บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ลักษณะของข้อมูลที่ใช้ในการทำแบบทดลอง

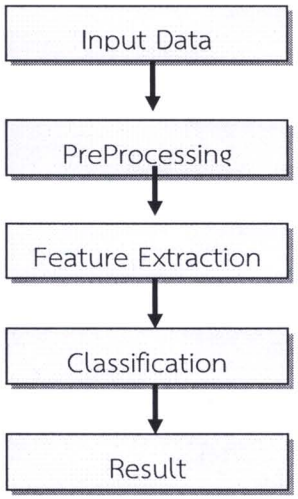
ข้อมูลที่นำมาใช้ในการทำแบบทดลอง เป็นข้อมูลที่ได้จากฐานข้อมูลความรู้ (Knowledge Discovery in Database (KDD) Cup data) ซึ่งเป็นชุดข้อมูลในปี 1999 โดยชุดข้อมูลนี้ได้มาจากความร่วมมือของโครงการวิจัยและพัฒนาเพื่อการทหารของประเทศสหรัฐอเมริกา (Defense Advanced Research Projects Agency: DARPA) ซึ่งร่วมมือกับทางมหาวิทยาลัยเมตซาชูเซตส์ สหรัฐอเมริกา ชุดข้อมูลนี้ถูกสร้างตามการจำลองการโจมตีของผู้บุกรุกจาก U.S. Air Force local area network ตั้งขึ้นที่ Lincoln Labs โดยข้อมูลนั้นมีระยะเวลาในการจัดทำนานถึง 9 สัปดาห์ จากการเก็บข้อมูลจากแพ็กเก็ต TCP ผ่านโปรแกรม TCP Dump ประกอบด้วยข้อมูลขนาดใหญ่มาก ซึ่งมี 41 แอทริบิวต์ (มิติ) ที่ได้จากแพ็กเก็ต TCP (Raw TCP Packet) รวมถึงชนิดของโปรโตคอลซึ่งมีค่า “TCP”, “ICMP”, “UDP” ซึ่งเป็นแอทริบิวต์ที่มีความต่อเนื่องเป็นในลักษณะข้อความ (Nominal) ซึ่งจะมีสถานะ (Label) กำกับไว้เสมอในแต่ละบรรทัด (Record) ว่าข้อมูลชุดนี้เป็นสถานะปกติ (Normal) หรือว่าเป็นชุดข้อมูลที่ถูกโจมตี (Malicious) ทางผู้จัดทำได้คัดเลือก ชุดข้อมูลที่มีมากถึง 5 ล้านบรรทัด ออกมาประมาณ 10% เท่านั้นเพื่อสะดวกในการจัดทำชุดข้อมูลเรียนรู้ (Training Set) และชุดข้อมูลทดสอบ (Test Set)

ชุดข้อมูลที่ได้นำมาทำแบบทดลองนี้ อยู่ในรูปแบบของเครื่องกลเรียนรู้ (Machine Learning Pattern) โดยสามารถแบ่งออกเป็นกลุ่มหลักๆ ได้ 5 กลุ่มคือ Normal, DoS, Probe, R2L และ U2R โดยแต่ละกลุ่มยังมีชนิดของข้อมูลย่อย ๆ อีก โดยแอทริบิวต์ขวามือสุดคือ คลาสแอทริบิวต์ (Class Attribute) ที่เป็นตัวบอกสถานะว่าถูกโจมตีหรือไม่ ซึ่งหากไม่ถูกโจมตีจะมีค่าเป็นปกติ

2.2 กระบวนการรู้จำทางคอมพิวเตอร์

โดยทั่วไปกระบวนการรู้จำทางคอมพิวเตอร์นั้น จะมีขั้นตอนการทำงานหลักๆ คือ การประมวลผลเบื้องต้น (Processing) การวิเคราะห์แยกองค์ประกอบเฉพาะของข้อมูล (Feature Extraction) และการจำแนกข้อมูล (Classification) แสดงดังรูปที่ 2-1 ซึ่งในแต่ละส่วนมีรายละเอียดเบื้องต้นดังนี้

