บทที่ 1

บทนำ (Introduction)

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันมีการใช้ internet เป็นจำนวนมาก และจำนวนผู้ใช้มีขนาดเพิ่มขึ้นในลักษณะก้าวกระโดด (exponential) การรักษาความปลอดภัยของข้อมูลถือว่าเป็นปัญหาที่สำคัญยิ่ง ระบบการตรวจจับการบุกรุกในเครือข่ายเป็นระบบที่คอยตรวจพฤติกรรมการใช้ที่ผิดปกติหรือผิดไปจากกฎเกณฑ์ที่ตั้งไว้จะสันนิฐานว่าระบบถูกบุกรุก จึงเป็นระบบที่สำคัญพื้นฐานสำหรับงานรักษาความปลอดภัยของข้อมูล เนื่องด้วยปริมาณการใช้งานบนระบบเครือข่ายได้มีปริมาณที่เพิ่มขึ้นอย่างรวดเร็ว ปัญหาคือต้องพยายามหาระบบการตรวจจับการบุกรุกในเครือข่ายที่มีการประมวลผลที่รวดเร็ว และมีความผิดพลาดในการแจ้งเตือนน้อยที่สุด งานวิจัยระบบการตรวจจับการบุกรุกในเครือข่ายส่วนใหญ่จะใช้คุณลักษณะทั้งหมดจำนวน 41 ตัวแปรในการหาวิธีใหม่ๆ ในการจัดประเภทของการบุกรุกในเครือข่ายโดยมีการปรับค่าน้ำหนักของแต่ละคุณลักษณะ ซึ่งงานวิจัยในระยะหลังเริ่มให้ความสนใจกระบวนการแปลงค่าคุณลักษณะก่อนดำเนินการจัดประเภทของการบุกรุกในเครือข่าย แสดงว่าคุณลักษณะบางตัวอาจเป็นตัวถ่วงต่อระยะเวลาการประมวลผลและลดประสิทธิภาพการตรวจจับการบุกรุกในเครือข่ายลง แต่มีงานวิจัยส่วนน้อยที่เริ่มสนใจคัดเลือกเฉพาะคุณลักษณะที่เด่นเท่านั้นเพื่อการลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพการตรวจจับการบุกรุกในเครือข่าย ซึ่งขั้นตอนวิธีในการคัดเลือกคุณลักษณะเด่นนี้ยังซับซ้อน และประมวลผลช้าอยู่ เปรียบเทียบได้กับการวิเคราะห์ผู้เป็นโรคเบาหวานจากการวิเคราะห์ผลเลือด ขณะที่เลือดประกอบด้วยสารเคมีหลายชนิดมาก จะเลือกพิจารณาเฉพาะสารเคมีตัวที่บ่งบอกภาวะน้ำตาลเท่านั้น การวิจัยการตลาดจะทำการค้นหาตัวแปรที่เป็นปัจจัยที่สำคัญก่อนจะสร้างตัวแบบแก้ปัญหาสิ่งที่สนใจ ดังนั้นการคัดเลือกคุณลักษณะที่เหมาะสม (อาจจะมีจำนวนน้อยกว่า 41 ตัวแปร) จึงเป็นสิ่งที่สำคัญก่อนการหาวิธีจัดประเภทชนิดของการบุกรุกในระบบการตรวจจับการบุกรุกในเครือข่าย ผู้วิจัยต้องการวิจัยค้นหาวิธีที่ดีที่สุดในการคัดเลือกคุณลักษณะที่เหมาะสมที่สุดสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย ที่สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพความแม่นยำการตรวจจับการบุกรุกในเครือข่าย

วิเคราะห์องค์ประกอบ (factor analysis) ใช้ในการลดตัวแปรที่ผู้วิจัยสนใจศึกษาโดยอาจกระทำการรวมตัวแปรย่อยๆ ให้เป็นตัวแปรใหญ่ขึ้นมาใหม่ ซึ่งตัวแปรย่อยที่ถูกรวมเป็นตัวแปร

