เทคนิคพาราเลลคอร์ดิเนตสำหรับตรวจสอบพฤติกรรมการกระทำผิดบนเครือข่าย

The Parallel Coordinates Methodology to Study Suspicious Behavior on a Computer Network

ณัฐโชติ พรหมฤทธิ์[1] และ อนิราช มิ่งขวัญ[2]

[1]ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[2]ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีและการจัดการอุตสาหกรรม
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

บทคัดย่อ – งานวิจัยนี้ทำการศึกษาพฤติกรรมการใช้งานเครือข่ายจากข้อมูลจราจรทางคอมพิวเตอร์ โดยตรวจจับพฤติกรรมที่มีการละเมิดนโยบายความปลอดภัยของเครือข่าย ซึ่งจะเป็นประโยชน์ในการสืบสวนหาตัวผู้กระทำความผิดมาลงโทษ ผู้วิจัยนำเสนอ User investigations with visualization time machine for network forensic (UIV) model ในการศึกษาพฤติกรรมการใช้งานเครือข่ายโดยใช้เทคนิค Parallel coordinates ซึ่งแสดงความสัมพันธ์ด้วยตัวแปร User, Source IP address, Time, Destination IP address, Destination service และ Domain name ในเบื้องต้นผู้วิจัยทำการทดลองโดยจำลองสถานการณ์การโจมตีเพื่อศึกษาว่าตัวแปรต่างๆ ที่นำเสนอจะสามารถบ่งชี้พฤติกรรมการใช้งานเครือข่ายได้หรือไม่ ซึ่งพบว่า (i) Attack signatures ที่ได้จากการทดลองจะมีความแตกต่างกันไปตามสถานการณ์การโจมตี และ (ii) ผู้วิเคราะห์สามารถติดตามพฤติกรรมการใช้งานเครือข่ายในแต่ละรายได้

คำสำคัญ – นิติวิทยาศาสตร์สำหรับเครือข่าย; การสร้างภาพข้อมูลจราจรทางคอมพิวเตอร์; Parallel coordinates

ABSTRACT – In this paper, the suspicious behavior on a computer network is used to analyze by detecting the violation behavior of network security policies. This paper proposed the user investigations with visualization time machine for network forensic (UIV) model. The proposed model is used parallel coordinates, which can be presented as the relationship of various parameters such as user, source ip address, time, destination ip address, destination service and domain name. For this system, the model is tested by simulated attack. The result of experiment shows that (i) the attacked signatures are different depended on situation attacked and (ii) the analyst are able to tracking individual behavior using UIV model.

KEYWORDS – Network forensic; Network traffic visualization; Parallel coordinates

1. บทนำ

เมื่อมีการกระทำความผิดบนเครือข่ายคอมพิวเตอร์ ข้อมูลจราจรทางคอมพิวเตอร์นับเป็นพยานหลักฐานสำคัญในการดำเนินคดีอันเป็นประโยชน์อย่างยิ่งต่อการสืบสวนสอบสวน เพื่อนำตัวผู้กระทำความผิดมาลงโทษ

การวิเคราะห์ข้อมูลจะต้องใช้เทคนิคทาง Network forensic [1], [2], [3], [4] ซึ่งในปัจจุบันยังไม่มีงานวิจัยที่ทำการศึกษาและแสดงพฤติกรรมการใช้งานเครือข่าย และตรวจจับผู้ใช้งานที่มีการละเมิดนโยบายความปลอดภัยของเครือข่ายจาก Flow data ด้วยเทคนิค Parallel coordinates ผู้วิจัยจึงนำเสนอ User investigations with visualization time machine for network forensic (UIV) model ซึ่งสามารถวิเคราะห์พฤติกรรมที่มีการละเมิดนโยบายความปลอดภัยของเครือข่ายในแบบ Visualization

ในส่วนถัดไปจะกล่าวถึงวรรณกรรมที่เกี่ยวข้อง, วิธีการดำเนินการวิจัย, ผลการดำเนินงาน และบทสรุป