Input Data เป็นข้อมูลที่มีลักษณะหลายรูปแบบตามความต้องการของระบบ เช่น ในการรู้จำตัวอักขระ อาจจะใช้ภาพที่มีลักษณะเป็นข้อความบรรทัดเดียว ข้อความหลายบรรทัด หรือ ภาพตัวอักขระจำนวน 1 อักขระ เป็นต้น ซึ่งข้อมูลอาจได้จากการเก็บข้อมูล หรือนำเอาเอกสารที่เป็นกระดาษไปสแกนเพื่อเปลี่ยนเป็นข้อมูลทางคอมพิวเตอร์ เช่น แฟ้มข้อมูลชนิด txt , xls, bitmap หรือ ได้จากการป้อนข้อมูลผ่านอุปกรณ์อินพุต เช่น เมาส์หรือปากกาอิเล็กทรอนิกส์



รูปที่ 2-1 กระบวนการรู้จำสำหรับปัญหา Intrusion Detection

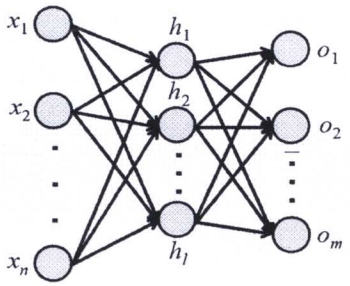
Preprocessing เป็นการประมวลผลเบื้องต้นเพื่อปรับเปลี่ยนลักษณะรูปแบบบางอย่างของข้อมูลอินพุต ทั้งนี้เพื่อปรับอินพุตให้มีความเหมาะสมและตรงตามที่ระบบต้องการ เช่น ปรับขนาด (Resize) ปรับลดจำนวนมิติ (Reduce Dimension) หรือการกำจัดสัญญาณรบกวน (Noise Remove)

Feature Extraction เป็นขั้นตอนของการสกัดเอาลักษณะเฉพาะของแต่ละอินพุตออกมาเป็นเวกเตอร์เพื่อนำไปใช้เป็นอินพุตในการเรียนรู้ระบบและทดสอบระบบ

Classification เป็นขั้นตอนในการจำแนกและตัดสินใจว่าอินพุตที่เข้ามานั้นเป็นการบุกรุกแบบใด โดยในขั้นตอนนี้มีหลายวิธีด้วยกัน เช่นการเปรียบเทียบอินพุตกับโครงสร้างของตัวต้นแบบการบุกรุกในฐานข้อมูลการเปรียบเทียบอินพุตกับกฎเพื่อการตัดสินใจ การใช้โครงข่ายประสาทเทียม หรือการใช้ตัวแบบฮิดเดนมาร์คอฟ เป็นต้น

2.2.1 ระบบโครงข่ายประสาทเทียมแบบวิธีการแพร่กระจายย้อนกลับ (Back propagation Algorithm)

ขั้นตอนวิธีการแพร่กระจายย้อนกลับ เป็นขั้นตอนวิธีที่ใช้ในการเรียนรู้ของเครือข่ายประสาทเทียมวิธีหนึ่งที่นิยมใช้ในโครงข่ายประสาทเทียมหลายชั้น (Multilayer neural network) เพื่อใช้ในการปรับค่าน้ำหนักในเส้นเชื่อมต่อระหว่างโหนดให้เหมาะสม โดยการปรับค่านี้จะขึ้นกับความแตกต่างของค่าเอาต์พุตที่คำนวณได้กับค่าเอาต์พุตที่ต้องการ พิจารณารูปต่อไปนี้ประกอบ



รูปที่ 2-2 ตัวอย่างข่ายงานประสาทเทียมแบบหลายชั้น

ตัวอย่างในรูปด้านบนแสดงข่ายงานป้อนไปหน้าแบบหลายชั้นซึ่งประกอบไปด้วยชั้นอินพุต ชั้นฮิดเดนหรือชั้นซ่อน และชั้นเอาต์พุต ในรูปแสดงชั้นฮิดเดนเพียงชั้นเดียวแต่อาจมีมากกว่าหนึ่งชั้นก็ได้ เส้นเชื่อมจะเชื่อมต่อเป็นชั้น ๆ ไม่ข้ามชั้นจากชั้นอินพุตไปชั้นฮิดเดน ถ้ามีชั้นฮิดเดนมากกว่าหนึ่งชั้นก็เชื่อมต่อกันไป และสุดท้ายจากชั้นฮิดเดนไปชั้นเอาต์พุต

