อัตราการยืนยันความเป็นมนุษย์กับตัวอักษรในชุดอักษรแคปต์ชาข้อความ

มณีรัตน์ ชาตธีรธรรม (2555) ได้พบว่าตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์สูงสุดที่สุดคือตัวอักษร f และตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์ต่ำที่สุดคือตัวอักษร v เมื่อพิจารณาจากข้อมูลในตารางที่ 4.15 พบว่าตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์สูงสุด ได้แก่ ตัวอักษร k, m, u, และ w และตัวอักษรที่มีอัตราการยืนยันความเป็นมนุษย์ต่ำสุด ได้แก่ ตัวอักษร e ความแตกต่างจากงานวิจัยในอดีตที่เกิดขึ้นนี้อาจเนื่องมาจากตัวแปรที่แตกต่างกัน โดยในการวิจัยของมณีรัตน์ได้กล่าวไว้ว่าสาเหตุที่หน่วยทดลองตอบแคปต์ชาข้อความไม่ถูกต้องนั้นอาจมาจากแคปต์ชาที่แสดงแบบอักษรตัวเขียน และในงานวิจัยนี้ไม่ได้วิเคราะห์แบบอักษรตัวเขียนแต่ได้เลือกวิเคราะห์เส้นรบกวนที่พาดข้ามกึ่งกลางของชุดอักษรแคปต์ชาข้อความจึงอาจทำให้อัตราการยืนยันความเป็นมนุษย์ของตัวอักษร e ต่ำที่สุด เนื่องจากเมื่อมีเส้นรบกวนพาดข้ามตรงกลางของตัวอักษรแล้วจะมีลักษณะคล้ายกับตัวอักษร c

5.13 การนำงานวิจัยไปใช้ (contribution)

5.13.1 การนำงานวิจัยไปใช้ในเชิงทฤษฎี (Theoretical Contribution)

ข้อค้นพบในงานวิจัยนี้สามารถใช้เป็นองค์ความรู้สำหรับการออกแบบแคปต์ชาข้อความในประเด็นต่างๆ ดังนี้

1. ระยะห่างระหว่างอักษรของชุดอักษรแคปต์ชาข้อความที่ใช้ในงานวิจัยนี้มีผลต่ออัตราการยืนยันความเป็นมนุษย์อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยผลการเปรียบเทียบยืนยันว่ามนุษย์สามารถอ่านและตอบชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นปกติและชิด (0 พอยต์) ได้ถูกต้องมากกว่าชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรแบบทับซ้อน (ลบ 3.75 พอยต์) สอดคล้องกับงานวิจัยในอดีต (Bursztein E. et al., 2011) ที่กล่าวว่าระยะห่างระหว่างอักษรไม่ควรต่ำกว่าลบ 3.75 พอยต์ (5 พิกเซล) หรือจนเกินไป เพราะจะทำให้ความสามารถในการเข้าใจแคปต์ชาข้อความของมนุษย์ลดลงมากจนถึงไม่สามารถเข้าใจแคปต์ชาข้อความนั้นได้เลย และมีผลต่ออัตราความทนทานของแคปต์ชาข้อความอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยผลการเปรียบเทียบยืนยันว่าชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบทับซ้อนเป็นแคปต์ชาข้อความที่สามารถป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ดีกว่าชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบชิด และชุดอักษรแคปต์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบชิดเป็นแคปต์ชาข้อความที่สามารถป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ดีกว่าชุดอักษร

แคปต์ชาข้อความที่มีระยะห่างอักษรเป็นปกติ สอดคล้องกับงานวิจัยในอดีต (Bursztein E. et al., 2014) ที่กล่าวว่า การลดระยะห่างระหว่างอักษรนั้นจะป้องกันการโจมตีด้านการแบ่งส่วน (anti-segmentation) ของโปรแกรมอัตโนมัติ

2. ความหนาของเส้นรบกวนที่พาดข้ามชุดอักษรแคปต์ชาข้อความที่ใช้ในงานวิจัยนี้มีผลต่ออัตราการยืนยันความเป็นมนุษย์อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยผลการเปรียบเทียบยืนยันว่า (1) ชุดอักษรแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 1 พิกเซล เป็นแคปต์ชาข้อความที่มนุษย์สามารถอ่านและตอบได้ถูกต้องมากกว่าชุดอักษรแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 3 และ 5 พิกเซล และ (2) ชุดอักษรแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 3 พิกเซล เป็นแคปต์ชาข้อความที่มนุษย์สามารถอ่านและตอบได้ถูกต้องมากกว่าชุดอักษรแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 5 พิกเซล สอดคล้องกับงานวิจัยในอดีต (Bursztein E. et al., 2014) ที่กล่าวว่าเมื่อเส้นรบกวนที่มีความหนา 1-2 พิกเซล พาดข้ามชุดอักษรแคปต์ชาข้อความแล้ว มนุษย์ยังสามารถอ่านและตอบแคปต์ชาข้อความได้ถูกต้องประมาณ 90%

นอกจากนี้ความหนาของเส้นรบกวนที่พาดข้ามชุดอักษรแคปต์ชาข้อความที่ใช้ในงานวิจัยนี้ยังมีผลต่ออัตราความทนทานของแคปต์ชาข้อความอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยผลการเปรียบเทียบยืนยันว่า (1) ชุดอักษรแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 5 พิกเซล เป็นแคปต์ชาข้อความที่สามารถป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ดีกว่าชุดอักษรแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 1 และ 3 พิกเซล และ (2) ชุดอักษรแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 3 พิกเซล เป็นแคปต์ชาข้อความที่สามารถป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ดีกว่าชุดอักษรแคปต์ชาข้อความที่ถูกรบกวนด้วยเส้นรบกวนที่มีความหนา 1 พิกเซล สอดคล้องกับงานวิจัยในอดีต (Yan J. & El Ahmad A. S., 2008) ที่กล่าวว่าโปรแกรมอัตโนมัติจะไม่สามารถตอบแคปต์ชาข้อความได้ถูกต้อง เมื่อแคปต์ชาข้อความถูกรบกวนด้วยเส้นรบกวนที่มีความหนามาก (thick)

3. สีของเส้นรบกวนที่พาดข้ามชุดอักษรแคปต์ชาข้อความที่ใช้ในงานวิจัยนี้ไม่มีผลต่ออัตราการยืนยันความเป็นมนุษย์ และอัตราความทนทานของแคปต์ชาข้อความ กล่าวคือการเปรียบเทียบเส้นรบกวนที่มีสีเดียวกับตัวอักษรของแคปต์ชาข้อความและเส้นรบกวนที่มีสีเดียวกับพื้นหลังของแคปต์ชาข้อความไม่พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ชุดอักษรแคปต์ชาข้อความที่ถูกพาดข้ามด้วยเส้นรบกวนที่มีสีเดียวกับตัวอักษรของแคปต์ชาข้อความ หรือสีเดียวกับพื้นหลังของแคปต์ชาข้อความต่างก็ทำให้เส้นของตัวอักษรไม่สมบูรณ์เช่นเดียวกัน ดังนั้น

การตอบแคปท์ชาข้อความของมนุษย์อาศัยการรับรู้ตามกฎความสมบูรณ์ของเกสตัลท์ที่กล่าวไว้ว่าแม้มีข้อมูลที่ขาดหายไป แต่มนุษย์สามารถรับรู้ให้เป็นภาพสมบูรณ์ได้โดยอาศัยประสบการณ์เดิม (Wasserstein et al., 1987)

5.13.2 การนำงานวิจัยไปใช้ในเชิงปฏิบัติ (Practical Contribution)

นักพัฒนาซอฟต์แวร์ทางธุรกิจสามารถประยุกต์ข้อค้นพบของงานวิจัยนี้ได้ดังนี้

1. นักพัฒนาซอฟต์แวร์ทางธุรกิจอาจเลือกพัฒนาแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบชิด (0 พอยต์) เพื่อให้แคปท์ชาข้อความที่พัฒนาขึ้นเป็นแคปท์ชาข้อความที่มนุษย์สามารถเข้าใจได้ง่าย โดยชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบชิดนั้นมนุษย์สามารถอ่านและตอบได้ถูกต้องเทียบเท่ากับชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบปกติ แต่ชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรเปิดชิดยังสามารถป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ดีกว่าชุดอักษรแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นปกติ
2. นักพัฒนาซอฟต์แวร์ทางธุรกิจอาจเลือกพัฒนาแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นตรงพาดข้ามชุดอักษรแคปท์ชาข้อความที่มีความหนา 3 พิกเซล เพื่อให้แคปท์ชาข้อความที่พัฒนาขึ้นมีสิ่งรบกวนแล้วมนุษย์ยังสามารถอ่านและตอบแคปท์ชาข้อความได้ง่ายกว่าชุดอักษรแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 5 พิกเซล และชุดอักษรแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นตรงพาดข้ามชุดอักษรแคปท์ชาข้อความที่มีความหนา 3 พิกเซลสามารถป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ดีกว่าชุดอักษรแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล
3. นักพัฒนาซอฟต์แวร์ทางธุรกิจไม่ควรพัฒนาแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบทับซ้อน (ลบ 3.75 พอยต์) ถึงแม้ว่าแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบทับซ้อนสามารถป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ในอัตราที่สูงที่สุดเมื่อเทียบกับแคปท์ชาข้อความที่มีระยะห่างระหว่างอักษรเป็นแบบปกติและชิด แต่มนุษย์ก็เข้าใจได้ยากเช่นกัน
4. นักพัฒนาซอฟต์แวร์ทางธุรกิจไม่ควรพัฒนาแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นตรงพาดข้ามชุดอักษรแคปท์ชาข้อความที่มีความหนา 5 พิกเซล ถึงแม้ว่าแคปท์ชาข้อความที่ถูกรบกวนด้วยเส้นตรงพาดข้ามชุดอักษรแคปท์ชาข้อความที่มีความหนา 5 พิกเซล สามารถป้องกันการโจมตีจากโปรแกรมอัตโนมัติได้ในอัตราที่สูงที่สุดเมื่อเทียบกับเส้นรบกวนที่มีความหนา 1 และ 3 พิกเซล แต่มนุษย์ก็เข้าใจได้ยากเช่นกัน

5.14 ข้อจำกัดของงานวิจัยและข้อเสนอแนะ

1. เนื่องจากงานวิจัยนี้เป็นงานวิจัยที่เก็บข้อมูลโดยการทดลองในห้องปฏิบัติการ และผู้วิจัยได้กำหนดขั้นตอนการทดลองให้หน่วยทดลองปฏิบัติตามอาจจะทำให้สภาวะนั้นแตกต่างจากสถานการณ์จริง อีกทั้งในงานวิจัยนี้ ผู้วิจัยได้กำหนดให้หน่วยทดลองตอบแคปช่าข้อความเพียงครั้งเดียวเท่านั้น และถึงแม้ว่าหน่วยทดลองจะตอบแคปช่าข้อความไม่ถูกต้องก็สามารถเข้าไปทำแบบทดสอบต่อบดสีได้ จึงอาจทำให้ผลที่ได้จากการทดลองอาจมีผลแตกต่างกับการใช้งานจริง

2. การกำหนดระยะห่างระหว่างอักษรเป็นแบบซิด ผู้วิจัยใช้วิธีปรับให้ระยะห่างด้านขวาของภาพอักษรแรกและระยะห่างด้านซ้ายของภาพอักษรหลังมีค่า 0 พอยต์ แต่เนื่องจากระยะห่างด้านขวาและด้านซ้ายของภาพอักษรของบางตัวอักษรนั้นมีค่าติดลบอยู่แล้ว (ดังข้อมูลในตารางที่ 3.1) ทำให้เมื่อคำนวณระยะห่างที่ต้องปรับระยะห่างระหว่างอักษรเป็น 0 พอยต์ แล้วยังไม่เพียงพอให้ตัวอักษรนั้นซิดติดกันพอดี งานในอนาคตจึงอาจเลือกปรับระยะห่างระหว่างอักษรให้ซิดโดยใช้ค่าระยะห่างที่แสดงในภาคผนวก ค เพื่อให้เส้นอักษรของชุดอักษรแคปช่าข้อความนั้นมีลักษณะที่ซิดติดกันพอดี

3. เส้นรบกวนที่พาดข้ามชุดอักษรแคปช่าข้อความในงานวิจัยนี้เป็นเส้นตรงที่พาดข้ามบริเวณตรงกลางของชุดอักษรแคปช่าข้อความ ทำให้มนุษย์อ่านตัวอักษรเป็นอีกตัวอักษรหนึ่ง จากข้อมูลในตารางที่ 4.15 พบว่าตัวอักษร e มีอัตราการยืนยันความเป็นมนุษย์ต่ำที่สุด อาจเนื่องมาจากตัวอักษร e มีเส้นอักษรตรงกลางของตัวอักษร เมื่อถูกพาดข้ามด้วยเส้นรบกวนบริเวณตรงกลางทำให้มนุษย์มีโอกาสอ่านได้ว่าเป็นตัวอักษร c งานในอนาคตจึงอาจเลือกวิเคราะห์เส้นรบกวนที่พาดข้ามตำแหน่งอื่นของชุดอักษรของแคปช่าข้อความ

4. การวิเคราะห์ความทนทานของแคปช่าข้อความ ผู้วิจัยเลือกใช้โปรแกรม GSA Captcha Breaker ซึ่งเป็นโปรแกรมแคปช่าโซลเวอร์ที่ใช้กระบวนการไอซีอาร์ขั้นสูงในการตอบแคปช่าข้อความ และมีอัตราการตอบได้ถูกต้องมากที่สุดจากแคปช่าข้อความหลายรูปแบบ (Beheshti and Liatsis, 2015) ผู้วิจัยใช้โปรแกรมรุ่นทดลอง โดยเลือกกำหนดการวิเคราะห์แคปช่าข้อความด้วยวิธีที่ดีที่สุด (find best solutions) ที่ OCR ของโปรแกรม GSA Captcha Breaker สามารถจะวิเคราะห์ได้ แต่โปรแกรม GSA Captcha Breaker แบบจ่ายค่าบริการจะสามารถใช้บริการจากผู้ให้บริการ Captcha Solver อื่นๆได้ โดยการวิเคราะห์ชุดอักษรแคปช่าข้อความของผู้ให้บริการแต่ละรายอาจใช้มนุษย์มาในการอ่านและตอบแคปช่าข้อความ หรือใช้ OCR ที่มีคุณสมบัติแตกต่างกันออกไปทำให้เพิ่มวิธีในการวิเคราะห์ชุดอักษรแคปช่ามากกว่าการใช้โปรแกรมรุ่นทดลอง จึงอาจจะมีผลที่ได้จากการทดลองอาจมีผลแตกต่างกับการใช้งานจริง

5. ผลการวิเคราะห์อัตราการยืนยันความเป็นมนุษย์และอัตราความทนทานของแคปTCHA ข้อความในลักษณะต่างๆ ได้แก่ ระยะห่างระหว่างอักษร ความหนาของเส้นรบกวน และสีของเส้นรบกวนนั้นเป็นข้อสรุปจากขอบเขตที่ได้กำหนดใช้ตัวอักษร Arial ขนาด 24 พอยต์ เท่านั้น หากนำไปปรับใช้กับแบบอักษรและขนาดอักษรอื่น อาจให้ข้อสรุปที่แตกต่างกันออกไป

6. การใช้งานเว็บไซต์ที่เป็นเครื่องมือในการทดลองนั้น เมื่อหน่วยทดลองเข้าสู่เว็บไซต์แล้ว หากต้องการเปลี่ยนชุดอักษรแคปTCHA หน่วยทดลองสามารถกดปุ่ม F5 เพื่อให้ระบบจะแสดงชุดอักษรแคปTCHAใหม่ โดยการเปลี่ยนชุดอักษรแคปชทานี้จะไม่นับว่าหน่วยทดลองตอบแคปชชาผิด ผู้วิจัยจะนับคำตอบของหน่วยทดลองเมื่อหน่วยทดลองได้ตอบแคปชชาแล้วเท่านั้น



รายการอ้างอิง

ภาษาไทย

- กัลยา วาณิชย์บัญชา. (2553). สถิติสำหรับงานวิจัย. กรุงเทพมหานคร: ธรรมสาร.
- ขวัญหทัย สันติบุตร. (2549). ผลกระทบของความล่าช้าของการแสดงผล รูปแบบของเมนู การใช้ข้อมูลป้อนกลับต่อประสิทธิภาพของการใช้งานบนเว็บแอปพลิเคชัน. (วิทยานิพนธ์ปริญญา มหาบัณฑิต), จุฬาลงกรณ์มหาวิทยาลัย,
- มณีรัตน์ ชาตรีงสรณ์. (2555). ผลของแบบอักษร การหมุนเอียง จำนวนอักษร และชุดอักขระต่อ อัตราการยืนยัน ความเป็นมนุษย์ และอัตราความทนทานของแคปชชา. (วิทยานิพนธ์ ปริญญา มหาบัณฑิต), จุฬาลงกรณ์มหาวิทยาลัย,
- วงศกร รัตติรุจิเสวต. (2557). ผลของระยะห่างระหว่างอักษร ตำแหน่งอักษร จำนวนอักษร และคู่สี อักษรและพื้นหลังต่ออัตราการยืนยันความเป็นมนุษย์ และอัตราความทนทานของแคปชชา ข้อความ. (วิทยานิพนธ์ปริญญา มหาบัณฑิต), จุฬาลงกรณ์มหาวิทยาลัย,
- สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์. (2556). รายงานผลการสำรวจพฤติกรรมผู้ใช้ อินเทอร์เน็ตในประเทศไทยปี 2556. Retrieved from กรุงเทพมหานคร:

ภาษาอังกฤษ

- Beheshti S. M. R. S., & Liatsis P. (2015). *CAPTCHA Usability and Performance, How to Measure the Usability Level of Human Interactive Applications Quantitatively and Qualitatively?* Paper presented at the Developments of E-Systems Engineering (DeSE).
- Bursztein E., Martin M., & Mitchell J. C. (2011). Text-based CAPTCHA Strengths and Weakness. *Paper presented at the Proceedings of the 18th ACM conference on Computer and Communications Security (CCS'11)*.
- Bursztein E., Moscicki A., Fabry C., Bethard S., Mitchell J. C., & Jurafsky D. (2014). Easy Does It: More Usable CAPTCHAs. *Paper presented at the Proceedings of the ACM conference on Human Factors in Computing Systems (CHI 2014)*.
- Corporation, M. (2014). Glyphs. Retrieved from [http://msdn.microsoft.com/en-us/library/ms753170\(v=vs.110\).aspx](http://msdn.microsoft.com/en-us/library/ms753170(v=vs.110).aspx)

- El Ahmand A. S., Yan J., & Tayara M. (2011). The robustness of Google CAPTCHA's. *UK: Computing Science*.
- Fontlab Ltd (Producer). (2006). Fontlab Studio 5. Retrieved from <http://www.dtpsoft.de/downloads/manuals/FLS5WinManual.pdf>
- Kazem Taghva, Julie Borsack, & Allen Condit. (1996). Evaluation of model-based retrieval effectiveness with OCR text. *ACM Transactions on Information Systems (TOIS)*, 14(1), 64-93.
- Kolupave A., & Ogijenko J. (2008). CAPTCHAs: Humans vs Bots. *Security and Privacy IEEE*.
- Kurt Alfred Kluever, & Richard Zanibbi. (2009). Balancing usability and security in a video CAPTCHA. *Proceedings of the 5th Symposium on Usable Privacy and Security*.
- Luis von Ahn, Manuel Blum, & John Langford. (2004). Telling humans and computers apart automatically. *Communications of the ACM*, 47(2), 56-60.
- Marti Motoyama, Kirill Levchenko, Chris Kanich, Damon McCoy, Geoffrey M. Voelker, & Stefan Savage (Producer). (2010). Re: CAPTCHAs-Understanding CAPTCHA-Solving Services in an Economic Context. Retrieved from http://static.usenix.org/event/sec10/tech/full_papers/Motoyama.pdf
- Montree Imsamai, & Suphakant Phimoltares. (2010). 3D CAPTCHA: A Next Generation of the CAPTCHA. In (07 June 2010 ed.): *Information Science and Applications (ICISA)*.
- Nielsen J., & Pernice K. (Producer). (2009). How to Conduct Eyetracking Studies. Retrieved from <http://www.nngroup.com/reports/how-to-conduct-eyetrackingstudies>
- Nuttanont Hongwarittorn, & Suttikiat Meelap. (2015). *Gestalt geometric CAPTCHA*. Paper presented at the Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia.
- Shirali-Shahreza M. H., & Shirali-Shahreza M. (2007). Multilingual CAPTCHA. *Paper presented at the Proceedings of the 5th IEEE International Conference on Computational Cybernetics (ICCC 2007)*.

