รายงานการวิจัยฉบับสมบูรณ์

การศึกษาและวิเคราะห์ประสิทธิภาพและความปลอดภัยและพัฒนาการใช้งานฟังค์ชันทางเดียวหรือข้อผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่า

The Study and Analysis of The Efficiency and Security of The Fuzzy Commitment with Value-Oriented Deviation

นายนล เปรมัษเฐียร

ได้รับทุนสนับสนุนงานวิจัยจากเงินงบประมาณแผ่นดินหรือรายได้ประจำปี

งบประมาณ 2555

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ชื่อโครงการ (ภาษาไทย) การศึกษาและวิเคราะห์ประสิทธิภาพและความปลอดภัยและพัฒนาการใช้งานฟังก์ชันทางเดียวหรือข้อผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่า

ชื่อโครงการ(ภาษาอังกฤษ) The Study and Analysis of The Efficiency and Security of The Fuzzy Commitment with Value-Oriented Deviation

แหล่งเงิน เงินงบประมาณแผ่นดิน/เงินรายได้

ประจำปีงบประมาณ 2555 **จำนวนเงินที่ได้รับการสนับสนุน** 50,000 **บาท**

ระยะเวลาทำการวิจัย 1 **ปี ตั้งแต่** ตุลาคม 2554 **ถึง** กันยายน 2555

ชื่อ-สกุล หัวหน้าโครงการ และผู้ร่วมโครงการวิจัย พร้อมระบุ หน่วยงานต้นสังกัดและ อีเมล์

ก. นล เปรมัษเฐียร หัวหน้าโครงการ คณะเทคโนโลยีสารสนเทศ สจล nol@it.kmitl.ac.th

คำสำคัญ ข้อผูกมัดแบบคลุมเครือ, การพิสูจน์ตัวตนแบบคลุมเครือ, การพิสูจน์เอกลักษณ์บุคคลแบบคลุมเครือ

(Keywords) fuzzy commitment, fuzzy authentication, fuzzy identity authentication

บทคัดย่อ

ฟังก์ชันทางเดียวแบบคลุมเครือเป็นส่วนสำคัญในระเบียบการพิสูจน์ตัวบุคคลด้วยการใช้ข้อมูลทางชีวภาพ โดยข้อมูลที่ใช้พิสูจน์สามารถมีค่าคลาดเคลื่อนได้ไม่เกินค่าที่กำหนด ความคลาดเคลื่อนที่มีได้อาจเป็นความคลาดเคลื่อนเป็นบิตหรือเป็นค่า ขึ้นกับฟังก์ชันที่ใช้ รายงานฉบับนี้ศึกษาประสิทธิการการใช้งานฟังก์ชันทางเดียวแบบคลุมเครือตัวหนึ่งที่ชื่อว่า ฟัซซีแมทชิง ซึ่งเป็นฟังก์ชันที่อนุญาตให้ความคลาดเคลื่อนเกิดขึ้นเป็นค่า และได้นำเสนอฟังก์ชันตัวใหม่ที่ใช้วิธีการมอดยูเลชันและไม่ต้องใช้การเข้ารหัสเพื่อแก้ไขความผิดพลาด และได้พัฒนาการใช้งานเบื้องต้นของฟังก์ชันดังกล่าวที่นำเสนอขึ้นใหม่

Abstract

Fuzzy commitment schemes are important components of a number of biometric authentication protocols. Inputs to the scheme can contain error up to some certain limit. The range of the error can be in the number of bits or value, depending on the scheme. This report studies the efficiency of one of such schemes, called fuzzy matching. This scheme allows an input to contain error by value. The study suggests a new scheme based on modulation and does not require the use of error correcting code. An demonstrative application is developed based on the new scheme.

กิตติกรรมประกาศ

โครงการวิจัยนี้สำเร็จเรียบร้อยได้ด้วยดี เพราะได้รับการสนับสนุนเป็นอย่างดีจากคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังและเจ้าหน้าที่ทุกท่าน โดยเฉพาะคุณวิมลลักษณ์ เทียนจิ้วที่ให้คำแนะนำเกี่ยวกับการขอทุนวิจัย การส่งรายงานและข้อมูลที่เป็นประโยชน์อื่นๆ คุณสุภา พิมพ์สวัสดิ์ และคุณกิริยา นิวิฐจรรยงค์ ที่ให้ความช่วยเหลือเกี่ยวกับการจ้างเขียนโปรแกรม การจัดซื้อวัสดุ และคุณพิจิตรา สุวรรณศรี ที่เป็นธุระด้านการเงินให้แก้โครงการ ขอขอบคุณอาจารย์พรฤดี เนติโสภากุล และอาจารย์สุภวรรณ อันนันหนับที่ให้คำแนะนำต่างๆ ที่เป็นประโยชน์ต่อโครงการ และขอขอบคุณทุกท่านๆ ที่ไม่ได้เอ่ยนาม ที่ช่วยให้งานวิจัยเรื่องนี้เสร็จสิ้นลงได้

สารบัญเรื่อง

หน้า

สารบัญตาราง

สารบัญภาพ

บทที่ 1 บทนำ

1.1 หลักการและเหตุผล

ในการทำข้อตกลงระหว่าง 2 ฝ่าย ผ่านช่องทางการสื่อสารใดๆ ความเที่ยงตรงของข้อตกลงเป็นหลักสำคัญว่าเมื่อมีการบรรลุข้อแล้ว ทั้งสองฝ่ายต้องยอมรับและจะไม่มีฝ่ายใดแก้ไขการตัดสินใจนั้นได้ ซึ่งวิธีการผูกมัด (Commitment scheme) สามารถตอบสนองความต้องการข้างต้นเพื่อป้องกันการบ่ายเบี่ยงโดยฝ่ายหนึ่งฝ่ายใดหรือจากทั้ง 2 ฝ่ายได้ ด้วยคุณสมบัติการผูกมัดเพื่อป้องกันการเปลี่ยนแปลงความลับและคุณสมบัติความปลอดภัยในการปิดบังความลับ ทำให้วิธีการผูกมัดถูกนำไปประยุกต์ใช้ในเทคนิคการเข้ารหัสต่างๆ มากมาย

ระบบตรวจพิสูจน์สิทธิ์ก็เป็นอีกเทคนิคการเข้ารหัสที่นำวิธีการผูกมัดมาใช้ เพื่อสร้างความเชื่อมั่นว่าเมื่อผู้ใช้ระบบได้กำหนดค่าใดค่าหนึ่งเพื่อเป็นรหัสลับแล้ว ค่าที่เลือกจะถูกเก็บไว้ในระบบอย่างเป็นความลับ โดยจะไม่มีผู้ใดสามารถแก้ไขความลับนั้นได้ จนกว่าเจ้าของรหัสลับจะพิสูจน์ความลับด้วยค่าที่ถูกต้องได้ก่อนจึงจะสามารถเปลี่ยนแปลงความลับได้สำหรับการใช้งานครั้งต่อไป โดยทั่วไประบบตรวจพิสูจน์สิทธิ์จะใช้วิธีการตรวจสอบคำตอบหรือรหัสผ่านของผู้อ้างสิทธิ์ว่าตรงกับรหัสลับของผู้อ้างสิทธิ์หรือไม่ ซึ่งระบบจะยอมรับคำตอบหรือรหัสผ่านที่ถูกต้องตรงกับข้อมูลลับเท่านั้น เมื่อมีการพัฒนารูปแบบของข้อมูลโดยนำข้อมูลชีวมิติ (Biometrics) ของบุคคลมาใช้งานร่วมกับระบบตรวจพิสูจน์สิทธิ์พบปัญหาที่เกิดในการใช้งาน คือข้อมูลที่ได้จากการอ่านค่ามักมีความคลาดเคลื่อนเกิดขึ้น ซึ่งความผิดพลาดที่เกิดขึ้นแม้เพียงเล็กน้อย ก็ทำให้ข้อมูลที่อ่านค่าได้ไม่ตรงกับข้อมูลลับที่เก็บไว้ ส่งผลให้ข้อมูลนั้นไม่สามารถใช้ยืนยันความเป็นเจ้าของสิทธิ์ได้

A. Jules และ M. Wattenberg ได้นำเสนอการใช้ข้อผูกมัดแบบคลุมเครือ (Fuzzy Commitment) เพื่อแก้ปัญหาที่เกิดจากความคลาดเคลื่อนของข้อมูลที่อ่านได้ โดยใช้เทคนิคการตรวจจับและแก้ไขข้อผิดพลาด (Error Correction) ช่วยให้ระบบตรวจพิสูจน์สิทธิ์ยอมรับข้อมูลที่มีความคลาดเคลื่อนเล็กน้อยได้ ด้วยการปรับข้อมูลที่มีความต่างกันเล็กน้อยให้สามารถเปรียบเทียบเข้ากันได้ ซึ่งความ คลาดเคลื่อนที่ยอมรับขึ้นอยู่กับอัตราการอนุญาตผิดพลาด (False Accept Rate) และอัตราการปฏิเสธผิดพลาด (False Reject Rate) อย่างไรก็ตามวิธีการผูกมัดแบบคลุมเครือยังมีจุดด้อยในการใช้งานอยู่ 2 ประการ คือข้อจำกัดในการใช้งานและความปลอดภัยในการใช้งาน ซึ่งอัลกอริทึมสามารถใช้ตรวจสอบข้อมูลที่มีลักษณะเป็นจุดเท่านั้น ไม่สามารถทำงานกับข้อมูลที่เป็นพื้นที่หรือรูปภาพได้ นอกจากนี้อัลกอริทึมยังสามารถถูกโจมตีได้ด้วยการใช้ข้อมูลในชุดรหัสสำหรับแก้ไขข้อผิดพลาด (Codeword) ทั้งหมดที่เป็นไปได้ทดสอบเพื่อหาข้อมูลที่เป็นความลับ

ต่อมา Ojha และ Sharm ได้นำเสนอวิธีการผูกมัดแบบคลุมเครืออีกรูปแบบหนึ่งซึ่งประยุกต์มาจากการเข้ารหัสแบบอสมมาตร (Asymmetric) วิธีการหนึ่งที่เรียกว่า McEliece's Cipher โดยการ