ในการปรับค่าน้ำหนักโดยขั้นตอนวิธีการแพร่กระจายย้อนกลับนั้น เราต้องนิยามค่าผิดพลาดสำหรับการเรียนรู้ของข่ายงาน $MSE(\vec{w})$ จากนั้นจะหาค่าน้ำหนักที่ให้ค่าผิดพลาดต่ำสุด นิยามค่าผิดพลาดดังนี้

$$MSE(\vec{w}) = \frac{1}{2}\sum_{p\in P}\sum_{k\in outputs}(d_{p,k} - o_{p,k})^2 \qquad ...(1)$$

โดยที่ *outputs* คือ เซตของเอาต์พุตโหนดในข่ายงานประสาทเทียม $d_{p,k}$ และ $o_{p,k}$ เป็นค่าเอาต์พุตเป้าหมายและเอาต์พุตที่ได้จากข่ายงานประสาทเทียมตามลำดับของเอาต์พุตโหนดที่ k ของตัวอย่างที่ p ขั้นตอนการแพร่กระจายย้อนกลับจะค้นหาค่าน้ำหนักที่ให้ค่าผิดพลาดกำลังสองเฉลี่ยต่ำสุด

ขั้นตอนของ Back-propagation Algorithm มีดังนี้

Algorithm Backpropagation:

Start with randomly chosen weights;

while MSE is unsatisfactory

and computational bounds are not exceeded, do

for each input pattern x_p, $1 \leq p \leq P$,

Compute hidden node inputs ($net_{p,j}^{(1)}$);

Compute hidden node outputs ($x_{p,j}^{(1)}$);

Compute inputs to the output nodes ($net_{p,k}^{(2)}$);

Compute the network outputs ($o_{p,k}$);

Modify outer layer weights:

$$\Delta w_{k,j}^{(2,1)} = \eta(d_{p,k} - o_{p,k})\mathcal{S}'(net_{p,k}^{(2)})x_{p,j}^{(1)}$$

Modify weights between input & hidden nodes:

$$\Delta w_{j,i}^{(1,0)} = \eta \sum_{k} \left((d_{p,k} - o_{p,k})\mathcal{S}'(net_{p,k}^{(2)})w_{k,j}^{(2,1)} \right) \mathcal{S}'(net_{p,j}^{(1)})x_{p,i}$$

end-for

end-while.

Note: if S is a logistic function, then
$S'(x) = S(x)(1 - S(x))$

2.2.2 วิธีการรู้จำแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine : SVM)

Support Vector Machine หรือ SVM จุดมุ่งหมายที่สำคัญของแนวคิด SVM คือการหาเส้นแบ่ง Hyperplane ซึ่งใช้แบ่งข้อมูลออกเป็นคลาส เพื่อให้ได้ผลลัพธ์ที่ดี โดยพิจารณาจากสมการเส้นตรง Hyper planes และ SVM จะทำการค้นหาจุดของข้อมูลที่อยู่ใกล้เส้นแบ่ง Hyper planes ซึ่งจุดนี้เรียกว่า "Support Vector" มีหลักการดังนี้