Tangmanee C., & Sujarit – Apirak P. (2008). *Attitude, Cognition, and awareness of Thai internet users and Thai E-commerce entrepreneurs towards CAPTCHA.*

(Master's Project), Chulalongkorn University,

Wasserstein, Zappulla, Rosen, Gerstman, & Rock. (1987). *In search of closure: Subjective contour illusions, gestalt completion tests, and implications.* Paper presented at the The annual meeting of the International Neuropsychological Society, Atlanta.

Yan J., & El Ahmad A. S. (2008). A Low-cost Attack on a Microsoft CAPTCHA. *Paper presented at the Proceeding of the 15th ACM conference on Computer and Communications Security (CCS'08).*



ภาคผนวก

ภาคผนวก ก

ใบงาน

ตัวอย่างใบงานกลุ่มที่ 1

1. โปรตเข้าเว็บไซต์เพื่อทำแบบทดสอบตาบอดสี โดยเข้าไปที่ www.colourblindness.com

2. เมื่อเข้ามาที่เว็บไซต์แล้ว กรุณาเข้าสู่ระบบ (login) โดยใช้

Username: student001

Password: student001

3. เมื่อเข้าสู่ระบบแล้วให้ตอบแบบสอบถามข้อมูลส่วนบุคคล

4. เมื่อตอบแบบสอบถามข้อมูลส่วนบุคคลเสร็จ ให้กดปุ่ม Next เพื่อทำแบบทดสอบตาบอดสี

5. เมื่อตอบแบบทดสอบตาบอดสีเสร็จ กรุณาติดต่อผู้วิจัย

ภาคผนวก ข
สมการคำนวณค่า z

สมการคำนวณค่า z (กัลยา วานิชย์บัญชา, 2553) สามารถคำนวณได้ดังนี้

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

เมื่อ \hat{p}_1 และ \hat{p}_2 คือ ค่าสัดส่วนที่ได้จากกลุ่มตัวอย่างที่ 1 และ 2 ตามลำดับ

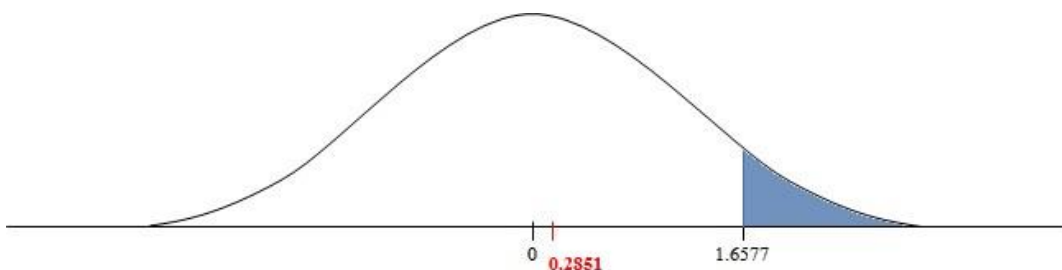
n_1 และ n_2 คือ ขนาดของกลุ่มตัวอย่างที่ 1 และ 2 ตามลำดับ

จาก x_1 และ x_2 คือ จำนวนคนในกลุ่มตัวอย่างที่ 1 และ 2 ตามลำดับ

จะได้ $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$

1. สมการเพื่อการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นปกติ และขีด ในหัวข้อที่ 4.4

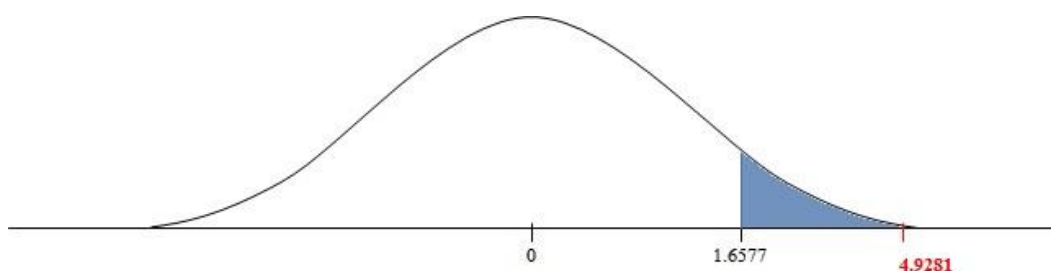
$$z = \frac{0.70000 - 0.6830}{\sqrt{0.6917(1 - 0.6917)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 0.2851$$



ภาพที่ ข. 1 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นปกติ และขีด

2. สมการเพื่อการเปรียบเทียบอัตราการเรียนรู้ยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นปกติ และทับซ้อน ในหัวข้อที่ 4.4

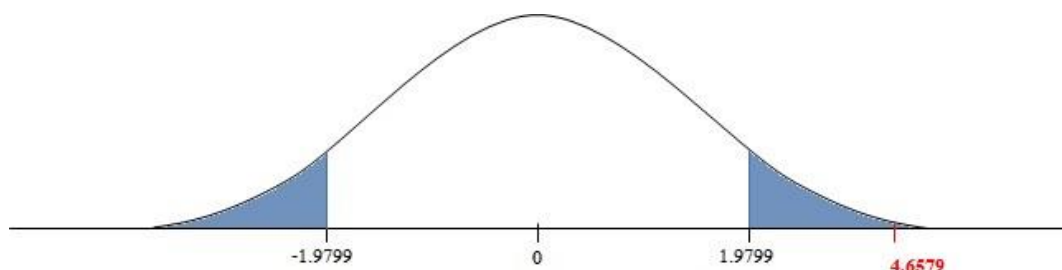
$$z = \frac{0.70000 - 0.3830}{\sqrt{0.5417(1 - 0.5417)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 4.9281$$



ภาพที่ ข. 2 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการเรียนรู้ยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นปกติ และทับซ้อน

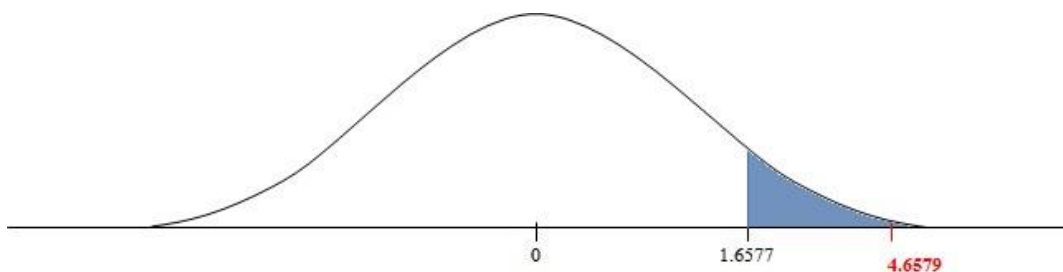
3. สมการเพื่อการเปรียบเทียบอัตราการเรียนรู้ยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นซิด และทับซ้อน ในหัวข้อที่ 4.4

$$z = \frac{0.6830 - 0.3830}{\sqrt{0.5333(1 - 0.5333)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 4.5679$$



ภาพที่ ข. 3 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการเรียนรู้ยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นซิด และทับซ้อนแบบสองทาง

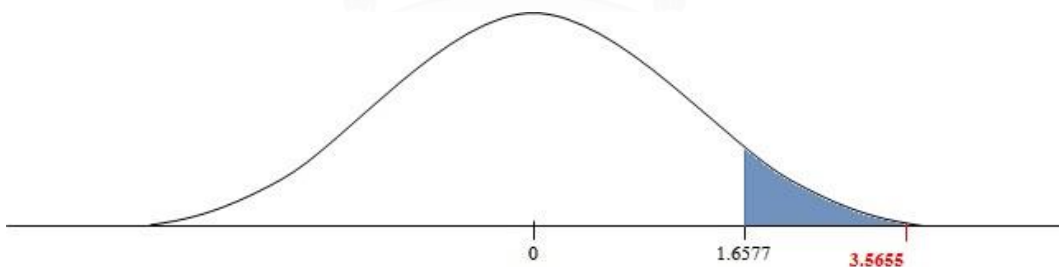
$$z = \frac{0.6830 - 0.3830}{\sqrt{0.5333(1 - 0.5333)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 4.5679$$



ภาพที่ ข. 4 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความมีระยะห่างระหว่างอักษรเป็นขีด และทับซ้อนแบบทางเดียว

4. สมการเพื่อการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 3 พิกเซล ในหัวข้อที่ 4.5

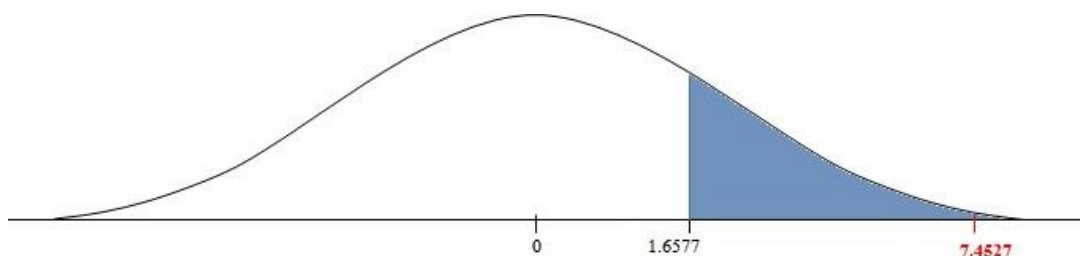
$$z = \frac{0.8170 - 0.6080}{\sqrt{0.7125(1 - 0.7125)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 3.5655$$



ภาพที่ ข. 5 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 3 พิกเซล

5. สมการเพื่อการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทซ์ข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 5 พิกเซล ในหัวข้อที่ 4.5

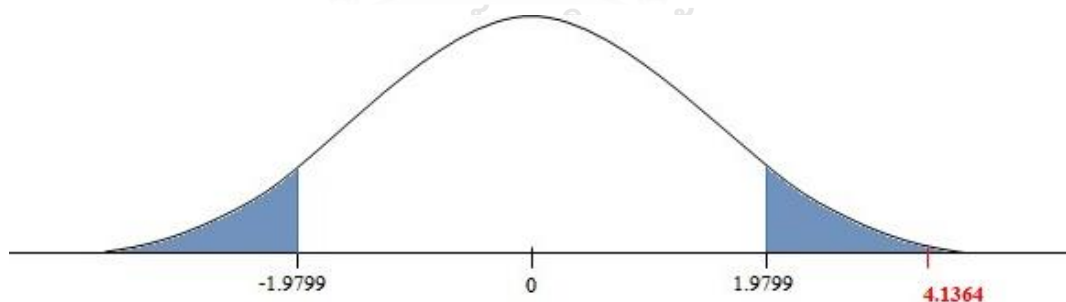
$$z = \frac{0.8170 - 0.3420}{\sqrt{0.5792(1 - 0.5792)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 7.4527$$



ภาพที่ ข. 6 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 5 พิกเซล

6. สมการเพื่อการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล ในหัวข้อที่ 4.5

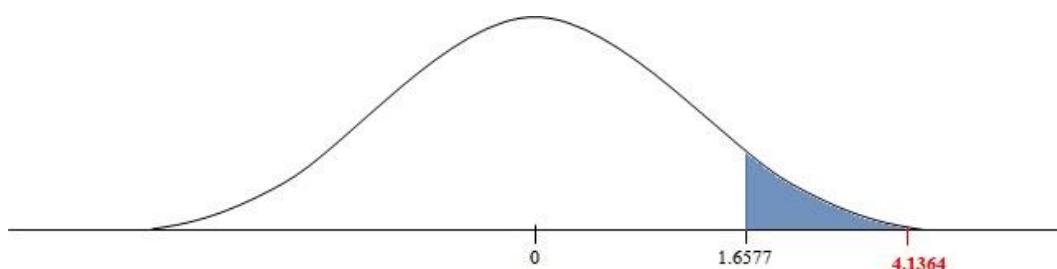
$$z = \frac{0.6080 - 0.3420}{\sqrt{0.4750(1 - 0.4750)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 4.1364$$



ภาพที่ ข. 7 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล แบบ

สองทาง

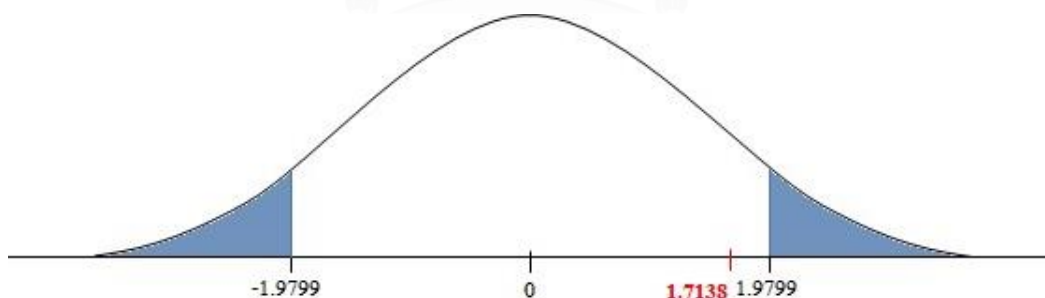
$$z = \frac{0.6080 - 0.3420}{\sqrt{0.4750(1 - 0.4750)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 4.1364$$



ภาพที่ ข. 8 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทชาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล แบบทางเดียว

7. สมการเพื่อการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปทชาข้อความ และสีเดียวกับพื้นหลังของแคปทชาข้อความ ในหัวข้อที่ 4.6

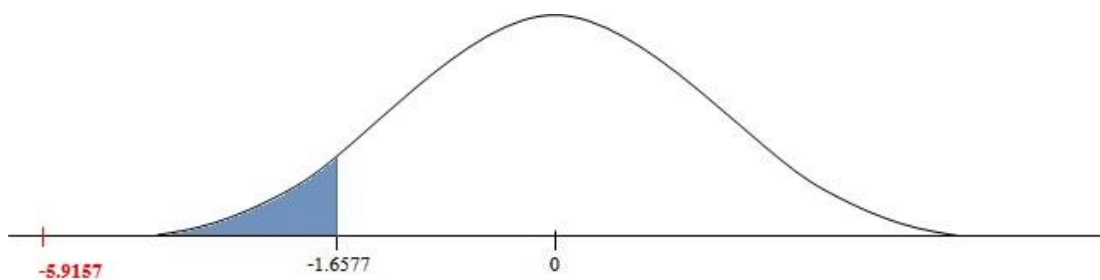
$$z = \frac{0.6330 - 0.5440}{\sqrt{0.5889(1 - 0.5889)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 1.7138$$



ภาพที่ ข. 9 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ เมื่อชุดอักษรแคปทชาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปทชาข้อความ และสีเดียวกับพื้นหลังของแคปทชาข้อความ

8. สมการเพื่อการเปรียบเทียบอัตราความทนทานของแคปทชาข้อความ เมื่อชุดอักษรแคปทชาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ และชิด ในหัวข้อที่ 4.8

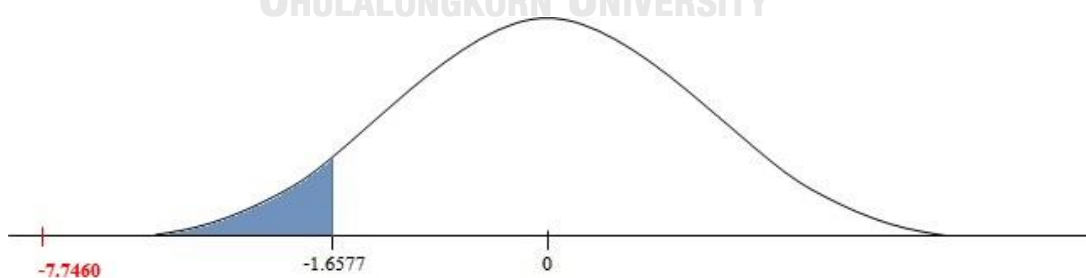
$$z = \frac{0.6000 - 0.9250}{\sqrt{0.7625(1 - 0.7625)\left(\frac{1}{120} + \frac{1}{120}\right)}} = -5.9157$$



ภาพที่ ข. 10 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ และชิด

9. สมการเพื่อการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ และทับซ้อน ในหัวข้อที่ 4.8

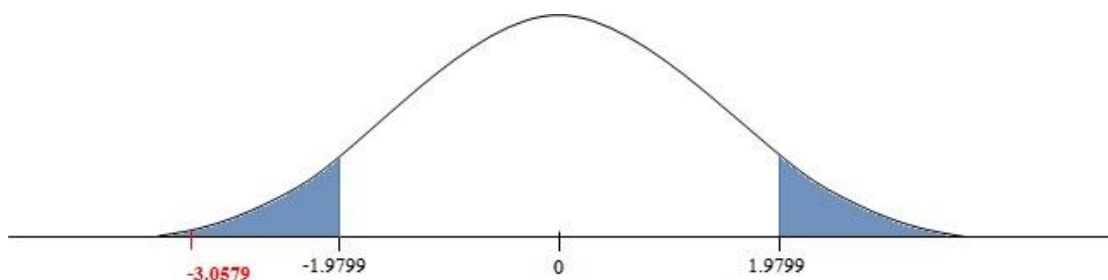
$$z = \frac{0.6000 - 1.0000}{\sqrt{0.8000(1 - 0.8000)\left(\frac{1}{120} + \frac{1}{120}\right)}} = -7.7460$$



ภาพที่ ข. 11 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นปกติ และทับซ้อน

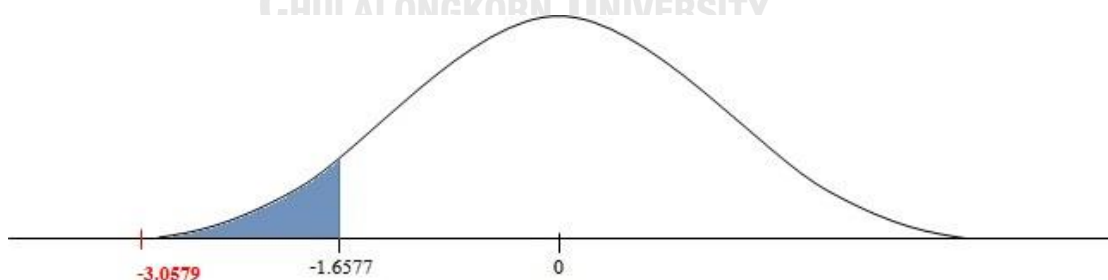
10. สมการเพื่อการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นซิด และทับซ้อน ในหัวข้อที่ 4.8

$$z = \frac{0.9250 - 1.0000}{\sqrt{0.9625(1 - 0.9625)\left(\frac{1}{120} + \frac{1}{120}\right)}} = -3.0579$$