2. วรรณกรรมที่เกี่ยวข้อง

Network forensic เป็นเครื่องมือที่สำคัญสำหรับการวิเคราะห์หาผู้ต้องสงสัยจากข้อมูลจราจรทางคอมพิวเตอร์ โดยงานทางด้าน Network forensic ส่วนใหญ่มักนำเสนอในแบบ Visualization [4], [5], [6], [7], [8], [9], [10]

Parallel coordinates เป็นเทคนิคหนึ่งทาง Visualization ซึ่งถูกนำเสนอเป็นครั้งแรกโดย Inselberg [11] และมีหลายงานวิจัยได้นำไปประยุกต์ใช้ในงานทางด้าน Network security [4], [7], [8], [9], [10] เช่น Hyunsang Choi และคณะ [4] ได้นำเสนอวิธีการตรวจจับการบุกรุกโดยใช้เทคนิค Misuse detection และแสดงผลแบบ Visualization โดยทำการเปรียบเทียบ Flow data บนเครือข่ายกับ Attack signatures ซึ่งแสดงความสัมพันธ์ของ Flow ด้วยเทคนิค Parallel coordinates จากตัวแปร 4 ตัว คือ 1. Source IP Address 2. Destination IP Address 3. Destination port และ 4. ขนาดเฉลี่ยของ Packet ในขณะที่ Sven Krasser และคณะ [9] ได้นำเสนอ Network traffic visualization โดยใช้เทคนิค Parallel coordinates และ Scatter plot สำหรับการวิเคราะห์ข้อมูลแบบ Real time และแบบ Forensic

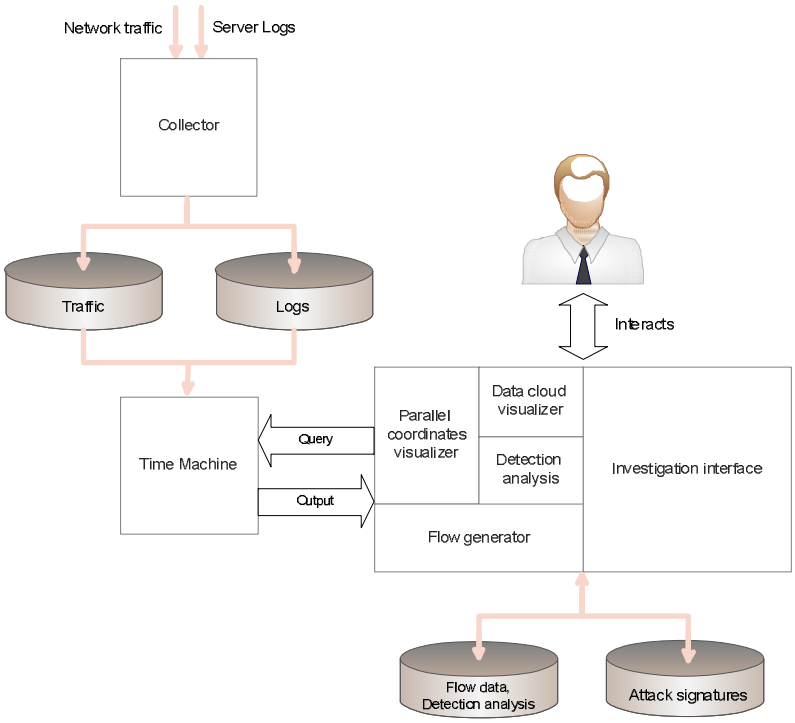
อย่างไรก็ตามงานวิจัยเหล่านี้ไม่ได้มุ่งศึกษาพฤติกรรมของผู้ใช้งานเครือข่าย ผู้วิจัยจึงนำเสนอ UIV Model ซึ่งสามารถรวบรวม จัดเก็บข้อมูล ตรวจจับผู้ต้องสงสัย สืบสวนเพื่อตอบคำถามที่สำคัญ เช่น ใคร ทำอะไร ที่ไหน เมื่อไหร่ และอย่างไรบนเครือข่ายคอมพิวเตอร์ และแสดงผลในแบบ Visualization โดยนำเสนอเทคนิค In -> Out Parallel coordinates สำหรับการศึกษาพฤติกรรมของผู้ใช้บริการเครือข่ายที่มีการติดต่อสื่อสารกับเครือข่ายภายนอกซึ่งแสดงความสัมพันธ์ของ Flow ด้วยตัวแปร 6 ตัว ได้แก่ 1. User 2. Source IP address 3. Time 4. Destination IP address 5. Destination service และ 6. Domain name และเทคนิค Out -> In Parallel coordinates สำหรับการศึกษาพฤติกรรมของ Domain name ที่มีการติดต่อสื่อสารกับเครือข่ายภายในซึ่งแสดงความสัมพันธ์ของ Flow ด้วยตัวแปร 5 ตัว ได้แก่ 1. Domain name 2. Source IP address 3. Time 4. Destination IP address และ 5. Destination service