เข้ารหัสประเภทนี้ตั้งอยู่บนพื้นฐานของเทคนิคการตรวจจับและแก้ไขข้อผิดพลาดอยู่แล้ว วิธีการนี้จึงใช้ประโยชน์จากความสามารถในการแก้ไขความคลาดเคลื่อนของรหัสที่มีอยู่ใน McEliece's Cipher ให้เกิดประโยชน์ ซึ่งวิธีการผูกมัดแบบคลุมเครือทั้ง 2 วิธีข้างต้นจะการตรวจจับและแก้ไขข้อผิดพลาดโดยดูจากความผิดเพี้ยนของบิตของข้อมูล ความผิดพลาดที่เกิดขึ้นจะต้องไม่เกินจำนวนบิตที่ยอมรับได้ เนื่องจากจุดประสงค์การใช้งานเทคนิคการแก้ไขความคลาดเคลื่อนรูปแบบนี้เพื่อให้สามารถส่งผ่านข้อมูลไปยังปลายทางโดยผ่านช่องทางที่มีสัญญาณรบกวนได้ ซึ่งสัญญาณรบกวนมักส่งผลให้บางบิตของข้อมูลผิดเพี้ยนหรือขาดหายไป

ในปี 2552 อรรคพล เสถียนจารุรัตน์และผู้วิจัยได้นำเสนอวิธีการผูกมัดแบบคลุมเครือที่สามารถแก้ไขความคลาดเคลื่อนของค่าของข้อมูล เนื่องจากข้อมูลบางลักษณะมีความผิดพลาดที่ค่าของข้อมูล ข้อมูลที่มีค่าคลาดเคลื่อนเพียงเล็กน้อยอาจทำให้จำนวนบิตในรูปฐานสองแตกต่างจากเดิมไปได้มาก เพราะบิตแต่ละตำแหน่งในระบบเลขฐานสองมีความสำคัญไม่เท่ากัน เช่น 7 และ 8 ซึ่งมีรูปฐานสองเป็น 0111 และ 1000 ตามลำดับ จึงมีการปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบเลขฐานหนึ่งเพื่อให้ความคลาดเคลื่อนของค่ากับความคลาดเคลื่อนของจำนวนบิตสัมพันธ์กัน แต่การเก็บข้อมูลในรูปของเลขฐานหนึ่งโดยตรงนั้นต้องใช้พื้นที่ในการเก็บมากและเวลาที่ใช้ในการประมวลผลเพื่อทดสอบข้อมูลที่อาจเป็นคำตอบได้นั้นอยู่ในรูปของฟังค์ชันพหุนามทำให้ระบบมีความปลอดภัยต่ำ จึงแก้ไขโดยการมอดุโล (Modulo) ข้อมูลลับด้วยค่ามอดุลัส (Modulus) จำนวนหนึ่งที่มีความเป็นจำนวนสัมพัทธ์ซึ่งกันและกัน แล้วเก็บค่าสมภาค (Congruence) ทั้งหมดไว้ในรูปของเลขฐานหนึ่ง เมื่อต้องการเปรียบ เทียบข้อมูลสามารถทำได้ด้วยการใช้ทฤษฎีบทเศษเหลือแบบจีน (Chinese Remainder Theorem) ในการคำนวณค่าตั้งต้นกลับคืนมา

แม้จะสามารถเก็บผลลัพธ์ในลักษณะฟังค์ชันทางเดียวและสามารถตรวจสอบข้อมูลโดยอนุญาตให้ความผิดพลาดเกิดขึ้นได้ตามค่าที่ยอมรับ วิธีการดังกล่าวยังมีข้อบกพร่องอยู่ กล่าวคือการใช้งานวิธีการผูกมัดแบบคลุมเครือดังกล่าวจำเป็นต้องเลือกเลขจำนวนต่างๆ ที่เป็นจำนวนเฉพาะสัมพัทธ์ซึ่งกันและกัน และนำจำนวนเหล่านั้นไปมอดุโลข้อมูลลับ ค่าสมภาคที่ได้ไม่ควรมีค่าใกล้เคียงกับข้อมูลลับและต้องไม่อยู่ใกล้ค่าศูนย์ด้วย เงื่อนไขดังกล่าวทำให้ความลับของข้อมูลที่เก็บนั้นมีความปลอดภัยน้อยลงจากการตัดค่าต่างๆ ที่เป็นไปไม่ได้เนื่องจากเงื่อนไขที่ระบุไว้ข้างต้นนั่นเอง ผู้วิจัยจึงมีแนวคิดที่จะทดสอบการใช้งานวิธีการผูกมัดแบบคลุมเครือด้วยทฤษฎีบทเศษเหลือของจีนในรูปแบบอื่น เช่น การไม่สนใจว่าค่าสมภาคที่ได้จะเป็นเท่าใด โดยยอมให้มีความผิดพลาดเกิดขึ้นได้แต่ไม่เกินค่าที่กำหนด หรือการยอมให้ค่าสมภาคตกอยู่ในขอบเขตที่มีปัญหาได้ไม่เกินจำนวนที่กำหนดแล้วใช้วิธีการตรวจสอบค่าความผิดพลาดส่วนใหญ่ที่เกิดขึ้นเพื่อใช้แก้ไขข้อมูลและตรวจสอบความถูกต้อง หรือการใช้วิธีการผูกมัดแบบคลุมเครือดังกล่าวโดยเก็บค่าสมภาคเพียงค่าเดียวเพื่อใช้ในการปรับปรุงค่าที่คลาดเคลื่อนของข้อมูลแทนการเก็บค่าสมภาคหลายค่า

1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาและวิเคราะห์ประสิทธิภาพและความปลอดภัยของการใช้งานวิธีการผูกมัดแบบคลุมเครือที่สามารถตรวจสอบและแก้ไขความคลาดเคลื่อนตามค่าได้ และค้นหารูปแบบการใช้งานวิธีการดังกล่าวที่เหมาะสมที่สุดสำหรับสถานการณ์ต่างๆ รวมทั้งการวิเคราะห์วิธีการเข้ารหัสเพื่อแก้ไขความคลาดเคลื่อนที่เหมาะสมต่อการประยุกต์ใช้วิธีการดังกล่าว

1.3 ขอบเขตการวิจัย

ก. หาประสิทธิภาพและความปลอดภัยของการใช้งานฟังค์ชันทางเดียวที่สามารถรับข้อมูลที่มีความคลาดเคลื่อนโดยค่าได้

- แบบไม่ยอมให้มีความผิดพลาดจากค่าสมภาค
- แบบยอมให้มีความผิดพลาดจากค่าสมภาคได้
- แบบที่เก็บค่าสมภาคเดียวแล้วไม่ใช้ทฤษฎีบทเศษเหลือแบบจีน

ข. วิเคราะห์ความเหมาะสมของการเข้ารหัสเพื่อแก้ไขความคลาดเคลื่อนแบบต่างๆ ต่อการใช้งานวิธีการผูกมัดแบบคลุมเครือที่ใช้ทฤษฎีบทเศษเหลือแบบจีน

1.4 แนวทางการวิจัย

แผนการดำเนินงานวิจัยมีดังนี้

1. ศึกษาวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่าและเทคนิคต่างๆ ที่เกี่ยวข้อง
2. กำหนดรูปแบบการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่าที่ต้องการจะศึกษาและวิเคราะห์
3. ศึกษาประสิทธิภาพและความปลอดภัยจากการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่าตามรูปแบบที่กำหนด
4. วิเคราะห์ผลที่ได้จากการศึกษาประสิทธิภาพและความปลอดภัยจากการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่าตามรูปแบบที่กำหนด
5. วิเคราะห์วิธีการเข้ารหัสเพื่อแก้ไขความคลาดเคลื่อนที่เหมาะสมต่อรูปแบบการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่า
6. สรุปผลการวิจัย

1.5 ประโยชน์ที่คาดว่าจะได้รับ

เพื่อพัฒนารูปแบบการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่า

บทที่ 2 การดำเนินงานวิจัย

1.1 ศึกษางานวิจัยที่เกี่ยวข้อง

ก. วิธีการผูกมัดแบบคลุมเครือ

วิธีการนี้ถูกนำเสนอโดย Juels และ Wattenberg เป็นครั้งแรกที่มีการนำเทคนิคการตรวจสอบและแก้ไขข้อผิดพลาดมาประยุกต์ใช้กับวิธีการผูกมัดเพื่อลดข้อจำกัดในการนำไปใช้กับข้อมูลประเภทชีวมิติซึ่งมีความคลาดเคลื่อนเกิดขึ้นได้ โดยหากข้อมูลที่ใช้ในการตรวจพิสูจน์มีบิตคลาดเคลื่อนจากข้อมูลลับไม่มากระบบจะทำการตรวจสอบและแก้ไขบิตที่ผิดพลาดให้ถูกต้องตรงกับข้อมูลลับได้ แต่หากมีบิตที่คลาดเคลื่อนมากเกินกว่าจำนวนที่กำหนด ระบบจะแปลงข้อมูลดังกล่าวจนแตกต่างจากข้อมูลลับโดยสิ้นเชิง ซึ่งวิธีการมีดังนี้

ขั้นตอนการลงทะเบียน

1. เมื่อระบบได้รับข้อมูลลับ s มาแล้ว จะกำหนดชุดรหัสข้อมูล C ซึ่งสมาชิกภายในเป็นรหัสข้อมูลเพื่อใช้สำหรับการตรวจสอบและแก้ไขข้อผิดพลาดได้ t บิต
2. สุ่มเลือกรหัสข้อมูล w จากชุดรหัสข้อมูล C
3. นำข้อมูลลับ s รวมกับของรหัสข้อมูล w โดยการ exclusive-or ได้ผลลัพธ์เป็น $s \oplus w$
4. หาค่าแฮช (Hash) ของรหัสข้อมูล w ได้ผลลัพธ์คือ $H(w)$
5. กำจัดค่า w แล้วเก็บไว้เพียงค่า $s \oplus w$ และ $H(w)$