ภาพที่ ข. 12 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นซิด และทับซ้อนแบบสองทาง

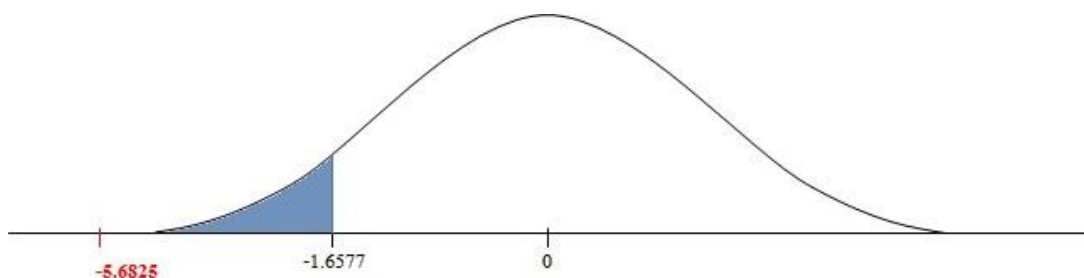
$$z = \frac{0.9250 - 1.0000}{\sqrt{0.9625(1 - 0.9625)\left(\frac{1}{120} + \frac{1}{120}\right)}} = -3.0579$$



ภาพที่ ข. 13 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความมีระยะห่างระหว่างอักษรเป็นซิดและทับซ้อนแบบทางเดียว

11. สมการเพื่อการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 3 พิกเซล ในหัวข้อที่ 4.9

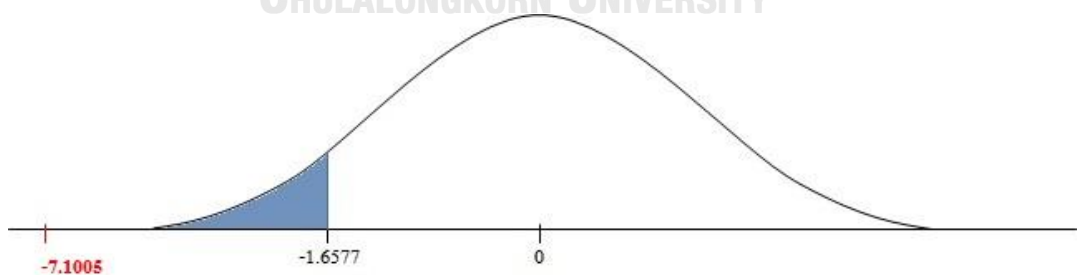
$$z = \frac{0.6170 - 0.9250}{\sqrt{0.7708(1 - 0.7708)\left(\frac{1}{120} + \frac{1}{120}\right)}} = -5.6825$$



ภาพที่ ข. 14 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 3 พิกเซล

12. สมการเพื่อการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 5 พิกเซล ในหัวข้อที่ 4.9

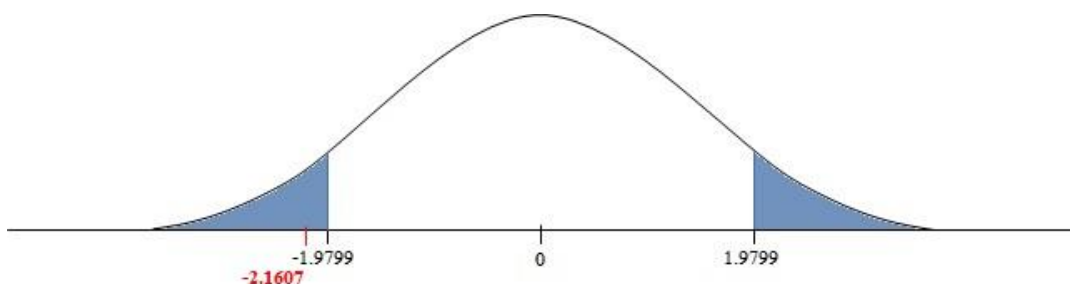
$$z = \frac{0.6170 - 0.9830}{\sqrt{0.8000(1 - 0.8000)\left(\frac{1}{120} + \frac{1}{120}\right)}} = -7.1005$$



ภาพที่ ข. 15 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 1 พิกเซล และ 5 พิกเซล

13. สมการเพื่อการเปรียบเทียบอัตราความทนทานของแคปต์ซาข้อความ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล ในหัวข้อที่ 4.9

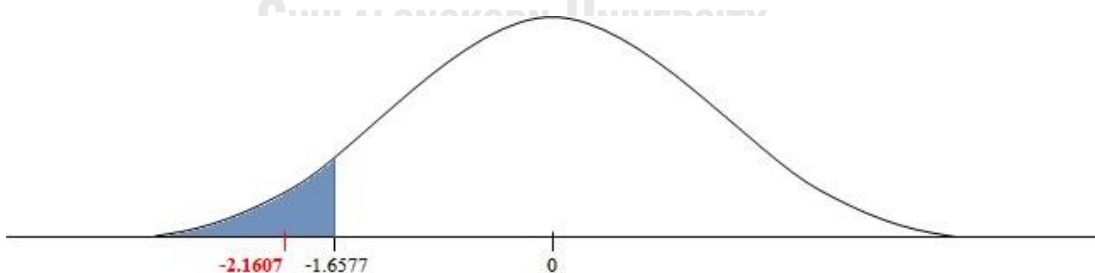
$$z = \frac{0.9250 - 0.9830}{\sqrt{0.9542(1 - 0.9542)\left(\frac{1}{120} + \frac{1}{120}\right)}} = -2.1607$$



ภาพที่ ข. 16 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปต์ซาข้อความ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล

แบบสองทาง

$$z = \frac{0.9250 - 0.9830}{\sqrt{0.9542(1 - 0.9542)\left(\frac{1}{120} + \frac{1}{120}\right)}} = -2.1607$$

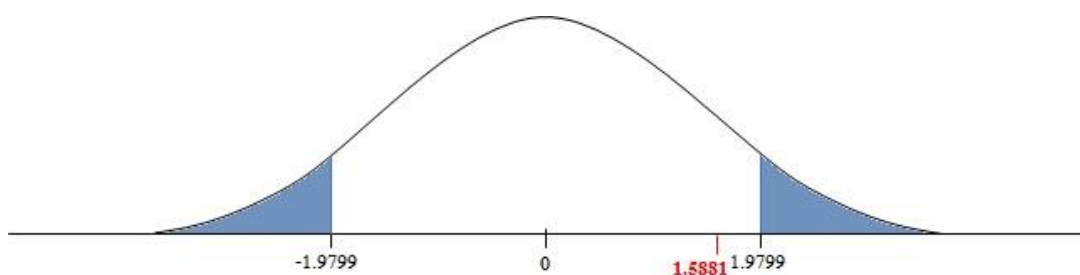


ภาพที่ ข. 17 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปต์ซาข้อความ เมื่อชุดอักษรแคปต์ซาข้อความถูกรบกวนด้วยเส้นที่มีความหนา 3 พิกเซล และ 5 พิกเซล

แบบทางเดียว

14. สมการเพื่อการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปท์ซาข้อความ และสีเดียวกับพื้นหลังของแคปท์ซาข้อความ ในหัวข้อที่ 4.10

$$z = \frac{0.8720 - 0.8110}{\sqrt{0.8417(1 - 0.8417)\left(\frac{1}{120} + \frac{1}{120}\right)}} = 1.5881$$



ภาพที่ ข. 18 กราฟแสดงผลการทดสอบสมมติฐานของการเปรียบเทียบอัตราความทนทานของแคปท์ซาข้อความ เมื่อชุดอักษรแคปท์ซาข้อความถูกรบกวนด้วยเส้นที่มีสีเดียวกับตัวอักษรแคปท์ซาข้อความ และสีเดียวกับพื้นหลังของแคปท์ซาข้อความ

ภาคผนวก ค

ค่าระยะห่างระหว่างอักษรแบบขีด (0 พอยต์)

ตารางที่ ค. 1 ผลรวมของระยะห่างระหว่างอักษรแต่ละคู่อักษร (หน่วย Ps)

		ตัวอักษรแรก												
		a	b	c	d	e	f	g	h	i	j	k	l	m
	ตัวอักษรหลัง	a	79	78	46	109	78	1	104	105	105	106	40	107
b		109	107	75	139	108	31	133	134	134	135	70	137	131
c		82	81	49	112	81	4	107	108	108	109	43	110	104
d		77	76	44	107	76	0	102	103	103	104	38	105	99
e		80	78	46	110	79	2	104	105	105	106	41	108	102
f		52	51	19	82	51	-26	77	78	78	79	13	80	74
g		75	74	42	105	74	-2	100	101	101	102	36	103	97
h		109	108	76	139	108	32	134	135	135	136	70	137	131
i		110	108	76	140	109	32	134	135	135	136	71	138	132
j		-3	-5	-37	27	-4	-81	21	22	22	23	-42	25	19
k		110	108	76	140	109	32	134	135	135	136	71	138	132
l		107	106	74	137	106	30	132	133	133	134	68	135	129
m		109	108	76	139	108	32	134	135	135	136	70	137	131
n		109	108	76	139	108	32	134	135	135	136	70	137	131
o		76	75	43	106	75	-1	101	102	102	103	37	104	98
p		109	108	76	139	108	32	134	135	135	136	70	137	131
q		78	77	45	108	77	0	103	104	104	105	39	106	100
r		108	107	75	138	107	31	133	134	134	135	69	136	130
s		74	72	40	104	73	-4	98	99	99	100	35	102	96
t		61	59	27	91	60	-17	85	86	86	87	22	89	83
u		107	106	74	137	106	30	132	133	133	134	68	135	129
v		56	54	22	86	55	-22	80	81	81	82	17	84	78
w		46	44	12	76	45	-32	70	71	71	72	7	74	68
x		50	49	17	80	49	-28	75	76	76	77	11	78	72
y		59	58	26	89	58	-19	84	85	85	86	20	87	81
z		63	61	29	93	62	-15	87	88	88	89	24	91	85

ตารางที่ ค.1 ผลรวมของระยะห่างระหว่างอักษรแต่ละคู่อักษร (หน่วย Ps) (ต่อ)

		ตัวอักษรแรก												
		n	o	p	q	r	s	t	u	v	w	x	y	z
		ตัวอักษรหลัง	a	106	74	77	109	23	75	44	109	48	44	44
b	135		103	106	138	52	105	73	138	78	74	73	75	88
c	109		77	80	112	26	78	47	112	51	47	47	48	61
d	104		72	75	107	21	73	42	107	46	42	42	43	56
e	106		74	77	109	23	76	44	109	49	45	44	46	59
f	79		47	50	82	-4	48	17	82	21	17	17	18	31
g	102		70	73	105	19	71	40	105	44	40	40	41	54
h	136		104	107	139	53	105	74	139	78	74	74	75	88
i	136		104	107	139	53	106	74	139	79	75	74	76	89
j	23		-9	-6	26	-60	-7	-39	26	-34	-38	-39	-37	-25
k	136		104	107	139	53	106	74	139	79	75	74	76	89
l	134		102	105	137	51	103	72	137	76	72	72	73	86
m	136		104	107	139	53	105	74	139	78	74	74	75	88
n	136		104	107	139	53	105	74	139	78	74	74	75	88
o	103		71	74	106	20	72	41	106	45	41	41	42	55
p	136		104	107	139	53	105	74	139	78	74	74	75	88
q	105		73	76	108	22	74	43	108	47	43	43	44	57
r	135		103	106	138	52	104	73	138	77	73	73	74	87
s	100		68	71	103	17	70	38	103	43	39	38	40	53
t	87		55	58	90	4	57	25	90	30	26	25	27	39
u	134		102	105	137	51	103	72	137	76	72	72	73	86
v	82		50	53	85	-1	52	20	85	25	21	20	22	34
w	72		40	43	75	-11	42	10	75	15	11	10	12	25
x	77		45	48	80	-6	46	15	80	19	15	15	16	29
y	86		54	57	89	2	55	24	89	28	24	24	25	38
z	89		57	60	92	6	59	27	92	32	28	27	29	41

ตารางที่ ค. 2 ระยะห่างระหว่างอักษรแต่ละคู่อักษรที่ควรปรับให้ลดลงเพื่อให้ชิดติดกันพอดี (หน่วย Ps)

		ตัวอักษรแรก												
ตัวอักษรหลัง		a	b	c	d	e	f	g	h	i	j	k	l	m
	a	79	78	46	109	78	90	104	105	105	106	80	90	101
	b	109	107	75	139	108	31	133	134	134	135	70	137	131
	c	82	81	49	112	81	90	107	108	108	109	120	110	104
	d	77	76	44	107	76	90	102	103	103	104	105	105	99
	e	80	78	46	110	79	90	104	105	105	106	105	108	102
	f	90	95	55	82	51	0	77	90	78	79	13	80	80
	g	75	74	42	105	74	80	100	101	101	102	100	103	97
	h	109	108	76	139	108	32	134	135	135	136	70	137	131
	i	110	108	76	140	109	32	134	135	135	136	71	138	132
	j	100	100	70	115	100	20	60	115	115	30	70	110	110
	k	110	108	76	140	109	32	134	135	135	136	71	138	132
	l	107	106	74	137	106	30	132	133	133	134	68	135	129
	m	109	108	76	139	108	32	134	135	135	136	70	137	131
	n	109	108	76	139	108	32	134	135	135	136	70	137	131
	o	76	75	43	106	75	70	101	102	102	103	110	104	98
	p	109	108	76	139	108	32	134	135	135	136	70	137	131
	q	78	77	50	108	77	50	103	104	104	105	50	106	100
	r	108	107	75	138	107	31	133	134	134	135	69	136	130
	s	74	72	40	104	73	30	98	99	99	100	80	102	96
	t	61	59	27	91	60	0	85	86	86	87	22	89	83
	u	107	106	74	137	106	30	132	133	133	134	68	135	129
	v	56	54	22	86	55	0	80	81	81	82	17	84	78
	w	46	44	12	76	45	0	70	71	71	72	7	74	68
	x	50	80	40	80	80	10	75	76	76	77	11	78	72
	y	59	58	26	89	58	0	84	85	85	86	20	87	81
z	63	61	29	93	62	10	87	88	88	89	24	91	85	

ตารางที่ ค.2 ระยะห่างระหว่างอักษรแต่ละคู่อักษรที่ควรปรับให้ลดลงเพื่อให้ชิดติดกันพอดี (หน่วย Ps) (ต่อ)

		ตัวอักษรแรก												
		n	o	p	q	r	s	t	u	v	w	x	y	z
ตัวอักษรหลัง	a	106	74	77	109	60	75	65	109	80	70	65	60	70
	b	135	103	106	138	52	105	73	138	78	74	73	75	88
	c	109	77	80	112	70	78	85	112	85	75	120	80	105
	d	104	72	75	107	60	73	80	107	80	75	105	80	90
	e	106	74	77	109	60	76	80	109	80	70	105	75	80
	f	95	100	95	82	0	100	17	82	21	17	17	18	31
	g	102	70	73	105	60	71	70	105	60	55	80	60	80
	h	136	104	107	139	53	105	74	139	78	74	74	75	88
	i	136	104	107	139	53	106	74	139	79	75	74	76	89
	j	110	90	90	20	50	100	50	100	70	60	60	60	70
	k	136	104	107	139	53	106	74	139	79	75	74	76	89
	l	134	102	105	137	51	103	72	137	76	72	72	73	86
	m	136	104	107	139	53	105	74	139	78	74	74	75	88
	n	136	104	107	139	53	105	74	139	78	74	74	75	88
	o	103	71	74	106	60	72	70	106	60	60	90	60	80
	p	136	104	107	139	53	105	74	139	78	74	74	75	88
	q	105	73	76	108	60	74	50	108	50	50	50	44	57
	r	135	103	106	138	52	104	73	138	77	73	73	74	87
	s	100	68	71	103	50	70	38	103	43	39	60	40	53
	t	87	55	58	90	4	57	25	90	30	26	25	27	39
	u	134	102	105	137	51	103	72	137	76	72	72	73	86
	v	82	60	53	85	0	52	20	85	25	21	20	22	34
	w	72	40	43	75	0	42	10	75	15	11	10	12	25
	x	77	80	80	80	0	46	15	80	19	15	15	16	29
	y	86	54	57	89	2	55	24	89	28	24	24	25	38
	z	89	57	60	92	6	59	27	92	32	28	27	29	41

