บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงคุณลักษณะทั่วไปของงานที่เกี่ยวข้องกับระบบการตรวจจับการบุกรุกในเครือข่าย รวมทั้งทฤษฎีการประยุกต์ข้อมูลทางสถิติโดยเฉพาะการวิเคราะห์องค์ประกอบที่สามารถนำมาพัฒนาและประยุกต์ขั้นตอนวิธีในการคัดเลือกคุณลักษณะที่เหมาะสมสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย

2.1 ระบบการตรวจจับการบุกรุกในเครือข่าย

Intrusion Detection คือการกลั่นกรองหรือตรวจสอบข้อมูลกิจกรรมต่าง ๆ ที่อาจจะเป็นภัยคุกคามต่อเครื่องหรือระบบเครือข่ายคอมพิวเตอร์ ดังนั้นระบบตรวจจับผู้บุกรุกก็คือซอฟต์แวร์หรือฮาร์ดแวร์ที่ใช้เพื่อให้บรรลุวัตถุประสงค์ดังกล่าว ซึ่งจะคอยทำหน้าที่เฝ้าตรวจสอบข้อมูลต่าง ๆ ที่วิ่งเข้าและออกบนเครือข่ายคอมพิวเตอร์ และเมื่อมีข้อมูลที่ต้องสงสัยระบบจะระบุว่าเป็นการบุกรุกพร้อมกับรายงานต่อไปยังผู้ดูแลระบบ (Administrator) และดำเนินการตามเงื่อนไขต่าง ๆ ตามที่ผู้ดูแลระบบได้กำหนดไว้ การบุกรุก (Intrusion) คือความพยายามที่จะกระทำการใดๆที่ส่งผลต่อความสมบูรณ์ (Integrity) ความลับ (Confidentiality) และความพร้อมใช้งาน (Availability) ของทรัพยากรต่าง ๆ บนระบบเครือข่าย การบุกรุกจะถูกกระทำโดยผู้ไม่ประสงค์ดีต่อระบบซึ่งอาจหมายถึงแฮกเกอร์ (Hacker), แคร็กเกอร์ (Cracker) หรือบุคคลกลุ่มอื่นที่ไม่หวังดีต่อระบบเครือข่าย โดยในระบบตรวจจับผู้บุกรุกจะเรียกบุคคลกลุ่มดังกล่าวว่า ผู้บุกรุก (Intruder) ซึ่งการบุกรุกของผู้บุกรุกสามารถกระทำได้จากภายในเครือข่าย (Inside) หมายถึงผู้บุกรุกที่มีสิทธิ์ในระบบเครือข่ายภายในที่พยายามจะใช้งานทรัพยากรระบบเกินอำนาจสิทธิ์ที่ตนได้รับ และผู้บุกรุกภายนอกเครือข่าย (Outside) หมายถึงผู้บุกรุกจากภายนอกเครือข่ายที่พยายามจะเข้าใช้งานทรัพยากรระบบทั้งที่ไม่ได้รับสิทธิ์ และบุคคลที่พยายามจะกระทำการใด ๆ ที่ไม่เป็นผลดีต่อระบบเครือข่าย จากภายนอก

ประเภทของระบบตรวจจับผู้บุกรุก

ในการจัดประเภทของระบบตรวจจับผู้บุกรุกสามารถแบ่งได้หลายประเภท โดยใช้หลักเกณฑ์ต่าง เช่น แหล่งข้อมูลที่นำมาใช้วิเคราะห์ หรือแนวทางในการใช้ตรวจจับการบุกรุก

รวมทั้ง ช่วงหรือระยะเวลาที่นำมาใช้เพื่อทำการวิเคราะห์การบุกรุก แต่ส่วนใหญ่แล้วจะแบ่งประเภทของระบบตรวจจับผู้บุกรุกออกเป็น 3 ประเภท ได้แก่ Host-Based Intrusion Detection System (HIDS), Network-Based Intrusion Detection System (NIDS) และ Application-Based Intrusion Detection System (AIDS)

