

ผลของแบบอักษร การหมุนเอียง และชุดอักขระ ต่ออัตราการยืนยันความเป็นมนุษย์ :

การศึกษาแบบที่หาข้อความ

Effects of Typefaces, Rotations and Sets of Characters of text-based CAPTCHA on Human Affirmative Rates

มณีนันท์ ชาติรังสรรค์¹ และ ชัชพงศ์ ตั้งมณี²

¹นิสิตปริญญาโท สาขาการพัฒนซอฟต์แวร์ด้านธุรกิจ คณะพาณิชยศาสตร์และการบัญชี

จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพฯ และอาจารย์ประจำภาควิชาบริหารธุรกิจ

คณะบริหารธุรกิจ เศรษฐศาสตร์และการสื่อสาร มหาวิทยาลัยนเรศวร พิษณุโลก

²ผู้ช่วยศาสตราจารย์ ประจำภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

กรุงเทพฯ

ABSTRACT – Text-based CAPTCHAs are still a problem for humans. One of the effectiveness indicators for CAPTCHAs, a human-affirmative rate is the proportion of those who are able to correctly decode the text-based CAPTCHAs to those who have made an attempt to do it. It is thus interesting to examine the effects of (1) rotations of characters (2) typefaces and (3) sets of characters of CAPTCHAs on the human affirmative rates. The analysis of a lab experiment indicated that the effect of typefaces and sets of characters of CAPTCHA are statistically significant on the human affirmative rates at the 0.05 level, but the effects of rotations of characters of CAPTCHA were not significant. In addition to extending knowledge of text-based CAPTCHA's design, developers could apply the findings to select typeface, rotations and sets of characters of text-based CAPTCHA that easy for human.

KEY WORDS -- CAPTCHA, Typefaces, Rotations of Characters, Sets of Characters, Human Affirmative Rates

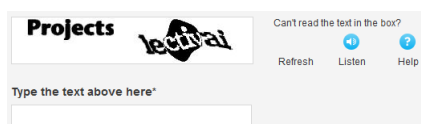
บทคัดย่อ -- ในปัจจุบันแบบที่หาข้อความยังคงเป็นปัญหาในการใช้งานของมนุษย์ หนึ่งในตัววัดประสิทธิภาพของแบบที่หาข้อความคือ อัตราการยืนยันความเป็นมนุษย์ อันหมายถึง อัตราที่ผู้ใช้สามารถระบุอักขระที่ปรากฏในแบบที่หาข้อความได้ถูกต้อง ดังนั้นจึงเป็นประเด็นที่น่าสนใจที่จะวิเคราะห์ถึงลักษณะต่างๆ ของแบบที่หาข้อความ การศึกษาครั้งนี้จึงมุ่งวิเคราะห์ผลของ (1) การหมุนเอียงอักขระ (2) แบบอักษร และ (3) ชุดของอักขระ ต่ออัตราดังกล่าว การวิเคราะห์ผลของการทดลองในห้องปฏิบัติการคอมพิวเตอร์ พบว่า แบบอักษร และชุดของอักขระ มีผลต่ออัตราการยืนยันความเป็นมนุษย์อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ทว่าผลของการหมุนเอียงอักขระต่ออัตราการยืนยันความเป็นมนุษย์ไม่มีนัยสำคัญ ข้อสรุปที่ได้จากการศึกษานี้เป็นแนวคิดในการพัฒนาแบบที่หาข้อความ อีกทั้งผู้พัฒนาแบบที่หาข้อความยังใช้เป็นข้อมูลในการเลือกระดับการเอียงตัวอักษร แบบตัวอักษร และชุดตัวอักษรให้เหมาะสมต่อการใช้งานของมนุษย์

คำสำคัญ --แบบที่หา, แบบอักษร, การหมุนเอียง, ชุดอักขระ, อัตราการยืนยันความเป็นมนุษย์

1. บทนำ

ในปัจจุบัน อินเทอร์เน็ตเข้ามามีบทบาทในชีวิตประจำวันของผู้คนเป็นอย่างมาก และมีจำนวนผู้เข้าใช้อินเทอร์เน็ตในประเทศไทยสูงขึ้น ในปี 2554 จำนวนผู้ใช้อินเทอร์เน็ตมีแนวโน้มที่จะขยายตัวเพิ่มสูงขึ้น [21] นอกจากนี้ ยังมีเว็บไซต์ต่างๆ เกิดขึ้นมากมายที่เปิดรับให้ผู้ใช้สามารถสมัครเป็นสมาชิกของเว็บไซต์ แสดงความคิดเห็น หรือส่งข้อความกลับไปยังเว็บไซต์ได้ จึงเป็นโอกาสให้ผู้ใช้ไม่ประสงค์ดี อาศัยช่องทางนี้พัฒนาโปรแกรมเพื่อส่งข้อความก่อกวนที่ผู้ใช้ไม่ได้ร้องขอเป็นจำนวนมาก หรือที่เรียกว่า “สแปม” (Spam) โดยสแปมเหล่านี้จะถูกส่งโดยโปรแกรมคอมพิวเตอร์อัตโนมัติ หรือรู้จักโดยทั่วไปว่าบอต [13] ดังนั้น จึงมีผู้คิดค้นวิธีการป้องกันการก่อกวนดังกล่าวด้วยประติษฐกรรมที่เรียกว่า “CAPTCHA” (Completely Automated Public Turing Test to Tell Computers and Humans Apart) โดยแคปทชา คือ การทดสอบเพื่อแยกระหว่างความเป็นมนุษย์กับความเป็นคอมพิวเตอร์ ทั้งนี้คำตอบต่อการทดสอบนั้นจะบ่งชี้ว่าผู้ใช้งานเป็นมนุษย์ ไม่ใช่โปรแกรมคอมพิวเตอร์อัตโนมัติ [1, 10] การศึกษานี้ได้เลือกแคปทชาข้อความมาเป็นประเด็นสำคัญของการวิจัย ดังแสดงในรูปที่ 1 เนื่องจากแคปทชาข้อความได้ถูกนำมาใช้งานมากที่สุด อีกทั้งยังสามารถสร้างแบบทดสอบของแคปทชาข้อความได้หลายลักษณะ [5]