ภาคผนวก ง
การเปรียบเทียบจำแนกตามคู่อักษร

ตารางที่ ง. 1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
ab	2	2	100.0%
ac	1	1	100.0%
ad	5	5	100.0%
ae	1	0	0.0%
af	3	3	100.0%
ag	6	6	100.0%
ah	3	3	100.0%
ai	3	2	66.7%
aj	3	2	66.7%
ak	4	3	75.0%
al	1	1	100.0%
am	4	4	100.0%
an	5	5	100.0%
ao	5	4	80.0%
ap	3	3	100.0%
aq	3	3	100.0%
ar	2	1	50.0%
as	4	2	50.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
at	2	2	100.0%
au	1	1	100.0%
av	3	3	100.0%
aw	1	1	100.0%
ax	1	1	100.0%
ay	3	3	100.0%
az	2	2	100.0%
ba	1	1	100.0%
bc	3	1	33.3%
bd	3	3	100.0%
be	5	3	60.0%
bf	4	4	100.0%
bg	2	2	100.0%
bh	1	1	100.0%
bi	3	2	66.7%
bj	5	5	100.0%
bk	4	3	75.0%
bl	3	3	100.0%
bm	3	3	100.0%
bn	4	4	100.0%
bo	3	0	0.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
bp	1	1	100.0%
bq	3	3	100.0%
br	1	1	100.0%
bs	4	3	75.0%
bt	6	5	83.3%
bu	5	5	100.0%
bv	5	5	100.0%
bw	1	1	100.0%
bx	3	3	100.0%
by	4	4	100.0%
bz	1	1	100.0%
ca	5	2	40.0%
cb	5	3	60.0%
cd	2	0	0.0%
ce	3	0	0.0%
cf	4	2	50.0%
cg	2	1	50.0%
ch	2	0	0.0%
ci	3	2	66.7%
cj	3	1	33.3%
ck	5	3	60.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
cl	5	3	60.0%
cm	3	0	0.0%
cn	2	2	100.0%
co	4	3	75.0%
cp	2	0	0.0%
cr	4	2	50.0%
cs	5	4	80.0%
ct	4	1	25.0%
cu	1	1	100.0%
cv	3	2	66.7%
cw	3	2	66.7%
cx	6	3	50.0%
cy	2	0	0.0%
cz	4	2	50.0%
da	1	1	100.0%
db	2	0	0.0%
dc	4	1	25.0%
de	1	0	0.0%
df	2	2	100.0%
dg	3	2	66.7%
dh	3	3	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
di	3	3	100.0%
dj	3	3	100.0%
dk	6	6	100.0%
dl	1	0	0.0%
dm	1	1	100.0%
dh	2	1	50.0%
do	2	1	50.0%
dp	2	2	100.0%
dq	1	1	100.0%
dr	3	2	66.7%
ds	2	1	50.0%
dt	4	2	50.0%
du	4	4	100.0%
dv	3	3	100.0%
dw	7	7	100.0%
dx	3	3	100.0%
dy	1	1	100.0%
dz	2	2	100.0%
ea	2	2	100.0%
eb	2	2	100.0%
ec	5	2	40.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
ed	2	0	0.0%
ef	1	0	0.0%
eg	2	1	50.0%
eh	5	1	20.0%
ei	3	0	0.0%
ej	4	4	100.0%
ek	2	2	100.0%
el	2	1	50.0%
em	2	1	50.0%
en	4	3	75.0%
eo	4	1	25.0%
ep	2	1	50.0%
eq	3	1	33.3%
er	5	3	60.0%
es	2	1	50.0%
et	2	1	50.0%
eu	5	3	60.0%
ev	4	2	50.0%
ex	3	1	33.3%
ey	1	0	0.0%
ez	1	1	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
fa	4	4	100.0%
fb	3	2	66.7%
fc	5	3	60.0%
fd	2	2	100.0%
fe	5	5	100.0%
fg	6	4	66.7%
fh	3	3	100.0%
fi	1	1	100.0%
fj	2	1	50.0%
fk	5	5	100.0%
fl	2	2	100.0%
fm	6	6	100.0%
fn	2	2	100.0%
fo	1	0	0.0%
fp	1	1	100.0%
fq	2	1	50.0%
fr	3	2	66.7%
fs	1	0	0.0%
ft	3	2	66.7%
fu	2	2	100.0%
fv	1	1	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
fw	3	3	100.0%
fx	6	5	83.3%
fy	5	4	80.0%
fz	2	2	100.0%
ga	1	1	100.0%
gb	6	6	100.0%
gc	1	1	100.0%
gd	2	2	100.0%
ge	3	1	33.3%
gf	3	3	100.0%
gh	1	1	100.0%
gi	3	2	66.7%
gj	3	2	66.7%
gk	3	3	100.0%
gl	2	2	100.0%
gm	5	5	100.0%
gn	2	2	100.0%
go	1	0	0.0%
gp	2	2	100.0%
gr	1	1	100.0%
gs	6	2	33.3%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
gt	1	1	100.0%
gu	6	6	100.0%
gv	2	2	100.0%
gw	2	2	100.0%
gx	3	3	100.0%
gy	1	1	100.0%
gz	2	2	100.0%
ha	2	2	100.0%
hb	2	2	100.0%
hc	5	2	40.0%
hd	3	2	66.7%
he	2	2	100.0%
hf	1	1	100.0%
hg	2	2	100.0%
hi	6	6	100.0%
hj	2	2	100.0%
hk	2	2	100.0%
hl	3	2	66.7%
hn	2	2	100.0%
ho	1	1	100.0%
hp	2	2	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
hq	2	2	100.0%
hr	3	3	100.0%
hs	1	1	100.0%
ht	3	3	100.0%
hu	5	5	100.0%
hv	2	2	100.0%
hw	2	2	100.0%
hx	5	5	100.0%
hz	5	5	100.0%
ia	5	5	100.0%
ib	1	1	100.0%
ic	4	2	50.0%
id	6	6	100.0%
ie	2	1	50.0%
if	2	1	50.0%
ih	2	2	100.0%
ij	1	1	100.0%
ik	4	3	75.0%
im	3	3	100.0%
in	3	3	100.0%
io	2	1	50.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
ip	1	1	100.0%
iq	4	4	100.0%
ir	2	2	100.0%
is	2	0	0.0%
it	5	4	80.0%
iu	2	2	100.0%
iv	1	1	100.0%
iw	3	3	100.0%
ix	4	4	100.0%
iy	4	4	100.0%
iz	3	2	66.7%
ja	7	7	100.0%
jb	4	2	50.0%
jc	6	5	83.3%
jd	1	1	100.0%
je	2	1	50.0%
jf	5	3	60.0%
kg	5	4	80.0%
jh	2	1	50.0%
ji	4	3	75.0%
jk	5	5	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
jl	1	0	0.0%
jm	1	1	100.0%
jo	7	6	85.7%
jp	2	2	100.0%
jr	3	3	100.0%
js	1	0	0.0%
jt	3	2	66.7%
ju	3	3	100.0%
jv	5	5	100.0%
jw	1	1	100.0%
jx	1	0	0.0%
jy	1	1	100.0%
jz	1	1	100.0%
ka	2	2	100.0%
kb	2	2	100.0%
kc	1	0	0.0%
kd	2	2	100.0%
ke	4	2	50.0%
kf	4	4	100.0%
kg	8	8	100.0%
kh	2	2	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
ki	4	4	100.0%
kj	2	2	100.0%
kl	1	1	100.0%
km	2	2	100.0%
kn	6	6	100.0%
ko	7	7	100.0%
kp	1	1	100.0%
kq	1	1	100.0%
kr	3	2	66.7%
ks	3	1	33.3%
kt	3	3	100.0%
ku	5	5	100.0%
kv	3	3	100.0%
kw	1	1	100.0%
kx	3	3	100.0%
ky	6	6	100.0%
kz	2	2	100.0%
la	6	4	66.7%
lb	2	2	100.0%
lc	2	1	50.0%
le	3	0	0.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
lf	2	2	100.0%
lg	1	1	100.0%
lh	2	2	100.0%
li	2	1	50.0%
lj	2	2	100.0%
lk	8	2	25.0%
lm	4	3	75.0%
ln	6	5	83.3%
lo	2	2	100.0%
lp	4	3	75.0%
lr	4	4	100.0%
ls	4	1	25.0%
lt	1	0	0.0%
lu	1	1	100.0%
lv	2	1	50.0%
lx	2	2	100.0%
lz	2	2	100.0%
ma	2	2	100.0%
mb	3	3	100.0%
mc	6	4	66.7%
md	4	4	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
me	5	3	60.0%
mf	1	1	100.0%
mg	3	3	100.0%
mh	2	2	100.0%
mi	4	4	100.0%
mj	5	5	100.0%
mk	3	3	100.0%
ml	2	2	100.0%
mn	2	2	100.0%
mo	7	4	57.1%
mp	2	2	100.0%
mq	2	2	100.0%
mr	2	2	100.0%
ms	1	1	100.0%
mt	2	2	100.0%
mu	4	4	100.0%
mv	1	1	100.0%
mw	3	3	100.0%
mx	2	2	100.0%
my	2	2	100.0%
mz	2	2	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
na	2	2	100.0%
nb	3	3	100.0%
nd	2	2	100.0%
nf	1	1	100.0%
ng	4	4	100.0%
nh	3	3	100.0%
ni	2	2	100.0%
nj	2	2	100.0%
nk	2	2	100.0%
nl	4	4	100.0%
nm	2	2	100.0%
np	3	3	100.0%
nq	4	4	100.0%
ns	1	1	100.0%
nt	3	2	66.7%
nu	8	8	100.0%
nv	3	3	100.0%
nw	3	3	100.0%
nx	2	2	100.0%
ny	3	3	100.0%
nz	5	5	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
oa	2	2	100.0%
ob	3	3	100.0%
oc	5	2	40.0%
od	2	2	100.0%
oe	3	2	66.7%
of	4	3	75.0%
og	1	1	100.0%
oh	3	1	33.3%
oi	3	2	66.7%
oj	3	3	100.0%
ok	6	5	83.3%
ol	3	1	33.3%
om	1	0	0.0%
on	3	2	66.7%
op	2	2	100.0%
oq	2	2	100.0%
or	1	0	0.0%
os	4	3	75.0%
ot	4	3	75.0%
ov	2	1	50.0%
ow	6	5	83.3%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
ox	2	1	50.0%
oy	2	1	50.0%
oz	9	7	77.8%
pa	1	1	100.0%
pb	6	5	83.3%
pc	4	3	75.0%
pd	2	2	100.0%
pe	2	0	0.0%
pf	1	1	100.0%
pg	3	3	100.0%
ph	2	2	100.0%
pi	1	0	0.0%
pj	2	2	100.0%
pk	2	2	100.0%
pm	6	6	100.0%
pn	1	1	100.0%
po	2	2	100.0%
pr	3	3	100.0%
ps	8	6	75.0%
pt	3	3	100.0%
pu	4	4	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
pv	3	3	100.0%
pw	1	1	100.0%
px	2	2	100.0%
py	3	3	100.0%
pz	5	5	100.0%
qb	3	3	100.0%
qc	3	1	33.3%
qd	3	3	100.0%
qe	1	1	100.0%
qf	2	1	50.0%
qg	1	1	100.0%
qh	1	0	0.0%
qi	1	1	100.0%
qj	4	2	50.0%
ql	2	1	50.0%
qm	3	3	100.0%
qn	1	1	100.0%
qo	5	4	80.0%
qp	1	1	100.0%
qs	1	0	0.0%
qu	1	1	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
qv	1	1	100.0%
qw	5	5	100.0%
qx	1	0	0.0%
qy	4	4	100.0%
qz	4	4	100.0%
ra	2	2	100.0%
rb	7	6	85.7%
rc	7	2	28.6%
rd	1	1	100.0%
rf	1	1	100.0%
rg	4	4	100.0%
rh	2	2	100.0%
ri	1	0	0.0%
rj	2	2	100.0%
rk	2	2	100.0%
rl	1	1	100.0%
rm	2	1	50.0%
rn	1	1	100.0%
ro	4	4	100.0%
rp	1	1	100.0%
rq	1	0	0.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
rs	4	3	75.0%
rt	3	1	33.3%
ru	2	2	100.0%
rv	2	2	100.0%
rw	4	4	100.0%
rx	1	0	0.0%
ry	1	0	0.0%
rz	3	2	66.7%
sa	5	3	60.0%
sb	8	3	37.5%
sc	3	2	66.7%
sd	1	1	100.0%
se	2	0	0.0%
sf	2	2	100.0%
sg	1	0	0.0%
sh	2	1	50.0%
si	6	4	66.7%
sj	3	2	66.7%
sk	4	0	0.0%
sl	1	1	100.0%
sm	1	1	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
sn	2	2	100.0%
so	1	1	100.0%
sp	1	1	100.0%
sq	3	1	33.3%
sr	2	0	0.0%
st	6	4	66.7%
su	4	4	100.0%
sv	4	3	75.0%
sw	5	2	40.0%
sx	2	2	100.0%
sy	5	3	60.0%
sz	2	1	50.0%
ta	7	7	100.0%
tb	1	0	0.0%
tc	1	1	100.0%
te	3	2	66.7%
tf	10	7	70.0%
tg	3	3	100.0%
th	3	2	66.7%
ti	2	2	100.0%
tj	1	1	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
tk	2	0	0.0%
tl	5	3	60.0%
tm	1	1	100.0%
tn	1	1	100.0%
tp	3	1	33.3%
tq	2	2	100.0%
tr	4	3	75.0%
ts	3	0	0.0%
tu	9	8	88.9%
tv	2	2	100.0%
tw	5	5	100.0%
tx	3	3	100.0%
ty	4	3	75.0%
tz	3	3	100.0%
ua	3	3	100.0%
ub	3	3	100.0%
uc	3	0	0.0%
ud	4	4	100.0%
ue	8	3	37.5%
uf	6	6	100.0%
ug	3	3	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
uh	3	3	100.0%
ui	2	2	100.0%
uj	3	3	100.0%
uk	1	1	100.0%
ul	1	1	100.0%
un	3	3	100.0%
uo	4	3	75.0%
up	5	5	100.0%
uq	6	6	100.0%
ur	6	6	100.0%
ut	6	6	100.0%
uv	2	1	50.0%
uw	6	6	100.0%
ux	1	1	100.0%
uy	3	3	100.0%
uz	4	4	100.0%
vb	7	7	100.0%
vc	2	2	100.0%
vd	4	4	100.0%
ve	3	2	66.7%
vf	3	3	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
vh	3	3	100.0%
vi	3	3	100.0%
vj	6	6	100.0%
vk	6	5	83.3%
vm	2	2	100.0%
vn	2	2	100.0%
vo	2	1	50.0%
vp	4	4	100.0%
vq	2	2	100.0%
vr	3	3	100.0%
vs	1	1	100.0%
vt	4	3	75.0%
vu	1	1	100.0%
vw	2	2	100.0%
vx	3	3	100.0%
vy	3	3	100.0%
vz	2	2	100.0%
wa	3	3	100.0%
wb	3	3	100.0%
wc	1	1	100.0%
wd	1	1	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
we	7	4	57.1%
wf	3	3	100.0%
wg	3	3	100.0%
wi	1	1	100.0%
wj	2	2	100.0%
wk	3	3	100.0%
wl	3	2	66.7%
wm	3	3	100.0%
wn	4	4	100.0%
wo	2	1	50.0%
wp	6	5	83.3%
wq	1	1	100.0%
wr	2	2	100.0%
ws	6	4	66.7%
wt	2	1	50.0%
wv	2	2	100.0%
wx	4	4	100.0%
wz	3	3	100.0%
xa	3	3	100.0%
xb	5	4	80.0%
xc	3	0	0.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
xd	2	2	100.0%
xe	4	2	50.0%
xf	2	2	100.0%
xg	1	1	100.0%
xh	2	2	100.0%
xi	3	3	100.0%
xk	1	1	100.0%
xl	2	1	50.0%
xm	2	2	100.0%
xn	2	2	100.0%
xo	3	3	100.0%
xp	2	2	100.0%
xq	2	1	50.0%
xr	3	3	100.0%
xs	4	4	100.0%
xu	3	3	100.0%
xw	5	5	100.0%
xy	1	1	100.0%
xz	2	2	100.0%
ya	1	1	100.0%
yb	1	1	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
yc	4	3	75.0%
yd	3	3	100.0%
ye	1	0	0.0%
yf	2	1	50.0%
yg	1	1	100.0%
yh	5	5	100.0%
yi	3	3	100.0%
yj	4	3	75.0%
yk	1	0	0.0%
yl	2	2	100.0%
ym	6	6	100.0%
yn	4	4	100.0%
yo	3	2	66.7%
yp	2	2	100.0%
yq	1	1	100.0%
yr	4	4	100.0%
ys	4	2	50.0%
yt	2	2	100.0%
yu	5	5	100.0%
yv	2	2	100.0%
yw	1	1	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่
อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
yx	2	2	100.0%
yz	3	3	100.0%
za	5	4	80.0%
zb	5	5	100.0%
zc	2	1	50.0%
zd	2	2	100.0%
ze	2	2	100.0%
zf	1	1	100.0%
zg	6	6	100.0%
zh	2	2	100.0%
zi	2	2	100.0%
zj	3	3	100.0%
zk	3	3	100.0%
zl	3	3	100.0%
zm	2	2	100.0%
zn	4	4	100.0%
zp	5	5	100.0%
zq	3	1	33.3%
zr	5	4	80.0%
zs	4	3	75.0%
zt	2	2	100.0%

ตารางที่ ง.1 อัตราการยืนยันความเป็นมนุษย์ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ตอบถูก	อัตราการยืนยันความเป็นมนุษย์ (%)
zu	2	2	100.0%
zv	4	4	100.0%
zw	2	2	100.0%
zx	4	4	100.0%
zy	3	3	100.0%