HIDS คือระบบนี้จะทำงานอยู่บนเครื่องคอมพิวเตอร์แต่ละเครื่อง โดยจะทำการตรวจสอบข้อมูลที่ผ่านเข้าและออกคอมพิวเตอร์แต่ละเครื่องเพื่อหาว่าโปรแกรมหรือผู้ใช้คนใดที่ทำให้เกิดการบุกรุกขึ้นบนระบบ

NIDS คือระบบนี้จะทำหน้าที่ตรวจจับข้อมูลทั้งหมดที่มีการไหลเข้าและออกบนระบบเครือข่ายแล้วทำการวิเคราะห์ว่ากิจกรรมใดบนเครือข่ายเป็นการบุกรุกหรือพยายามที่จะบุกรุกหรือไม่ โดยอาศัยค่าต่าง ๆ อาทิเช่น ปริมาณข้อมูลบนเครือข่าย, ลักษณะของเพ็กเก็จที่ส่งเข้ามาในเครือข่าย เป็นต้น

AIDS ระบบนี้จะทำงานคล้ายกับ HIDS แต่จะมีส่วนของการรวบรวมข้อมูลจากการทำงานของโปรแกรมประยุกต์ที่ทำงานบนเครื่องคอมพิวเตอร์เพิ่มเข้ามา เพื่อใช้สำหรับการวิเคราะห์และตรวจสอบหาพฤติกรรมที่น่าสงสัยหรือผิดปกติที่เกิดขึ้น

แนวทางในการตรวจจับการบุกรุก

แนวทางในการใช้ในการตรวจสอบการบุกรุกมีอยู่ 2 แนวทางได้แก่ Anomaly Detection และ Misuse Detection โดยแนวทางแรกสำหรับ Anomaly Detection นั้นเป็นการตรวจจับและค้นหาพฤติกรรมต่าง ๆ ของผู้ใช้ที่เปลี่ยนแปลงไปจากสภาวะการใช้งานปกติ ตัวอย่างเช่นการที่ผู้ใช้ส่วนใหญ่จะมีการเข้าถึงข้อมูลหรือทรัพยากรต่าง ๆ เฉพาะช่วงกลางวันเท่านั้น แต่ถ้าหากวันใดได้มีการใช้งานในช่วงกลางคืนของผู้ใช้บางคนก็จะถูกสงสัยว่าเป็นการบุกรุก เป็นต้น ขณะที่แนวทางที่สองสำหรับ Misuse Detection จะตรวจจับโดยอาศัยรูปแบบที่ถูกกำหนดหรือสร้างขึ้นไว้แล้วว่าพฤติกรรมรูปแบบใดถูกระบุว่าเป็นผู้บุกรุก ซึ่งรูปแบบดังกล่าวจะถูกนำมาเปรียบเทียบค้นหากับเหตุการณ์ต่างๆที่เกิดขึ้นในระบบซึ่งถ้าพฤติกรรมใดที่เกิดขึ้นตรงกันกับรูปแบบพฤติกรรมที่ได้กำหนดไว้ก็จะถูกระบุว่าเป็นการบุกรุก แนวทางนี้เป็นแนวทางที่จะต้องอาศัยผู้เชี่ยวชาญทางด้านนี้เป็นผู้กำหนดรูปแบบพฤติกรรมการบุกรุกไว้ล่วงหน้าแล้ว ดังนั้นแนวทางนี้อาจจะถูกเรียกอีกชื่อหนึ่งว่า Signature-Based Detection

งานวิจัยระบบการตรวจจับการบุกรุกในเครือข่ายส่วนใหญ่จะใช้คุณลักษณะทั้งหมดจำนวน 41 ตัวแปรในการหาวิธีใหม่ๆ เพื่อจัดประเภทของการบุกรุกในเครือข่ายโดยมีการปรับค่าน้ำหนักของแต่ละคุณลักษณะ ซึ่งงานวิจัยในระยะหลังเริ่มให้ความสนใจกระบวนการแปลงค่าคุณลักษณะก่อนดำเนินการจัดประเภทของการบุกรุกในเครือข่าย แสดงว่าคุณลักษณะบางตัวอาจเป็นตัวถ่วงต่อระยะเวลาการประมวลผลและลดประสิทธิภาพการตรวจจับการบุกรุกในเครือข่ายลง แต่มีงานวิจัยส่วนน้อยที่เริ่มสนใจคัดเลือกเฉพาะคุณลักษณะที่เด่นเท่านั้นเพื่อการลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพการตรวจจับการบุกรุกในเครือข่าย ซึ่งขั้นตอนวิธีในการคัดเลือกคุณลักษณะเด่นนี้ยังซับซ้อน