3. วิธีการดำเนินการวิจัย

เนื้อหาในส่วนนี้จะกล่าวถึงการออกแบบ UIV ภาพรวมของการทดลอง วัตถุประสงค์ในการทดลอง และขั้นตอนการดำเนินการวิจัย ดังต่อไปนี้

3.1. สถาปัตยกรรม

ผู้วิจัยนำเสนอ UIV โดยมีจุดประสงค์สำคัญเพื่อศึกษาพฤติกรรมการใช้งานเครือข่ายจาก Flow data และตรวจจับพฤติกรรมที่มีการละเมิดนโยบายความปลอดภัยของเครือข่าย โดยแสดงผลในแบบ Visualization [12] ซึ่งสถาปัตยกรรมของ UIV ประกอบด้วย



รูปที่ 1. สถาปัตยกรรมของ UIV

Collector ทำหน้าที่รวบรวม Network traffic บนเครือข่าย และ Logs จาก Server Logs

Time Machine ทำหน้าที่ดึงข้อมูล Traffic และ Logs ตามช่วงเวลาที่ระบุ

Flow generator ทำหน้าที่สร้าง Flow ของ data จากข้อมูล Network traffic และ Logs ด้วยเทคนิค Parallel coordinates

Detection analysis ทำหน้าที่วิเคราะห์พฤติกรรมการละเมิดนโยบายความปลอดภัยของเครือข่ายจาก Flow data และ Attack signatures

Parallel coordinates visualizer แสดง Flow data ในรูปแบบกราฟิคด้วยเทคนิค Parallel coordinates

Data cloud visualizer แสดงผลการตรวจจับด้วย Data cloud

Investigation interface แสดงการใช้งานเครือข่ายแบบ Timeline

จากรูปที่ 1. Collector จะรวบรวม Traffic และ Logs ลง Database เพื่อเป็นข้อมูลสำหรับการตรวจจับพฤติกรรมที่มีการละเมิดนโยบายความปลอดภัยของเครือข่ายใน Phase 1 : Detection of suspect และทำการสืบสวนใน Phase 2 : Investigation โดย Flow generator จะสร้าง Flow data ด้วยเทคนิค Parallel coordinates จาก Traffic และ Logs ซึ่งทำการดึงข้อมูลตามช่วงเวลาที่ระบุด้วย Time Machine และแสดงผล Flow data แบบ Parallel coordinates ด้วย Parallel coordinates visualizer ในขณะที่ Detection analysis จะนำ Flow data และ Attack signatures มาวิเคราะห์พฤติกรรมที่มีการละเมิดนโยบายความปลอดภัยของเครือข่าย และแสดงผลลัพธ์ด้วย Data cloud visualizer นอกจากนี้ผู้ดูแลระบบเครือข่าย นักวิเคราะห์ และพนักงานเจ้าหน้าที่สามารถสืบสวนเพิ่มเติมด้วย Investigation interface