ตารางที่ ง. 2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
ab	2	2	100.0%
ac	1	1	100.0%
ad	5	5	100.0%
ae	1	1	100.0%
af	3	3	100.0%
ag	6	4	66.7%
ah	3	3	100.0%
ai	3	3	100.0%
aj	3	2	66.7%
ak	4	3	75.0%
al	1	1	100.0%
am	4	4	100.0%
an	5	2	40.0%
ao	5	3	60.0%
ap	3	3	100.0%
aq	3	3	100.0%
ar	2	2	100.0%
as	4	4	100.0%
at	2	1	50.0%
au	1	1	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
av	3	2	66.7%
aw	1	1	100.0%
ax	1	0	0.0%
ay	3	3	100.0%
az	2	1	50.0%
ba	1	1	100.0%
bc	3	3	100.0%
bd	3	2	66.7%
be	5	5	100.0%
bf	4	3	75.0%
bg	2	2	100.0%
bh	1	0	0.0%
bi	3	1	33.3%
bj	5	5	100.0%
bk	4	3	75.0%
bl	3	1	33.3%
bm	3	2	66.7%
bn	4	3	75.0%
bo	3	3	100.0%
bp	1	1	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
bq	3	3	100.0%
br	1	1	100.0%
bs	4	4	100.0%
bt	6	6	100.0%
bu	5	5	100.0%
bv	5	3	60.0%
bw	1	1	100.0%
bx	3	2	66.7%
by	4	4	100.0%
bz	1	0	0.0%
ca	5	5	100.0%
cb	5	4	80.0%
cd	2	2	100.0%
ce	3	3	100.0%
cf	4	2	50.0%
cg	2	2	100.0%
ch	2	2	100.0%
ci	3	3	100.0%
cj	3	2	66.7%
ck	5	5	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
cl	5	3	60.0%
cm	3	2	66.7%
cn	2	2	100.0%
co	4	4	100.0%
cp	2	2	100.0%
cr	4	3	75.0%
cs	5	4	80.0%
ct	4	3	75.0%
cu	1	0	0.0%
cv	3	1	33.3%
cw	3	0	0.0%
cx	6	5	83.3%
cy	2	1	50.0%
cz	4	3	75.0%
da	1	0	0.0%
db	2	2	100.0%
dc	4	4	100.0%
de	1	1	100.0%
df	2	1	50.0%
dg	3	3	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
dh	3	1	33.3%
di	3	2	66.7%
dj	3	1	33.3%
dk	6	5	83.3%
dl	1	1	100.0%
dm	1	1	100.0%
dn	2	2	100.0%
do	2	2	100.0%
dp	2	1	50.0%
dq	1	1	100.0%
dr	3	2	66.7%
ds	2	1	50.0%
dt	4	2	50.0%
du	4	4	100.0%
dv	3	2	66.7%
dw	7	4	57.1%
dx	3	3	100.0%
dy	1	0	0.0%
dz	2	2	100.0%
ea	2	2	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
eb	2	2	100.0%
ec	5	5	100.0%
ed	2	2	100.0%
ef	1	1	100.0%
eg	2	1	50.0%
eh	5	5	100.0%
ei	3	3	100.0%
ej	4	2	50.0%
ek	2	1	50.0%
el	2	2	100.0%
em	2	1	50.0%
en	4	4	100.0%
eo	4	3	75.0%
ep	2	1	50.0%
eq	3	3	100.0%
er	5	3	60.0%
es	2	2	100.0%
et	2	2	100.0%
eu	5	3	60.0%
ev	4	2	50.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
ex	3	3	100.0%
ey	1	1	100.0%
ez	1	1	100.0%
fa	4	4	100.0%
fb	3	1	33.3%
fc	5	5	100.0%
fd	2	1	50.0%
fe	5	3	60.0%
fg	6	6	100.0%
fh	3	2	66.7%
fi	1	1	100.0%
fj	2	2	100.0%
fk	5	4	80.0%
fl	2	2	100.0%
fm	6	6	100.0%
fn	2	2	100.0%
fo	1	1	100.0%
fp	1	1	100.0%
fq	2	2	100.0%
fr	3	2	66.7%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
fs	1	1	100.0%
ft	3	2	66.7%
fu	2	1	50.0%
fv	1	1	100.0%
fw	3	3	100.0%
fx	6	6	100.0%
fy	5	4	80.0%
fz	2	2	100.0%
ga	1	0	0.0%
gb	6	5	83.3%
gc	1	0	0.0%
gd	2	2	100.0%
ge	3	3	100.0%
gf	3	3	100.0%
gh	1	1	100.0%
gi	3	3	100.0%
gj	3	3	100.0%
gk	3	2	66.7%
gl	2	1	50.0%
gm	5	3	60.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
gn	2	1	50.0%
go	1	1	100.0%
gp	2	2	100.0%
gr	1	1	100.0%
gs	6	6	100.0%
gt	1	1	100.0%
gu	6	5	83.3%
gv	2	2	100.0%
gw	2	2	100.0%
gx	3	3	100.0%
gy	1	1	100.0%
gz	2	2	100.0%
ha	2	2	100.0%
hb	2	1	50.0%
hc	5	3	60.0%
hd	3	2	66.7%
he	2	1	50.0%
hf	1	1	100.0%
hg	2	1	50.0%
hi	6	6	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
hj	2	1	50.0%
hk	2	2	100.0%
hl	3	3	100.0%
hn	2	1	50.0%
ho	1	1	100.0%
hp	2	1	50.0%
hq	2	2	100.0%
hr	3	3	100.0%
hs	1	1	100.0%
ht	3	2	66.7%
hu	5	4	80.0%
hv	2	2	100.0%
hw	2	2	100.0%
hx	5	4	80.0%
hz	5	4	80.0%
ia	5	4	80.0%
ib	1	1	100.0%
ic	4	4	100.0%
id	6	4	66.7%
ie	2	1	50.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
if	2	2	100.0%
ih	2	2	100.0%
ij	1	0	0.0%
ik	4	4	100.0%
im	3	2	66.7%
in	3	3	100.0%
io	2	2	100.0%
ip	1	1	100.0%
iq	4	4	100.0%
ir	2	1	50.0%
is	2	2	100.0%
it	5	5	100.0%
iu	2	2	100.0%
iv	1	1	100.0%
iw	3	3	100.0%
ix	4	4	100.0%
iy	4	4	100.0%
iz	3	3	100.0%
ja	7	6	85.7%
jb	4	3	75.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
jc	6	3	50.0%
jd	1	1	100.0%
je	2	2	100.0%
jf	5	4	80.0%
yg	5	3	60.0%
jh	2	2	100.0%
ji	4	3	75.0%
jk	5	5	100.0%
jl	1	1	100.0%
jm	1	1	100.0%
jo	7	5	71.4%
jp	2	1	50.0%
jr	3	3	100.0%
js	1	1	100.0%
jt	3	3	100.0%
ju	3	3	100.0%
jv	5	3	60.0%
jw	1	1	100.0%
jx	1	1	100.0%
jy	1	1	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
jz	1	1	100.0%
ka	2	1	50.0%
kb	2	1	50.0%
kc	1	1	100.0%
kd	2	1	50.0%
ke	4	4	100.0%
kf	4	4	100.0%
kg	8	6	75.0%
kh	2	2	100.0%
ki	4	4	100.0%
kj	2	2	100.0%
kl	1	1	100.0%
km	2	2	100.0%
kn	6	5	83.3%
ko	7	7	100.0%
kp	1	1	100.0%
kq	1	1	100.0%
kr	3	3	100.0%
ks	3	3	100.0%
kt	3	3	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
ku	5	5	100.0%
kv	3	3	100.0%
kw	1	1	100.0%
kx	3	3	100.0%
ky	6	4	66.7%
kz	2	1	50.0%
la	6	5	83.3%
lb	2	1	50.0%
lc	2	2	100.0%
le	3	3	100.0%
lf	2	2	100.0%
lg	1	1	100.0%
lh	2	1	50.0%
li	2	2	100.0%
lj	2	1	50.0%
lk	8	8	100.0%
lm	4	4	100.0%
ln	6	4	66.7%
lo	2	1	50.0%
lp	4	3	75.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
lr	4	3	75.0%
ls	4	4	100.0%
lt	1	1	100.0%
lu	1	1	100.0%
lv	2	2	100.0%
lx	2	1	50.0%
lz	2	1	50.0%
ma	2	2	100.0%
mb	3	3	100.0%
mc	6	4	66.7%
md	4	3	75.0%
me	5	5	100.0%
mf	1	1	100.0%
mg	3	3	100.0%
mh	2	2	100.0%
mi	4	3	75.0%
mj	5	4	80.0%
mk	3	3	100.0%
ml	2	1	50.0%
mn	2	2	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุ่มอักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุ่ม เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุ่ม	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
mo	7	6	85.7%
mp	2	2	100.0%
mq	2	1	50.0%
mr	2	2	100.0%
ms	1	1	100.0%
mt	2	2	100.0%
mu	4	3	75.0%
mv	1	1	100.0%
mw	3	2	66.7%
mx	2	2	100.0%
my	2	2	100.0%
mz	2	1	50.0%
na	2	2	100.0%
nb	3	3	100.0%
nd	2	0	0.0%
nf	1	1	100.0%
ng	4	2	50.0%
nh	3	2	66.7%
ni	2	2	100.0%
nj	2	1	50.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
nk	2	2	100.0%
nl	4	4	100.0%
nm	2	2	100.0%
np	3	2	66.7%
nq	4	3	75.0%
ns	1	1	100.0%
nt	3	2	66.7%
nu	8	7	87.5%
nv	3	2	66.7%
nw	3	3	100.0%
nx	2	1	50.0%
ny	3	3	100.0%
nz	5	5	100.0%
oa	2	1	50.0%
ob	3	2	66.7%
oc	5	3	60.0%
od	2	2	100.0%
oe	3	2	66.7%
of	4	4	100.0%
og	1	1	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
oh	3	2	66.7%
oi	3	2	66.7%
oj	3	3	100.0%
ok	6	5	83.3%
ol	3	3	100.0%
om	1	1	100.0%
on	3	3	100.0%
op	2	1	50.0%
oq	2	2	100.0%
or	1	1	100.0%
os	4	3	75.0%
ot	4	2	50.0%
ov	2	2	100.0%
ow	6	5	83.3%
ox	2	1	50.0%
oy	2	2	100.0%
oz	9	6	66.7%
pa	1	1	100.0%
pb	6	5	83.3%
pc	4	2	50.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
pd	2	1	50.0%
pe	2	2	100.0%
pf	1	1	100.0%
pg	3	2	66.7%
ph	2	2	100.0%
pi	1	1	100.0%
pj	2	1	50.0%
pk	2	1	50.0%
pm	6	5	83.3%
pn	1	1	100.0%
po	2	1	50.0%
pr	3	2	66.7%
ps	8	5	62.5%
pt	3	1	33.3%
pu	4	2	50.0%
pv	3	3	100.0%
pw	1	0	0.0%
px	2	2	100.0%
py	3	3	100.0%
pz	5	4	80.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
qb	3	2	66.7%
qc	3	2	66.7%
qd	3	2	66.7%
qe	1	1	100.0%
qf	2	1	50.0%
qg	1	1	100.0%
qh	1	1	100.0%
qi	1	1	100.0%
qj	4	4	100.0%
ql	2	2	100.0%
qm	3	3	100.0%
qn	1	0	0.0%
qo	5	5	100.0%
qp	1	1	100.0%
qs	1	1	100.0%
qu	1	1	100.0%
qv	1	1	100.0%
qw	5	4	80.0%
qx	1	1	100.0%
qy	4	3	75.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
qz	4	4	100.0%
ra	2	1	50.0%
rb	7	5	71.4%
rc	7	6	85.7%
rd	1	1	100.0%
rf	1	1	100.0%
rg	4	4	100.0%
rh	2	1	50.0%
ri	1	1	100.0%
rj	2	1	50.0%
rk	2	0	0.0%
rl	1	1	100.0%
rm	2	2	100.0%
rn	1	1	100.0%
ro	4	1	25.0%
rp	1	1	100.0%
rq	1	1	100.0%
rs	4	4	100.0%
rt	3	3	100.0%
ru	2	0	0.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
rv	2	2	100.0%
rw	4	3	75.0%
rx	1	1	100.0%
ry	1	1	100.0%
rz	3	1	33.3%
sa	5	5	100.0%
sb	8	7	87.5%
sc	3	3	100.0%
sd	1	0	0.0%
se	2	2	100.0%
sf	2	1	50.0%
sg	1	1	100.0%
sh	2	2	100.0%
si	6	6	100.0%
sj	3	3	100.0%
sk	4	4	100.0%
sl	1	1	100.0%
sm	1	1	100.0%
sn	2	2	100.0%
so	1	1	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
sp	1	0	0.0%
sq	3	3	100.0%
sr	2	2	100.0%
st	6	4	66.7%
su	4	3	75.0%
sv	4	3	75.0%
sw	5	5	100.0%
sx	2	2	100.0%
sy	5	5	100.0%
sz	2	2	100.0%
ta	7	6	85.7%
tb	1	1	100.0%
tc	1	1	100.0%
te	3	3	100.0%
tf	10	8	80.0%
tg	3	1	33.3%
th	3	2	66.7%
ti	2	2	100.0%
tj	1	1	100.0%
tk	2	2	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
tl	5	4	80.0%
tm	1	1	100.0%
tn	1	1	100.0%
tp	3	3	100.0%
tq	2	2	100.0%
tr	4	2	50.0%
ts	3	3	100.0%
tu	9	8	88.9%
tv	2	1	50.0%
tw	5	5	100.0%
tx	3	2	66.7%
ty	4	4	100.0%
tz	3	2	66.7%
ua	3	1	33.3%
ub	3	3	100.0%
uc	3	3	100.0%
ud	4	2	50.0%
ue	8	8	100.0%
uf	6	4	66.7%
ug	3	3	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
uh	3	3	100.0%
ui	2	2	100.0%
uj	3	3	100.0%
uk	1	1	100.0%
ul	1	1	100.0%
un	3	2	66.7%
uo	4	3	75.0%
up	5	2	40.0%
uq	6	3	50.0%
ur	6	4	66.7%
ut	6	5	83.3%
uv	2	2	100.0%
uw	6	6	100.0%
ux	1	1	100.0%
uy	3	2	66.7%
uz	4	3	75.0%
vb	7	5	71.4%
vc	2	1	50.0%
vd	4	2	50.0%
ve	3	2	66.7%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
vf	3	3	100.0%
vh	3	3	100.0%
vi	3	3	100.0%
vj	6	5	83.3%
vk	6	4	66.7%
vm	2	1	50.0%
vn	2	2	100.0%
vo	2	1	50.0%
vp	4	4	100.0%
vq	2	2	100.0%
vr	3	2	66.7%
vs	1	0	0.0%
vt	4	4	100.0%
vu	1	1	100.0%
vw	2	1	50.0%
vx	3	3	100.0%
vy	3	3	100.0%
vz	2	1	50.0%
wa	3	2	66.7%
wb	3	3	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
wc	1	1	100.0%
wd	1	1	100.0%
we	7	5	71.4%
wf	3	2	66.7%
wg	3	3	100.0%
wi	1	1	100.0%
wj	2	2	100.0%
wk	3	3	100.0%
wl	3	3	100.0%
wm	3	2	66.7%
wn	4	2	50.0%
wo	2	2	100.0%
wp	6	4	66.7%
wq	1	1	100.0%
wr	2	1	50.0%
ws	6	6	100.0%
wt	2	2	100.0%
wv	2	2	100.0%
wx	4	4	100.0%
wz	3	3	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
xa	3	1	33.3%
xb	5	5	100.0%
xc	3	2	66.7%
xd	2	1	50.0%
xe	4	3	75.0%
xf	2	2	100.0%
xg	1	1	100.0%
xh	2	2	100.0%
xi	3	3	100.0%
xk	1	1	100.0%
xl	2	2	100.0%
xm	2	2	100.0%
xn	2	1	50.0%
xo	3	3	100.0%
xp	2	2	100.0%
xq	2	2	100.0%
xr	3	3	100.0%
xs	4	4	100.0%
xu	3	2	66.7%
xw	5	5	100.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
xy	1	1	100.0%
xz	2	0	0.0%
ya	1	0	0.0%
yb	1	1	100.0%
yc	4	3	75.0%
yd	3	3	100.0%
ye	1	1	100.0%
yf	2	1	50.0%
yg	1	1	100.0%
yh	5	4	80.0%
yi	3	2	66.7%
yj	4	3	75.0%
yk	1	1	100.0%
yl	2	1	50.0%
ym	6	6	100.0%
yn	4	4	100.0%
yo	3	3	100.0%
yp	2	1	50.0%
yq	1	1	100.0%
yr	4	3	75.0%

ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
ys	4	4	100.0%
yt	2	2	100.0%
yu	5	4	80.0%
yv	2	2	100.0%
yw	1	0	0.0%
yx	2	1	50.0%
yz	3	3	100.0%
za	5	5	100.0%
zb	5	4	80.0%
zc	2	1	50.0%
zd	2	2	100.0%
ze	2	0	0.0%
zf	1	1	100.0%
zg	6	4	66.7%
zh	2	2	100.0%
zi	2	2	100.0%
zj	3	2	66.7%
zk	3	3	100.0%
zl	3	3	100.0%
zm	2	0	0.0%

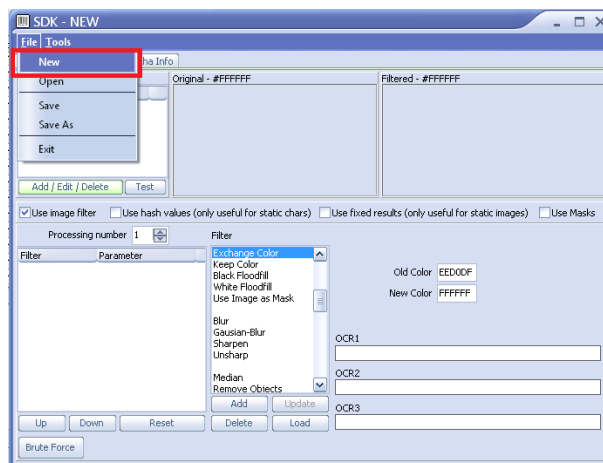
ตารางที่ ง.2 อัตราความทนทานของแคปท์ซาข้อความ จำแนกตามคู่อักษร โดยไม่มีการสุม่อักษรซ้ำ และบางคู่อักษรที่ไม่ถูกสุม่อ เช่น cq และ ew จะไม่ได้แสดงในตาราง (ต่อ)

คู่อักษร	จำนวนครั้งที่ถูกสุม่อ	จำนวนครั้งที่ GSA ตอบผิด	อัตราการความทนทานของคู่อักษร (%)
zn	4	4	100.0%
zp	5	2	40.0%
zq	3	2	66.7%
zr	5	3	60.0%
zs	4	4	100.0%
zt	2	1	50.0%
zu	2	2	100.0%
zv	4	3	75.0%
zw	2	2	100.0%
zx	4	3	75.0%
zy	3	2	66.7%

ภาคผนวก จ

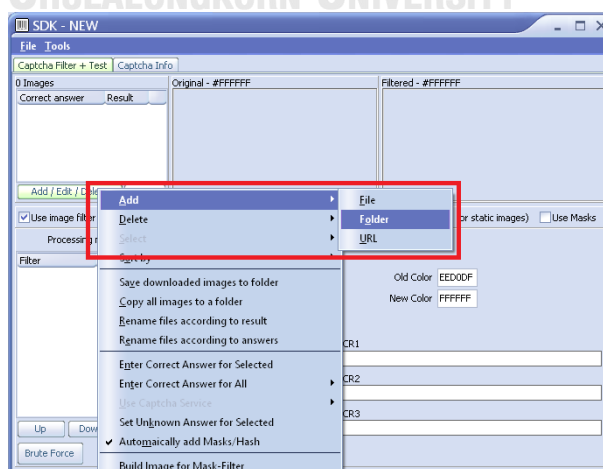
การกำหนด Filter ของโปรแกรม GSA Captcha Breaker

1. กดปุ่ม New บริเวณกรอบสีแดงในภาพที่ จ.1 เพื่อสร้าง SDK และกำหนด Filter สำหรับการวิเคราะห์แคปต์ชาข้อความในแต่ละรูปแบบ



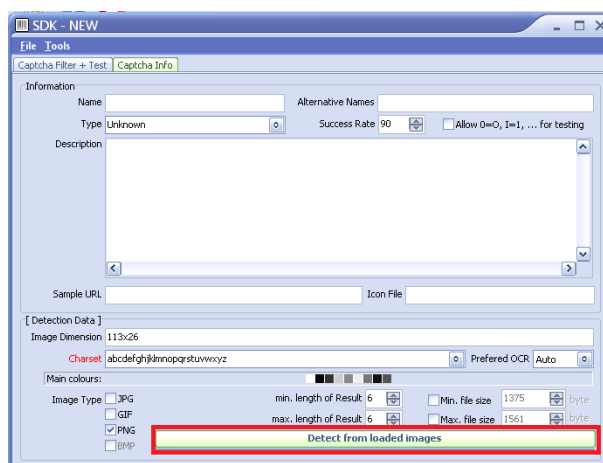
ภาพที่ จ. 1 สร้าง SDK ใหม่สำหรับแคปต์ชาข้อความแต่ละลักษณะ

2. Tab Captcha Filter + Test กดปุ่ม Add/Edit/Delete เพื่อปรับรูปแคปต์ชาข้อความที่ต้องการวิเคราะห์ลงในโปรแกรม GSA Captcha Breaker โดยเลือก Add และเลือก Folder บริเวณกรอบสีแดงในภาพที่ จ.2



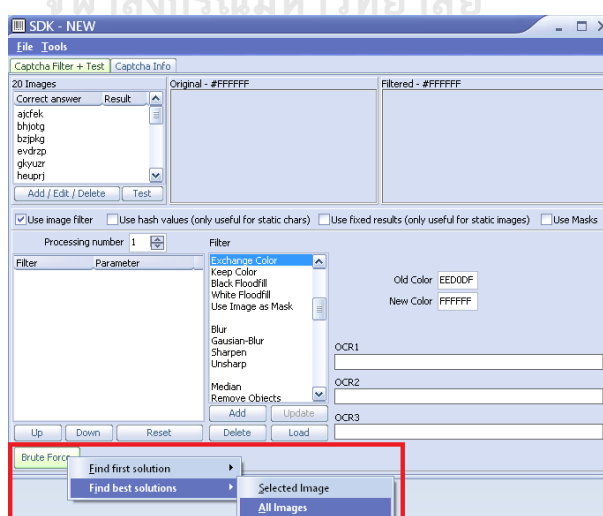
ภาพที่ จ. 2 เลือกโฟลเดอร์รูปแคปต์ชาข้อความเพื่อปรับรูปแคปต์ชาข้อความลงในโปรแกรม

3. Tab Captcha Info กดปุ่ม Detect from loaded images บริเวณกรอบสีแดงในภาพที่ จ.3 เพื่อให้โปรแกรมตรวจสอบตัวอักษรที่เป็นไปได้สำหรับคำตอบของแคปต์ชาข้อความ (Charset) โปรแกรมจะตรวจสอบจากชื่อไฟล์ของรูปแคปต์ชาข้อความ (ต้องตั้งชื่อรูปแคปต์ชาข้อความแต่ละรูปเป็นคำตอบที่ถูกต้องสำหรับแคปต์ชาข้อความนั้น)