ใหม่ก็จะมีการให้น้ำหนักแก่ตัวแปรย่อยเหล่านั้น หรือการเลือกแต่เฉพาะตัวแปรเด่นที่มีความสำคัญมากสุดในแต่ละกลุ่ม โดยตัดตัวแปรย่อยๆ ในแต่ละกลุ่มออก ซึ่งวิธีการดังกล่าวก็น่าจะสามารถนำมาใช้เป็นขั้นตอนวิธีในการคัดเลือกเฉพาะคุณสมบัติที่เด่น ผู้วิจัยจึงอยากศึกษาว่าวิธีการวิเคราะห์องค์ประกอบเป็นวิธีที่ดีที่สุดในการคัดเลือกคุณลักษณะที่เหมาะสมที่สุดสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย ที่สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพความแม่นยำการตรวจจับการบุกรุกในเครือข่ายหรือไม่

1.2 วัตถุประสงค์ของการวิจัย

งานวิจัยนี้ต้องการหาวิธีที่ดีที่สุดในการคัดเลือกคุณลักษณะที่เหมาะสมที่สุดสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย ต้องการขั้นตอนวิธีที่ง่ายไม่ซับซ้อนที่สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพความแม่นยำการตรวจจับการบุกรุกในเครือข่าย

1.3 สมมติฐานของการศึกษา

วิเคราะห์องค์ประกอบสามารถใช้ในการคัดเลือกคุณลักษณะที่เหมาะสมที่สุดสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย ขั้นตอนวิธีต้องง่ายไม่ซับซ้อนที่สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพความแม่นยำการตรวจจับการบุกรุกในเครือข่าย

1.4 ขอบเขตการวิจัย

ศึกษาวิธีการวิเคราะห์องค์ประกอบว่าเป็นวิธีที่ดีที่สุดในการคัดเลือกคุณลักษณะที่เหมาะสมที่สุดสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย ที่สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพความแม่นยำการตรวจจับการบุกรุกในเครือข่ายได้หรือไม่ โดยทำการทดลองกับฐานข้อมูลจำนวน 494,020 รายการ จาก KDD Cup 1999 โดยทำการเปรียบเทียบผลทดลองที่ได้จากวิธีที่นำเสนอกับผลการทดลองที่ได้จากวิธีที่ได้จากงานวิจัยที่ผ่านมาพร้อมทั้งเปรียบเทียบผลการทดลองจากคุณสมบัติทั้งหมดทุกตัว

1.5 ขั้นตอนการศึกษาและดำเนินงานวิจัย

งานวิจัยนี้มีขั้นตอนการศึกษาและดำเนินงานวิจัยดังนี้

1. ศึกษางานวิจัยที่เกี่ยวข้องด้านการคัดเลือกคุณลักษณะที่เหมาะสมสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย
2. ศึกษาการวิเคราะห์องค์ประกอบ
3. ตั้งสมมติฐานโดยคาดว่าวิธีวิเคราะห์องค์ประกอบสามารถใช้ในการคัดเลือกคุณลักษณะที่เหมาะสมที่สุดสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย ขั้นตอนวิธีต้องง่ายไม่ซับซ้อนที่สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพความแม่นยำการตรวจจับการบุกรุกในเครือข่าย
4. ฐานข้อมูลจะใช้ฐานข้อมูลจำนวน 494,020 รายการ จาก KDD Cup 1999
5. ทำการเปรียบเทียบผลทดลองที่ได้จากวิธีที่นำเสนอกับผลการทดลองที่ได้จากวิธีที่ได้จากงานวิจัยที่ผ่านมาพร้อมทั้งเปรียบเทียบผลการทดลองจากคุณสมบัติทั้งหมดทุกตัว

1.6 ประโยชน์ที่คาดว่าจะได้รับ

สามารถหาขั้นตอนวิธีใช้ในการคัดเลือกคุณลักษณะที่เหมาะสมที่สุดสำหรับระบบการตรวจจับการบุกรุกในเครือข่าย ขั้นตอนวิธีต้องง่ายไม่ซับซ้อนที่สามารถลดเวลาในการประมวลผลและเพิ่มประสิทธิภาพความแม่นยำการตรวจจับการบุกรุกในเครือข่าย