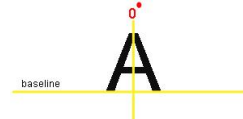
งานวิจัยของไมโครซอฟท์ [3] ได้ตรวจสอบวิธีการบิดเบือนอักขระที่ปรากฏในแคปทชา และพบว่า การหมุนเอียงอักขระ (Rotation) ตั้งแต่ลบ 45 ถึง 45 องศา เป็นการบิดเบือนที่ร้อยละ 99 ของหน่วยทดลองยังคงตอบได้ถูกต้อง นอกจากนี้ นักวิจัยที่ไมโครซอฟท์ได้ทดลองนำตัวอักษรหนึ่งตัวที่ได้บิดเบือนมาให้คอมพิวเตอร์อัตโนมัติโจมตี พบว่าการหมุนเอียงอักขระ เป็นการบิดเบือนที่คอมพิวเตอร์อัตโนมัติโจมตีผิดพลาดทั้งหมด



รูปที่ 1. ตัวอย่างแคปทชาข้อความจากเว็บไซต์ Skype.com [12]

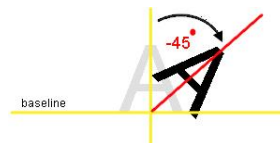
การวิจัยนี้จึงสนใจการหมุนเอียงอักขระ โดยที่การหมุนเอียงอักขระ หมายถึง การเอียงอักขระไปในมุมตามต้องการ สำหรับแสดงเป็นแคปทชาข้อความ โดยการวัดมุมจะวัดจาก

อักขระที่ตั้งตรงกับเส้นบรรทัด (baseline) ทั้งนี้เส้นบรรทัด คือ เส้นตรงแนวนอนที่ตัวอักขระวางอยู่ในแนวเดียวกัน ปกติอักขระจะตั้งบนเส้นบรรทัดโดยไม่ลอยสูงขึ้นหรือต่ำกว่าเส้นบรรทัด การตั้งตรงของอักขระอยู่ในรูปที่ 2 คือการแสดงที่ศูนย์องศาหรือไม่หมุนเอียงนั่นเอง

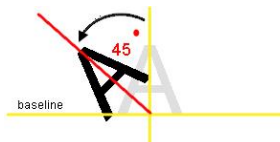


รูปที่ 2. ตัวอย่างการเอียงอักขระระดับศูนย์องศา

ทั้งนี้การศึกษาในอดีต [3, 6] ได้กำหนดระดับการหมุนเอียงของอักขระไว้ตั้งแต่ลบ 45 องศา ถึงบวก 45 องศา และอ้างว่ามีประสิทธิภาพมาก ทว่ายังไม่มีการเปรียบเทียบเชิงประจักษ์ยืนยันความสามารถของการแสดงอักขระที่ส่องมุมการหมุน ดังนั้น ผู้วิจัยจึงสนใจที่จะเปรียบเทียบการหมุนเอียงอักขระในแคปทชาที่ (1) ระดับลบ 45 องศา โดยอักขระจะหมุนเอียงทิศทางตามเข็มนาฬิกา หรือทางขวา ทั้งนี้รูปที่ 3 แสดงอักขระ A ที่เอียงทิศทางตามเข็มนาฬิกา หรือทำมุม 45 องศา กับเส้นตรงที่ตั้งฉากกับเส้นบรรทัด และ (2) ที่ระดับบวก 45 องศา โดยอักขระจะหมุนเอียงทิศทางทวนเข็มนาฬิกา หรือทางซ้าย ทั้งนี้รูปที่ 4 แสดงอักขระ A ที่เอียงทิศทางทวนเข็มนาฬิกา หรือทำมุม 45 องศา กับเส้นตรงที่ตั้งฉากกับเส้นบรรทัด



รูปที่ 3. ตัวอย่างการหมุนเอียงอักขระในระดับลบ 45 องศา



รูปที่ 4. ตัวอย่างการหมุนเอียงอักขระในระดับบวก 45 องศา

มีงานวิจัยที่ได้พัฒนาแคปทชาประเภทข้อความแบบลายมือมนุษย์ พบว่า อัตราการตอบของมนุษย์นั้นตอบได้ถูกต้องมากกว่าอัตราการตอบได้ถูกต้องของโปรแกรมถอดรหัสข้อความ [8, 9] และยังไม่พบงานในอดีตที่ได้พัฒนาแคปทชาข้อความด้วยการใช้ภาษาอังกฤษแบบตัวเขียน ดังนั้น ประเด็นที่

ผู้วิจัยสนใจคือ แบบตัวอักษรที่ปรากฏในแคปทชา อันจำแนกได้สองแบบคือ (1) แบบตัวเขียน (Cursive Letter) หรือแบบที่คล้ายลายมือมนุษย์ และ (2) แบบตัวพิมพ์ (Typed Letter)

ตัวแปรอิสระอีกหนึ่งตัวของงานนี้คือ ชุดอักขระ (Sets of Characters) ของแคปทชาข้อความ ในที่นี้ ชุดของอักขระหมายถึง ลักษณะของตัวอักษรที่ปรากฏในแคปทชาข้อความ โดยทั่วไปแล้วการสร้างแคปทชาข้อความมีทั้งนำตัวอักษรอังกฤษ และตัวเลข มาใช้ และเมื่อถูกทำให้บิดเบี้ยวด้วยแล้ว จะยิ่งเพิ่มโอกาสสับสนกับมนุษย์ในการโต้ตอบกับแคปทชา เนื่องด้วยตัวอักษรบางตัวอักษรกับตัวเลขมีลักษณะคล้ายกัน เช่น ตัวเลข 5 กับ อักษร S หรือ ตัวเลข 2 กับ อักษร Z เป็นต้น ดังนั้นผู้วิจัยจึงเลือกที่จะใช้อักษรเท่านั้น ไม่รวมตัวเลข ในการพัฒนาแคปทชาสำหรับการศึกษานี้