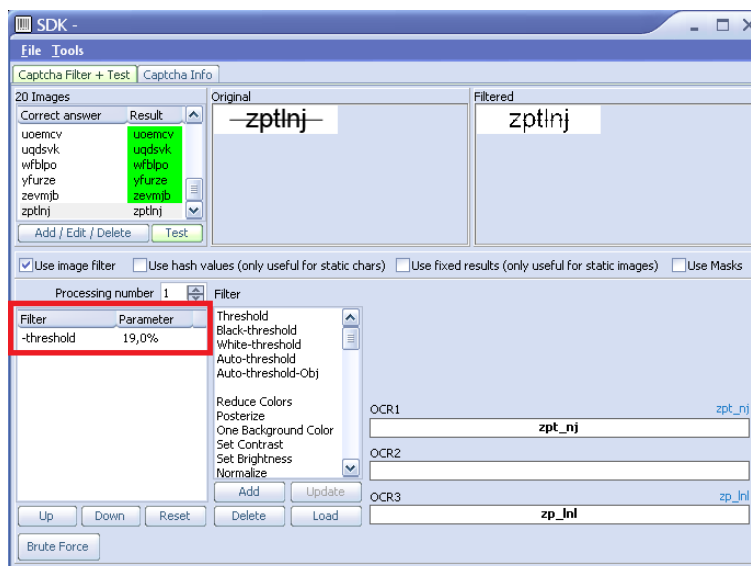
ภาพที่ จ. 3 ตรวจสอบตัวอักษรที่เป็นไปได้สำหรับคำตอบของแคปต์ชาข้อความ

4. Tab Captcha Filter + Test กดปุ่ม Brute Force เลือก Find best solutions สำหรับทุกภาพ (All Images) บริเวณกรอบสีแดงในภาพที่ จ.4 เพื่อให้โปรแกรมเลือก Filter ที่เหมาะสมที่สุดสำหรับแคปต์ชาข้อความรูปแบบนั้น

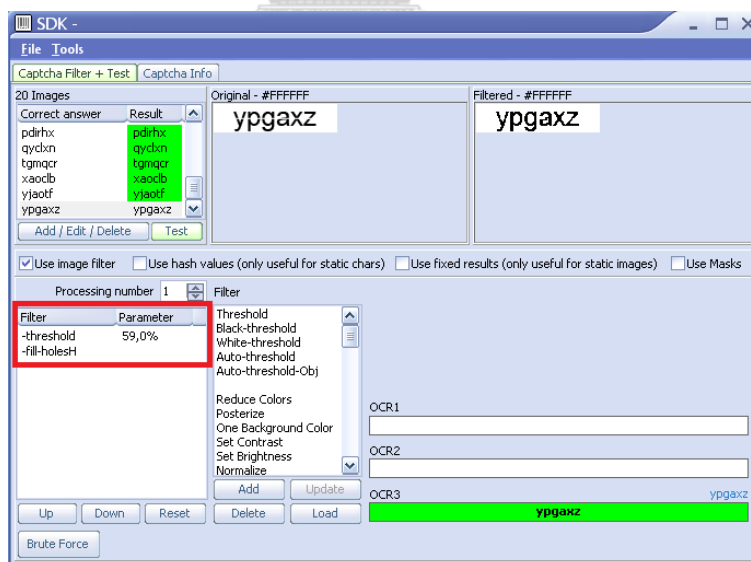


ภาพที่ จ. 4 วิเคราะห์ Filter ที่เหมาะสมที่สุดสำหรับรูปแบบของแคปต์ชาข้อความนั้น

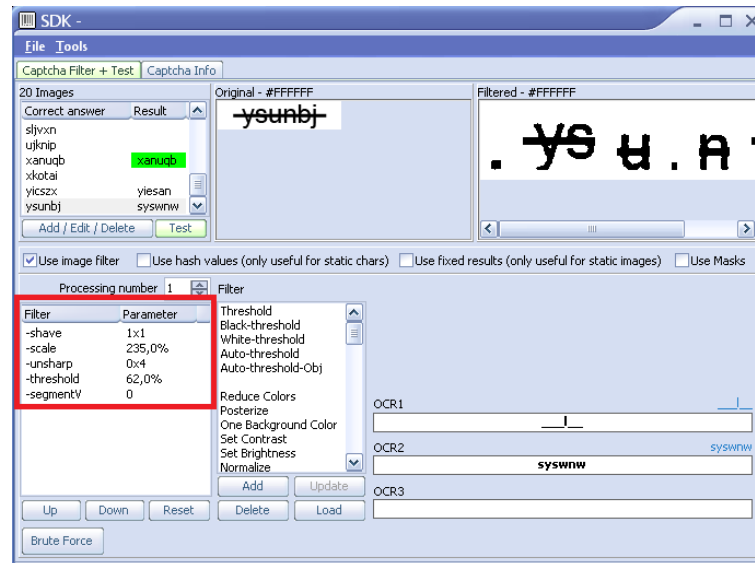
5. เนื่องจากรูปแบบของแคปต์ชาข้อความมีลักษณะแตกต่างกันไปตามตัวแปรที่กำหนด สำหรับงานวิจัยนี้มีรูปแบบของแคปต์ชาข้อความทั้งหมด 18 แบบ Filter ที่เหมาะสมสำหรับแต่ละรูปแบบของแคปต์ชาข้อความจึงแตกต่างกันไป ดังแสดงบริเวณกรอบสีแดงในภาพที่ จ.5 ถึง จ.39



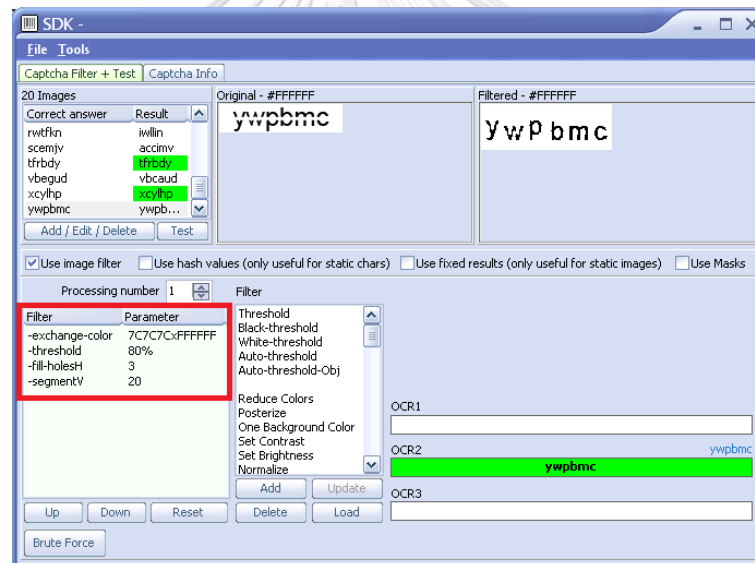
ภาพที่ จ. 5 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 1



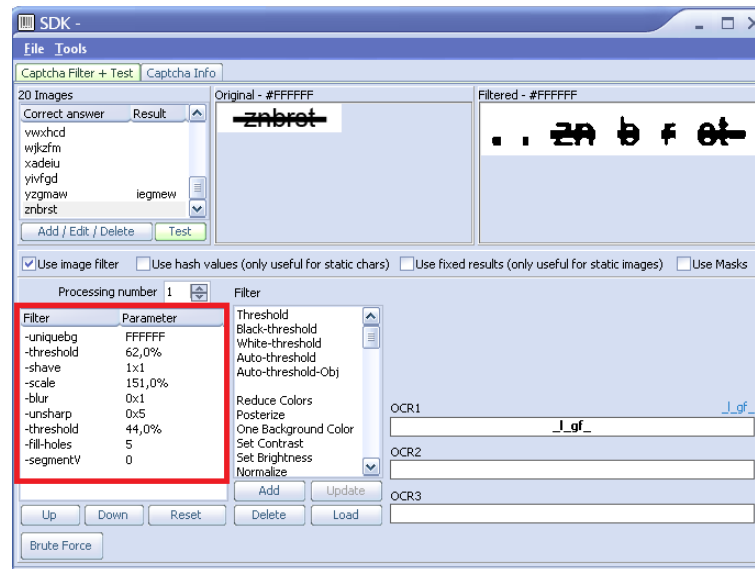
ภาพที่ จ. 6 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 2



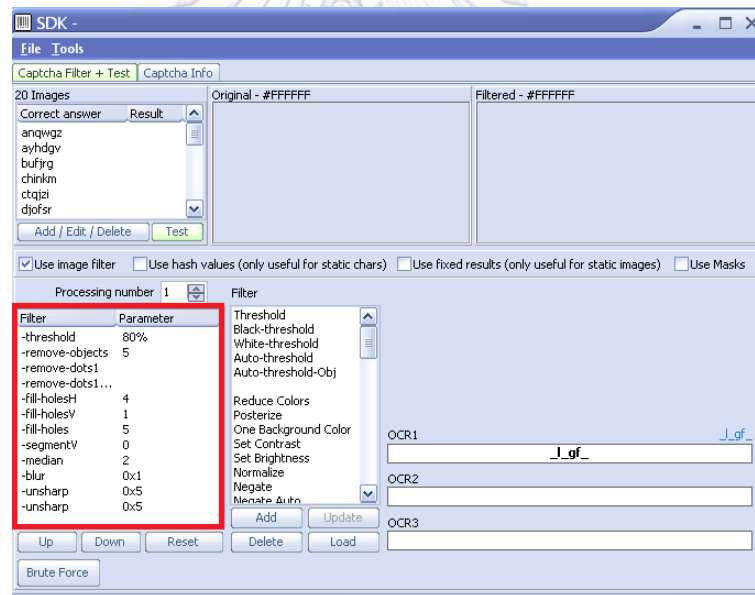
ภาพที่ จ. 7 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 3



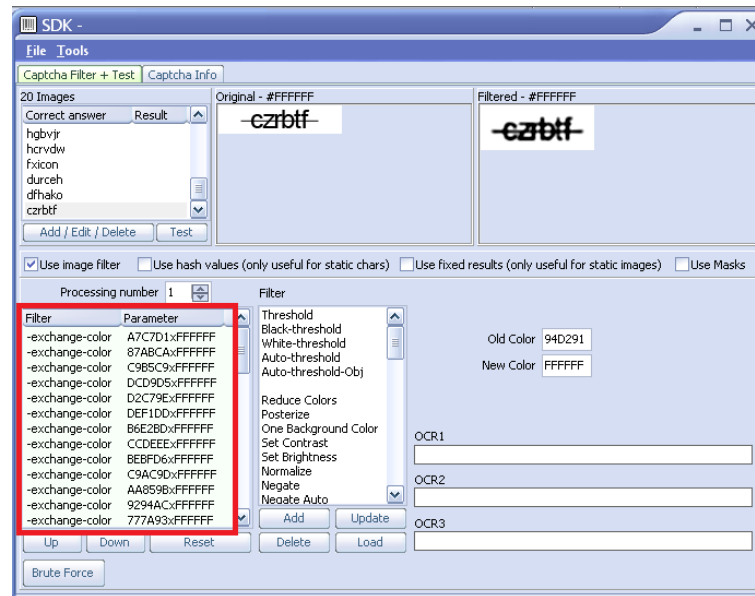
ภาพที่ จ. 8 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 4



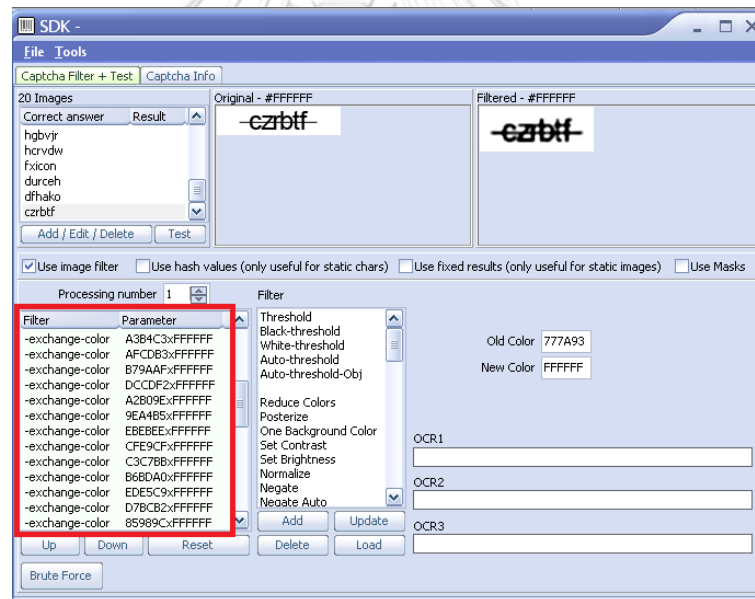
ภาพที่ จ. 9 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 5



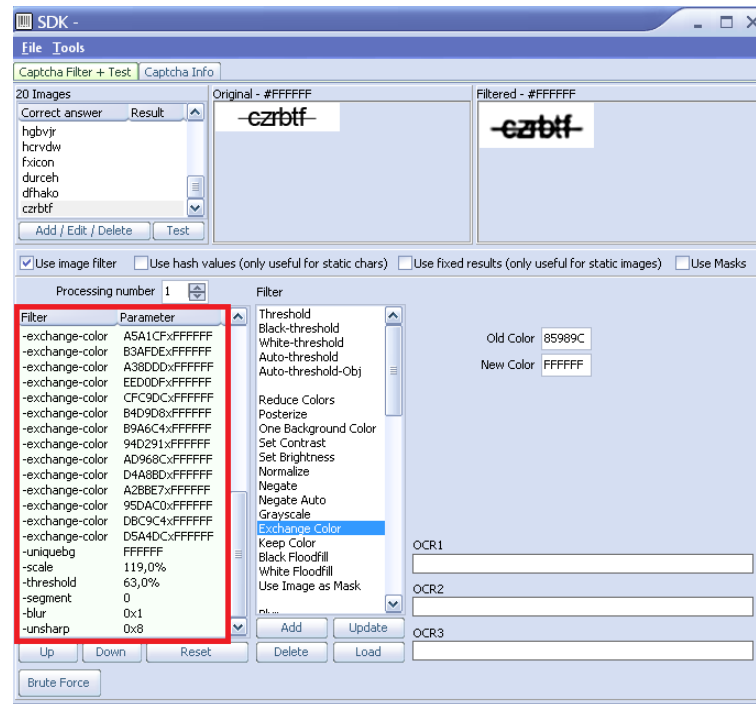
ภาพที่ จ. 10 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 6



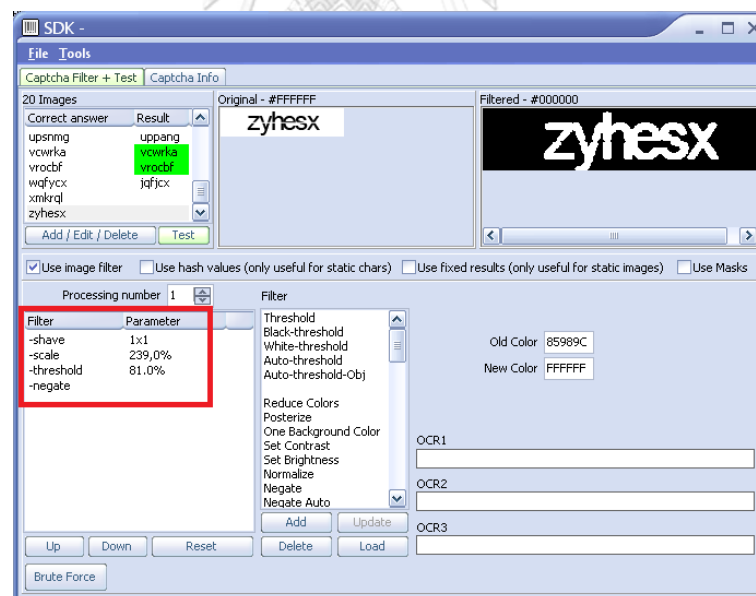
ภาพที่ จ. 11 Filter ที่เหมาะสมสำหรับแคปช่าข้อความรูปแบบที่ 7 (1)



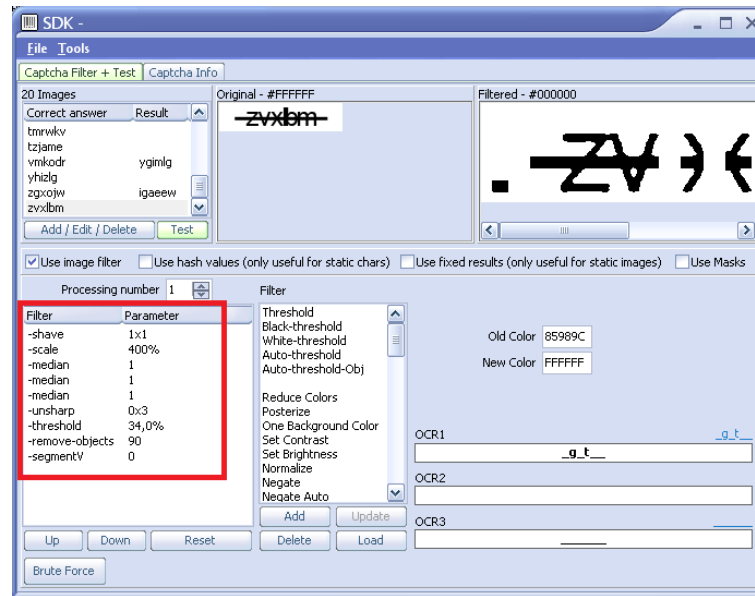
ภาพที่ จ. 12 Filter ที่เหมาะสมสำหรับแคปช่าข้อความรูปแบบที่ 7 (2)



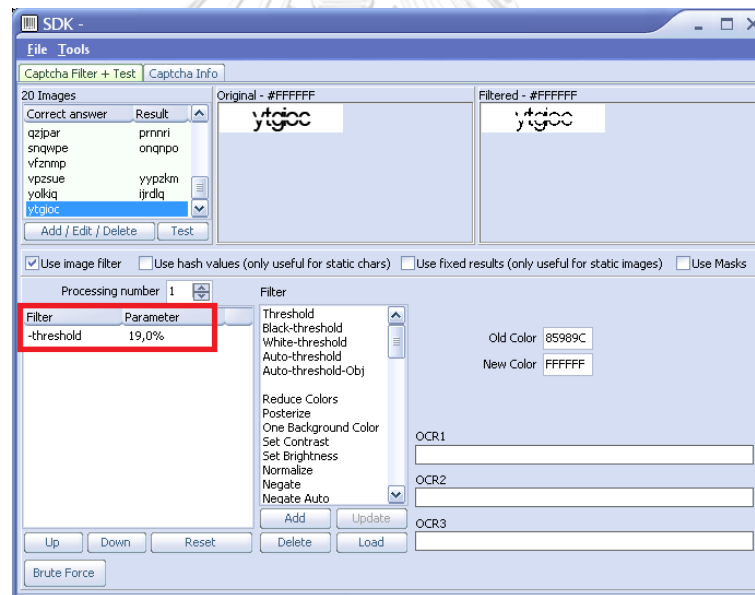
ภาพที่ จ. 13 Filter ที่เหมาะสมสำหรับแคปช่าข้อความรูปแบบที่ 7 (3)