นำข้อมูลมาคำนวณหาค่า y ซึ่งค่า $y \in \{-1,1\}$ จากสมการ

$$y = w^T x + b \qquad \text{...(2)}$$

คำนวณหาเส้นแบ่ง ซึ่งเรียกว่าเส้น Optimal Hyperplane จากสมการ

$$w^T x + b = 0 \qquad \text{...(3)}$$

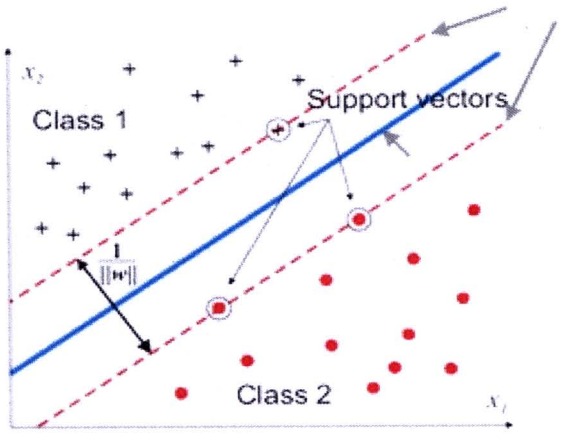
ระยะทาง (d) หรือ maximum margin จากเส้นขอบ ณ จุด x_i ไปยัง hyperplane แสดงดังสมการ

$$d = \frac{|w^T x_i + b|}{\|w\|} \quad \text{...(4)}$$

w คือ เวกเตอร์น้ำหนัก (Weight Vector)

x_i คือ Input

b คือ ค่าคงที่ที่กำหนดขึ้นเพื่อให้เหมาะสมกับการจัดกลุ่ม



รูปที่ 2-3 การแบ่งกลุ่มข้อมูลโดย Support Vector Machine

เลือกจุดที่อยู่ใกล้เส้นตรง Optimal Hyperplane ทั้งเหนือเส้นซึ่งเรียกว่า “ขอบล่าง” ซึ่งเป็นขอบล่างสุดของคลาสเอกสารที่อยู่เหนือเส้นตรง Optimal Hyperplane และใต้เส้นเรียกว่า “ขอบบน” ซึ่งเป็นขอบบนสุดของคลาสเอกสารที่อยู่ใต้เส้นตรง Optimal Hyperplane เพื่อที่จะหาระยะทางระหว่างเส้นขอบทั้งสองโดยจะเลือกเอาค่าระยะทางที่ห่างจากเส้นตรงOptimal Hyperplane ที่น้อยที่สุดเป็นตัวเลือกในการจัดกลุ่มเอกสาร

2.3 วิธีการวิเคราะห์องค์ประกอบหลัก

วิธีการวิเคราะห์องค์ประกอบหลัก Principal Component Analysis (PCA) เป็นวิธีการทางสถิติ เพื่อใช้ในการสกัดปัจจัยที่อาศัยหลักความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรที่ใช้เป็นข้อมูล องค์ประกอบหลักตัวแปร คือ การผสมเชิงเส้นตรง (Linear Combination) ของตัวแปรที่อธิบายการผันแปรของข้อมูลได้มากที่สุด จากนั้นหาการผสมเชิงเส้นครั้งที่สองที่สามารถอธิบายการผันแปรได้มากที่สุดเป็นอันดับที่สอง โดยที่ไม่สัมพันธ์กับการผสมครั้งแรก การวิเคราะห์องค์ประกอบหลักถูกนำไป

ประยุกต์ใช้งานต่างๆ เช่น การบีบอัดข้อมูล , การสร้างภาพใบหน้าไอเกนเพื่อใช้ในระบบจดจำ และ การลบออกของพื้นหลังโดยใช้ไอเกน เป็นต้นวิธีการวิเคราะห์องค์ประกอบหลักสามารถนำมาใช้ในการลดมิติของข้อมูลโดย การวิเคราะห์ข้อมูลและเลือกเฉพาะข้อมูลที่มีความสำคัญเท่านั้น ส่วนข้อมูลที่ไม่สำคัญจะถูกตัดทิ้งไป ดังนั้นเมื่อข้อมูลผ่านกระบวนการ *PCA* แล้ว จะได้ผลลัพธ์เป็นไอเกนเวกเตอร์และค่าไอเกน ซึ่งไอเกนเวกเตอร์ที่มีค่าสมนัยกับค่าไอเกนที่มีค่าสูงๆ จะเป็นการดึงข้อมูลที่มีความถี่ต่ำ ส่วนไอเกนเวกเตอร์ที่สมนัยกับค่าไอเกนที่ต่ำๆ จะเป็นการดึงข้อมูลที่มีความถี่สูง

2.3.1 การหาค่าไอเกน และไอเกนเวกเตอร์ (Eigen Value and Eigen Vector)