3.2. ภาพรวมของการทดลอง และวัตถุประสงค์ในการทดลอง

งานวิจัยนี้ทำการศึกษาพฤติกรรมของผู้ใช้บริการเครือข่ายและ Domain name จาก Flow data โดยใช้เทคนิค Parallel coordinates ซึ่งผู้วิจัยคาดว่าการศึกษาในมิติของเวลาจะสามารถบ่งชี้ถึงพฤติกรรมการใช้งานเครือข่าย จึงกำหนดให้ User, Time และ Domain name เป็นตัวแปรหนึ่งใน Parallel coordinates

ดังนั้นเมื่อพบการโจมตีจากผู้ใช้บริการ หรือจาก Domain name ผู้วิเคราะห์จะสามารถสืบสวนเพื่อหาตัวผู้กระทำความผิดได้ เช่น เมื่อพบการทำ Host scan ของผู้ใช้รายหนึ่ง ผู้วิเคราะห์สามารถทำการศึกษาพฤติกรรมการใช้งานเครือข่ายของผู้ใช้รายนี้ ซึ่งพบว่ามีการทำ Port scan และ Single source DoS ในลำดับถัดมา ด้วยเหตุนี้ผู้วิเคราะห์จึงกลับไปตรวจสอบประวัติการใช้งานเครือข่ายย้อนหลังในอีกหนึ่งวัน หรือในหนึ่งเดือน ทำให้สามารถศึกษาพฤติกรรมการใช้งานเครือข่ายที่ผ่านมา และทราบว่าผู้ใช้รายนี้มีความถี่ในการละเมิดนโยบายความปลอดภัยเครือข่ายเป็นอย่างไร

เพื่อตรวจสอบสมติฐานว่าเทคนิค Parallel coordinates ที่นำเสนอจะสามารถตอบคำถามดังกรณีที่กล่าวมาได้ ในเบื้องต้นผู้วิจัยจึงจำลองสถานการณ์การโจมตี แบบ Host scan, Port scan, Single source DoS และ Back scatter โดยมีวัตถุประสงค์ในการทดลองดังต่อไปนี้

1. เพื่อศึกษา Attack signatures โดยใช้เทคนิค In -> Out Parallel coordinates และ Out -> In Parallel coordinates

2. เพื่อศึกษาพฤติกรรมการละเมิดนโยบายความปลอดภัยเครือข่ายของผู้ใช้บริการและจาก Domain name

3.3. ขั้นตอนการดำเนินการวิจัย

ผู้วิจัยได้จำลองสถานการณ์การโจมตีในแบบต่างๆ ด้วยโปรแกรม MATLAB และระบบจัดการฐานข้อมูล MySQL โดยมีขั้นตอนการดำเนินการวิจัยทั้งหมด 5 ขั้นตอน ได้แก่ 1. การเตรียมข้อมูลเพื่อทำการทดลอง 2. รวบรวมข้อมูลลง Database 3. ดึงข้อมูลจาก Database ตามช่วงเวลาที่กำหนด 4. สร้าง Flow data และ 5. แสดง Flow data ในรูปกราฟิคด้วยเทคนิค Parallel coordinates

1. การเตรียมข้อมูลเพื่อทำการทดลอง

ผู้วิจัยได้ทำการสร้าง Datasets เป็น Text file เพื่อจำลองสถานการณ์การโจมตีทั้งหมด 10 กรณี ได้แก่

Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	51341	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	63726	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	38434	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	50989	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	47690	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	42662	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	43556	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	61523	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	35067	temp.gamefanc.com
Charlie	215.123.153.189	30/Jun/2010 13:27:04	12.90.27.73	26112	temp.gamefanc.com

รูปที่ 2. ตัวอย่าง Datasets สำหรับการทำ Port scan

1. การทำ Host scan จากผู้ใช้งานเครือข่าย
2. การทำ Port scan จากผู้ใช้งานเครือข่าย
3. การทำ Single source DoS จากผู้ใช้งานเครือข่าย
4. การทำ Back scatter จากผู้ใช้งานเครือข่าย
5. การโจมตีแบบ Host scan, Port scan, Single source DoS จากผู้ใช้งานเครือข่าย
6. การทำ Host scan จากเครือข่ายภายนอก
7. การทำ Port scan จากเครือข่ายภายนอก
8. การทำ Single source DoS จากเครือข่ายภายนอก