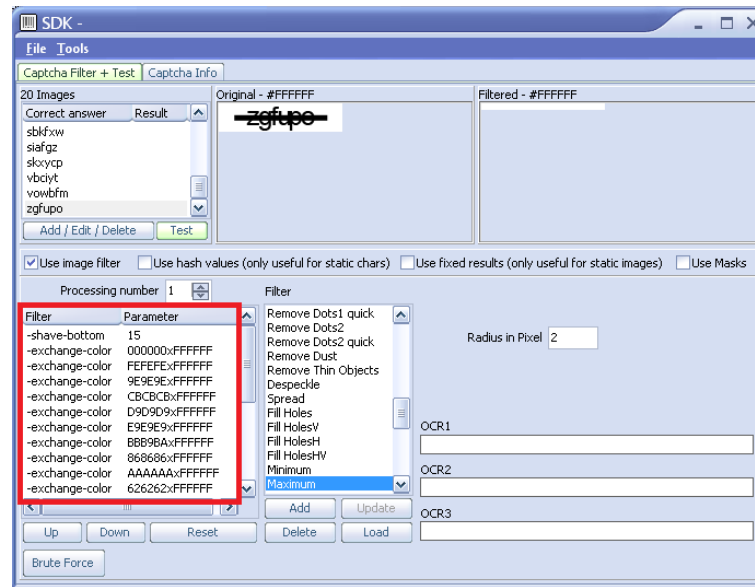
ภาพที่ จ. 14 Filter ที่เหมาะสมสำหรับแคปช่าข้อความรูปแบบที่ 8



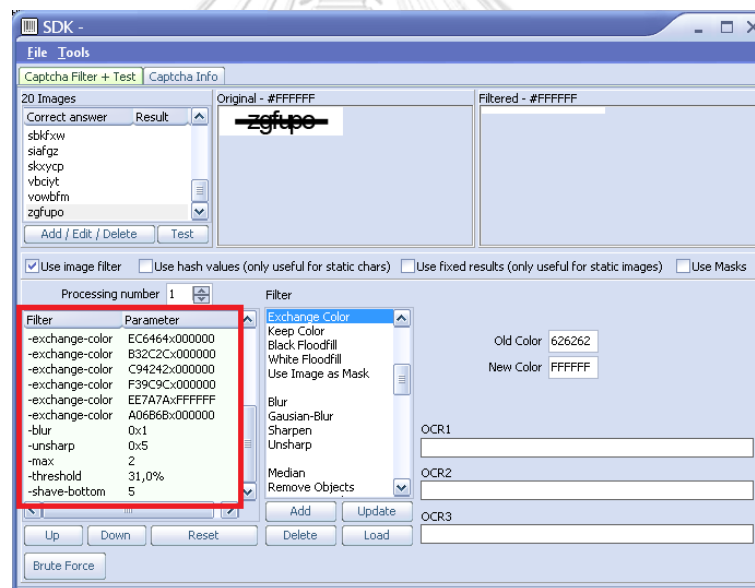
ภาพที่ จ. 15 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 9



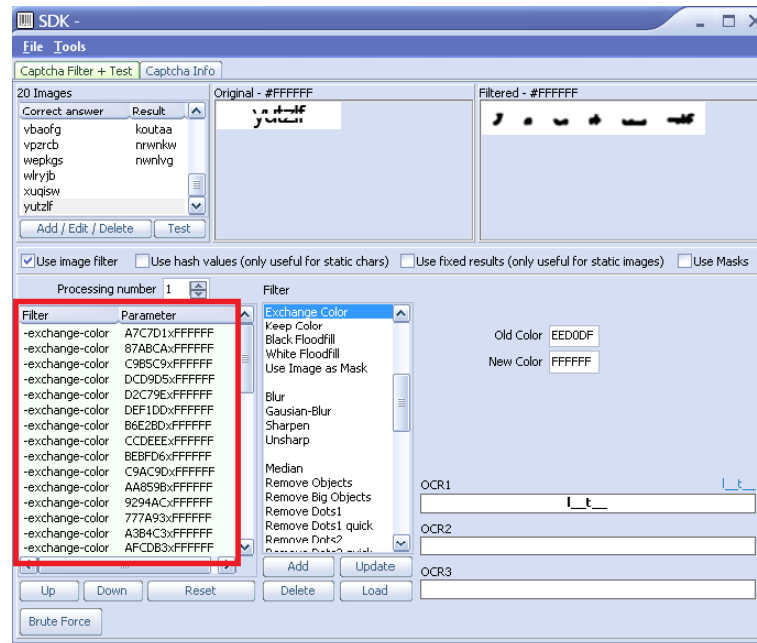
ภาพที่ จ. 16 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 10



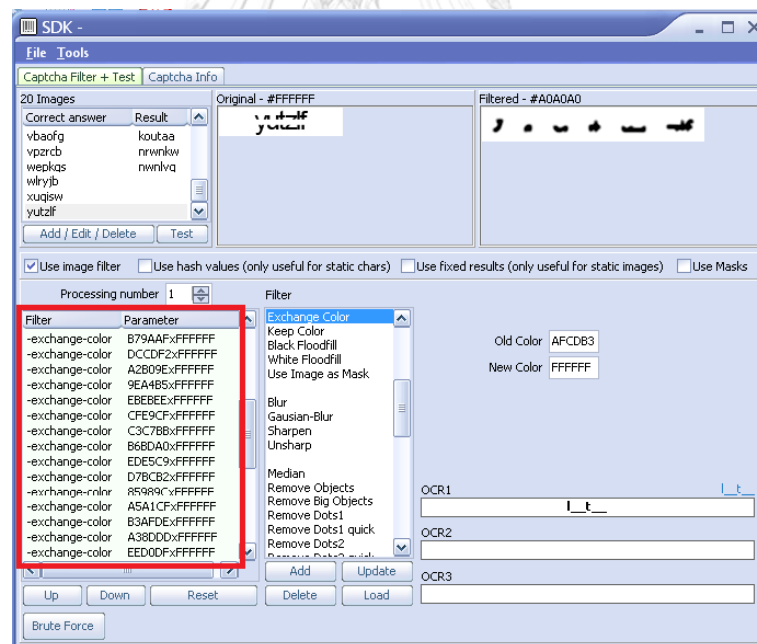
ภาพที่ จ. 17 Filter ที่เหมาะสมสำหรับแคปที่ชาข้อความรูปแบบที่ 11 (1)



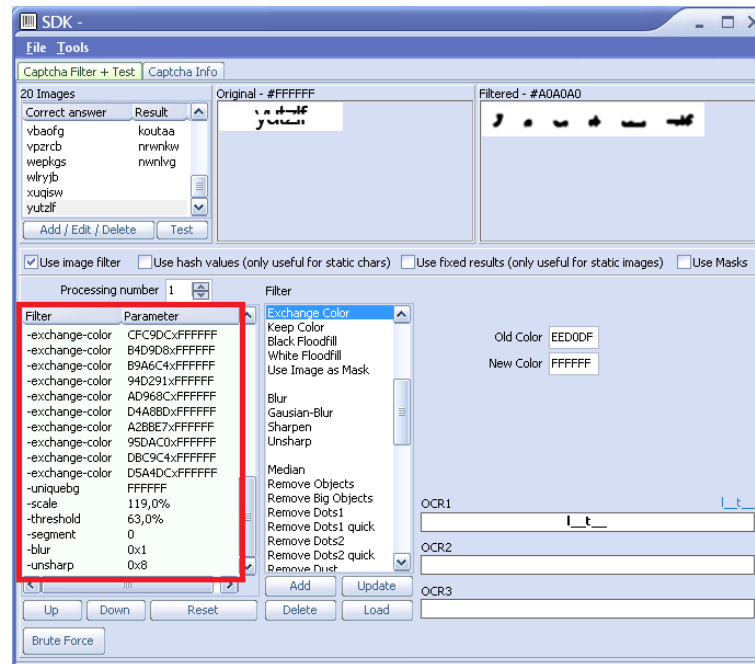
ภาพที่ จ. 18 Filter ที่เหมาะสมสำหรับแคปที่ชาข้อความรูปแบบที่ 11 (2)



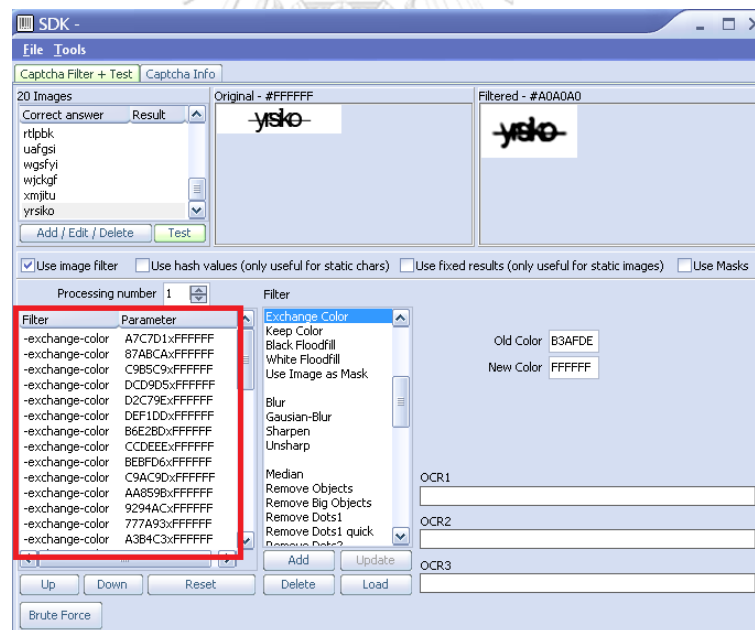
ภาพที่ จ. 19 Filter ที่เหมาะสมสำหรับแคปที่ชาข้อความรูปแบบที่ 12 (1)



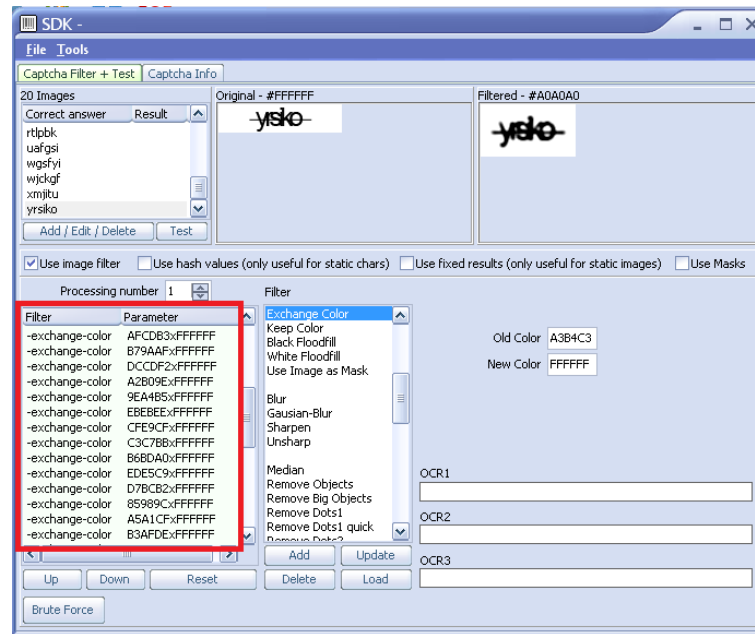
ภาพที่ จ. 20 Filter ที่เหมาะสมสำหรับแคปที่ชาข้อความรูปแบบที่ 12 (2)



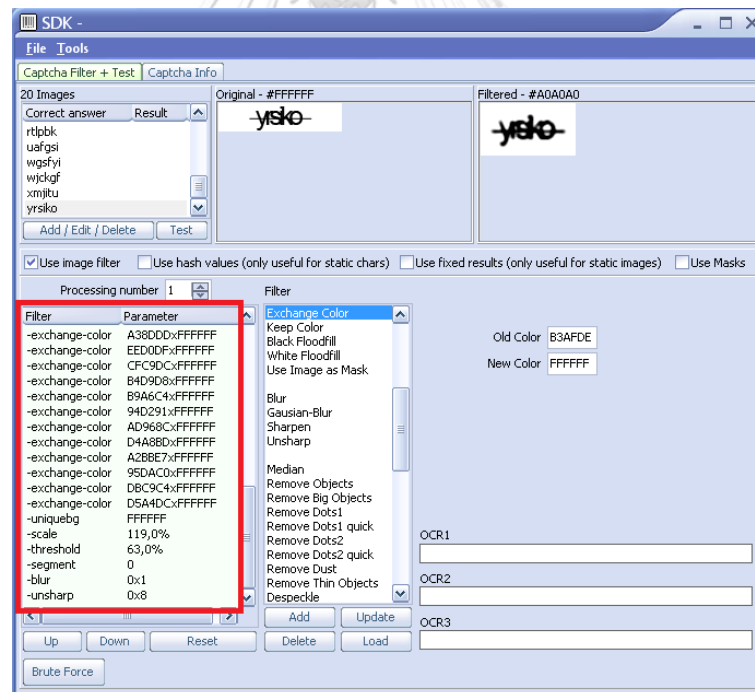
ภาพที่ จ. 21 Filter ที่เหมาะสมสำหรับแคปที่ชาข้อความรูปแบบที่ 12 (3)



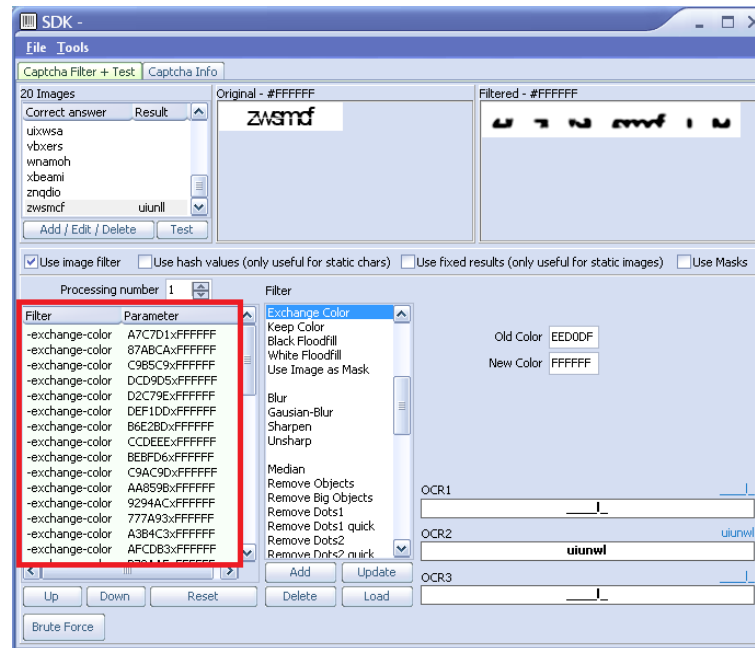
ภาพที่ จ. 22 Filter ที่เหมาะสมสำหรับแคปที่ชาข้อความรูปแบบที่ 13 (1)



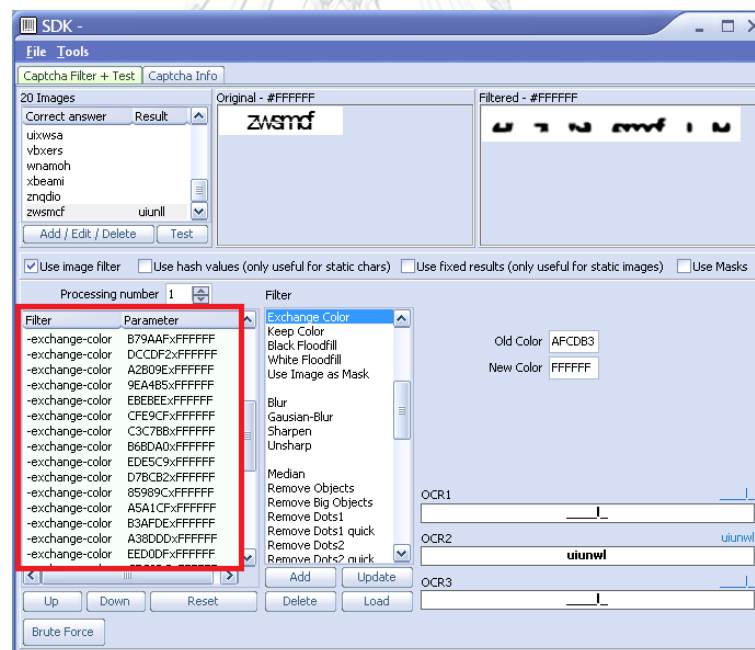
ภาพที่ จ. 23 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 13 (2)



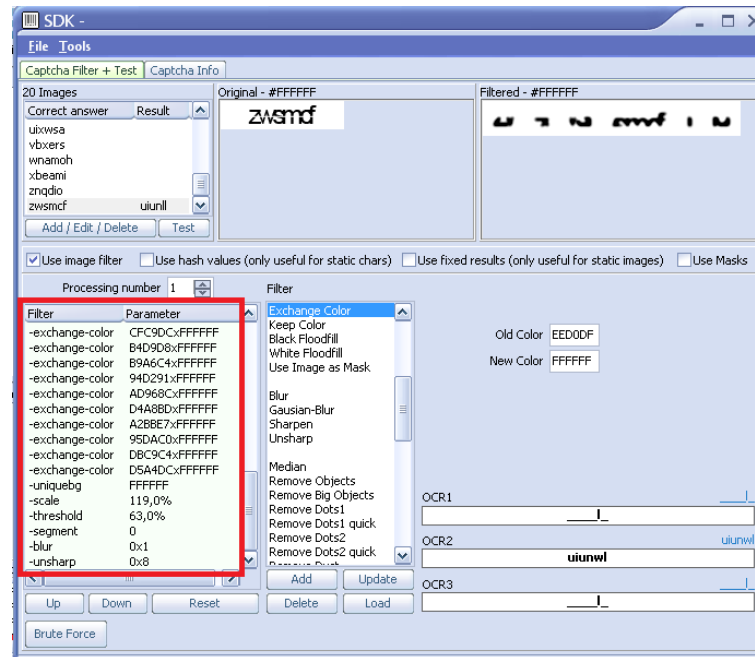
ภาพที่ จ. 24 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 13 (3)



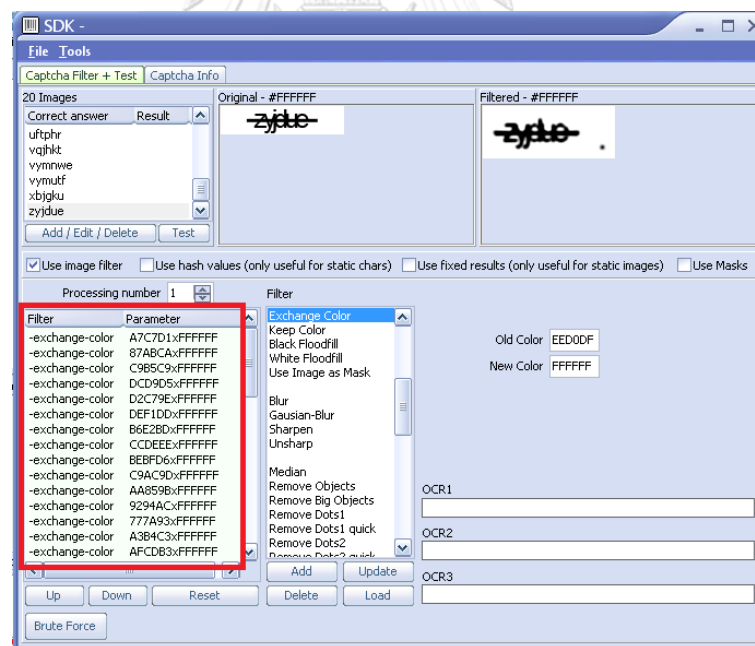
ภาพที่ จ. 25 Filter ที่เหมาะสมสำหรับแคปที่ชาข้อความรูปแบบที่ 14 (1)