ขั้นตอนการตรวจพิสูจน์สิทธิ์

1. ผู้ที่ต้องการตรวจพิสูจน์สิทธิ์ป้อนข้อมูล s' เข้าไปในระบบ แล้วระบบจะทำการคำนวณหาค่า $s' \oplus s \oplus w$ หากข้อมูล s' แตกต่างจากข้อมูลลับ s ไม่เกิน t บิต ระบบจะสามารถตรวจสอบและแก้ไขข้อผิดพลาดแล้วคืนผลลัพธ์จากการทำ $s' \oplus s \oplus w$ เป็นรหัสข้อมูล w กลับมา
2. นำค่า w ที่ได้ไปคำนวณหาค่าแฮชแล้วเปรียบเทียบกับค่า $H(w)$ ที่เก็บไว้ หากค่าแฮชทั้งสองถูกต้องตรงกันนั่นคือสามารถตรวจพิสูจน์สิทธิ์ได้สำเร็จ

ข. วิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่า

ในการใช้งานวิธีการผูกมัดแบบคลุมเครือบางกรณีข้อมูลมีความผิดพลาดที่ค่าของข้อมูล ข้อมูลที่มีบิตคลาดเคลื่อนเพียงเล็กน้อยอาจส่งผลให้ค่าของข้อมูลในรูปเลขฐานสองมีค่าแตกต่างจากเดิมไปมาก เช่น 00011111 และ 10011111 มีบิตที่แตกต่างกันเพียงตำแหน่งเดียวแต่ค่ากลับแตกต่างกันมากถึง 128 แนวคิดวิธีการผูกมัดแบบคลุมเครือในลักษณะที่สามารถรองรับข้อมูลที่มีค่าคลาดเคลื่อนได้ โดยใช้ประโยชน์จาก

รูปเลขฐานหนึ่ง ช่วยให้ค่าที่คลาดเคลื่อนกับความคลาดเคลื่อนของบิตสัมพันธ์กัน แต่การใช้งานข้อมูลที่เป็นเลขฐานหนึ่งมีจุดด้อยที่ง่ายต่อการโจมตี จึงใช้แนวคิดการแบ่งความลับและการนำทฤษฎีบทเศษเหลือแบบจีนมาประยุกต์ใช้เพื่อเพิ่มความปลอดภัยให้กับวิธีการ ซึ่งวิธีการมีดังนี้

<u>ขั้นตอนการลงทะเบียน</u>

1. สำหรับข้อมูลลับ s ให้เลือกค่า $n_1, n_2, \ldots, n_k$ ทั้งหมดจำนวน k ค่า โดยค่าที่เลือกทั้งหมดต้องเป็นจำนวนเฉพาะสัมพัทธ์ซึ่งกันและกัน นอกจากนี้ค่าทั้งหมดต้องมีค่าน้อยกว่าค่าของข้อมูลลับ เพราะหากนำ n_i ที่มีค่ามากกว่าข้อมูลลับมามอดุโลข้อมูลลับ s แล้วผลลัพธ์ที่ได้ก็คือค่า s ของข้อมูลลับ ซึ่งจะส่งผลต่อความปลอดภัยของข้อมูลลับเอง
2. คำนวณหาค่าสมภาค $r_i = s \bmod n_i$ สำหรับ $1 \le i \le k$ ทั้งหมด ซึ่งค่า r_i ที่ได้จะถูกเก็บอยู่ในรูปของเลขฐานหนึ่งและค่า r_i ต้องมีค่ามากกว่าค่าความคลาดเคลื่อนที่ยอมรับได้ e และน้อยกว่าค่ามอดุลัสลบด้วยค่าความคลาดเคลื่อนที่ยอมรับได้ $e \le r_i \le n_i - e$ หากค่า r_i ที่ได้ผิดไปจากเงื่อนไขนี้จะส่งผลให้การตรวจสอบและแก้ไขข้อผิดพลาดไม่สามารถคืนค่าที่ถูกต้องได้
3. สำหรับทุกค่า r_i ให้สุ่มเลือกรหัสข้อมูล w_i จากชุดรหัสข้อมูล C เพื่อใช้สำหรับการตรวจสอบและแก้ไขข้อผิดพลาดได้ t บิตของเลขฐานหนึ่ง
4. นำค่า r_i รวมกับค่าของรหัสข้อมูล w_i โดยการ exclusive-or ได้ผลลัพธ์เป็น $r_i \oplus w_i$
5. แทนการเก็บค่า $H(w_i)$ ระบบจะทำลายค่า w_i แต่เก็บค่า $r_i \oplus w_i$ ทั้งหมดและเก็บ $H(s)$ ซึ่งค่า s ที่ใช้ในการคำนวณอยู่ในรูปเลขฐานสอง

<u>ขั้นตอนการตรวจพิสูจน์สิทธิ์</u>

1. เมื่อผู้ที่ต้องการตรวจพิสูจน์สิทธิ์ป้อนข้อมูล s' เข้าไปในระบบ แล้วระบบจะทำการคำนวณหาค่าสมภาค $r'_i = s' \bmod n_i$ สำหรับ $1 \le i \le k$ ทั้งหมด
2. แล้วจะทำการคำนวณหาค่า $w'_i = r'_i \oplus r_i \oplus w_i$
3. ระบบจะตรวจสอบและแก้ไขข้อผิดพลาดของค่า w'_i ได้ผลลัพธ์เป็น w''_i ซึ่งหากค่าของ r'_i มีความใกล้เคียงกับค่า r_i แล้ว w''_i จะเท่ากับ w_i
4. แล้วระบบจะหาค่าของ $r''_i = w''_i \oplus r'_i \oplus w_i$
5. คำนวณค่า s'' โดยใช้ค่า r''_i และ n_i ทั้งหมด คำนวณด้วยทฤษฎีบทเศษเหลือแบบจีน
6. เปรียบเทียบค่าแฮช $H(s'')$ กับ $H(s)$ หากค่าที่ได้ถูกต้องตรงกันแสดงว่าการตรวจพิสูจน์สิทธิ์สำเร็จ

2.2 ศึกษารูปแบบการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่า

ด้วยเงื่อนไขว่าค่าสมภาคที่ได้จากการมอดุโลข้อมูลลับด้วยจำนวนเฉพาะสัมพัทธ์ ไม่ควรมีค่าใกล้เคียงกับข้อมูลลับและต้องไม่อยู่ใกล้ค่าศูนย์ด้วย เพราะค่าสมภาคลักษณะนั้นอาจส่งผลให้การคืนค่าข้อมูลลับเกิดความผิดพลาดขึ้นได้ แต่ในขณะเดียวกันก็ทำให้ความลับของข้อมูลที่เก็บนั้นมีความปลอดภัยน้อยลงจากการตัดค่าต่างๆ ที่เป็นไปไม่ได้เนื่องจากข้อจำกัดที่ระบุไว้ข้างต้นนั่นเอง ดังนั้น ผู้วิจัยจึงมีแนวคิดที่จะศึกษาการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่าในรูปแบบอื่นดังนี้

ก. แบบไม่ยอมให้มีความผิดพลาดจากค่าสมภาค

การใช้งานรูปแบบแรกยึดตามข้อจำกัดของวิธีการเป็นสำคัญ โดยจะทำการเลือกค่ามอดุลัสที่ใช้เพื่อให้ได้ค่าสมภาคตรงตามเงื่อนไขคือ ไม่ใกล้เคียงกับค่าศูนย์และไม่ใกล้เคียงค่าข้อมูลลับ ส่งผลให้โดเมนของข้อมูลลับมีขนาดลดลง ซึ่งสังเกตได้จากตัวเลขที่ใช้สำหรับการมอดุเลชั่นมีจำนวนน้อยลง

ในการศึกษานี้เราพิจารณาในกรณีที่ค่ามอดุลัสทั้งหมดที่ใช้มีขนาดใหญ่มาก จำนวนเฉพาะสัมพัทธ์ $n_1, n_2, \ldots, n_k$ มีขนาดเท่ากัน นอกจากนี้ค่าที่เลือกใช้ต้องไม่มีข้อผิดพลาดที่ส่งผลต่อขั้นตอนการคำนวณด้วยทฤษฎีบทเศษเหลือแบบจีน

สำหรับค่า n_i ใดๆสำหรับข้อมูลลับ s และชุดรหัสข้อมูล C เพื่อใช้สำหรับการตรวจสอบและแก้ไขข้อผิดพลาดได้ t บิต ค่าสมภาคที่ได้ $r_i = s \bmod n_i$ ต้องอยู่ในช่วง

$$t \le r_i \le n_i - t \qquad (1)$$

หากค่า n_j ใดที่เลือกใช้ไม่สามารถให้ผลลัพธ์ดังคุณสมบัติในข้อ (1) แสดงว่า

$$r_j < t \text{ หรือ } r_j > n_j - t - 1 \qquad (2)$$

หากผลลัพธ์ $r_j < t$ และข้อมูลลับ s มีความคลาดเคลื่อน t บิต เมื่อผ่านขั้นตอนการตรวจสอบและแก้ไขข้อผิดพลาดแล้วผลลัพธ์ r'_j ที่ได้จะถูกเปลี่ยนค่าเป็น $n_j - r_j + t$ ไม่สามารถคืนค่าที่ถูกต้องได้เช่นเดียวกับกรณีที่ผลลัพธ์ $r_j > n_j - t - 1$

ข. แบบยอมให้มีความผิดพลาดจากค่าสมภาคได้

การใช้งานรูปแบบนี้ไม่จำกัดค่าสมภาคที่ได้ ส่งผลให้เอนโทรปีของข้อมูลลับไม่ถูกลดขนาดลง ทำให้การเลือกค่า $n_1, n_2, \ldots, n_k$ ทำได้อิสระตามต้องการ เน้นเพียงค่าทั้งหมดต้องถูกต้องตามหลักการของทฤษฎีบทเศษเหลือแบบจีนคือเป็นจำนวนเฉพาะสัมพัทธ์ซึ่งกันและกัน อย่างไรก็ตามการใช้งานนอกกรอบเงื่อนไขย่อมส่งผลให้มีโอกาสเกิดข้อผิดพลาดขึ้นได้มากกว่าแบบแรก ซึ่งข้อผิดพลาดที่เกิดขึ้นคือกรณีที่ไม่สามารถคืนกลับค่าข้อมูลลับ s ได้ถูกต้อง

