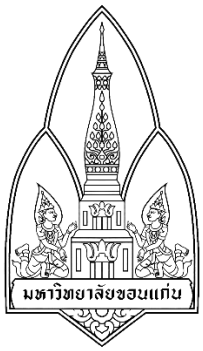


CSC1/2562

เอกสารเค้าโครงโครงการวิจัยทางวิทยาการคอมพิวเตอร์

การวิเคราะห์รูปแบบการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งด้วยเทคนิคการเรียนรู้ของเครื่อง

Pattern Analysis of Denial of Service Attack in Internet of Things Networks using Machine Learning

โดย

593021270-9 นางสาวฑิตยา ศรีวุฒิทรัพย์

อาจารย์ที่ปรึกษา : ผศ.ดร.ชิตสุธา สุ่มเล็ก

ตำแหน่ง ผู้ช่วยศาสตราจารย์

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

รายงานนี้เป็นส่วนหนึ่งของการศึกษาวิชา 322 498 โครงการวิจัยทางวิทยาการคอมพิวเตอร์

ภาคเรียน 1 ปีการศึกษา 2562

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

มหาวิทยาลัยขอนแก่น

(เดือน กันยายน พ.ศ. 2562)

การเสนอเค้าโครงโครงงานคอมพิวเตอร์

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

ชื่อ นางสาวทิตยา ศรีวุฒิทรัพย์ รหัสประจำตัว 593021270-9

Miss Titaya Sriwuttisap

นักศึกษาระดับปริญญาตรี วิทยาการคอมพิวเตอร์

อาจารย์ที่ปรึกษาโครงงาน ผศ.ดร.ชิตสุธา สุ่มเล็ก

Project Advisor Asst. Prof. Chitsutha Soomlek

1. ชื่อหัวข้อโครงงาน

ภาษาไทย การวิเคราะห์รูปแบบการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งด้วยเทคนิคการเรียนรู้ของเครื่อง

ภาษาอังกฤษ Pattern Analysis of Denial of Service Attack in Internet of Things Networks using Machine Learning

2. หลักการและเหตุผล

ปัจจุบันอยู่ในยุคที่เทคโนโลยีกำลังพัฒนาไปได้อย่างก้าวไกล การเชื่อมต่ออินเทอร์เน็ตไม่จำกัดอยู่เพียงบนคอมพิวเตอร์หรือโทรศัพท์อีกต่อไป สังเกตได้ว่าอุปกรณ์ที่เชื่อมต่อกับอินเทอร์เน็ตนอกจากคอมพิวเตอร์และโทรศัพท์มีจำนวนมากขึ้น เช่น กล้องวงจรปิด นาฬิกาข้อมือ รถยนต์ เป็นต้น เพื่ออำนวยความสะดวกต่อผู้ใช้งาน แต่ทว่าเมื่ออุปกรณ์เหล่านี้มีมากขึ้นก็มักนำมาซึ่งภัยคุกคามต่าง ๆ อุปกรณ์ IoT เคยถูกควบคุมเพื่อใช้ในการโจมตีทางไซเบอร์อยู่หลายครั้ง สาเหตุหลักเกิดจากอุปกรณ์จำนวนมากถูกติดตั้งโดยใช้รหัสผ่านที่มาจากโรงงาน ทำให้ผู้ประสงค์ร้ายสามารถล็อกอินเข้าไปติดตั้งมัลแวร์ในอุปกรณ์ดังกล่าวเพื่อควบคุมมาใช้ในการโจมตีได้ หนึ่งในเหตุการณ์ที่เคยเกิดขึ้นคือมีการใช้มัลแวร์ชื่อ Mirai ควบคุมอุปกรณ์ IoT ไปโจมตีแบบ DDoS ความรุนแรงสูงถึง 1.1 Tbps [1] จากเหตุการณ์ดังกล่าวทำให้รู้ว่าควรให้ความสำคัญกับความปลอดภัยบนอุปกรณ์ IoT อย่างมาก

ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงทำการนำชุดมูลจากการโจมตีแบบ Mirai บนอุปกรณ์ IoT มาวิเคราะห์รูปแบบการโจมตีเพื่อหาวิธีตรวจจับที่แม่นยำขึ้น

3. วัตถุประสงค์ของโครงงาน

3.1 เพื่อวิเคราะห์รูปแบบการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งด้วยเทคนิคการเรียนรู้ของเครื่อง

3.2 เพื่อสร้างโมเดลการตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งด้วยเทคนิคการเรียนรู้ของเครื่อง

4. ทฤษฎีและผลงานวิจัยที่เกี่ยวข้อง

4.1 Internet of Things (IoT)

IoT คือ การที่อุปกรณ์ต่าง ๆ เชื่อมต่อกับอินเทอร์เน็ต และอุปกรณ์ยังสามารถสื่อสารกันเองผ่านทางโปรโตคอล IP ซึ่งอุปกรณ์จำเป็นต้องทำงานร่วมกับ Radio Frequency Identification (RFID), Quick Response (QR) และเซ็นเซอร์เพื่อเปิดใช้การสื่อสารระหว่างอุปกรณ์ [2]

ลักษณะของ Internet of Things (IoT) มี 6 ประการดังนี้ [3]

1. เชื่อมโยงถึงกัน (Interconnected) คือ อำนวยความสะดวกให้มนุษย์กับอุปกรณ์ และอุปกรณ์เชื่อมต่อระหว่างอุปกรณ์

2. การตรวจจับอัจฉริยะ (Smart Sensing) คือ อุปกรณ์ที่เชื่อมต่อกับ IoT จะมีความสามารถในการตรวจจับอัจฉริยะ เช่น การตรวจจับการเคลื่อนไหว หรืออุณหภูมิจากนั้นทำการตอบสนองต่อสิ่งเหล่านี้ตามที่ได้ถูกตั้งโปรแกรม

3. สติปัญญา (Intelligence) คือ อุปกรณ์ที่เชื่อมต่อ IoT สามารถมีสติปัญญาโดยอุปกรณ์อาจถูกติดตั้งปัญญาประดิษฐ์ (Artificial Intelligence – AI) เพื่อเพิ่มความฉลาดของอุปกรณ์ให้มากขึ้น

4. ประหยัดพลังงาน (Save Energy) คือ อุปกรณ์จะพยายามประหยัดพลังงานอย่างมีประสิทธิภาพ เช่น การที่เปิดไฟเมื่อมีการเคลื่อนไหว ทำให้ช่วงเวลาที่ไม่มีคนอยู่ไฟจะปิดเองอัตโนมัติเป็นการลดการใช้พลังงานลง