อย่างไรก็ตาม ยังมีงานในอดีตที่นำเสนอเกี่ยวกับอักษรภาษาอังกฤษที่สะดวกต่อการอ่านของมนุษย์ โดย Paterson และ Tinker [14] พบว่าบุคคลสามารถเข้าใจตัวพิมพ์เล็กได้ง่ายและเร็วกว่าตัวพิมพ์ใหญ่ อีกทั้ง Arditi และ Cho [2] ได้สรุปไว้ว่าระยะเวลาการอ่านโดยเฉลี่ยของข้อความที่เป็นตัวพิมพ์ใหญ่ทั้งหมด (All Capital) เร็วกว่าข้อความที่มีตัวพิมพ์ใหญ่ผสมตัวพิมพ์เล็ก (Mixed) แต่ขัดแย้งกับ Tinker [15] ที่กล่าวไว้ว่าตัวพิมพ์เล็กเป็นอักษรที่มนุษย์อ่านง่ายและเร็วกว่าตัวพิมพ์ใหญ่ ดังนั้นผู้วิจัยจึงสนใจที่จะเปรียบเทียบลักษณะของชุดอักขระของแคปทชาข้อความที่เป็น อักษรตัวใหญ่ทั้งหมด และอักษรตัวเล็กทั้งหมด

เนื่องจากแคปทชาข้อความที่ดีจะต้องง่ายต่อการตอบของมนุษย์ ดังนั้นการตรวจสอบความสามารถของแคปทชาข้อความเมื่อมีลักษณะต่างๆ กับตามสี่ตัวแปรที่ได้อธิบายข้างต้น จึงสามารถกระทำกับอัตราการยืนยันความเป็นมนุษย์ (Human Affirmative Rate) โดยที่อัตราการยืนยันความเป็นมนุษย์ หมายถึง การที่ผู้ใช้สามารถตอบแคปทชาข้อความได้ถูกต้อง

2. วัตถุประสงค์

วัตถุประสงค์ของงานการศึกษานี้มีอยู่สามข้อ ดังนี้ (1) เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ ระหว่างแคปทชาที่แสดงการหมุนเอียงอักษรในระดับลบ 45 องศา และบวก 45 องศา (2) เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ ระหว่างแคปทชาที่แสดงแบบอักษรตัวพิมพ์ กับที่แสดงแบบ

อักษรตัวเขียน และ (3) เพื่อเปรียบเทียบอัตราการยืนยันความเป็นมนุษย์ระหว่างแคปทชาที่แสดงชุดอักขระเป็นอักษรตัวใหญ่ล้วน กับที่แสดงเป็นอักษรตัวเล็กล้วน

3. ระเบียบวิธีวิจัย

งานวิจัยนี้เป็นการวิจัยเชิงทดลอง (Experiment Research) ที่ต้องการวิเคราะห์ผลของตัวแปรที่คุมค่าให้เปลี่ยนต่ออีกตัวแปร [18, 20] โดยเลือกใช้การทดลองในห้องปฏิบัติการ (Laboratory Experiment) และผู้วิจัยจะพยายามอย่างที่สุดในการควบคุมตัวแปรอื่นๆ ให้คงที่ เช่น สีอักษรและสีพื้นหลังของแคปทชาข้อความ เว็บไซต์ที่ใช้แสดงแคปทชา เพื่อให้เกิดความแตกต่างอัตราการยืนยันความเป็นมนุษย์ที่อาจเกิดขึ้นเพราะ (1) การหมุนเอียงอักษรของแคปทชาข้อความ (2) แบบอักษรของแคปทชาข้อความ และ (3) ชุดอักขระของแคปทชาข้อความ เท่านั้น โดยการนำเสนอรายละเอียดของระเบียบวิธีวิจัย เป็นดังนี้ การทดสอบสมมติฐาน ตัวแปรที่ศึกษา เครื่องมือในการเก็บข้อมูล การเลือกหน่วยทดลอง ขั้นตอนการเก็บข้อมูล ประเด็นความเชื่อถือและความถูกต้อง และกรอบการวิเคราะห์ข้อมูล

3.1 การทดสอบสมมติฐาน

สมมติฐานสามข้อที่การศึกษามุ่งทดสอบ คือ (1) อัตราการยืนยันความเป็นมนุษย์ควรต้องแตกต่างกัน ระหว่างแคปทชาที่แสดงการหมุนเอียงอักษรในระดับที่ลบ 45 องศาและบวก 45 องศาอย่างมีนัยสำคัญทางสถิติ (2) อัตราการยืนยันความเป็นมนุษย์ของแคปทชาที่แสดงแบบอักษรตัวพิมพ์จะสูงกว่าแบบอักษรตัวเขียนอย่างมีนัยสำคัญทางสถิติ และ (3) อัตราการยืนยันความเป็นมนุษย์ของแคปทชาข้อความที่แสดงชุดอักขระเป็นอักษรตัวเล็กจะสูงกว่าเมื่อแสดงเป็นอักษรตัวใหญ่อย่างมีนัยสำคัญทางสถิติ ทั้งนี้สมมติฐานที่กำหนดมาจากรรณกรรมในอดีต [2, 3, 7, 8, 9, 11, 14, 15]

3.2 ตัวแปรที่ศึกษา

ตัวแปรอิสระในงานนี้มีจำนวนสามตัวแปร ได้แก่ (1) การหมุนเอียงอักษรของแคปทชาข้อความ คือ การเอียงอักษรไปในมุมตามต้องการที่ปรากฏในแคปทชาข้อความ โดยจำแนกค่าของตัวแปรได้สองค่า คือ เอียงลบ 45 องศา และ เอียงบวก 45