ค. แบบที่เก็บค่าสมภาคเพียงค่าเดียวและไม่ใช้ทฤษฎีบทเศษเหลือแบบจีน

การใช้งานรูปแบบนี้มีความเรียบง่ายกว่า 3 รูปแบบข้างต้น เพราะมีการดัดแปลงวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่า ให้ใช้ค่าสมภาคเพียงค่าเดียวโดยไม่ต้องแปลงค่าที่ได้ไปอยู่ในรูปของเลขฐานหนึ่ง และไม่มีการนำเทคนิคการตรวจสอบและแก้ไขข้อผิดพลาดมาใช้ ส่งผลให้ลดปริมาณการคำนวณได้โดยไม่ไม่ลดความปลอดภัยของวิธีการลง และยังทำให้วิธีการนี้สามารถใช้ได้แม้กับกรณีที่มีความคลาดเคลื่อนเกิดขึ้นมากเกินกว่าที่การตรวจสอบและแก้ไขข้อผิดพลาดจะทำได้

รูปแบบการใช้งานนี้เมื่อข้อมูลลับ s ถูกป้อนเข้าสู่ระบบ ค่าความคลาดเคลื่อนสูงสุดที่เป็นไปได้ t จะถูกกำหนด ค่ามอดุลัส n จะถูกเลือกด้วยขั้นตอนดังนี้

1. คำนวณหาค่า $n = 2^{*}t$
2. คำนวณค่า $r = s \bmod n$
3. หากค่า $r < t$ หรือ $s - r < t$ ให้ $n = n + 1$ แล้วคำนวณในข้อ 2. อีกครั้ง
4. ค่า n ที่สามารถให้ผลลัพธ์ r ในช่วงที่ต้องการได้จะถูกใช้เป็นค่ามอดุลัสสำหรับข้อมูลลับ s นั้น

หลังจากที่กำหนดค่ามอดุลัส n ได้แล้วจะทำการคำนวณข้อมูลลับดังนี้

1. คำนวณค่า $r = s \bmod n$
2. หาค่าแฮชของข้อมูลลับ $h = \mathrm{H}(s)$
3. ข้อมูลลับที่จะถูกเก็บไว้ในระบบคือ (n, r, h)

ในการตรวจพิสูจน์สิทธิ์เมื่อมีการป้อนข้อมูล s' เข้าสู่ระบบ มีขั้นตอนดังนี้

1. คำนวณค่า $r' = s' \bmod n$
2. หาค่า $s'' = s' + r - r'$
3. หาค่าแฮชของข้อมูล $\mathrm{H}(s')$
4. การตรวจพิสูจน์สิทธิ์จะสำเร็จในกรณีที่ $\mathrm{H}(s') = h$ เท่านั้น

บทที่ 3 วิเคราะห์ผลการวิจัย

จากการศึกษาการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่าในรูปแบบต่างๆ ได้พบประโยชน์และข้อจำกัดในการใช้งานต่างกันดังนี้

3.1 แบบไม่ยอมให้มีความผิดพลาดจากค่าสมภาค

จากการศึกษารูปแบบการใช้งานวิธีนี้พบว่า การกำหนดให้ค่า r_i มีคุณสมบัติ $t \le r_i \le n_i - t$ ส่งผลให้ขอบเขตความเป็นไปได้ของข้อมูลลับถูกจำกัดแคบลงจาก n เหลือ $(n_1\text{-}2t)\cdot(n_2\text{-}2t)\cdot\ldots(n_k\text{-}2t)$ หากค่าที่มากที่สุดของข้อมูลลับเท่ากับ n สำหรับความสามารถในการตรวจสอบและแก้ไขข้อผิดพลาดขนาด t บิต แล้วขนาดของมอดุลัสที่ใหญ่ที่สุดและส่งผลต่อการลดขอบเขตความเป็นไปได้ของข้อมูลลับน้อยที่สุดพบว่ามีจำนวนไม่มาก ซึ่งมอดุลัสที่มีขนาดใหญ่มักต้องใช้พื้นที่ในการจัดเก็บค่า $r_i \oplus w_i$ มากด้วยเพราะค่าที่เก็บอยู่ในรูปของเลขฐานหนึ่ง จากกาตรวจวัดจำนวนบิตของเลขฐานหนึ่งที่ถูกจัดเก็บจากการคำนวณข้อมูลลับขนาด n ด้วยมอดุลัสจำนวน k ค่า ซึ่งสามารถตรวจสอบและแก้ไขข้อผิดพลาดได้ t บิต พบว่าขนาดพื้นที่ที่ต้องการสำหรับการเก็บข้อมูลเลขฐานหนึ่งเท่ากับ

$$\text{จำนวนบิตของเลขฐานหนึ่ง} = k\cdot 2^{(n/k)} \qquad (3)$$

และขอบเขตความเป็นไปได้ของข้อมูลลับที่ถูกลดขนาดลงสามารถคำนวณได้จาก

$$\text{เอนโทรปีของข้อมูลลับที่ลดลง} = \log_2(2^{(n/k)}\text{-}2t)^k \qquad (4)$$

จากการคำนวณการใช้งานด้วยรูปแบบนี้โดยทดสอบกับข้อมูลลับที่มีขนาด 256 บิตและ 512 บิต ด้วยความสามารถในการตรวจสอบและแก้ไขข้อผิดพลาดได้ 50 บิต ผลที่ได้ดังแสดงในตาราง

k	จำนวนบิต	ขนาดเอนโทรปีที่เหลือ
2	6.81E+38	ประมาณ 256
4	7.38E+19	ประมาณ 256
8	3.44E+10	255.9999997
16	1048576	255.9647511
32	8192	233.132871

ตารางที่ 3.1 ขนาดเอนโทรปีที่ลดลงของข้อมูลลับขนาด 256 บิต

*จำนวนบิตอยู่ในรูปเลขฐานหนึ่ง

k	จำนวนบิต	ขนาดเอนโทรปีที่เหลือ
2	2.32E+77	ประมาณ 512
4	1.36E+39	ประมาณ 512
8	1.48E+20	ประมาณ 512
16	6.87E+10	511.9999995
32	2097152	511.9295021
64	16384	466.265742

ตารางที่ 3.2 ขนาดเอนโทรปีที่ลดลงของข้อมูลลับขนาด 512 บิต

จากตารางสามารถสรุปได้ว่าการใช้มอดุลัสจำนวนน้อยค่ากลับมีความต้องการพื้นที่ในการจัดเก็บข้อมูลที่อยู่ในรูปของเลขฐานหนึ่งมากกว่าการเลือกใช้มอดุลัสหลายค่า และการเพิ่มขึ้นของจำนวนมอดุลัสที่ใช้ส่งผลต่อการลดขนาดของเอนโทปีลง โดยสำหรับข้อมูลลับขนาด 256 บิต ที่สามารถตรวจสอบและแก้ไขข้อผิดพลาดได้ 50 บิต จำนวนมอดุลัสที่ใช้ไม่ควรมากเกินกว่า 38 ค่า เพราะจะทำให้ค่ามอดุลัสน้อยกว่า $2t$ ซึ่งจะไม่เข้าเงื่อนไขในข้อ (1) ส่วนข้อมูลลับขนาด 256 บิต ที่สามารถตรวจสอบและแก้ไขข้อผิดพลาดได้ 50 บิต จำนวนมอดุลัสที่ใช้ไม่ควรมากเกินกว่า 77 ค่า นอกจากนี้มีข้อเสนอแนะที่สำคัญอีกประการคือ ไม่ควรใช้ค่ามอดุลัสที่มีค่าใกล้ค่า t ซึ่งเป็นจำนวนบิตที่สามารถตรวจสอบและแก้ไขข้อผิดพลาดได้ เนื่องจากจะทำให้ขนาดเอนโทรปีของข้อมูลลับลดลงอย่างมากและทำให้ไม่สามารถยอมรับความปลอดภัยของวิธีการได้

3.2 แบบยอมให้มีความผิดพลาดจากค่าสมภาค

จากการศึกษาการใช้งานรูปแบบนี้พบว่า แม้ว่าค่า r_i ที่ได้จะมีค่าอยู่ในช่วง 0 ถึง $t-1$ หรือ $n-t+1$ ถึง $n-1$ แต่ไม่ได้หมายความว่าการคืนกลับค่าข้อมูลลับ s จะไม่สามารถทำได้ ทั้งนี้ขึ้นอยู่กลับค่าความคลาดเคลื่อน t บิตที่เกิดขึ้น บางกรณีค่าความคลาดเคลื่อน t บิตไม่ส่งผลต่อการคืนค่าข้อมูลลับแม้ว่า r_i จะมีค่าอยู่ในช่วง 0 ถึง $t-1$ หรือ $n-t+1$ ถึง $n-1$ ก็ตาม ทั้งนี้เนื่องจากการใช้ค่ามอดุลัสจำนวนหลายค่า แม้ว่าอาจจะมีบางค่าไม่สามารถให้ผลลัพธ์ที่ถูกต้องได้ แต่หากผลจากการคำนวณส่วนใหญ่มีค่าอยู่ในช่วงที่กำหนด ผลลัพธ์ที่ได้ก็จะสามารถใช้ในการคืนค่าข้อมูลลับได้ อาจกล่าวได้ว่าการใช้งานวิธีการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่ารูปแบบนี้มีความปลอดภัยตราบใดที่ผลส่วนใหญ่ที่ได้จากการคำนวณโดยใช้มอดุลัสหลายค่ามีค่าอยู่ในช่วงที่กำหนด คือ อยู่ในช่วง t ถึง $n-t$

สำหรับข้อมูลลับขนาด n บิต กำหนดจำนวนบิตที่สามารถตรวจสอบและแก้ไขข้อผิดพลาดได้ t บิตโดยใช้มอดุลัสจำนวน k ค่า สามารถคำนวณหาความน่าจะเป็นที่จะเกิดความผิดพลาดจากการใช้งานในรูปแบบนี้ได้ดังนี้

$$p(\text{error}) = 2(\sum_{i=1 \text{ to } k/2}(C(k,k\text{-}1)(2^{(n/k)}\text{-}t)^i t^{(k-i)}/2^n) \quad (5)$$

ซึ่ง C(n, r) เป็นฟังก์ชัน Combinatorial Coefficient โดยคำนวณได้จาก n!/((n-r)!r!)