5. การแสดงออก (Expressing) คือ อุปกรณ์ที่เชื่อมต่อ IoT มีความสามารถพิเศษในการบอกสถานะปัจจุบันไปยังอุปกรณ์ที่เชื่อมต่ออื่น ๆ ในบริเวณโดยรอบ มันอำนวยความสะดวกในการสื่อสารที่ดีขึ้นระหว่างมนุษย์และเครื่องจักร

6. ความปลอดภัย (Safety) คือ อุปกรณ์ที่เชื่อมต่อ IoT สามารถช่วยรับรองความปลอดภัยของชีวิตบุคคล เช่น การแจ้งเตือนผ่านระบบ เมื่ออุปกรณ์มีการทำงานที่ผิดปกติ

4.2 Denial of Service (DoS) และ Distributed Denial of Service (DDoS)

Denial of Service คือ การพยายามทำให้บริการของเครื่องเหยื่อไม่สามารถให้บริการได้ทั้งแบบชั่วคราวหรือถาวร [4]

Distributed Denial of Service คือ การพยายามทำให้บริการของเครื่องเหยื่อไม่สามารถให้บริการได้ทั้งแบบชั่วคราวหรือถาวร โดยแทนที่จะเป็นคอมพิวเตอร์หนึ่งเครื่องกลับใช้เป็นคอมพิวเตอร์จำนวนมากแทน ซึ่งผู้โจมตีจะใช้เครื่องมือที่จะใช้ในการโจมตีไปติดตั้งบนคอมพิวเตอร์ที่ถูกควบคุมไว้แล้ว จากนั้นทำการออกคำสั่งให้คอมพิวเตอร์จำนวนมากโมตีไปยังเป้าหมาย [4]

ซึ่งสามารถแบ่งการโจมตีแบบ Denial of Service (DoS) และ Distributed Denial of Service (DDoS) ได้เป็น 5 ประเภทดังนี้

1. SYN flood คือ ประเภทของการโจมตี DoS ที่ผู้โจมตีส่งชุดของการร้องขอ SYN ไปยังระบบเป้าหมายในความพยายามที่จะใช้ทรัพยากรเซิร์ฟเวอร์จำนวนมากเพื่อให้ระบบไม่ตอบสนองต่อการรับส่งข้อมูลที่ถูกกฎหมาย [4]
2. Teardrop Attacks คือ การส่งแพ็กเก็ตที่ไม่สามารถประกอบได้ไปให้เครื่องเป้าหมายเพื่อให้เกิดความสับสน ตามปกติหามีข้อมูลขนาดใหญ่ต้องการส่งผ่านเราเตอร์จะต้องทำการแบ่งเป็นส่วนย่อยโดยการ Fragment ก่อน จากนั้นเมื่อปลายทางได้ครบจะทำการประกอบเข้าด้วยกันเพื่อให้ได้ข้อมูลที่สมบูรณ์ แต่การจะรวมได้ต้องใช้ค่า Offset ที่ปรากฏอยู่ในแพ็กเก็ตแรกและแพ็กเก็ตถัดไป สำหรับการโจมตีแบบ Teardrop นี้ ผู้โจมตีจะส่งค่า Offset ในแพ็กเก็ตที่สองและต่อ ๆ ไปที่จะทำให้เครื่องรับปลายทางเกิดความสับสน หากระบบปฏิบัติการไม่สามารถรับมือกับปัญหานี้ก็จะทำให้ระบบหยุดการทำงานในทันที [5]

3. Low-rate Denial-of-Service attacks (LDoS) คือ การออกแบบมาเพื่อใช้ประโยชน์จากอัตราส่วนของเวลาในระดับช้าของ TCP ความสามารถในการดำเนินการกลไกการหมดเวลาส่งสัญญาณ (RTO) เพื่อลดปริมาณงาน TCP ในระยะสั้น ผู้โจมตีสามารถสร้าง TCP overflow ได้โดยการเข้าสู่สถานะ RTO ซ้ำ ๆ ผ่านการส่งอัตราสูงและจำนวนมาก ในขณะที่ RTO ทำการจับเวลาช้า ปริมาณข้อมูล TCP ที่โหนดของเป้าหมายจะลดลงอย่างมากในขณะที่ผู้โจมตีจะมีอัตราเฉลี่ยต่ำจึงทำให้ตรวจจับได้ยาก [4]
4. Internet Control Message Protocol (ICMP) flood คือ โปรโตคอลไร้การเชื่อมต่อที่ใช้สำหรับการดำเนินงาน IP การวินิจฉัยและข้อผิดพลาด ICMP Flood การส่งแพ็คเก็ต ICMP จำนวนมากอย่างผิดปกติทุกประเภท (โดยเฉพาะการทดสอบแพ็คเก็ต "ping") สามารถแฝงเครือข่ายเซิร์ฟเวอร์เป้าหมายที่พยายามประมวลผลคำขอ ICMP ที่เข้ามาทั้งหมด และอาจส่งผลให้เกิดการปฏิเสธ เงื่อนไขการให้บริการสำหรับเซิร์ฟเวอร์เป้าหมาย [4]
5. Peer-to-peer attacks คือ เครือข่ายแบบกระจายซึ่งแต่ละโหนดในเครือข่าย (peer) ทำหน้าที่เป็นทั้งผู้จัดหา (seed) และผู้บริโภค (leeches) ของทรัพยากร ตรงกันข้ามกับโมเดลไคลเอนต์ - เซิร์ฟเวอร์ส่วนกลางที่ไคลเอ็นต์ – เซิร์ฟเวอร์หรือโหนดระบบปฏิบัติการร้องขอการเข้าถึงทรัพยากรจากเซิร์ฟเวอร์ส่วนกลาง [4]