องศา (โปรคดูตัวอย่างในรูปที่ 3 และรูปที่ 4) (2) แบบอักษรของแผนที่ชาข้อความ คือ แบบอักษรที่ปรากฏในแผนที่ชาข้อความ โดยจำแนกค่าของตัวแปรได้สองลักษณะ คือ อักษรตัวพิมพ์ โดยใช้แบบอักษรในกลุ่ม San Serif คือ อักษร Arial ดังรูปที่ 5 และ อักษรตัวเขียน ใช้แบบอักษรในกลุ่ม Script คือ อักษร Cursive standard ดังรูปที่ 6 เนื่องจากเป็นอักษรที่มีการดาวน์โหลดจำนวนมาก [4] และเป็นอักษรที่หน่วยตัวอย่างในช่วงอายุ 20 – 29 ปี อ่านง่ายที่สุด และ (3) ชุดอักขระของแผนที่ชาข้อความ คือ ลักษณะอักขระที่ปรากฏในแผนที่ชาข้อความ โดยสามารถจำแนกออกได้เป็นสองลักษณะ คือ อักษรตัวใหญ่ทั้งหมด และ อักษรตัวเล็กทั้งหมด

A B C D a b c d

รูปที่ 5. ตัวอย่างอักษรตัวพิมพ์ที่ใช้ในงานนี้

A B C D a b c d

รูปที่ 6. ตัวอย่างอักษรตัวเขียนที่ใช้ในงานนี้

ตัวแปรตามมีหนึ่งตัวแปร คือ อัตราการยืนยันความเป็นมนุษย์ (Human Affirmative Rates) หมายถึง อัตราผู้ใช้สามารถตอบแผนที่ชาข้อความได้ถูกต้อง ในที่นี้ การตอบแผนที่ชาข้อความ คือ การที่ผู้ใช้สามารถระบุอักษรที่ปรากฏในแผนที่ชาข้อความได้ถูกต้องทั้งหมด และสามารถเข้าใช้บริการของเว็บไซต์ได้ [16] โดยวัดจากอัตราการตอบแผนที่ชาข้อความได้ถูกต้องของผู้ใช้ คือ ร้อยละของจำนวนการตอบแผนที่ชาข้อความได้ถูกต้องของผู้ใช้ ต่อจำนวนการตอบแผนที่ชาข้อความทั้งหมดของผู้ใช้

$$P = (n/N) \times 100$$

เมื่อ P คือ อัตราการยืนยันความเป็นมนุษย์ (ร้อยละ)

n คือ จำนวนการตอบแผนที่ชาข้อความได้ถูกต้องของหน่วยทดลอง

N คือ จำนวนการตอบแผนที่ชาข้อความทั้งหมดของหน่วยทดลอง

3.3 เครื่องมือในการเก็บข้อมูล

ผู้วิจัยได้พัฒนาเว็บไซต์ให้บริการดาวน์โหลดเพลงสำหรับการทดลองนี้โดยเฉพาะ และกำหนดให้หน่วยทดลองแวะชมเว็บไซต์ที่พัฒนา เมื่อกำหนดให้หน่วยทดลองเลือกดาวน์โหลด

เพลง หน่วยทดลองต้องได้ตอบกับแผนที่ชาในเงื่อนไขต่างๆของการทดลอง อันเป็นการเก็บข้อมูลของอัตราการยืนยันความเป็นมนุษย์ ทั้งนี้การออกแบบเว็บไซต์จะมีข้อมูลและรายละเอียดเหมือนกันทุกประการ แต่จะมีความแตกต่างเฉพาะรูปแบบของแผนที่ชาข้อความอยู่ 8 รูปแบบ โดยระบบดาวน์โหลดเพลงจะสุ่มตัวอักษร A-Z หรือตัวอักษร a-z จำนวน 4 ตัวอักษรขึ้นมาแสดงเป็นแผนที่ชาข้อความตามรูปแบบต่างๆ ที่กล่าวข้างต้นดังรูปที่ 7

↶↷↶↷

รูปที่ 7. ตัวอย่างแผนที่ชาข้อความที่ใช้ในงานนี้

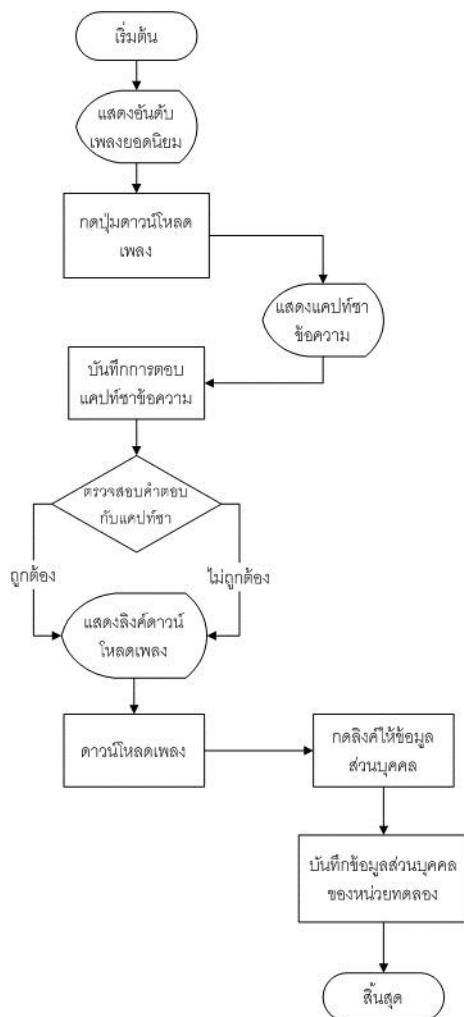
อีกทั้ง ผู้วิจัยได้พัฒนางานเพื่อกำหนดประเภทเพลงให้กับหน่วยทดลองดาวน์โหลดเพลง โดยที่งานจะเป็นตัวกำหนดให้หน่วยทดลองพบแผนที่ชาข้อความที่ต่างกัน นอกจากนี้ผู้วิจัยได้ตั้งคำถามเพื่อเก็บข้อมูลส่วนบุคคลของหน่วยทดลอง โดยคำถามนี้จะแสดงหลังจากหน่วยทดลองได้ดาวน์โหลดเพลงเป็นที่เรียบร้อยแล้ว