การทดสอบความเป็นไปได้ของการเกิดความผิดพลาดจากการใช้งานกับข้อมูลลับขนาด 256 บิตและ 512 บิต โดยมีความสามารถในการตรวจสอบและแก้ไขข้อผิดพลาด t= 100, 1000 และ 10000 ตามลำดับเป็นดังนี้

จำนวนมอดุลัสที่ใช้	ข้อมูลลับขนาด 256 บิต	ข้อมูลลับขนาด 512 บิต
2	1.18E-36	3.45E-75
4	3.53E-34	1.04E-72
8	4.11E-29	1.21E-67
16	5.81E-23	1.73E-61
32	1.82E-10	3.38E-45

ตารางที่ 3.3 ความน่าจะเป็นของการเกิดความผิดพลาดในการใช้งาน โดย t = 100

จำนวนมอดุลัสที่ใช้	ข้อมูลลับขนาด 256 บิต	ข้อมูลลับขนาด 512 บิต
2	1.18E-35	3.45E-74
4	3.53E-32	1.04E-70
8	4.11E-25	1.21E-63
16	5.81E-15	1.73E-53
32	n/a	2.76E-29

ตารางที่ 3.4 ความน่าจะเป็นของการเกิดความผิดพลาดในการใช้งาน โดย t = 1000

จำนวนมอดุลัสที่ใช้	ข้อมูลลับขนาด 256 บิต	ข้อมูลลับขนาด 512 บิต
2	1.18E-34	3.45E-73
4	3.53E-30	1.04E-68
8	4.11E-21	1.21E-59
16	1.83E-07	1.73E-45
32	n/a	2.49E-14

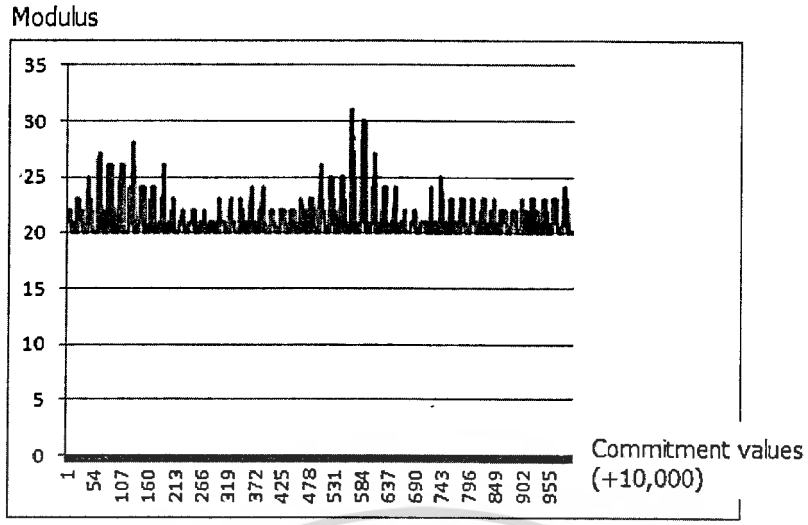
ตารางที่ 3.5 ความน่าจะเป็นของการเกิดความผิดพลาดในการใช้งาน โดย t = 10000

จากตารางพบว่าเมื่อจำนวนบิตที่สามารถตรวจสอบและแก้ไขข้อผิดพลาดได้มีขนาดมากขึ้นอัตราการเกิดความผิดพลาดมีเพิ่มมากขึ้นตามไปด้วย โดยความเป็นไปได้ที่จะเกิดความผิดพลาดเมื่อใช้วิธีการนี้กับข้อมูลลับขนาด 256 บิต จำนวนมอดุลัส 32 ค่า ที่สามารถตรวจสอบและแก้ไขข้อผิดพลาดได้ 1000 บิตและ10000 บิต จะไม่สามารถคำนวณได้เนื่องจากค่า $2^{256/32}$ เล็กกว่าค่า 1000 และ 10000 มาก ค่าที่แสดงเป็นขอบเขตค่าความคลาดเคลื่อนที่อาจเกิดขึ้น แม้ว่าค่ามอดุลัสที่เลือกใช้บางค่าจะให้ผลลัพธ์ที่มีปัญหาแต่ก็ไม่ได้หมายความว่าการตรวจพิสูจน์จะไม่สามารถทำได้ เช่นกรณีที่ข้อมูลที่ต้องการตรวจพิสูจน์มีบิตที่คลาดเคลื่อนน้อยกว่า t บิต หรือกรณีที่บิตที่คลาดเคลื่อนไม่มากพอที่จะส่งผลเสียต่อการตรวจพิสูจน์ได้ นัยยะสำคัญของการใช้การผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่ารูปแบบนี้เพื่อลดความจำเป็นของการสร้างชุดมอดุลัสใหม่ทุกครั้งสำหรับข้อมูลลับแต่ละค่า

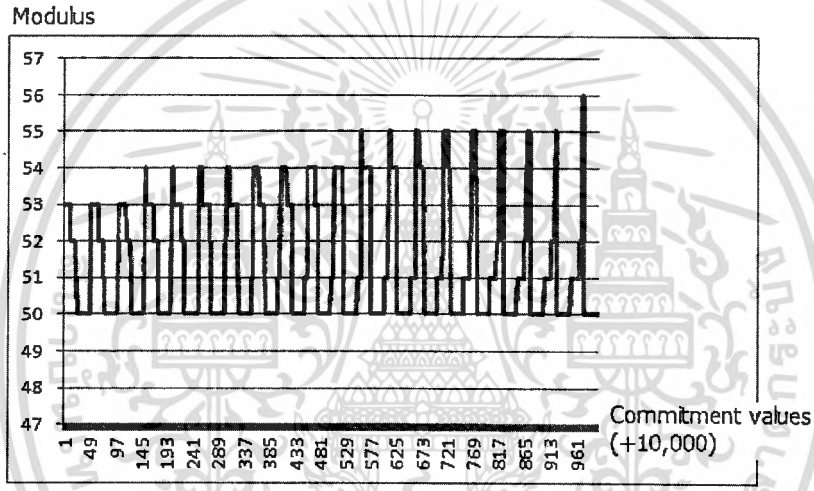
3.3 แบบที่เก็บค่าสมภาคเพียงค่าเดียวและไม่ใช้ทฤษฎีบทเศษเหลือแบบจีน

ข้อดีที่พบจากการศึกษาการใช้งานรูปแบบนี้ซึ่งไม่มีการใช้เทคนิคตรวจสอบและแก้ไขข้อผิดพลาด ทำให้สามารถใช้งานได้กับข้อมูลที่มีข้อผิดพลาดได้ไม่จำกัด แม้จะไม่มีการใช้เทคนิคนี้แต่ในการทำงานยังคงไว้ซึ่งขั้นตอนของการคำนวณสำหรับความผิดพลาดที่อาจจะเกิดขึ้น ทำให้วิธีการนี้ยังคงลักษณะของการผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนเป็นค่าได้โดย เมื่อรับข้อมูลที่ต้อง การพิสูจน์เข้ามาในระบบแล้วในเบื้องต้นจะทำการคำนวณหาค่าสมภาค $r' = s' \bmod n$ และผลพลอยได้จากการใช้ค่ามอดุลัสเดียวทำให้ลดการใช้พื้นที่จัดเก็บข้อมูลและลดขั้นตอนการทำงานได้ แม้ว่าผลที่ได้จากการคำนวณจะให้ข้อมูลบางอย่างแก่ผู้โจมตีระบบแต่ถือได้ว่าเป็นเรื่องปกติของวิธีการผูกมัดแบบคลุมเครือ

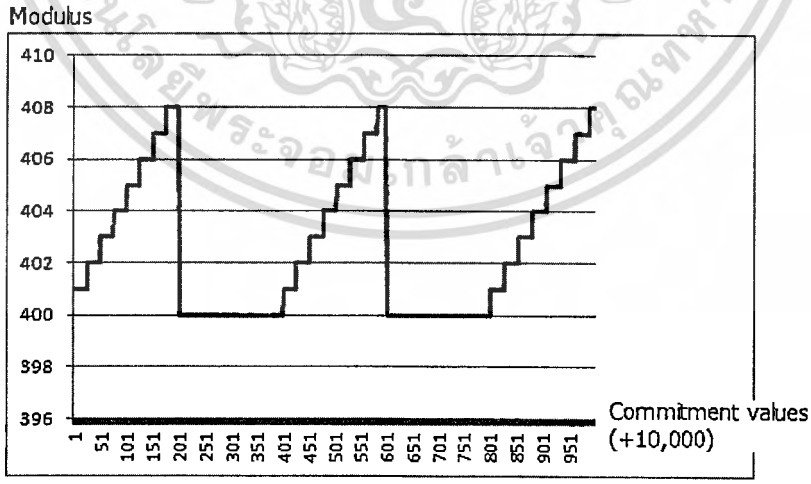
จากการตรวจวัดความซับซ้อน (Complexity) ของการหาค่ามอดุลัส พบว่าการเพิ่มขนาดของข้อมูลลับหรือการเพิ่มค่าคลาดเคลื่อนมากขึ้นจะส่งผลให้การหาค่ามอดุลัสทำได้เร็วยิ่งขึ้น จากการทดสอบกับข้อมูลลับที่มีค่าตั้งแต่ 10,000 ถึง 11,000 โดยมีค่าความผิดพลาดที่ 10, 25 และ200 ได้ผลดังแสดงในแผนภาพด้านล่างตามลำดับ