4.3 Naive Bayes

หลักการของวิธี Naive Bayes จะใช้การคำนวณความน่าจะเป็นซึ่งจะมีการใช้สมการที่เรียกว่า Bayes' theorem [6] หรือทฤษฎีของเบย์โดยมีสมการดังนี้

$$P(h \mid D) = P(D \mid h)*P(h) / P(D) \quad (1)$$

จากสมการข้างต้นสามารถสร้างตำนวณความน่าจะเป็นของการจำแนกประเภทอย่างง่ายได้ดังนี้

$$P(d \mid h) = P(a ,\ldots,a \mid h) = P(a \mid h) \quad (2)$$

4.4. Generalize Linear Model (GLM)

GLM เป็นวิธีการทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลที่มีการติดตามระยะยาว (Longitudinal data) และมีการวัดซ้ำค่าสังเกตของตัวแปรตามที่สนใจที่เกิดขึ้นภายใต้หน่วยศึกษาเดียวกัน (Cluster/Subject) ที่มักพบว่าค่าสังเกตดังกล่าวมีความสัมพันธ์กันเอง (Correlated data) โดยได้ขยายวิธีการดังกล่าวมาจากตัวแบบเชิงเส้นนัยทั่วไป (Generalized Linear Model: GLM) [7] ด้วยการนำรูปแบบความสัมพันธ์ของค่าสังเกตที่สนใจ (Working Correlation Structure) ไปพิจารณาในการวิเคราะห์สมการหรือตัวแบบ อีกทั้งยังมีข้อดีกว่า GLM คือค่าสังเกตหรือตัวแปรตามไม่จำเป็นต้องมีการแจกแจงในวงเลขชี้กำลัง (exponential family) ซึ่งทำให้สามารถนำไปวิเคราะห์ข้อมูลที่ไม่ใช่ค่าต่อเนื่องได้ โดยเฉพาะข้อมูลเชิงกลุ่ม (Count data)ที่มีการเก็บรวบรวมไว้ในด้านต่าง ๆ

ไม่ว่าจะเป็นทางการเกษตร การแพทย์ การสาธารณสุข หรือแม้แต่ในด้านธุรกิจ เช่น การศึกษาติดตามระยะยาวจำนวนการเกิดอุบัติเหตุบนทางหลวง การศึกษาติดตามระยะยาวจำนวนการเรียกค่าสินไหมทดแทนการประกันภัย เป็นต้น

4.5 Logistic Regression

Logistic Regression [8] เป็นการพยากรณ์ความน่าจะเป็นของผลลัพธ์ที่จะเกิดขึ้นซึ่งสามารถเป็นได้เพียง 2 ค่า โดยการพยากรณ์จะขึ้นกับตัวแปรที่ส่งผลต่อเหตุการณ์นั้นๆซึ่งอาจจะมีเพียงหนึ่งตัวหรือมากกว่า

4.6 Neural Network

Neural Network มีโครงสร้างและการทำงานของการประมวลผลเหมือนกับสมองของสิ่งมีชีวิตซึ่งมีปรับเปลี่ยนตัวเองต่อการตอบสนองของอินพุตตามกฎของการเรียนรู้ (Learning rule) หลังจากที่โครงข่ายได้เรียนรู้สิ่งที่ต้องการแล้ว โครงข่ายนั้นจะสามารถทำงานที่กำหนดไว้ได้โครงข่ายประสาทเทียมได้ถูกพัฒนาคิดค้นจากการทำงานของสมองมนุษย์โดยสมองมนุษย์ประกอบไปด้วยหน่วยประมวลผลเรียกว่า นิวรอน (เซลล์ประสาท หรือ Neuron) จำนวนนิวรอลในสมองมนุษย์มีอยู่ประมาณและมีการเชื่อมต่อกันอย่างมากมาย สมองมนุษย์จึงสามารถกล่าวได้ว่าเป็นคอมพิวเตอร์ที่มีการปรับตัวเอง (Adaptive) ไม่เป็นเชิงเส้น (Nonlinear) และทำงานแบบขนาน (Parallel) ในการดูแลจัดการการทำงานร่วมกันของนิวรอนในสมอง การคำนวณเชิงนิวรอลเป็นการคำนวณที่เลียนแบบมาจากการทำงานของสมองมนุษย์นั่นเอง [9]

4.7 k-Nearest Neighbors (k-nn)

หนึ่งในวิธีสำหรับการจำแนกรูปแบบคือ k- Nearest Neighbors (k-NN) [10] แยกประเภทของวัตถุที่ไม่มีชื่อแต่ละอันตามป้ายกำกับ k-Nearest Neighbors แม้จะมีความเรียบง่าย แต่กฎของ k-NN มักให้ผลการแข่งขันและในบางโดเมนเมื่อรวมเข้ากับความรู้ก่อนหน้านี้อย่างชาญฉลาดจะมีความก้าวหน้าอย่างมีนัยสำคัญขั้นสูง

ผลของ k-NN จัดหมวดหมู่ขึ้นอย่างมีนัยสำคัญเกี่ยวกับตัวชี้วัดที่ใช้ในการคำนวณระยะทางระหว่างเวกเตอร์คุณลักษณะที่แตกต่างกัน [10] ในการศึกษานี้ใช้ระยะทางแบบยุคลิด ระยะห่างระหว่างจุดสอง X(x1,x2,x3,...,xn) และ Y(y1,y2,y3,...,yn) อธิบายได้ดังนี้

$$d = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{3}$$

4.8 Decision Tree

หลักการของวิธี Decision Tree [11] เป็นการเรียนรู้โดยการจำแนกประเภท (Classification) ข้อมูลออกเป็นกลุ่ม (Class) ต่าง ๆ โดยใช้คุณลักษณะ (Attribute) ข้อมูลในการจำแนกประเภท ซึ่งทำให้สามารถทราบได้ว่าคุณลักษณะมาเป็นตัวกำหนดการจำแนกประเภท และคุณลักษณะแต่ละตัวมีความสำคัญมากน้อยเท่าใด

4.9 Random Forest

หลักการของวิธี Random Forest [12] ใช้การสุ่มข้อมูล และสุ่มเลือกคุณลักษณะต่าง ๆ ออกมาเป็นหลาย ๆ ชุดและจากนั้นนำมาสร้างโมเดลด้วยเทคนิค Decision Tree หลาย ๆ ต้น

4.10 Receiver Operating Characteristic (ROC)

การวิเคราะห์ประเมินความถูกต้อง และประสิทธิภาพของแบบจำลองที่ใช้ทำนายข้อมูลด้วย ROC [13] (Receiver Operating Characteristic) เป็นการทดสอบประสิทธิภาพแบบจำลองที่ทำการพิจารณาในรูปแบบของกราฟ โดยวิธีการสร้าง ROC curve มาใช้เพื่อเลือกจุดตัดที่เหมาะสม เป็นความสัมพันธ์ระหว่าง True positive rate (Sensitivity) กับ False positive rate (1–Specificity) ด้วยการแปรค่าจุดตัด (Cut – off point) ซึ่งอาจจะยังไม่สามารถให้ผลลัพธ์ที่ได้ดีนัก โดยรวมเหมาะกับการวิเคราะห์ปัญหาที่มีคำตอบเป็นค่าบวกหรือลบ