3.4 หน่วยทดลองและการเก็บข้อมูล

การเก็บข้อมูลของงานวิจัยนี้ทำในห้องปฏิบัติการของศูนย์คอมพิวเตอร์ โดยกลุ่มหน่วยทดลองจากนิสิตปริญญาตรี คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย จำนวน 240 คน สำหรับเป็นหน่วยทดลองอย่างน้อยกลุ่มละ 30 คน สำหรับทั้งสิ้น 8 เงื่อนไขของการทดลอง (2 การหมุนเอียง x 2 แบบอักษร x 2 ชุดอักขระ) ก่อนเริ่มการทดลอง ผู้วิจัยจะอธิบายโดยสรุปเกี่ยวกับระบบและขั้นตอนการทำงานให้กับหน่วยทดลองทราบ แต่ไม่อธิบายวัตถุประสงค์ที่แท้จริงของการทดลอง โดยผู้วิจัยจะแจกใบงานให้กับหน่วยทดลองเพื่อกำหนดให้หน่วยทดลองเลือกดาวน์โหลดเพลงตามประเภทเพลงที่กำหนดไว้

เมื่อหน่วยทดลองคลิกปุ่มดาวน์โหลดเพลง ระบบดาวน์โหลดเพลงจะแสดงหน้าเว็บที่มีแผนที่ชาข้อความ เพื่อให้หน่วยทดลองกรอกข้อความตามที่ปรากฏในแผนที่ชาข้อความ จากนั้นระบบดาวน์โหลดเพลงจะแสดงหน้าเว็บไซต์ที่ปรากฏลิงค์ดาวน์โหลดเพลง เพื่อให้หน่วยทดลองดาวน์โหลดเพลงได้ โดยระบบจะตรวจสอบผลการตอบแผนที่ชาข้อความของหน่วยทดลองว่าถูกต้องหรือไม่ถูกต้อง แต่ระบบจะไม่แจ้งให้หน่วย

ทดลองทราบ จากนั้นระบบจะบันทึกผลการตอบของหน่วยทดลองลงในฐานข้อมูล และหลังจากหน่วยทดลองดาวน์โหลดเพลงเรียบร้อยแล้ว จะมีลิงก์ให้ข้อมูลส่วนบุคคลปรากฏขึ้นมา เพื่อให้หน่วยทดลองตอบคำถามเกี่ยวกับข้อมูลส่วนบุคคล โดยขั้นตอนการเก็บข้อมูลการยืนยันความเป็นมนุษย์ สามารถแสดงเป็นแผนภาพดังรูปที่ 8



รูปที่ 8. แผนภาพแสดงขั้นตอนการเก็บข้อมูลจากหน่วยทดลอง

3.5 ประเด็นความเชื่อถือได้ และความถูกต้อง

ผู้วิจัยได้พยายามอย่างที่สุด เพื่อตอบวัตถุประสงค์ทั้งหมดของงานวิจัยนี้ให้ถูกต้องและน่าเชื่อถือ ดังนั้นสิ่งสำคัญ คือการควบคุมการทดลองให้ได้ข้อมูลที่ดีสำหรับนำไปสู่ผลสรุปที่ถูกต้องและน่าเชื่อถือ จึงต้องควบคุมปัจจัยต่างๆ ดังนี้ (1) เทคโนโลยีการสื่อสาร อันเป็นตัวแปรสำคัญที่มีผลต่อการทดลอง โดยเฉพาะการส่งข้อมูลระหว่างกัน ผู้วิจัยได้พยายามลด

ความคลาดเคลื่อนในเรื่องของความเร็วที่แตกต่างกัน ที่อาจเกิดขึ้นขณะที่หน่วยทดลองกำลังทำงานที่ได้รับมอบหมายให้น้อยที่สุด ดังนั้นจึงจำเป็นต้องทดลองในห้องปฏิบัติการ เพื่อให้ผลการทดลองที่ออกมาอยู่บนพื้นฐานของระบบที่เหมือนกันและน่าเชื่อถือ ตลอดจนถูกต้องในระดับที่ยอมรับได้ (2) การกำหนดหน่วยทดลอง ผู้วิจัยเลือกนิสิตปริญญาตรี ของคณะพาณิชยศาสตร์และการบัญชี เนื่องจากมีลักษณะความสามารถที่ทัดเทียมกัน กล่าวคือ หน่วยทดลองมีรายวิชาที่มีลักษณะเหมือนหรือคล้ายคลึงกัน ทำให้หน่วยทดลองมีความใกล้เคียงกันมากที่สุด (3) การพัฒนาเว็บไซต์เพื่อทดลอง ผู้วิจัยได้เลือกพัฒนาเว็บไซต์ให้บริการดาวน์โหลดเพลงขึ้นมาใหม่ เพื่อไม่ให้หน่วยทดลองเกิดความความคุ้นเคยกับระบบนั้นมาก่อน ซึ่งอาจจะส่งผลต่อความถูกต้องและน่าเชื่อถือได้ และงานที่กำหนดให้หน่วยทดลองทำ พร้อมกับกำกับการที่ไม่บอกวัตถุประสงค์ น่าที่จะส่งผลให้ข้อมูลที่เก็บมีความถูกต้องและน่าเชื่อถือ [19]

3.6 กรอบการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลจะกระทำดังนี้ คือ (1) การวิเคราะห์ที่เกี่ยวข้องกับข้อมูลส่วนตัวของหน่วยทดลอง จะกระทำด้วยสถิติเชิงพรรณนา เพื่อวิเคราะห์คำตอบจากคำถามส่วนบุคคล รวมถึงประเภทของสาขา และ (2) การตอบวัตถุประสงค์ทั้งสามข้อ ผู้วิจัยเลือกใช้วิธีการทางสถิติเชิงพารามิเตอร์ คือการทดสอบ Z แบบสองทาง เพื่อทดสอบสมมติฐานของผลต่างระหว่างสัดส่วนของประชากรสองประชากร สำหรับตัวแปรอิสระ การหมุนเวียนอักษรของแคปทช่าข้อความ ส่วนตัวแปรอิสระ แบบอักษร และชุดอักขระของแคปทช่าข้อความ จะใช้การทดสอบ Z แบบทางเดียว เพื่อทดสอบสมมติฐานของผลต่างระหว่างสัดส่วนของประชากรสองประชากร [18]