ความหมายของคุ่าไอเกน และไอเกนเวกเตอร์ กำหนดให้ A เป็นค่าเมทริกซ์จัตุรัส

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

และ *v* เป็นเวกเตอร์หลัก (Column Vector) และ λ เป็นค่าคงที่ใดๆ โดยที่

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_3 \end{bmatrix}$$

ที่ทำให้

$$Av = \lambda v \qquad \text{...(5)}$$

เมื่อ A แทน ค่าเมตทริกซ์

λ แทน เป็นค่าคงที่ใดๆ เป็นสเกลาร์

v แทน ค่าไอเกนเวกเตอร์

จากสมการจะเห็นว่า $v = 0$ ที่ทำให้สมการ เป็นจริงทุกๆ ค่าของ λ สมการที่ (5) อาจเขียนให้อยู่ในอีกรูปหนึ่งคือ

$$(\lambda I - A)v = \bar{0} \qquad \text{...(6)}$$

เมื่อ A แทน ค่าเมทริกซ์

I แทน เมทริกซ์เอกลักษณ์

λ แทน เป็นค่าคงที่ใดๆ เป็นสเกลาร์

v แทน ค่าไอเกนเวกเตอร์

เราจะคำนวณค่าไอเกน และเวคเตอร์ไอเกน ของสมการ (6) โดย

$$\det(\lambda I - A) = 0 \qquad \ldots(7)$$

จากนั้นก็ใช้วิธีแก้สมการแบบปกติ

2.4 การสกัดคุณลักษณะสำคัญ (*Feature Extraction)*

การสกัดคุณลักษณะสำคัญเป็นอีกขบวนการหนึ่งที่สำคัญมาก ตำราส่วนใหญ่จะแยกส่วนนี้ออกจากการประมวลผลเบื้องต้น คือจะอยู่ระหว่างขั้นตอนการประมวลผลเบื้องต้นกับขั้นตอนการรู้จำ การสกัดคุณลักษณะสำคัญเป็นการดึงเอาโครงสร้างพื้นฐานที่สำคัญของข้อมูลนั้นออกมา โดยโครงสร้างพื้นฐานที่ว่าจะต้องมีการกำหนดไว้ก่อนว่าจะมีอะไรบ้าง มีการนิยามอย่างไร ตัวอย่างเช่น สำหรับภาษาไทยเราอาจกำหนดว่าตัวอักษรภาษาไทยทั้งหมดประกอบด้วยโครงสร้างพื้นฐานคือ เส้นตรง (แนวตั้ง/นอน) เส้นเอียง หัว (วงกลม) ส่วนโค้ง ส่วนเว้า จุดแตกกิ่ง จุดตัด เป็นต้น เมื่อเราสามารถแยกเอาองค์ประกอบของตัวอักษรแต่ละตัวออกมาได้แล้ว จากนั้นเราก็นำเสนอรูปภาพของตัวอักษรนั้นในรูปแบบของรายการขององค์ประกอบพื้นฐานต่างๆ แทน ซึ่งจะถูกส่งต่อเป็นอินพุตสำหรับขั้นตอนการรู้จำต่อไป

2.5 การทบทวนวรรณกรรม/สารสนเทศ (*Information)* ที่เกี่ยวข้อง

Dong Seong Kim, Ha-nam Nguyen, Thanda Thein และ Jong Sou Park (2005) ได้นำเสนองานวิจัยเรื่อง An Optimized Intrusion Detection System Using PCA and BNN โดยได้นำเสนอการหาค่าที่เหมาะสมสำหรับการตรวจจับการบุกรุกโดยอาศัยการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA) และโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Backpropagation Neural Network: BNN) โดยมุ่งเน้นในการแก้ปัญหา 2 ปัญหาด้วยกันคือ การกำหนดจำนวนของ Hidden Layer และการจัดการค่าของน้ำหนัก เพื่อใช้ในการกำหนดรูปแบบของโครงข่ายประสาทเทียม และการประมวลผลข้อมูลที่ตรวจสอบที่มีปริมาณมาก โดยพิจารณาถึงการเพิ่มอัตราการตรวจจับและลดเวลาการประมวลผล โดยนำข้อดีของ Genetic Algorithm (GA) มาใช้ โดยการทำงานของ GA จะทำงานบนการทำงานที่รวมกันระหว่าง PCA และ BNN แต่ผลการทดลองยัง