4.11 Extreme Learning Machine (ELM)

ELM เป็นขั้นตอนวิธีที่มีความเร็ว มีโครงสร้างการทำงานแบบ Single-hidden layer feed forward neural networks (SLFNs) แต่ทว่า ELM [14] ยังคงใช้หลักการสุ่มเพื่อทำการกำหนดค่าน้ำหนักในชั้นนำเข้าข้อมูล และค่าเอนเอียงในชั้นซ่อน ทำให้มีโอกาสที่จะนำไปสู่ค่าเหมาะสมเฉพาะที่ (Local optimal)

4.12 Confusion Matrix

Confusion Matrix คือ การประเมินผลลัพธ์การทำนาย (หรือผลลัพธ์จากโปรแกรม) เปรียบเทียบกับผลลัพธ์จริง ๆ ที่หาโดยคน [15]

True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าจริง และคนบอกว่ามันจริง

True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าไม่จริง และคนบอกว่ามันไม่จริง

False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าจริง แต่คนบอกว่าไม่จริง

False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าไม่จริง แต่คนบอกว่าจริง

การหาค่าต่าง ๆ จาก Confusion Matrix มีดังนี้ [16]

1. Accuracy คือ ค่าที่บอกว่าโปรแกรมสามารถทำนายได้แม่นยำขนาดไหน หาได้จาก (TP+TN)/(TP+TN+FP+FN)

2. Recall (True Positive Rate) คือ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าจริง เป็นอัตราส่วนเท่าไรของจริงทั้งหมด หาได้จาก TP/(TP+FN)

3. True Negative Rate (TNR) คือ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าไม่จริง เป็นอัตราส่วนเท่าไรของจริงทั้งหมด หาได้จาก TN/(TN+FP)

4. False Positive Rate (TPR) คือ ค่าที่บอกว่าโปรแกรมทำนายว่าจริง เป็นอันตราส่วนเท่าไรของไม่จริงทั้งหมด หาได้จาก FP/(TN+FP)

5. False Negative Rate (FNR) คือ ค่าที่บอกว่าโปรแกรมทำนายว่าไม่จริง เป็นอันตราส่วนเท่าไรของจริงทั้งหมด หาได้จาก FN/(TP+FN)

6. Precision คือ ค่าที่บอกว่าโปรแกรมทำนายว่าจริง ถูกต้องเท่าไร หาได้จาก TP/(TP+FP)

4.13 Area Under <ROC> Curve (AUC)

AUC [16] เป็นอีกหนึ่ง Metric ยอดนิยมที่ใช้กันแทบทุกงานเลย AUC มีค่าอยู่ระหว่าง 0-1 ยิ่งเข้าใกล้ 1 แปลว่าโมเดลในภาพรวมสามารถทำนาย y ได้ดีมาก

AUC = 0.50 ไม่ต่างอะไรกับการเดาสุ่มเลย

AUC > 0.70 คือเกณฑ์มาตรฐานสำหรับโมเดลส่วนใหญ่

AUC > 0.80 โมเดลทำงานได้ดี

AUC > 0.90 โมเดลทำงานได้ดีมาก

4.14 F-measure

ค่าประสิทธิภาพโดยรวม (F-measure) หมายถึง การวัดประสิทธิภาพโดยรวมของทั้งสองค่าระหว่างค่าความแม่นยำ และค่าความครบถ้วนซึ่งนำ ค่าทั้งสองมาค่านวณร่วมกัน [17]

F-Measure = 2 x Precision x Recall / (Precision + Recall) (4)

4.15 Survey on Android Malware Detection and Protection using Data Mining Algorithms [18]

ในงานวิจัยนี้ศึกษาเกี่ยวกับประสิทธิภาพของอัลกอริธึมการทำเหมืองข้อมูลเช่น Naive Bayes, J48, Multiclass Classifier, Random Tree, SVM, Decision Tree อัลกอริทึมแต่ละอันถูกประเมินโดยเกณฑ์ต่าง ๆ เพื่อระบุว่าอันไหนที่เหมาะสมในการตรวจจับซอฟต์แวร์ที่เป็นอันตราย ชุดข้อมูลจะถูกเปรียบเทียบและประเมินผลโดยใช้ Confusion matrix เพื่อหาค่าความถูกต้อง (Accuracy)

สรุปผลการทดลองได้ว่า Multiclass ทำงานได้ดีกว่าอัลกอริธึมอื่น แต่ Naive Bayes ช่วยลดความซับซ้อนของเวลาในกระบวนการตรวจจับ

4.16 Machine learning-based recommendation trust model for machine-to-machine communication [19]

ในงานวิจัยนี้ทำการวิเคราะห์เปรียบเทียบจะดำเนินการกับห้าวิธีการเรียนรู้ของเครื่องที่แตกต่างกัน คือ Naive Bayes (NB), Decision Tree (DT), Linear and Radial Support Vector Machine (SVM), KNearest Neighbor (KNN) และ Random Forest (RF) เพื่อตรวจสอบประสิทธิภาพการทำงานของโมเดลเหล่านี้มีการใช้มาตรการการคำนวณความน่าเชื่อถือสองแบบ Receiver Operating Characteristics (ROCs), Precision and Recall ซึ่งโปรแกรมที่ใช้ในการทดลอง คือ Matlab กำหนดโหนดที่เป็นอันตรายมีการเปลี่ยนแปลงในช่วง 10%, 20%, 30%, 40%

4.17 Intrusion Detection for IoT Based on Improved Genetic Algorithm and Deep Belief Network Computing [20]

ในงานวิจัยนี้เสนอกลไกในการตรวจจับการโจมตีที่ไม่เหมาะสมใน IoT LLNs โดยใช้วิธีการMultilayer Perceptron (MLP) เป็นเครื่องมือในการจำแนกประเภท MLP จะแบ่งประเภทข้อมูลเป็นเครือข่ายตามปกติหรือภายใต้การโจมตี ซึ่งกลไกยังระบุถึงโหนดที่ได้รับผลกระทบจากการโจมตีและระบุโหนดของผู้โจมตี

กลไกที่เสนอเพื่อตรวจจับ Misappropriation Attack ใน IoT

1. สร้างคุณสมบัติในการตรวจจับ Misappropriation Attack ใน IoT คือ คุณสมบัติที่เลือกเป็นอินพุตสำหรับโมเดลการตรวจจับใด ๆ ควรมีความสัมพันธ์อย่างมากกับเป้าหมาย