4. ผลการวิเคราะห์ข้อมูล

ผู้วิจัยได้พัฒนาเว็บไซต์ดาวน์โหลดเพลงสำหรับเก็บข้อมูลอัตราการยืนยันความเป็นมนุษย์ในห้องปฏิบัติการคอมพิวเตอร์ของคณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ทั้งนี้เก็บจากหน่วยทดลองจำนวนทั้งสิ้น 240 คน

รายละเอียดข้อมูลส่วนบุคคลของหน่วยทดลอง จำแนกตามเพศ อายุ หลักสูตรที่ศึกษา ประเภทสาขาของหน่วยทดลอง และการใช้อุปกรณ์ทางสาขาของหน่วยทดลอง โดยการ

จำแนกตามเพศ พบว่า จำนวนหน่วยทดลองเป็น หญิงจำนวน 176 คน และชาย 64 คน การจำแนกตามประเภทสายตา พบว่า กลุ่มที่มีสายตาสั้นมีจำนวนสูงที่สุด คิดเป็นร้อยละ 59.58 รองลงมาคือ กลุ่มสายตปกติ คิดเป็นร้อยละ 37.50 กลุ่มสายตาเอียงคิดเป็นร้อยละ 2.92 และกลุ่มสายตาวายคิดเป็นร้อยละ 0.00

หลังจากการเก็บรวบรวมข้อมูลเรียบร้อยแล้ว ผู้วิจัยได้นำข้อมูลมาวิเคราะห์เพื่อนำเสนออัตราการยืนยันความเป็นมนุษย์ โดยจำแนกตาม (1) การหมุนเอียงอักษร (2) แบบอักษร และ (3) ชุดอักขระ ของแคปทช่าข้อความ ดังแสดงในตารางที่ 1 ถึง 3

ตารางที่ 1. แสดงอัตราการยืนยันความเป็นมนุษย์ จำแนกตามการหมุนเอียงอักษร

การหมุนเอียงอักษร (องศา)	จำนวนผู้ตอบ แคปทช่าได้ถูกต้อง (คน)	จำนวนผู้ตอบ แคปทช่าทั้งหมด (คน)	อัตราการยืนยัน ความเป็นมนุษย์ (%)
ลบ 45	76	120	63.33
บวก 45	80	120	66.66

ตารางที่ 2. แสดงอัตราการยืนยันความเป็นมนุษย์ จำแนกตามแบบอักษร

แบบอักษร	จำนวนผู้ตอบ แคปทช่าได้ถูกต้อง (คน)	จำนวนผู้ตอบ แคปทช่าทั้งหมด (คน)	อัตราการยืนยัน ความเป็นมนุษย์ (%)
ตัวพิมพ์	111	120	92.50
ตัวเขียน	45	120	37.50

ตารางที่ 3. แสดงอัตราการยืนยันความเป็นมนุษย์ จำแนกตามชุดอักขระ

ชุดอักขระ	จำนวนผู้ตอบ แคปทช่าได้ถูกต้อง (คน)	จำนวนผู้ตอบ แคปทช่าทั้งหมด (คน)	อัตราการยืนยัน ความเป็นมนุษย์ (%)
ตัวใหญ่	69	120	57.50

ตัวเล็ก	87	120	72.50
---------	----	-----	-------

การทดสอบสมมติฐานเกี่ยวกับผลต่างของสัดส่วนระหว่างแคปทช่าที่แสดงการหมุนเอียงอักษรในระดับลบ 45 องศา และบวก 45 องศา ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เพราะค่า z ที่คำนวณได้คือ -0.540786 มีค่าน้อยกว่า 1.960

ส่วนการทดสอบสมมติฐานเกี่ยวกับผลต่างของสัดส่วนระหว่างแคปทช่าที่แสดงเป็นอักษรตัวพิมพ์และตัวเขียน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ เพราะค่า z ที่คำนวณได้คือ 8.931906 มีค่ามากกว่า 1.645 ดังนั้นผลการทดลองสรุปได้ว่า อัตราการยืนยันความเป็นมนุษย์ของแคปทช่าที่แสดงแบบอักษรตัวพิมพ์สูงกว่าแบบอักษรตัวเขียน

ส่วนการทดสอบสมมติฐานของสัดส่วนระหว่างแคปทช่าที่แสดงชุดอักขระเป็นอักษรตัวใหญ่ทั้งหมด และอักษรตัวเล็กทั้งหมด มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ เพราะค่า z ที่คำนวณได้คือ 2.435974 มีค่ามากกว่า 1.645 ดังนั้นผลการทดลองสรุปได้ว่า อัตราการยืนยันความเป็นมนุษย์ของแคปทช่าที่แสดงชุดอักขระเป็นอักษรตัวเล็กสูงกว่าที่แสดงด้วยอักษรตัวใหญ่

5. สรุปและอภิปรายผล

การเก็บข้อมูลของงานวิจัยนี้ทำในห้องปฏิบัติการของศูนย์คอมพิวเตอร์ โดยสุ่มหน่วยทดลองจากนิสิตปริญญาตรี คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย จำนวนทั้งสิ้น 240 คน

5.1 อัตราการยืนยันความเป็นมนุษย์กับการหมุนเอียงอักษร

ผลการทดสอบสมมติฐานพบว่า ความแตกต่างของอัตราการยืนยันความเป็นมนุษย์ระหว่างแคปทช่าที่มีการหมุนเอียงระดับลบ 45 องศา และบวก 45 องศา ไม่มีนัยสำคัญทางสถิติ ผลการวิเคราะห์ไม่เป็นไปตามที่ผู้วิจัยคาด แต่ยังคงคล้องกับความคิดของ Simion และคณะ [11] พบว่าการหมุนเอียงอักษรไปทางขวา (ทิศทางตามเข็มนาฬิกา) หรือทางซ้าย (ทิศทางทวนเข็มนาฬิกา) ค่าเฉลี่ยเวลาการตอบสนองของมนุษย์ไม่แตกต่างกันอย่างมีนัยสำคัญ แต่การไม่พบความแตกต่างในครั้งนี้ อาจเนื่องมาจากจำนวนอักษรที่จะถูกทำให้หมุนเอียงมีจำนวนน้อย จึงไม่ส่งผลต่อการรับรู้ของหน่วยทดลอง และทำให้ไม่พบความแตกต่างของอัตราการยืนยันความเป็นมนุษย์