2.2 การวิเคราะห์องค์ประกอบ (Factor Analysis)

เป็นการวิเคราะห์หลายตัวแปรเทคนิคหนึ่งเพื่อการสรุปรายละเอียดของตัวแปรหลายตัวหรือเรียกว่าเป็นเทคนิคที่ใช้ในการลดจำนวนตัวแปรเทคนิคหนึ่งโดยการศึกษาถึงโครงสร้างความสัมพันธ์ของตัวแปร และสร้างตัวแปรใหม่เรียกว่า องค์ประกอบ โดยองค์ประกอบที่สร้างขึ้นจะเป็นการนำตัวแปรที่มีความสัมพันธ์กันหรือมีความร่วมกันสูงมารวมกันเป็นองค์ประกอบเดียวกัน ส่วนตัวแปรที่อยู่คนละองค์ประกอบมีความร่วมกันน้อย หรือไม่มีความสัมพันธ์กันเลย ในเทคนิคนี้จะใช้ค่าสัมประสิทธิ์สหสัมพันธ์วัดความสัมพันธ์ ระหว่างตัวแปร ดังนั้นการวิเคราะห์องค์ประกอบจึงเป็นเทคนิคการลดจำนวนตัวแปร จากจำนวนตัวแปรมากๆ ให้เหลือเพียงไม่กี่ปัจจัย หรือ ตัวแปร

2.3 งานวิจัยที่เกี่ยวข้อง

งานวิจัยทาง NIDS ก็พยายามพัฒนาขั้นตอนวิธี เพื่อเพิ่มค่าความแม่นยำในการทำนาย โดยงานวิจัยส่วนใหญ่เลือกใช้ data mining และ machine learning มีดังนี้

คุณ Chi-Hoon Lee และคณะ [12] ได้นำเสนองานวิจัยเรื่อง “Network Intrusion Detection through Genetic Feature Selection” งานวิจัยนี้เสนอวิธีใหม่ในการคัดเลือกคุณสมบัติร่วมสูงสุดระหว่าง รูปแบบพฤติกรรมปกติและรูปแบบพฤติกรรมที่เป็นการบุกรุกระบบเครือข่ายงาน โดยงานวิจัยนี้นำเสนอวิธีการคัดเลือกคุณสมบัติที่ดีที่สุดจากการใช้หลักการคัดเลือกพันธุกรรม ซึ่งสามารถเพิ่มระดับอัตราความแม่นยำของระบบการตรวจจับการบุกรุกเครือข่าย ซึ่งได้ทำการทดลองกับชุดข้อมูล KDD CUP 1999 จาก MIT Lincoln Labs ผลการ

ทดลองได้เลือกคุณลักษณะเด่นจำนวน 21 ตัว จากทั้งหมด 41 ตัว (คิดเป็น 51.22%) แต่ข้อเสียของการใช้หลักการคัดเลือกพันธุกรรมในการคัดเลือกคุณลักษณะเด่นใช้เวลาในการประมวลผลนานมาก

คุณ Kok-Chine Khor และคณะ [24] ได้นำเสนองานวิจัยเรื่อง “A Feature Approach for Network Intrusion Detection” งานวิจัยนี้เสนอวิธีใหม่ในการคัดเลือกคุณสมบัติโดยใช้ขบวนการวิธีทาง data mining ที่ซับซ้อน สามารถเพิ่มระดับอัตราความแม่นยำของระบบการตรวจจับการบุกรุกเครือข่าย ซึ่งได้ทำการทดลองกับชุดข้อมูล KDD CUP 1999 จาก MIT Lincoln Labs ผลการทดลองได้เลือกคุณลักษณะเด่นจำนวน 7 ตัว จากทั้งหมด 41 ตัว (คิดเป็น 17.07%) แต่ข้อเสียคือไม่สามารถตรวจจับการบุกรุกชนิด R2L ได้