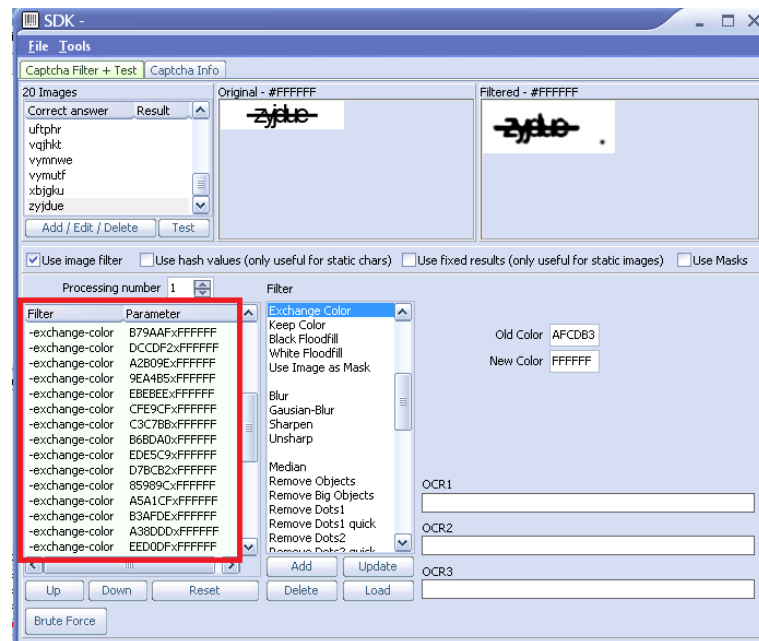
ภาพที่ จ. 26 Filter ที่เหมาะสมสำหรับแคปที่ชาข้อความรูปแบบที่ 14 (2)



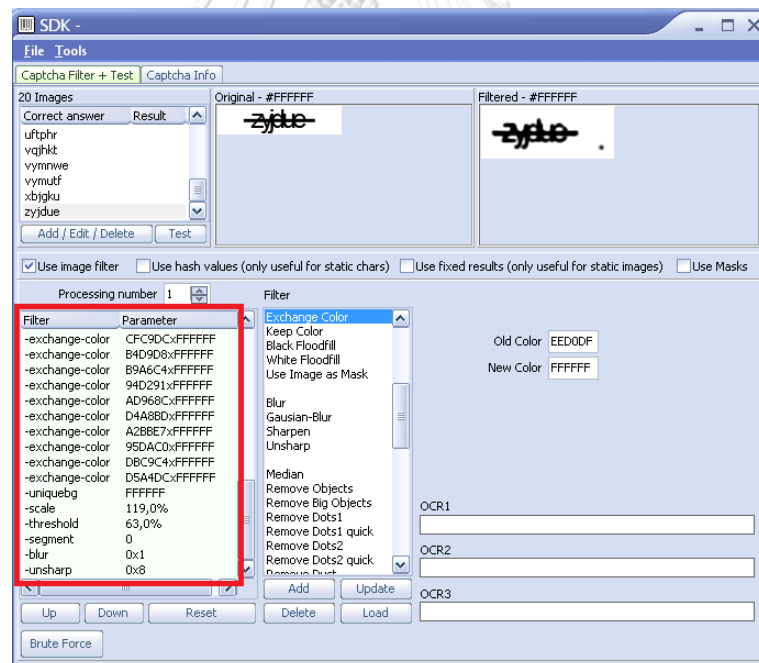
ภาพที่ จ. 27 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 14 (3)



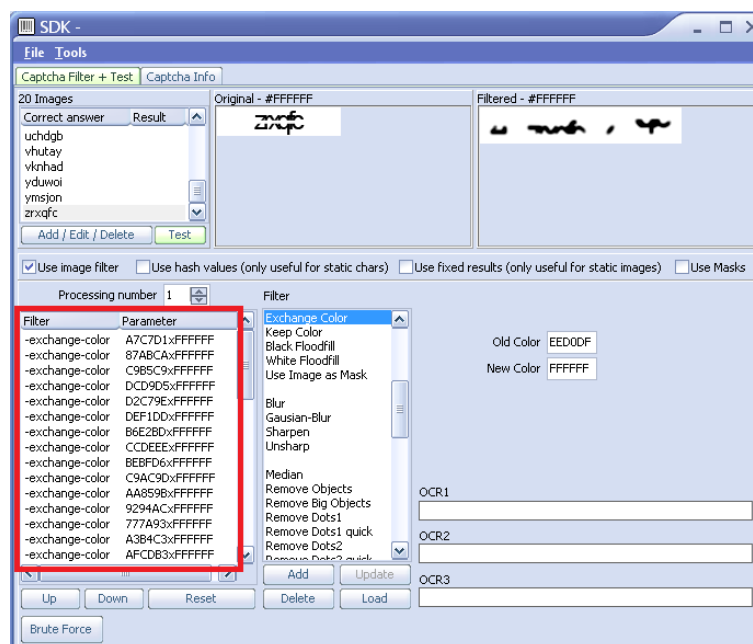
ภาพที่ จ. 28 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 15 (1)



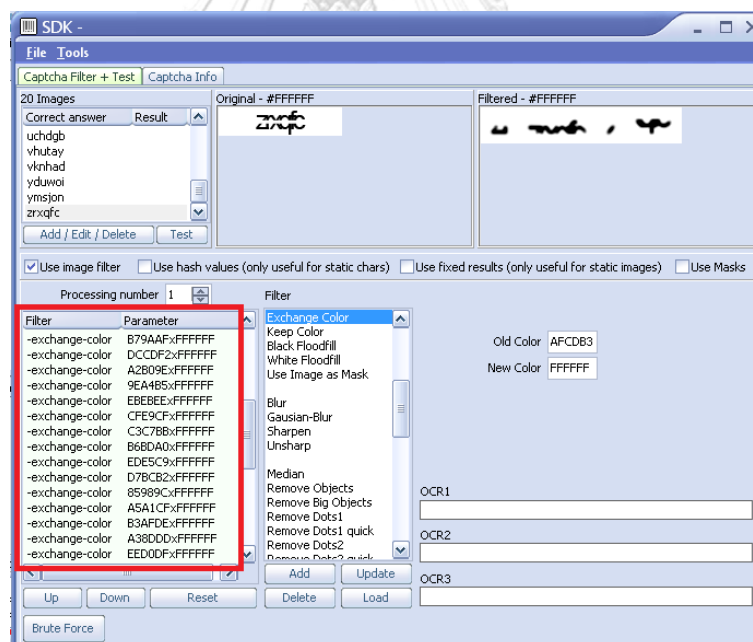
ภาพที่ จ. 29 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 15 (2)



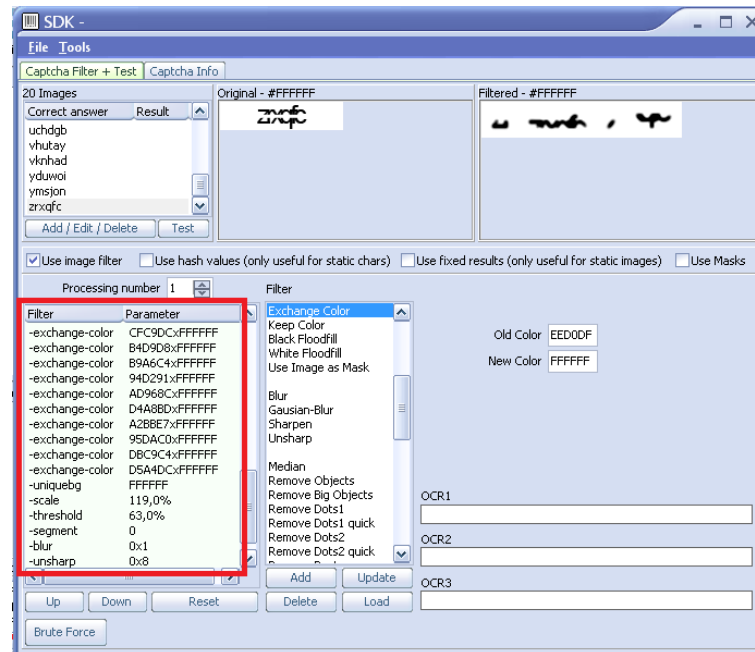
ภาพที่ จ. 30 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 15 (3)



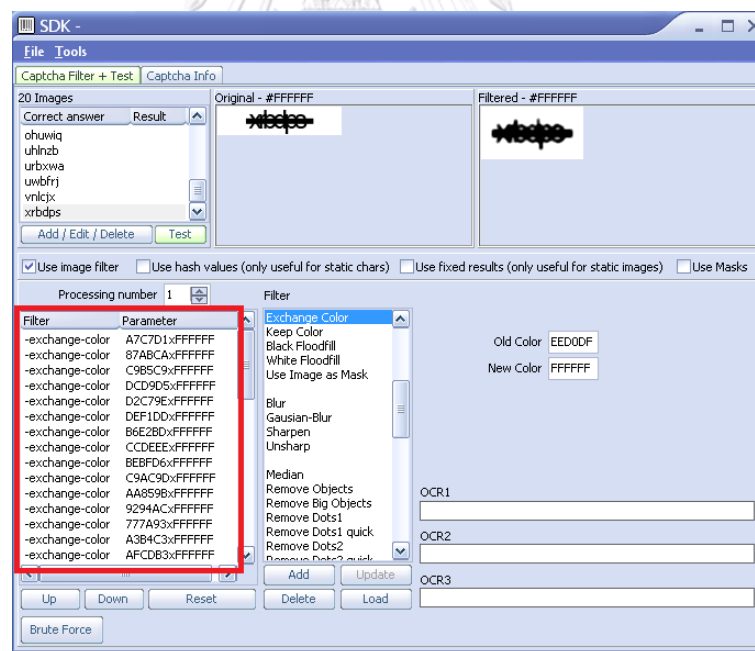
ภาพที่ จ. 31 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 16 (1)



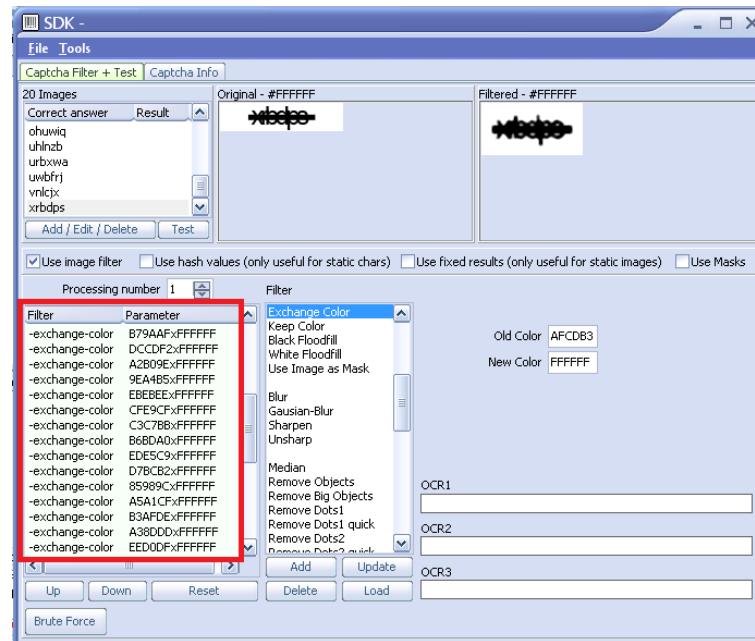
ภาพที่ จ. 32 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 16 (2)



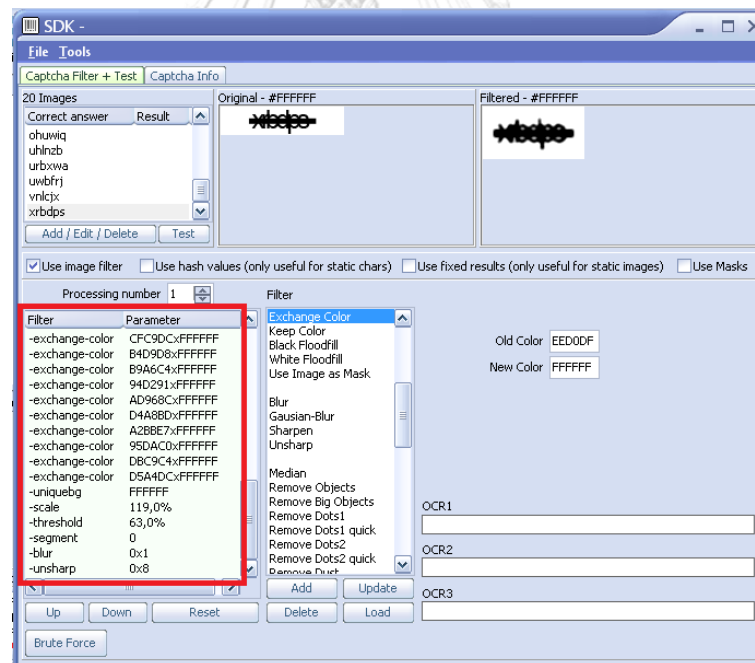
ภาพที่ จ. 33 Filter ที่เหมาะสมสำหรับแคปช่าข้อความรูปแบบที่ 16 (3)



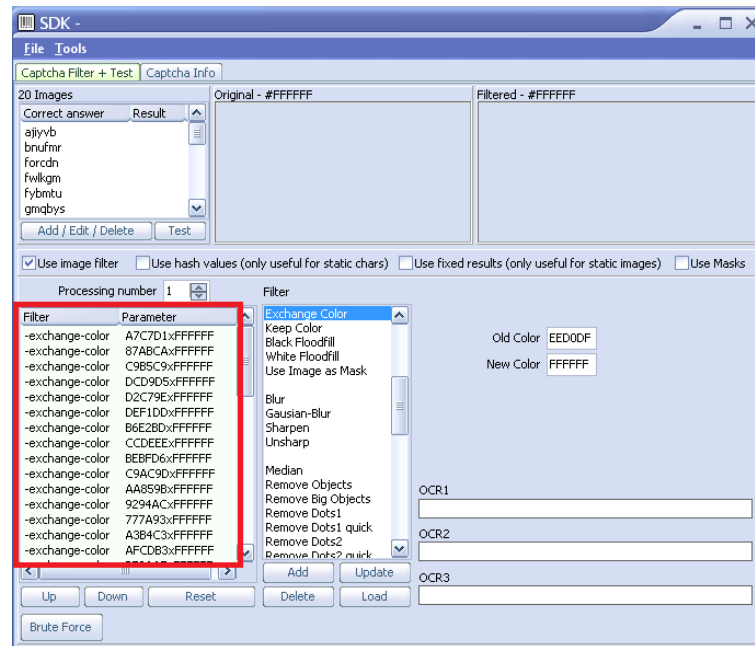
ภาพที่ จ. 34 Filter ที่เหมาะสมสำหรับแคปช่าข้อความรูปแบบที่ 17 (1)



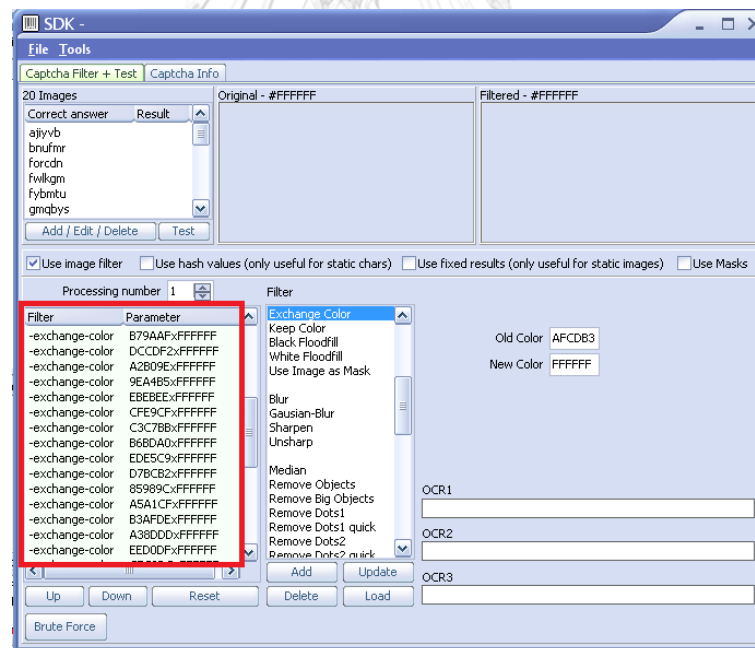
ภาพที่ จ. 35 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 17 (2)



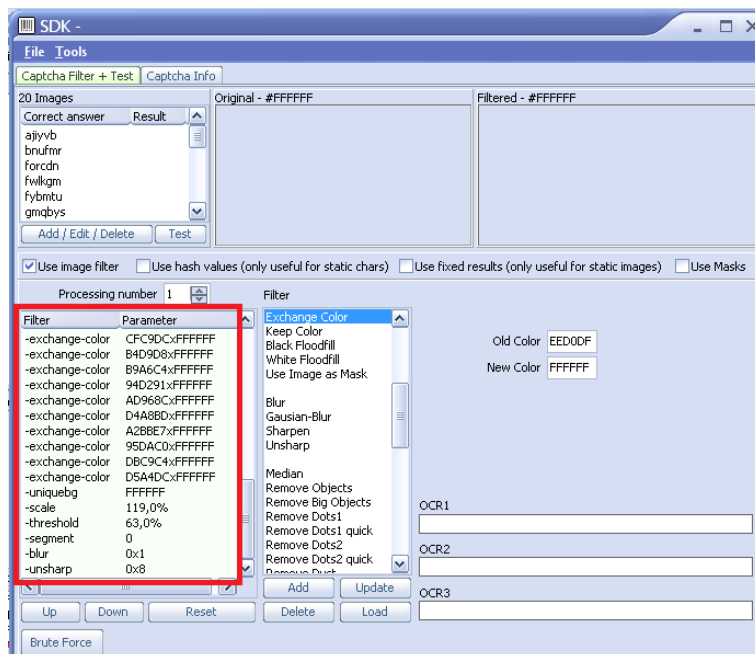
ภาพที่ จ. 36 Filter ที่เหมาะสมสำหรับแคปต์ชาข้อความรูปแบบที่ 17 (3)



ภาพที่ จ. 37 Filter ที่เหมาะสมสำหรับแคปท์ชาข้อความรูปแบบที่ 18 (1)



ภาพที่ จ. 38 Filter ที่เหมาะสมสำหรับแคปท์ชาข้อความรูปแบบที่ 18 (2)



ภาพที่ จ. 39 Filter ที่เหมาะสมสำหรับแคปช่าข้อความรูปแบบที่ 18 (3)



ประวัติผู้เขียนวิทยานิพนธ์

นางสาวศุภาพิชญ์ ชูชาติ เกิดเมื่อวันที่ 8 ตุลาคม พ.ศ. 2530 สำเร็จการศึกษาระดับปริญญาตรี หลักสูตรเทคโนโลยีสารสนเทศเพื่อธุรกิจ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปี พ.ศ. 2553 และได้เข้าศึกษาต่อในระดับปริญญาโท หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศทางธุรกิจ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปี พ.ศ. 2557