5.2 อัตราการยืนยันความเป็นมนุษย์ต่อแบบอักษร

ผลการทดสอบสมมติฐานพบว่า อัตราการยืนยันความเป็นมนุษย์ระหว่างแบบอักษรที่แสดงแบบอักษรตัวพิมพ์สูงกับแบบอักษรที่แสดงเป็นแบบอักษรตัวเขียน อย่างมีนัยสำคัญทางสถิติที่ 0.05 โดยผลการทดสอบดังกล่าวเป็นไปตามที่ผู้วิจัยคาดไว้ว่าแบบอักษรของแบบอักษรจะมีผลต่ออัตราการยืนยันความเป็นมนุษย์ โดยสอดคล้องกับงานวิจัยในอดีตในบริบทของความยากง่ายในการอ่านของมนุษย์ คือมนุษย์อ่านอักษรตัวพิมพ์ได้ง่ายกว่าอักษรที่เขียนด้วยลายมือมนุษย์ [7]

5.3 อัตราการยืนยันความเป็นมนุษย์ต่อชุดอักขระ

ผลการทดสอบสมมติฐานพบว่า อัตราการยืนยันความเป็นมนุษย์ของแบบอักษรที่แสดงชุดอักขระเป็นอักษรตัวเล็กทั้งหมดสูงกว่าที่แสดงเป็นอักษรตัวใหญ่ทั้งหมดอย่างมีนัยสำคัญทางสถิติที่ 0.05 โดยผลการทดสอบดังกล่าวเป็นไปตามที่ผู้วิจัยคาดไว้ว่า ชุดอักขระของแบบอักษรข้อความจะมีผลต่ออัตราการยืนยันความเป็นมนุษย์ โดยสอดคล้องกับ Tinker [15] ที่กล่าวว่าตัวพิมพ์เล็กเป็นอักษรที่มนุษย์อ่านง่ายและเร็วกว่าตัวพิมพ์ใหญ่ และยังสอดคล้องกับงานวิจัยของ Paterson และ Tinker [14] พบว่าบุคคลสามารถเข้าใจตัวพิมพ์เล็กได้ง่ายและเร็วกว่าตัวพิมพ์ใหญ่

5.4 การนำงานวิจัยไปประยุกต์ใช้

การนำงานวิจัยนี้ไปใช้ในเชิงทฤษฎี กล่าวคือ ข้อค้นพบในการศึกษานี้ ช่วยต่อยอดองค์ความรู้สำหรับการพัฒนาแบบอักษรต่อไป และช่วยยืนยันเชิงประจักษ์ให้ชัดเจนมากขึ้น อีกทั้งผู้ประกอบการทางเว็บไซต์ สามารถเป็นนำข้อมูลจากการศึกษานี้ไปเป็นแนวทางสำหรับการเลือกแบบอักษรข้อความให้กับเว็บไซต์ของผู้ประกอบการเอง เพื่อให้เหมาะสมต่อการใช้งานของมนุษย์ นอกจากนี้ยังสามารถนำไปใช้ในเชิงประยุกต์ได้ด้วย กล่าวคือนักพัฒนาแบบอักษรข้อความ สามารถนำข้อค้นพบจากการศึกษานี้ไปใช้เป็นแนวทางในการพัฒนาแบบอักษรข้อความ ในประเด็นดังต่อไปนี้ (1) การหมุนเอียงอักษร (2) แบบอักษร และ (3) ชุดอักขระ ของแบบอักษรข้อความ

5.5 ข้อจำกัดของงานวิจัยและโอกาสการศึกษาต่อในอนาคต
งานวิจัยนี้เป็นงานวิจัยที่ทดลองโดยการเก็บข้อมูลในห้องปฏิบัติการ โดยควบคุมปัจจัยทางสภาพแวดล้อมให้คงที่

ได้แก่ สถานที่ อุปกรณ์คอมพิวเตอร์ที่ใช้งาน และความเร็วการเข้าถึงระบบอินเทอร์เน็ต ทำให้สภาวะดังกล่าวแตกต่างกับการได้ตอบแบบอักษรข้อความแบบสถานการณ์จริง อีกทั้งสถานการณ์จริงสำหรับการได้ตอบกับแบบอักษรนั้น ผู้ใช้งานเว็บไซต์จะมีความสนใจในการได้ตอบแบบอักษรข้อความเพื่อให้เข้าใช้งานเว็บไซต์ได้ ขณะที่หน่วยทดลองอาจไม่มีความสนใจที่จะได้ตอบกับแบบอักษรเพื่อเข้าใช้งานเว็บไซต์ที่ใช้ในการศึกษาครั้งนี้ แต่เป็นเพียงการดาวน์โหลดเพลงที่ผู้วิจัยกำหนดไว้ ทำให้ผลที่ได้จากการทดลองอาจมีผลแตกต่างกับการใช้งานจริง อีกทั้งการเลือกหน่วยทดลองเพื่อเก็บข้อมูลของงานวิจัยนี้ ผู้วิจัยเลือกจากผู้ที่อยู่ในกลุ่มผู้ใช้อินเทอร์เน็ตส่วนใหญ่ในรอบสิบปี ดังนั้นการนำผลสรุปไปใช้ให้ค่านึงด้วยว่าในโลกอินเทอร์เน็ตจริงๆ นั้นประกอบด้วยผู้ใช้หลากหลายกลุ่มไม่เฉพาะที่เป็นนิสิต นักศึกษา และจากการทดลองครั้งนี้ หน่วยทดลองส่วนใหญ่มีสายตาสั้น โดยอาจส่งผลต่อความยากง่ายในการอ่านข้อความของแต่ละหน่วยทดลอง [17] จึงอาจทำให้ผลสรุปที่แตกต่างออกไปจากงานวิจัยนี้