แผนภาพที่ 3.1 ค่ามอดุลัสเมื่อกำหนดค่าความคลาดเคลื่อนที่ยอมรับได้เท่ากับ 10

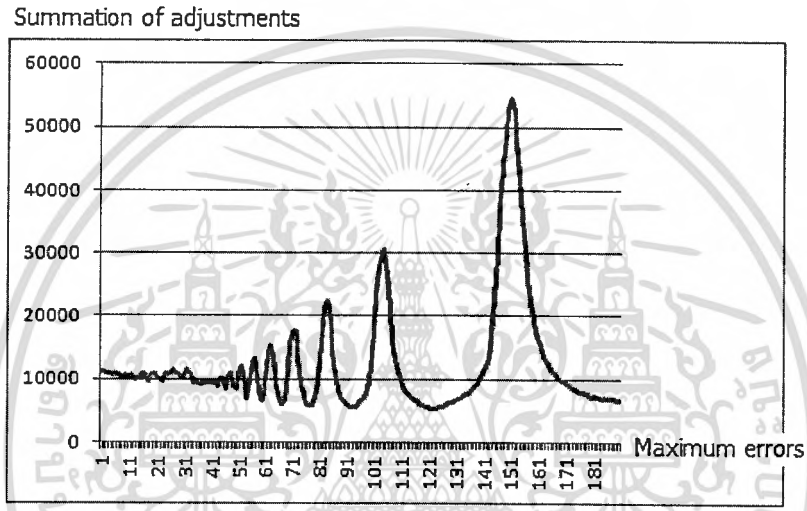


แผนภาพที่ 3.2 ค่ามอดุลัสเมื่อกำหนดค่าความคลาดเคลื่อนที่ยอมรับได้เท่ากับ 25



แผนภาพที่ 3.3 ค่ามอดุลัสเมื่อกำหนดค่าความคลาดเคลื่อนที่ยอมรับได้เท่ากับ 200

จากแผนภาพจะเห็นว่าจำนวนครั้งที่เพิ่มขึ้นในการหาค่ามอดุลัสจะมากสุดไม่เกิน 200 ครั้ง โดยอัตราเฉลี่ยสูงสุดระหว่างค่ามอดุลัสที่เลือกกับค่าตั้งต้น (ค่าคลาดเคลื่อน x 2) เท่ากับ 1.95 ในแต่ละความคลาดเคลื่อนได้เก็บผลรวมจำนวนครั้งที่ใช้ในการหาค่ามอดุลัสที่เหมาะสมสำหรับข้อมูลที่มีค่าตั้งแต่ 10,000 ถึง 11,000 พบว่ามีผลรวมจำนวนครั้งมากสุดไม่เกิน 60,000 ครั้ง หรือเฉลี่ย 60 ครั้งต่อ 1 ข้อมูลลับ กล่าวได้ว่าเมื่อมีการรับข้อมูลลับเข้ามาในระบบจะมีการทำซ้ำขั้นตอนหาค่ามอดุลัสที่เหมาะสมถึง 60 ครั้งจึงจะได้ค่ามอดุลัสที่ต้องการ โดยค่าคลาดเคลื่อนที่เพิ่มขึ้นจะส่งผลให้ต้องใช้จำนวนครั้งในการหาค่ามอดุลัสเพิ่มขึ้น ซึ่งจำนวนที่เพิ่มขึ้นมีลักษณะความซับซ้อนแบบเส้นตรง (Linear Complexity) ดังแผนภาพที่ 3.4



แผนภาพที่ 3.4 ผลรวมจำนวนครั้งในการหาค่ามอดุลัสของแต่ละค่าความคลาดเคลื่อน

บทที่ 4 สรุปผลการวิจัย

งานวิจัยนี้ได้ทำการศึกษาการใช้งานฟังก์ชันทางเดียวแบบคลุมเครือที่รับค่าความคลาดเคลื่อนเป็นค่า โดยทำการศึกษาสองส่วน ส่วนแรกเป็นการศึกษาการใช้งานฟังก์ชันทางเดียวแบบคลุมเครือที่รับค่าความคลาดเคลื่อนเป็นค่าที่มีอยู่แล้ว คือฟัชชีแมทชิง ซึ่งใช้วิธีการแบ่งค่าของตัวเลขออกเป็นส่วนๆ ด้วยวิธีการของทฤษฎีเศษเหลือแบบจีน แล้วแปลงค่านั้นเป็นเลขฐานหนึ่งเพื่อให้มีความสามารถในการแก้ไขความผิดพลาดแบบเป็นค่าได้ ซึ่งส่วนนี้ได้แยกการศึกษาออกเป็นสองส่วนย่อย คือส่วนย่อยที่หนึ่งเป็นการใช้งานที่ไม่ยอมให้เกิดความผิดพลาดขึ้นด้วยการกำหนดค่าตัวหารที่เหมาะสม ซึ่งการทำเช่นนี้จะทำให้ผู้โจมตีสามารถทราบข้อมูลบางอย่างเกี่ยวกับความลับที่เก็บไว้ในระบบ ในส่วนนี้ได้ทำการศึกษาปริมาณข้อมูลที่รั่วไหลจากการเลือกใช้วิธีการดังกล่าวว่าควรมีการใช้งานในลักษณะใดที่จะไม่ทำให้ข้อมูลรั่วไหลออกมามากเกินไปจนไม่ปลอดภัย และส่วนย่อยที่สองเป็นการใช้งานที่ยอมให้เกิดความผิดพลาดได้ ทำให้สามรถเลือกค่าตัวหารเป็นค่าใดๆ ก็ได้ ทำให้ไม่มีการรั่วไหลของข้อมูลเช่นในกรณีแรก แต่การเลือกตัวหารด้วยวิธีการดังกล่าวทำให้มีโอกาสที่ระบบจะทำงานผิดพลาดได้ การศึกษาในส่วนย่อยนี้ได้ศึกษาโอกาสที่จะเกิดความผิดพลาดหากเลือกใช้วิธีการดังกล่าว ว่าควรใช้งานในลักษณะใดที่จะไม่เกิดความผิดพลาดขึ้นมากเกินไป ในส่วนที่สองเป็นการนำเสนอฟังก์ชันทางเดียวแบบคลุมเครือที่รับค่าความคลาดเคลื่อนเป็นค่าแบบใหม่ที่ไม่ใช้การแปลงเป็นเลขฐานหนึ่งและไม่ใช้ทฤษฎีเศษเหลือแบบจีน ซึ่งทำให้มีรูปแบบการใช้งานที่ง่ายและยืดหยุ่นกว่า เนื่องจากไม่มีปัญหาการใช้งานหนึ่งสองอย่างของฟังก์ชันเดิมที่ได้กล่าวข้างต้น ไม่จำเป็นต้องใช้การเข้ารหัสเพื่อแก้ไขความผิดพลาด และได้มีการพัฒนาการใช้งานของฟังก์ชันที่นำเสนอใหม่

บรรณานุกรม

[1] [1] B. Schneier, Applied Cryptography, Wiley, 1996

[2] A. Juels and M. Wattenberg, "A Fuzzy Commitment Scheme," in Proceedings of the 6th ACM Conference on Computer and Communications Security (ACM CCS '99), pp. 28-36, November 1-4 1999, Singapore.

[3] A. Satienjarurat and N. Premasathian, "Fuzzy Matching of Objects using Fuzzy Commitment," 2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology (ECTI-CON 2009), pp. 622-625, May 6-9 2009, Chonburi, Thailand.

[4] A. A. Al-saggaf and H. S. Acharya, "A Fuzzy Commitment Scheme," in Proc. of IEEE International conference on Advances in Computer Vision and Information Technology (ACVIT), November 28-30, 2007, Aurangabad, India.

[5] D. B. Ojha and A. Sharm, "A Fuzzy Commitment Scheme with McEliece's Cipher," Surveys in Mathematics and its Applications Vol. 5, 2010, pp. 73–82.

[6] T. Ignatenko and F. M. J. Willems, "Information Leakage in Fuzzy Commitment Scheme," IEEE Transactions on Information Forensics and Security, Vol. 5, No. 2, June 2010, pp.337-348

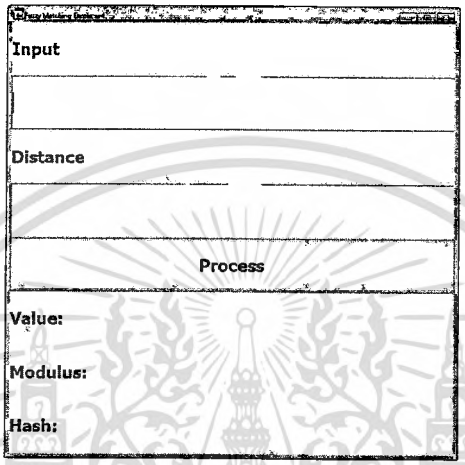
ภาคผนวก ก

แสดงภาพตัวอย่างหน้าโปรแกรมที่ใช้ในการศึกษาและวิเคราะห์ประสิทธิภาพและความปลอดภัยและพัฒนาการใช้งานฟังค์ชันทางเดียวหรือข้อผูกมัดแบบคลุมเครือที่รับความคลาดเคลื่อนได้เป็นค่าเป็น java application โดยการพัฒนาการใช้งานนี้ได้ใช้ฟังก์ชันแบบที่เก็บค่าสมภาคเพียงค่าเดียว

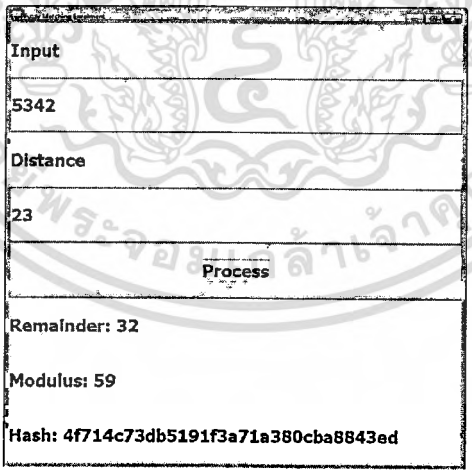
การพัฒนาการใช้งานจะประกอบด้วย application ย่อยสองตัว คือส่วนที่ทำหน้าที่สร้าง commitment และส่วน authentication