ออกมาไม่เป็นที่น่าพอใจตามที่คาดหวังไว้ ในส่วนงานในอนาคตได้มีการชี้ถึงประเด็นว่า ถ้ามีการปรับเปลี่ยนตัว PCA และ BPN น่าจะทำให้ได้ผลการทดลองที่ดีขึ้น

Hai-Hua Gao, Hui-Hua Yang และ Xing-Yu Wang (2005) ได้นำเสนองานวิจัยเรื่อง Kernel PCA Based Network Intrusion Feature Extraction and Detection Using SVM โดยได้นำเสนอวิธีการใหม่ในการตรวจจับการบุกรุกด้วยการประยุกต์ Kernel Principal Component Analysis: KPCA สำหรับการสกัดคุณลักษณะและใช้ Support Vector Machine: SVM ในการแบ่งประเภท โดยทำการเปรียบเทียบผลกับข้อมูลที่ไม่ได้ผ่านการสกัดคุณลักษณะ และการสกัดคุณลักษณะด้วยวิธีการ PCA โดยผลการทดลองชี้ให้เห็นว่าการสกัดคุณลักษณะของข้อมูลสามารถลดขนาดของข้อมูลนำเข้าโดยไม่ทำให้ประสิทธิภาพในการแบ่งกลุ่มลดลง ซึ่งการทดลองด้วย SVM ใช้ข้อมูลเพียง 4 คุณลักษณะหลักที่สกัดได้จาก KPCA ก็ทำให้ได้ผลลัพธ์ที่ดีกว่าชุดข้อมูลที่ไม่ผ่านการสกัด และชุดข้อมูลที่ผ่านการสกัดด้วย PCA

Hai-Hua Gao, Hui-Hua Yang และ Xing-Yu Wang (2005) ได้นำเสนองานวิจัยเรื่อง Principal Component Neural Networks Based Intrusion Feature Extraction and Detection Using SVM โดยได้นำเสนอวิธีการใหม่ในการสกัดคุณลักษณะชุดข้อมูลการบุกรุก โดยการประยุกต์ใช้ Principal Component Neural Network: PCNN และนำผลลัพธ์ที่ได้จาการสกัดคุณลักษณะ มาทำการแบ่งกลุ่มด้วย SVM โดยที่ใช้อัลกอริทึม Adaptive Principal Component Extraction: APEX มาดัดแปลงให้เหมาะสมในการทำงานของ PCNN โดยผลที่ได้จากการทดลองนำมาเปรียบเทียบกับ SVM ที่ไม่ได้ทำการสกัดคุณลักษณะชุดข้อมูล ผลการทดลองแสดงให้เห็นชัดว่า การสกัดคุณลักษณะด้วย PCNN สามารถลดจำนวนมิติของข้อมูลนำเข้า และไม่ทำให้ประสิทธิภาพในการตรวจจับการบุกรุกลดลง

Zhu Xiaorong, Wang Dianchun และ Ye Changguo (2009) ได้นำเสนองานวิจัยเรื่อง A New Feature Extraction Method of Intrusion Detection โดยได้นำเสนอวิธีการนำเอา Kernel Principal Component Analysis: KPCA มาทำการสกัดคุณลักษณะจากการตัวอย่างของข้อมูลการบุกรุกที่จะใช้ฝึกฝน โดยที่วิธีการนี้สกัดคุณลักษณะและลดจำนวนมิติของข้อมูลได้อย่างมีประสิทธิผล โดยได้นำเสนอวิธีการนำ Reduce SVM: RSVM ร่วมกับวิธีการ nonlinear proximal SVM ซึ่งวิธีการที่นำเสนอนี้สามารถลดความซับซ้อนในการคำนวณของ Kernel Matrix ได้ และยังส่งผลให้ความเร็วในการฝึกฝนและผลลัพธ์ของการแบ่งกลุ่มดีขึ้น