9. การทำ Back scatter จากเครือข่ายภายนอก และ

10. การโจมตีแบบ Host scan, Port scan, Single source DoS จากเครือข่ายภายนอก

จากรูปที่ 2. แสดงตัวอย่าง Datasets ของการทำ Port scan โดยกำหนดชื่อผู้ใช้เป็น Charlie กำหนด Source IP address และ Destination IP address อย่างละ 1 หมายเลขจากการสุ่ม กำหนด Domain name 1 Domain จากการสุ่ม และทำการสุ่ม Destination service จนครบทั้งหมด โดยจะบันทึกเวลาการ Scan port แต่ละครั้งจากเวลาในระบบ

2. รวบรวมข้อมูลลง Database

Collector จะทำหน้าที่รวบรวมข้อมูลจาก Datasets, Transform ข้อมูล และจัดเก็บข้อมูลใน Relational database โดย Database ของ UIV จะประกอบด้วยตารางทั้งหมด 6 ตาราง ได้แก่

1. USER(userid, email)
2. SUSPECT_USER(suspectid, userid, suspect_level, attack_name, detect_time, attack_time)
3. INOUT_TRAFFIC(trafficid, userid, source_ipaddress, destination_ipaddress, domain_name, time, destination_service, cite)
4. DOMAIN_NAME(domain_name)
5. OUTIN_TRAFFIC(trafficid, source_ipaddress, destination_ipaddress, domain_name, time, destination_service, cite)
6. SUSPECT_DOMAIN(suspectid, domain_name, suspect_level, attack_name, detect_time, attack_time)

จากตาราง INOUT_TRAFFIC และตาราง OUTIN_TRAFFIC ผู้วิเคราะห์สามารถดูข้อมูล Traffic ว่ามาจากไฟล์ใด และเรคคอร์ดไหนใน Datasets ได้จากฟิลด์ cite

3. ดึงข้อมูลจาก Database ตามช่วงเวลาที่กำหนด

Time Machine จะทำการดึงข้อมูลของผู้ใช้งานเครือข่ายหรือ Domain name จาก Database ตามช่วงเวลาที่ระบุ ซึ่งจะทำให้ผู้วิเคราะห์สามารถตรวจสอบประวัติการใช้งานเครือข่ายย้อนหลังได้

4. สร้าง Flow data

Flow generator จะ Transform ข้อมูลจากใน Database เป็น Matrix โดยจะแทน Flow data ของผู้ใช้งานเครือข่ายด้วย In->Out Flow data matrix และแทน Flow data ของ Domain name ด้วย Out->In Flow data matrix

กำหนดให้ In->Out Flow data matrix $A = [a_{ij}]$

เมื่อ $i = 1$ ถึง $Max_{UserTrafficIn24H}$

และ $j = 1$ ถึง 6 โดย

$$a_{i1} = \alpha \cdot \frac{Max_{UserTrafficIn24H}}{Max_{TrafficIn24H}} \quad (1)$$

$$a_{i2} = \alpha \cdot \frac{\text{int}(SourceIPaddress)}{\text{int}(255.255.255.255)} \quad (2)$$

$$a_{i3} = \alpha \cdot (SerialDateNumber(Time) - floor(SerialDateNumber(Time))) \quad (3)$$

$$a_{i4} = \alpha \cdot \frac{\text{int}(DestinationIPaddress)}{\text{int}(255.255.255.255)} \quad (4)$$

$$a_{i5} = \alpha \cdot \frac{DestinationService}{65535} \quad (5)$$

$$a_{i6} = \alpha \cdot \frac{Order(DomainName)}{\text{Max}_{DomainName}} \quad (6)$$