เมื่อเปิด application ส่วนสร้าง commitment จะปรากฏหน้าต่างดังภาพข้างล่าง

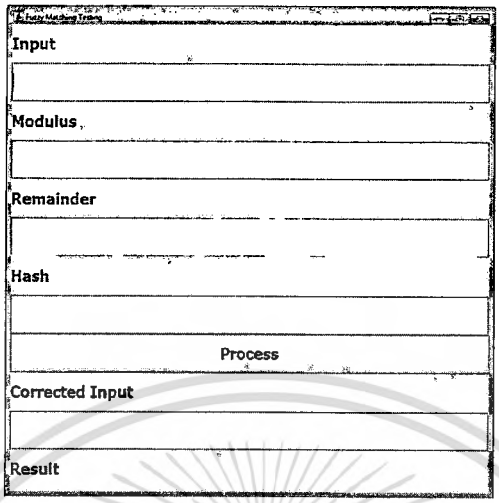
ผู้ใช้กรอกข้อมูลที่ต้องการสร้าง commitment ในช่อง Input และกรอกค่าความคลาดเคลื่อนสูงสุดที่ต้องการในช่อง Distance



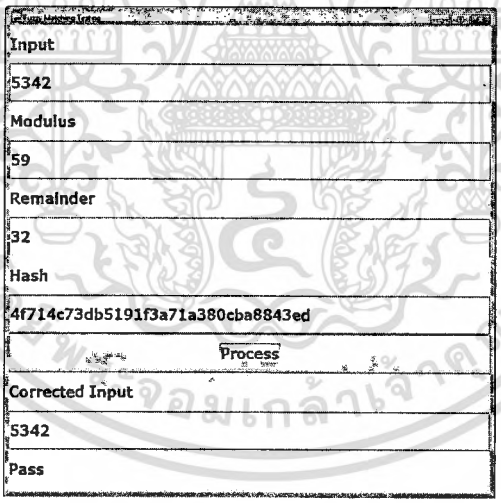
หลังจากนั้นผู้ใช้กดปุ่ม Process และ application ก็จะทำการคำนวณค่า Remainder, Modulus และ Hash มาให้



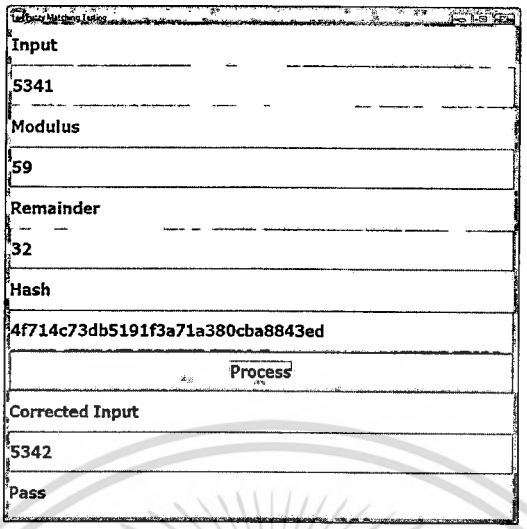
สำหรับ application ที่ใช้สำหรับ authentication จะมีหน้าต่างดังภาพข้างล่าง



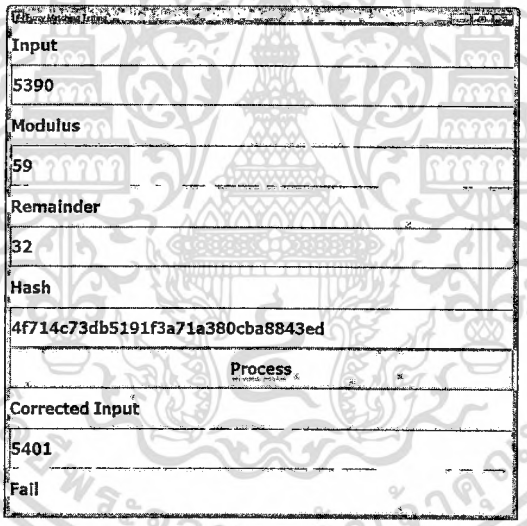
ผู้ใช้กรอกข้อมูลในช่องต่างๆ แล้วกด Process โดยที่ข้อมูล Input สามารถมีค่าคลาดเคลื่อนได้ไม่เกิน ค่าที่ตั้งไว้



ระบบแสดงข้อความ authentication ผ่าน



หากข้อมูลเข้ามีค่าคลาดเคลื่อนมากเกินไปก็จะไม่ผ่านการ authentication



ภาคผนวก ข

ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่

1. Nol Premasathian and Watcharee Tantikittipisut , "Analysis of Fuzzy Matching Scheme," International Conference on Emerging Trends of Computers & Information Technology 2012, pp 39-45, Coimbatore, India, 13-14 May 2012.
2. Nol Premasathian and Watcharee Tantikittipisut , "Fuzzy Matching Using One Congruence," International Conference on Computer Science and Medical Engineering, pp 34-37, Cochin, India, 31 August - 1 September 2012.

Analysis of Fuzzy Matching Scheme

Nol Premasathian
Faculty of Information Technology,
King Mongkut's Institute of Technology Ladkrabang,
Bangkok, Thailand
nol@it.kmitl.ac.th

Wacharee Tantikittipisut
Independent Scholar
Bangkok, Thailand
charlieante@hotmail.com

Abstract— **This paper analyzes the usage of fuzzy matching scheme, a fuzzy commitment scheme that allows input to be slightly different by value, not by bits, in two approaches, the one with the moduli are carefully selected to accommodate the secret value and the other one with the moduli are freely selected. The first approach is not subject to error in the reconstruction of the value using Chinese remainder theorem but the secret domain is reduced by the protocol while the second approach has a small chance in failing to reconstruct the secret value but does not suffer the reduction of the secret domain.**

Keywords-component; fuzzy commitment; fuzzy matching; fuzzy authentication

I. INTRODUCTION

Commitment schemes are important components of a number of cryptographic protocols [1]. In a commitment scheme, a user commits to a value while keeping it a secret. Later, an attempt is made to guess the value or to prove the possession of it. Then committed value is revealed and compared with the attempted one. The attempt is successful if and only if both values are exactly the same. A string with a single bit change will fail the verification. The committed value cannot be changed after the commitment is made. The original commitment scheme is not suitable for a biometric secrecy system due to the random noise in the data. The fuzzy commitment scheme introduced by Juels and Wattenberg allows the commitment to be made in a fuzzy manner [2]. That means the attempt is successful when the attempted input is the same as or is slightly different from the committed value. Satienjarurat and Premasathian introduced a scheme called fuzzy matching which compares the attempted input with the committed information by value, not by bit [3]. This paper analyzes the efficiency and security of the fuzzy matching scheme and introduced some improvement to it.

II. EXISITING WORKS

A. Commitment Schemes

In a commitment scheme, a party can commit to a value while keeping it hidden and can reveal it later on. The value cannot be altered by anyone once it has been committed. A commitment scheme can be constructed in several ways [1]. One using symmetric encryption can be constructed as follows.

1. Bob generates a random-bit string, R and sends it to Alice.
2. Alice, wanting to commit a bit b, creates a session key k and encrypt (R,b) using k and sends the result $E_k(R,b)$to Bob.
3. To reveal the committed value, Alice sends the key k to Bob.
4. Bob decrypts the cipher to reveal the bit. He also verifies the integrity of his random string.

Alice can cheat if she can find another key that can decrypt the cipher into another message with a different committed bit value with exactly the same random string. If the encryption algorithm is secure enough, the chance is negligible.

Bit commitment can be constructed using One-Way functions as follows.

1. Alice creates a random strings R.
2. Alice compute the one-way function of the committed bit b along with the string created in step 1.
3. Alice sends the result of the one-way function $H(b,R)$ to Bob.
4. To reveal the committed value, Alice sends the original message (b,R) to Bob.
5. Bob computes the one-way function to verify the protocol.

In contrast to the previous protocol, this protocol does not require Bob to send a message to Alice. If the one-way function is collision-resistance, the protocol is secure. This protocol can be further enhanced by splitting the random string into two parts, R_1 and R_2. Alice also sends the unencrypted R_1 to Bob in step 3 and sends both R_1 and R_2 along with the committed value b to Bob in step 4.

B. Fuzzy Commitment Scheme

The fuzzy commitment scheme introduced by Juels and Wattenberg makes use of error correction to correct the attempted input to the committed value if the attempted input is close enough to the committed value. If the attempted input is not close enough to the committed value, the correction will

transform it to a value completely different to the committed one. The scheme works as follows.

1. A codeword w is randomly chosen from a set of error correcting code C, which contains codes capable of correcting up to t bits of error.
2. The committed value s is added with the codeword w using exclusive-or operation. The result is $s \oplus w$.
3. Hash the code word w to get $H(w)$. Destroy w and keep only $s \oplus w$ and $H(w)$.
4. An attempting user enters s' to the scheme, which calculates $s' \oplus s \oplus w$. If s' does not differ from s by more than t bits, then it is possible to correct $s' \oplus s \oplus w$ back to w. This can be verified by computing the digest of corrected $s' \oplus s \oplus w$ and compared it with the stored $H(w)$.

There are a number of fuzzy matching techniques proposed such as the one by Al-saggaf and Acharya [4], which employs a technique similar to the one proposed by Juels and Wattenberg. Ojha and Sharm proposed a fuzzy commitment scheme with McEliece's cipher, an asymmetric encryption based on error correction [5]. All of the aforementioned algorithms are fuzzy by bit, not by value.

C. *Fuzzy Matching Scheme*

Two values close to each other by bit are not implicitly close to each other by value and vice versa. For example, the bit string 01111111 is close to the bit string 11111111 by bit as there is only one bit difference between them but they greatly differ by value, 127 and 255. The first bit string 01111111 is totally different from the bit string 10000000 by bit but these two are very close by value, 127 and 128. In contrast to the fuzzy commitment scheme, which accounts errors by bit, the fuzzy matching scheme allows input to be slightly different by value. The fuzzy matching scheme makes use of unary digits due to the fact that in unary representation, if two values are close to each other by bit, they must also be close to each other by value and vice versa. Substituting unary digits for the binary one straightaway in the fuzzy commitment scheme would make the new scheme cumbersome and insecure as unary digits requires immensely large spaces and the scheme will be broken in polynomial time. Therefore the scheme separates the commitment into parts and uses the Chinese remainder theorem to reconstruct the secret. By doing this the scheme can use a given space to store a larger commitment while preserving its unary representative character. The scheme works as follows.