ดังนั้น เพื่อเป็นข้อเสนอแนะในการศึกษาต่อไปในอนาคต อาจพิจารณาถึงตัวแปรอื่นเพิ่มเติม เช่น การนำโปรแกรมคอมพิวเตอร์อัตโนมัติมาทดสอบเพื่อเปรียบเทียบกับการได้ตอบแบบอักษรข้อความของมนุษย์ หรือให้ความสนใจในประเด็นรายอักขระว่าอักษรตัวไหนในแบบอักษรข้อความที่มนุษย์สามารถได้ตอบได้อย่างถูกต้อง หรือไม่ถูกต้อง โดยอาจควบคู่กับการใช้เทคนิคการติดตามการมองเห็น (Eye-Tracking Technique) ไปด้วย เพื่อดูว่าหน่วยตัวอย่างมองที่ตำแหน่งอักษรตัวไหนในแบบอักษรข้อความใช้เวลานาน หรือใช้เวลาไม่นาน ยังส่งผลให้ทราบว่าตัวอักษรไหนอ่านยากหรือง่าย นอกจากนี้ยังสามารถที่จะลองปรับเปลี่ยนการทดลองโดยเก็บข้อมูลในโลกอินเทอร์เน็ตจริง เพื่อผู้วิจัยจะได้เก็บข้อมูลที่เกิดขึ้นในสถานการณ์จริง และผลสรุปที่ได้จะใกล้เคียงกับความเป็นจริงมากที่สุด

เอกสารอ้างอิง

- [1] Ahn, L. V., Blum, M., and Langford, J. "Telling Humans and Computers Apart (Automatically) or How Lazy Cryptographers do AI," Communications of the ACM 47(2) (2004): 57-60.

- [2] Arditi, A., and Cho, J. "Letter case and text legibility in normal and low vision," *Vision Research* 47 (2007): 2499–2505.
- [3] Chellapilla, K., Larson, K., Simard, P., and Czerwinski, M. "Designing Human Friendly Human Interaction Proofs (HIPs)," *Conference on Human factors in computing systems* (2005): 711-720.
- [4] Dafont.com. "Script > School" <http://www.dafont.com>, 2012
- [5] Kolupave, A., and Ogijenko, J. "CAPTCHAs: Humans vs bots," *Security and Privacy IEEE* (2008): 68-70.
- [6] Imsamai, M., and Phimoltares, S. "3D CAPTCHA A Next Generation of the CAPTCHA," Master's Thesis, Department of Mathematics Faculty of Science Chulalongkorn University, 2010.
- [7] Quant, L. "Factors Affecting the Legibility of Hand Writing," *Journal of Experimental Education* 14 (1964) : 297-376.
- [8] Rusu, A., and Govindaraju, V. "Handwritten CAPTCHA: Using the difference in the abilities of humans and machines in reading handwritten words," *Proceedings of the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004)*, IEEE Computer Society, 2004.
- [9] Rusu, A., and Govindaraju, V. "A Human Interactive Proof Algorithm Using Handwriting Recognition," *Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR'05)*, IEEE Computer Society, 2005.
- [10] Shirali-Shahreza, M. H., and Shirali-Shahreza, M. "Multilingual CAPTCHA," *Proceeding of 5th IEEE International Conference on Computational Cybernetics*, 2007.
- [11] Simion, F., Bagnara, S., Bisiacchi, P., Roncato, S., and Umiltà, C. "Laterality Effects, Levels of Processing, and Stimulus Properties," *Journal of Experimental Psychology: Human Perception and Performance* 6(1) (1980): 184-195.
- [12] Skype. "Create an account" . [https://login.skype.com/account/ signup-form](https://login.skype.com/account/signup-form) [2012, January 20].
- [13] Tangmanee, C., and Sujarit - Apirak, P. "Attitude, Cognition, and awareness of Thai internet users and Thai E-commerce entrepreneurs towards CAPTCHA," Master's Project, Department of Commerce Faculty of Commerce and Accountancy Chulalongkorn University, 2551.
- [14] Tinker, M. A. and Paterson, D. G. "Readability of newspaper headlines printed in capitals and lower case," *Journal of Applied Psychology* (1946) : 161-168.
- [15] Tinker, M.A. *Legibility of print*. Ames, Iowa: Iowa State University Press, 1963
- [16] Yan, J., and Ahmad, A. S. E. "Usability of CAPTCHAs Or usability issues in CAPTCHA design," *Symposium On Usable Privacy and Security (SOUPS'08)*, July 23-25, 2008, USA, pp. 44-52.
- [17] กำธร สติรกุล. "หนังสือและการพิมพ์". กรุงเทพฯ : มหาวิทยาลัยรามคำแหง, 2515.
- [18] กัลยา วาณิชย์บัญชา. "สถิติสำหรับงานวิจัย". พิมพ์ครั้งที่ 5. กรุงเทพฯ: ชรรรมสาร, 2553.
- [19] ขวัญหทัย สันติบุตร. "ผลกระทบของความล่าช้าของการแสดงผล รูปแบบของเมนู การใช้ข้อมูล ป้อนกลับต่อประสิทธิภาพของการใช้งานบนเว็บแอปพลิเคชัน". วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ, สาขาการพัฒนาระบบสารสนเทศ, วิทยาลัยการพัฒนาระบบสารสนเทศ มหาวิทยาลัย, 2549.
- [20] ศิริวรรณ เสรีรัตน์, สมชาย หิรัญภักดี, จิรศักดิ์ จิยะจันทร์, ขวดี ประภาวรรณ, ณดา จันทร์สม และ วลัยลักษณ์ อัคริธวงศ์. "การวิจัยธุรกิจ". พิมพ์ครั้งที่ 1. กรุงเทพฯ: A.N. การพิมพ์ 304, 2541
- [21] ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. "รายงานผลการสำรวจกลุ่มผู้ใช้อินเทอร์เน็ตในประเทศไทย". สำนักงานเลขานุการคณะกรรมการเทคโนโลยีสารสนเทศแห่งชาติ, 2553.