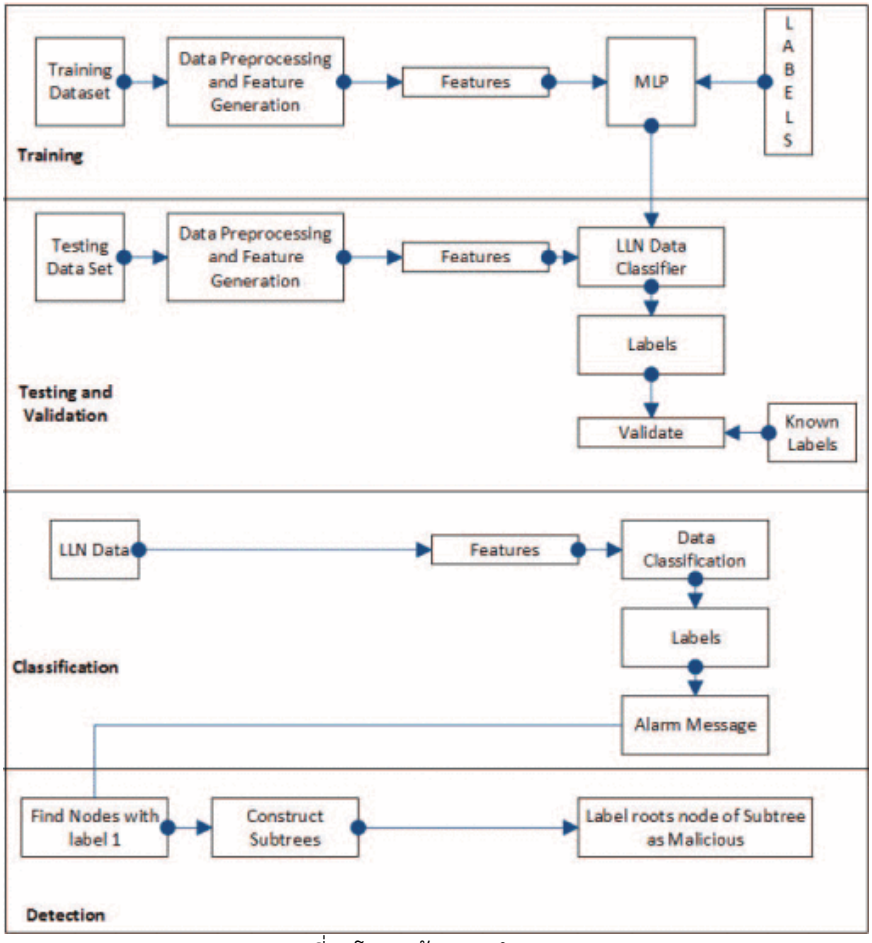
2. แบบจำลอง มี 3 องค์ประกอบหลักคือ

- สภาพแวดล้อมการจำลอง Cooja : การจำลอง Cooja เป็นแพลตฟอร์มการจำลองที่จัดทำโดย Contiki OS เพื่อจำลองเครือข่ายเซ็นเซอร์ไร้สาย Cooja จัดเตรียมไบนารีเพื่อเรียกใช้งานโหนดเซ็นเซอร์ประเภทต่างๆ ในการจำลองของเราเรามีมลทินท้องฟ้า

- Mote output Collector : Mote output ของการจำลองถูกรวบรวมเพื่อวิเคราะห์และตรวจจับการโจมตี

- หน่วยตรวจจับการโจมตี : นี้เป็นเครือข่าย Multilayer Perceptron ซึ่งจะมีความแตกต่างของจุดข้อมูลสำหรับโหนดทั้งหมดใน DODAG และจัดประเภทพวกเขาเป็นปกติ (0) หรือภายใต้การโจมตี (1)

3. หน่วยตรวจจับการโจมตีโดยอิงจาก Multilayer Perceptron



ภาพที่ 1 โครงสร้างการทำงาน [20]

5. วิธีดำเนินการวิจัย

5.1 การรวบรวมข้อมูลจากงานวิจัยที่เกี่ยวกับหัวข้อ

รวบรวมบทวิจัยจากแหล่งข้อมูลงานวิชาการต่าง ๆ ที่เป็นประโยชน์งานที่ศึกษา และนำมาใช้ในการวิเคราะห์รูปแบบการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งด้วยเทคนิคการเรียนรู้ของเครื่อง

5.2 การศึกษาข้อมูลที่ได้งานวิจัยและทดลองใช้เครื่องมือ

ศึกษากระบวนการทำงานของการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่ง โดยศึกษาวิธีการทำงานแต่ละโมเดลในการจำแนกและทำนายประเภทของข้อมูล จากนั้นทดลองทำและนำมาวิเคราะห์ข้อมูลในการทำนาย

5.3 การสรุปข้อมูลจากการศึกษางานวิจัย

สรุปข้อมูลเนื้อหาต่าง ๆ ที่ได้จากการศึกษา เพื่อให้เข้าใจหลักการทำงานของวิธีการทำเหมืองข้อมูลเพื่อหาวิธีตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งจากนั้นเลือกใช้เครื่องมือและรูปแบบตัวโมเดลที่ใช้ในการวิเคราะห์ข้อมูลตามความเหมาะสม

5.4 ทำเหมืองข้อมูลเพื่อหาวิธีตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่ง

นำชุดข้อมูลมาทำเหมืองข้อมูลเพื่อหาความสัมพันธ์ของข้อมูลต่าง ๆ แล้วนำมาวิเคราะห์เพื่อหาปัจจัยที่มีผลต่อการป้องกันการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่ง ซึ่งจะสามารถใช้หารูปแบบที่เหมาะสมในการป้องกันได้

5.5 ทดลอง ทดสอบและปรับปรุง

ทดลองหาวิธีการป้องกันใหม่ ๆ ที่มีประสิทธิภาพดีกว่าวิธีการเดิม จากนั้นทดลองใช้งานโดยการป้อนข้อมูลใส่แบบจำลองเพื่อทดสอบว่าสามารถให้ผลลัพธ์ที่ถูกต้องได้ทั้งหมดร้อยละเท่าใด และทำการปรับปรุงแบบจำลองเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด

5. 6 สรุปผลการทำทำเหมืองข้อมูลเพื่อหาวิธีตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่ง