1. For a committed value s, select k number, $n_1, n_2, \ldots, n_k$, all of which being relatively prime to one another and $n_1 . n_2 \ldots n_k$ is greater than the largest possible commitment. Calculate $s_i = s \bmod n_i$, for all $1 \le i \le k$.
2. For each remainder s_i, randomly select a codeword w_i from a set of error correcting code C, which contains codes capable of correcting up to t bits of error. Each s_i is added with its corresponding codeword w_i using exclusive-or operation. The result is $s_i \oplus w_i$.
3. Instead of individually hashing each w_i, the scheme hashes the secret s, destroys all w_i and keeps all of the computed $s_i \oplus w_i$ and $H(s)$ computed from s in binary.
4. An attempting user enters s' to the scheme, which calculates $s'_i = s' \bmod n_i$, for all $1 \le i \le k$. Next, compute $w_i' = s'_i \oplus s_i \oplus w_i$ and correct the error in w_i' to get w_i''. If s'_i is close to s_i enough, w_i''will be the same as w_i. Then compute $s_i'' = w_i'' \oplus s'_i \oplus w_i$ and use all s_i'' and n_i to reconstruct s'' by the Chinese remainder theorem. Compute $H(s'')$ and compared with the stored $H(s)$. The attempt is successful if the two digest exactly match each other.

III. Analysis of the Scheme

This section conducts the analysis of fuzzy matching scheme in two parts. As the scheme reconstructs the secret using the Chinese remainder theorem, we suggest two approaches in applying the scheme. The first approach is to select a codeword carefully so that the secret is not close a multiple of any number by which it is modulated. The domain of the secret can then be reduced by checking integers being used in the modulation. The analysis of this reduction is in the first part. The second one does not has such restriction and allows integers to be chosen deliberately as long as they are relatively prime to one another as required by the Chinese remainder theorem. This can result in an erroneous fuzzy verification and the second analysis addresses this problem together with its probability.

A. *The reduction in the secret domain*

We consider only the problem when all numbers are large. Therefore we consider all integers $n_1, n_2, \ldots, n_k$ in the Chinese remainder theorem equal in size, though they must differ according to the theorem. So the largest commitment possible is broken in to $n_1 . n_2 \ldots n_k$, with all of the factors being roughly equal. We will use the term "size of the modulus" to refer to the size of each integer.

In order to reconstruct the secret correctly, we must make sure that any error will not affect the Chinese remainder process. For any integer n_i, the secret s, and the set of error correcting code C, which contains codes capable of correcting up to t bits of error, we require that $s_i = s \bmod n_i$ must have the following property,

$$t \le s_i \le n_i - t. \quad (1)$$

If an integer n_j does not meet this requirement, it means that either $s_j < t$ or $s_j > n_j - t$-1. If $s_j < t$ and s has an error of $-t$, the error correction will not correct s_j' back to s_j as the value is now $n_j - s_j + t$. The problem of $s_j > n_j - t - 1$ can be proven in a similar way. If a scheme opts to enforce this requirement, the domain possible secret is reduced from n to $(n_1 - 2t).(n_2 - 2t)\ldots(n_k - 2t)$. For a given error correction capability t and the largest possible value of secret n, larger modulus size results in a small

number of moduli and less effect on the reduction of possible secret domain. However, large modulus sizes also requires a large space to hold the value $s_i \oplus w_i$, which can be as large as the size of the modulus in unary representation. We measure the number of unary bits required for a secret of size n, the number of moduli k, and the error correction capability t, using the following equation,

$$Number\ of\ unary\ bits = k.2^{(n/k)}, \quad (2)$$

and the reduced domain of the secret in the term of entropy using the following equation,

$$The\ reduced\ entropy\ of\ the\ secret\ = \log_2(2^{(n/k)}\text{-}2t)^k \quad (3)$$

The results of secret of 256 and 512 bits, with t = 50 are show in the following tables.

TABLE I. THE REDUCED ENTROPY OF THE SECRET (256 BITS)

k	Number of bits	The reduced entropy
2	6.81E+38	About 256
4	7.38E+19	About 256
8	3.44E+10	255.9999997
16	1048576	255.9647511
32	8192	233.132871

Number of bits are in unary.

TABLE II. THE REDUCED ENTROPY OF THE SECRET (512 BITS)

k	Number of bits	The reduced entropy
2	2.32E+77	About 512
4	1.36E+39	About 512
8	1.48E+20	About 512
16	6.87E+10	511.9999995
32	2097152	511.9295021
64	16384	466.265742

Number of bits are in unary.

From the tables, we can conclude that using a small number of moduli would require a prohibitively large space for the unary bits. As the number of moduli grows, the space required is reduced as is the entropy of the secret. There is a limit of the number of moduli for each size of secret. For a 256-bit secret and t= 50, k cannot be greater than 38 or each modulus will be smaller than $2t$, contradicting the requirement in (1). For a 512-bit secret, the limit is increased to 77. Anyway it is strongly recommended not to use anything close to t as it will greatly reduced the secret entropy and make the scheme unacceptably insecure.

B. *The rate of error using fuzzy matching scheme*

If we remove the restriction of (1), the entropy of the secret in the scheme is not reduced and therefore we have the freedom of choosing a larger k. However this comes with the penalty of the possibility of error due to the nature of modulation mentioned in the previous section. We define an error as a situation that we cannot reconstruct s correctly. With some s_i being in the region of 0 to $t-1$ or $n-t+1$ to $n-1$ does not necessarily mean that we are unable to reconstruct s. As the entered input can be deviated by up to the value of t, the deviation can be positive or negative. A positive deviation up to up to the value of t will not affect the case when s_i is in the region of 0 to $t-1$ as the input is not different from s_i in the troubling direction. Similarly a negative deviation up to the value of t will not affect the case when s_i is in the region of $n-t+1$ to $n-1$. As we have a number of modulations and some may not give the correct answer, we can use the majority voting scheme to count which deviation occurs most. So we can say that the scheme is safe as long as the majority of do not fall in the region of 0 to $t-1$ or the majority of do not fall in the region of $n-t+1$ to $n-1$, instead of requiring them to satisfy both at the same time. For a secret of size n, the number of moduli k, and the error correction capability t, we calculate the probability of an error in the scheme as in

$$p(error) = 2(\Sigma_{i=1\ to\ k/2}(C(k, k\text{-}i)(2^{(n/k)}\text{-}t)^i t^{k-i}/2^n). \quad (4)$$

Where C(n,r) is the combinatorial coefficient function, $n!/((n-r)!r!)$. The probability of error when t = 100 and 1000 and n = 256 and 512 is shown in the tables below.

TABLE III. THE PROBABILITY OF ERROR (256 AND 512 BITS)

Number of moduli	256 bits	512 bits
2	1.18E-36	3.45E-75
4	3.53E-34	1.04E-72
8	4.11E-29	1.21E-67
16	5.81E-23	1.73E-61
32	1.82E-10	3.38E-45

t=100

TABLE IV. THE PROBABILITY OF ERROR (256 AND 512 BITS)

Number of moduli	256 bits	512 bits
2	1.18E-35	3.45E-74
4	3.53E-32	1.04E-70
8	4.11E-25	1.21E-63
16	5.28E-15	1.73E-53
32	n/a	2.76E-29

t=1000

TABLE V. THE PROBABILITY OF ERROR (256 AND 512 BITS)

Number of moduli	256 bits	512 bits
2	1.18E-34	3.45E-73
4	3.53E-30	1.04E-68
8	4.11E-21	1.21E-59
16	1.83E-07	1.73E-45
32	n/a	2.49E-14

t=10000

From the table, the error rate increases when the correction capability t becomes larger. The probability of 256 bits with 32

moduli and t=1000 and 10000 cannot be computed as $2^{256/32}$ is smaller than 1000 and 10000. The figure is actually the bound of error. Even if the selection of the moduli falls in to the area of trouble, it does not mean that all of the verification will face a problem. For example, some attempt may not contain the deviation as large as t or the deviation may not be large enough to cross the border, or the deviation may be in the direction that does not cause a problem. This is therefore an alternative to the first method where the problem is completely prevented by carefully selecting proper moduli, and comes at the cost of the reduction in the secret domain, in which the second method does not suffer. The second approach significantly reduces the probability of requiring a new set of moduli for each secret, compared to the first approach. The error rate in the method is generally much lower than the FRR in a biometric authentication system.

C. *Making the scheme more obscure*

In order to make the scheme more obscure to an attacker, we may use a set error correcting code that contains codewords capable of correcting information more than the maximum deviation value allowed. By using a codeword capable of correcting up to t' bits, with the original error correcting requirement of t bits, we can add up to t'-t error bits to the value $s_i \oplus w_i$. This obscurity will reduce the pattern of unary format from the value

IV. Conclusions and Future Works

In this paper we have analyzed the fuzzy matching scheme. The scheme can be used in two approaches. The first one selects only suitable moduli to prevent errors while the second one has no restriction and let the error occurs. We investigated the outcome of these two options in terms of the reduction of the secret domain and the probability of error.

The future work of this research is the analysis of the effect of the set of the codeword on the security of the scheme, the study of constructing a fuzzy matching scheme correcting error by value, not by bit, without using a random oracle. The leakage of information will be analyzed as in [6] as well.

References

[1] B. Schneier, Applied Cryptography, Wiley, 1996

[2] A. Juels and M. Wattenberg, "A Fuzzy Commitment Scheme," in Proceedings of the 6th ACM Conference on Computer and Communications Security (ACM CCS '99), pp. 28-36, November 1-4 1999, Singapore.

[3] A. Satienjarurat and N. Premasathian, "Fuzzy Matching of Objects using Fuzzy Commitment," 2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology (ECTI-CON 2009), pp. 622-625, May 6-9 2009, Chonburi, Thailand.

[4] A. A. Al-saggaf and H. S. Acharya, "A Fuzzy Commitment Scheme," in Proc. of IEEE International conference on Advances in Computer Vision and Information Technology (ACVIT), November 28-30, 2007, Aurangabad, India.

[5] D. B. Ojha and A. Sharm, "A Fuzzy Commitment Scheme with McEliece's Cipher," Surveys in Mathematics and its Applications Vol. 5, 2010, pp. 73–82

[6] T. Ignatenko and F. M. J. Willems, "Information Leakage in Fuzzy Commitment Schemes," IEEE Transactions on Information Forensics and Security, Vol. 5, No. 2, June 2010, pp. 337-348.