สรุปผลการทำเหมืองข้อมูลเพื่อหาวิธีตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งที่สามารถได้ผลลัพธ์ถูกต้องและผิดพลาดว่ามีร้อยละเป็นเท่าใดจากข้อมูลที่ได้ทำการทดสอบทั้งหมด

5.7 สร้างโมเดล

จัดทำโมเดลตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่ง

6. ขอบเขตและข้อจำกัดของการวิจัย

ชุดข้อมูลที่ 1 [21]

เครือข่ายการเฝ้าระวังวิดีโอกล้อง IP จริงประกอบด้วยการตั้งค่ากล้องวงจรปิด HD สองแบบ จำนวนแบบละ 4 ตัว ใช้พลังงานผ่าน PoE และเชื่อมต่อกับ DVR ผ่านอุโมงค์ VPN ในแต่ละสถานที่ DVR ที่ไซต์ระยะไกลช่วยให้ผู้ใช้สามารถเข้าถึงสตรีมวิดีโอไปทั่วโลกผ่านการเชื่อมต่อ VPN แบบไคลเอ็นต์ถึงไซต์ กล้องที่ใช้ในเครือข่ายและการกำหนดค่าต่างๆได้อธิบายไว้ในภาพที 3

การโจมตีจำนวนหนึ่งซึ่งสามารถดำเนินการบนเครือข่ายกล้องวงจรปิดได้ การโจมตีที่สำคัญที่สุดส่งผลกระทบต่อความพร้อมใช้งานและความสมบูรณ์ของอัปลิงค์วิดีโอ เช่น SYN flood ในกล้องเป้าหมาย man in the middle attack เกี่ยวข้องกับการฉีดวิดีโอลงในสตรีมวิดีโอสด ดังนั้นในการประเมินจึงมุ่งเน้นไปที่การโจมตีประเภทนี้ ภาพที่ 2 สรุปชุดข้อมูลการโจมตีที่ใช้ในการทดลอง

การตั้งค่า wiretap ที่ใช้งานอยู่ใช้ Raspberry PI 3B เป็นบริดจ์เครือข่ายทางกายภาพ PI ได้รับอะแดปเตอร์ USB ต่ออีเธอร์เน็ตเพื่อข้ามพอร์ตอีเธอร์เน็ตที่สองจากนั้นในทางกายภาพได้วางไว้ตรงกลางของสายเคเบิล เครือข่ายเพิ่มเติมคือเครือข่าย Wi-Fi ที่มีอุปกรณ์ 9 IoT และพีซีสามเครื่อง อุปกรณ์ IoT ประกอบไปด้วย Thermostat , มอนิเตอร์สำหรับจับตาดูเด็กทารก webcam กระดิ่งประตู 2 ตัว และกล้องรักษาความปลอดภัยราคาถูก 4 ตัว จากนั้นติดตั้งหนึ่งในกล้องรักษาความปลอดภัยด้วยตัวอย่างจริงของ Mirai botnet malware

ไฟล์ pcap มีข้อมูลจำนวน 764,137 แถว เป็นปกติ 121,621 แถว และโจมตี 642,516 แถว

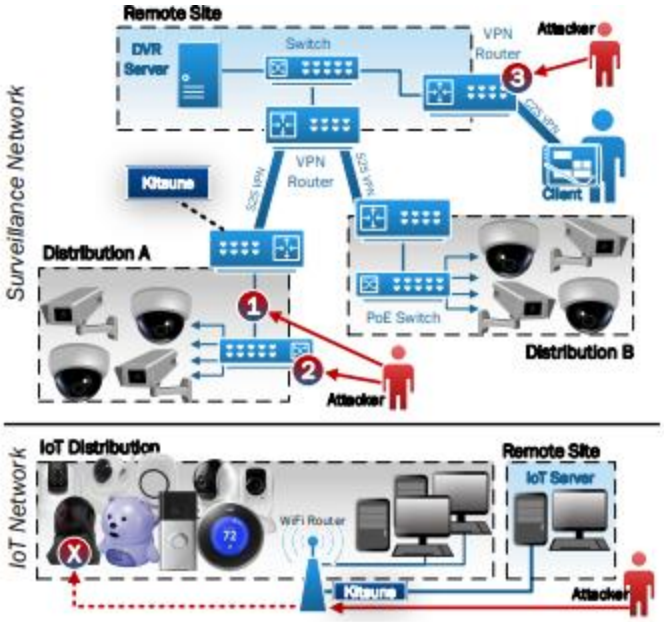
Attack Type	Attack Name	Tool	Description: *The attacker…*	Violation	Vector	# Packets	Train [min.]	Execute [min.]
Recon.	**OS Scan**	Nmap	*…scans the network for hosts, and their operating systems, to reveal possible vulnerabilities.*	**C**	1	1,697,851	33.3	18.9
	Fuzzing	SFuzz	*…searches for vulnerabilities in the camera's web servers by sending random commands to their cgis.*	**C**	3	2,244,139	33.3	52.2
Man in the Middle	**Video Injection**	Video Jack	*…injects a recorded video clip into a live video stream.*	**C, I**	1	2,472,401	14.2	19.2
	ARP MitM	Ettercap	*…intercepts all LAN traffic via an ARP poisoning attack.*	**C**	1	2,504,267	8.05	20.1
	Active Wiretap	Raspberry PI 3B	*…intercepts all LAN traffic via active wiretap (network bridge) covertly installed on an exposed cable.*	**C**	2	4,554,925	20.8	74.8
Denial of Service	**SSDP Flood**	Saddam	*…overloads the DVR by causing cameras to spam the server with UPnP advertisements.*	**A**	1	4,077,266	14.4	26.4
	SYN DoS	Hping3	*…disables a camera's video stream by overloading its web server.*	**A**	1	2,771,276	18.7	34.1
	SSL Renegotiation	THC	*…disables a camera's video stream by sending many SSL renegotiation packets to the camera.*	**A**	1	6,084,492	10.7	54.9
Botnet Malware	**Mirai**	Telnet	*…infects IoT with the Mirai malware by exploiting default credentials, and then scans for new vulnerable victims network.*	**C, I**	X	764,137	52.0	66.9

ภาพที่ 2 ชุดข้อมูล [21]

	SNC-EM602RC	SNC-EM600	SNC-EB600	SNC-EB602R
Resolution	1280x720			
Codec	H.264/MPEG4			
Frames/Sec	15			
Avg. Packets/Sec	195	350	290	320
Avg. Bandwidth	1.8 Mbit/s	1.4 Mbit/s	1.8 Mbit/s	1.8 Mbit/s
Protocol	RTP	RTP	Https *(TLSv1)*	Http/TCP

ภาพที่ 3 ข้อมูลจำเพาะและสถิติของกล้องที่ใช้ในการทดลอง [21]

ภาพที่ 4 ซ้าย SNC-EM602RC. ขวา SNC-EB602R [21]



ภาพที่ 5 เครือข่ายเฝ้าระวัง (บนสุด) และเครือข่าย IoT (ด้านล่าง) [21]

ชุดข้อมูลที่ 2 [22]

ไฟล์เหล่านี้ถูกสร้างขึ้นในห้องปฏิบัติการ Stratosphere ซึ่งเป็นส่วนหนึ่งของโครงการ Aposemat เพื่อรวบรวมมัลแวร์ IoT ที่ถูกจับที่ Done ในมหาวิทยาลัย CVUT ที่ปราก สาธารณรัฐเช็ก เพื่อจัดเก็บปริมาณการใช้มัลแวร์ IoT ที่มีอายุการใช้งานยาวนานและเพื่อสร้างไฟล์ netflows ที่มีป้ายกำกับ

- Infected device: 192.168.1.195
- Default GW: 192.168.1.1

ไฟล์ pcap มีข้อมูลจำนวน 233,865 แถว เป็นปกติ 5,396 แถว และโจมตี 228,469 แถว

ขอบเขตงาน

วิเคราะห์ข้อมูลโดยใช้โมเดลดังนี้ Decision Tree ,Random Forest ,Neural Networks, k nearest neighbor, Logistic Regression ,Naïve Bayes ,Support Vector Machine และใช้ชุดข้อมูลจากสองแหล่ง [21] [22] ที่อธิบายจากข้างต้น

เมื่อได้โมเดลแล้วจะวัดผลโมเดลโดยใช้ Accuracy Recall Precision และ F-measure

7. สถานที่ทำวิจัย

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น หรืออื่นๆ

8. ประโยชน์ที่คาดว่าจะได้รับ

โมเดลสามารถตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งได้

9. แผนและระยะเวลาดำเนินการ

ตารางการดำเนินงานและระยะเวลา

กิจกรรม	กรกฎาคม		สิงหาคม					กันยายน					ตุลาคม				
	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1. การรวบรวมข้อมูลจากงานวิจัยที่เกี่ยวกับหัวข้อ																	
2. การศึกษาข้อมูลที่ได้งานวิจัยและทดลองใช้เครื่องมือ																	
3. การสรุปข้อมูลจากการศึกษางานวิจัย																	
4. ทำเหมืองข้อมูลเพื่อหาวิธีตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่ง																	
5. ทดลอง ทดสอบและปรับปรุง																	
6. สรุปผลการทำทำเหมืองข้อมูลเพื่อหาวิธีตรวจจับการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่ง																	
7. สร้างโมเดล																	

10. งบประมาณ

หมวดวัสดุอุปกรณ์

- ค่าวัสดุสำนักงาน (กระดาษ ปากกา และอื่นๆ) 20 บาท
- ค่าวัสดุคอมพิวเตอร์ (แผ่นดิสก์ ซีดี และอื่นๆ) 20 บาท

หมวดค่าใช้สอย

- ค่าถ่ายเอกสาร 20 บาท
- ค่าจัดรูปเล่ม 100 บาท

หมวดค่าใช้จ่ายอื่นๆ

11. เอกสารอ้างอิง

[1] “ThaiCERT ไทยเซิร์ต - Botnet of Things - ภัยคุกคามจาก Internet of Things และแนวทางการรับมือ.” Accessed September 10, 2019. https://www.thaicert.or.th/papers/general/2016/pa2016ge001.html.

[2] Singh, S., and N. Singh. “Internet of Things (IoT): Security Challenges, Business Opportunities Reference Architecture for E-Commerce.” In 2015 International Conference on Green Computing and Internet of Things (ICGCIoT), 1577–81, 2015. https://doi.org/10.1109/ICGCIoT.2015.7380718.

[3] “Internet of Things: Six Key Characteristics.” Frog (blog), August 23, 2014. https://designmind.frogdesign.com/2014/08/internet-things-six-key-characteristics/.

[4] “5 Major Types of DOS Attack | Learn & Know Them.” Concise Courses, December 13, 2018. https://www.concise-courses.com/5-major-types-of-dos-attack/.

[5] “ตัวอย่างประเภทการโจมตีในระบบเครือข่าย Internet.” sssangtopo (blog), October 19, 2012. https://sssangtopo.wordpress.com/about/ตัวอย่างประเภทการโจมตี/

[6] Richard, J. R. and Geatz, M. W. (2003). Data Mining a Tutorial-Based Primer.Pearson Education Inc.

[7] “สมการการประมาณค่านัยทั่วไปสำหรับการวิเคราะห์ข้อมูลเชิงกลุ่ม.” Accessed August 15, 2019. http://www.stat.mju.ac.th/nucs2018/Stat_Html/notice_3.php.

[8] จิราภา เลาหะวรนันท์, รชต ลิ้มสุทธิวันภูมิ และ บัณฑิต ฐานะโสภณ. [ม.ป.ป]. การใช้เทคนิคการทำเหมืองข้อมูลในการจำแนกและคัดเลือกแขนงวิชาสำหรับนักศึกษาคณะเทคโนโลยีสารสนเทศ. กรุงเทพฯ : คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

[9] “มหาวิทยาลัยเทคโนโลยีมหานคร.” Accessed August 15, 2019. http://www.mut.ac.th/research-detail-92.

[10] Udovychenko, Y., A. Popov, and I. Chaikovsky. “Ischemic Heart Disease Recognition by K-NN Classification of Current Density Distribution Maps.” In 2015 IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO), 402–5, 2015. https://doi.org/10.1109/ELNANO.2015.7146919.

[11] จิราพร วังหอม, ญาดา พรภักดี, เนรัญชิรา สาครเจริญ, วิราวรรณ พุทธมาตย์, สนธยา เลิศสงคราม, อุทุมพร บุญจำรูญ. [ม.ป.ป.]. ค้นเมื่อ 20 พฤศจิกายน 2561, จาก http://home.kku.ac.th/wichuda/DMining/Sudent/DT.ppt

[12] ศริญญา มาปลูก. (2559). การวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining โดยซอฟต์แวร์ RapidMiner Studio 6. ค้นเมื่อ 20 พฤศจิกายน 2561, จาก https://erp.mju.ac.th/acticleDetail.aspx?qid=579

[13] Swets, J. (1988). Measuring the accuracy of diagnostic systems. Science, 240, 1285–1293. http://wixtedlab.ucsd.edu/publications/Psych%20218/Swets_1988.pdf)

[14] มุสิกะวัน, ภัคราช, สิรภัทร เชี่ยวชาญวัฒนา, and คำรณ สุนัติ. “เครื่องจักรเรียนรู้เอ็กซ์ตรีมเชิงวิวัฒนาการโดยใช้พื้นฐานของขั้นตอนการค้นหาค่าที่เหมาะสมแบบสเต็บไซส์แรนดอมและขั้นตอนวิธีแบบหิ่งห้อย.” Information Technology Journal 8, no. 2 (2012): 26–32.

[15] “Confusion Matrix.” Plagad’s Blog (blog), August 26, 2010. https://plagad.wordpress.com/2010/08/26/confusion-matrix/.

[16] Satangmongkol, Kasidis. “อธิบาย 10 Metrics พื้นฐานสำหรับวัดผลโมเดล Machine Learning.” DataRockie, March 30, 2019. https://datarockie.com/2019/03/30/top-ten-machine-learning-metrics/.

[17] รวมทรัพย์, ธวัช. “การจำแนกใบพืชโดยการใช้คุณลักษณะรูปทรงและพื้นผิวของใบพืชด้วย ขั้นตอนวิธีเชิงพันธุกรรม.” Thesis, ธวัช รวมทรัพย์, 2560. http://dspace.spu.ac.th/handle/123456789/5155.

[18] Uma, K. V., and E. S. Blessie. “Survey on Android Malware Detection and Protection Using Data Mining Algorithms.” In 2018 2nd International Conference on 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 209–12, 2018. https://doi.org/10.1109/I-SMAC.2018.8653720.

[19] Eziama, E., L. M. S. Jaimes, A. James, K. S. Nwizege, A. Balador, and K. Tepe. “Machine Learning-Based Recommendation Trust Model for Machine-to-Machine Communication.” In 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 1–6, 2018. https://doi.org/10.1109/ISSPIT.2018.8705147..

[20] Sahay, R., G. Geethakumari, K. Modugu, and B. Mitra. “Traffic Convergence Detection in IoT LLNs: A Multilayer Perceptron Based Mechanism.” In 2018 IEEE Symposium Series on Computational Intelligence (SSCI), 1715–22, 2018. https://doi.org/10.1109/SSCI.2018.8628921.

[21] Mirsky, Yisroel, Tomer Doitshman, Yuval Elovici, and Asaf Shabtai. “Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection.” ArXiv:1802.09089 [Cs], February 25, 2018. http://arxiv.org/abs/1802.09089.

[22] “Malware on IoT Dataset.” Stratosphere IPS. Accessed September 25, 2019. https://www.stratosphereips.org/datasets-iot.

ชื่อ นางสาวฑิตยา ศรีวุฒิทรัพย์ รหัสประจำตัว 593021270-9
Miss Titaya Sriwuttisap
นักศึกษาระดับปริญญาตรี วิทยาการคอมพิวเตอร์
อาจารย์ที่ปรึกษาโครงงาน ผศ.ดร.ชิตสุธา สุ่มเล็ก
Project Advisor Asst. Prof. Chitsutha Soomlek

ชื่อหัวข้อโครงงาน

ภาษาไทย การวิเคราะห์รูปแบบการโจมตีโดยปฏิเสธการให้บริการบนเครือข่ายอินเทอร์เน็ตประสานสรรพสิ่งด้วยเทคนิคการเรียนรู้ของเครื่อง

ภาษาอังกฤษ Pattern Analysis of Denial of Service Attack in Internet of Things Networks using Machine Learning

ผลลัพธ์การทดลอง

นำข้อมูลแหล่งที่ 1 [1] และแหล่งที่ 2 [2] มาเตรียมการดังนี้

1. ไฟล์ที่ได้จากต้นฉบับเป็นไฟล์ Pcap ดังนั้นทำการส่งออกไฟล์เป็นไฟล์ CSV โดยแอตตริบิวท์ที่นำมาใช้ คือ Label ,Time ,Protocol ,Length ,Sourceport ,Winsize ,SEQ ,FLAGACK ,TCPACK , timestamptsecr ,sack_re ,FLAG
   - Label 0 หมายถึง การใช้งานทั่วไป
     1 หมายถึง การโจมตี
   - FLAGACK 1 หมายถึง not set
     2 หมายถึง set
2. ข้อมูลสำหรับฝึก แหล่งที่ 1 และ 2 นำข้อมูลการใช้งานทั่วไปมาจำนวน 2500 นำข้อมูลการโจมตีมาจำนวน 2500
3. ข้อมูลสำหรับทดสอบ แหล่งที่ 1 และ 2 นำข้อมูลการใช้งานทั่วไปมาจำนวน 1250 นำข้อมูลการโจมตีมาจำนวน 1250
4. ทำการนอร์มอลไลซ์ข้อมูลให้อยู่ระหว่าง 0 – 1
5. ช่องที่ว่างทั้งหมดทำการเติม 2
6. ทำการฝึกข้อมูลโดยวิธี k-fold คือนำข้อมูลแบ่งเป็น 10 ส่วน 9 ส่วนใช้เทรนนิ่ง อีกหนึ่งส่วนใช้ทดสอบ ทำจนครบสิบรอบจะได้ค่าเฉลี่ย Train accuracy
7. ทำการทดสอบโมเดลที่ได้โดยใช้ข้อมูลที่เตรียมไว้ค่าเฉลี่ย Test accuracy

	Decision Tree	Random Forest	Neural Networks	knn	LogisticRegression	Naïve Bayes	Support Vector Machine	ELM
Time	0.11359357833862305	0.07579660415649414	7.83453822135925 3	0.11783933639526367	0.44630789756774 9	0.00299072265625	1.26262402534484 86	0.03723073005676 2695
accuracy	94.67000000000002	94.72	76.78	92.94	76.51	74.36	76.66999999999999	49.84999999999994

อ้างอิง

[1] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, “Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection,” in NDSS, 2018.

[2] “CTU-13 Dataset.” Stratosphere IPS. Accessed September 10, 2019. https://www.stratosphereips.org/datasets-ctu13.