และกำหนดให้ Out->In Flow data matrix $B = [b_{ij}]$

เมื่อ $i = 1$ ถึง $Max_{DomainNameTrafficIn24H}$

และ $j = 1$ ถึง 5 โดย

$$b_{i1} = \alpha \cdot \frac{Max_{\ DomainName\ \ TrafficIn\ \ 24\ H}}{Max_{\ TrafficIn\ \ 24\ H}} \quad (7)$$

$$b_{i2} = \alpha \cdot \frac{\text{int}(SourceIPaddress)}{\text{int}(255.255.255.255)} \quad (8)$$

$$b_{i3} = \alpha \cdot (SerialDateNumber(Time) - floor(SerialDateNumber(Time))) \quad (9)$$

$$b_{i4} = \alpha \cdot \frac{\text{int}(DestinationIPaddress)}{\text{int}(255.255.255.255)} \quad (10)$$

$$b_{i5} = \alpha \cdot \frac{DestinationService}{65535} \quad (11)$$

จาก Matrix A และ Matrix B เราสามารถทำการปรับค่าของ $[a_{ij}]$ และ $[b_{ij}]$ ให้มีค่าระหว่าง 0 ถึง α โดยการคูณด้วยจำนวนเต็มบวก α

5. แสดง Flow data ในรูปกราฟิคด้วยเทคนิค Parallel coordinates

Parallel coordinates visualizer จะนำ In->Out Flow data matrix และ Out->In Flow data matrix มา Plot เป็นกราฟิคด้วยเทคนิค Parallel coordinates ดังรูปที่ 3.

รูปที่ 3. แสดง Flow ของการโจมตีแบบ Host scan, Port scan, Single source DoS, Backscatter และ Flow ของ Trudy ด้วยเทคนิค In->Out Parallel coordinates ซึ่งแต่ละ Flow จะแสดงถึงพฤติกรรมต่างๆ ดังต่อไปนี้

(a) Host scan ผู้บุกรุกจะพยายามส่ง Traffic ไปยัง Host ในเครือข่ายเป้าหมาย โดยการกระจาย Destination IP address ภายในช่วงของ Subnet ซึ่งเหตุการณ์จะดำเนินต่อเนื่องอยู่ในช่วงระยะเวลาหนึ่ง ทำให้เรามองเห็นเส้นที่ลากไปยังแกน Time และแกน Destination IP address กระจายภายในช่วงหนึ่งของแกนทั้งสอง ในขณะที่แกนที่เหลือ เส้นจะตกลงบนจุดเพียงจุดเดียว

(b) Port scan ผู้บุกรุกจะพยายามส่ง Traffic กระจายไปยัง Port ต่างๆ ของ Host เป้าหมาย ซึ่งเหตุการณ์จะดำเนินต่อเนื่องอยู่ในช่วงระยะเวลาหนึ่ง ทำให้เรามองเห็นเส้นที่ลากไปยังแกน Time และแกน Destination service กระจายภายในช่วงหนึ่งของแกนทั้งสอง ในขณะที่แกนที่เหลือ เส้นจะตกลงบนจุดเพียงจุดเดียว

(c) Single source DoS เป็นการโจมตีที่ทำให้ Host ปลายทางหยุดการให้บริการ โดยผู้บุกรุกจะพยายามส่ง Traffic จำนวนมากไปยัง Host เป้าหมาย ซึ่งเหตุการณ์จะดำเนินต่อเนื่องอยู่ในช่วงระยะเวลาหนึ่ง ทำให้เรามองเห็นเส้นที่ลากไปยังแกน Time กระจายภายในช่วงหนึ่งของแกนนี้ ในขณะที่แกนที่เหลือ เส้นจะตกลงบนจุดเพียงจุดเดียว

(d) Backscatter เป็นการโจมตีที่ผู้บุกรุกจะทำ Host scan และ Port scan พร้อมกัน ซึ่งเหตุการณ์จะดำเนินต่อเนื่องอยู่ในช่วงระยะเวลาหนึ่ง ทำให้เรามองเห็นเส้นที่ลากไปยังแกน Time, แกน Destination IP address และแกน Destination service กระจายภายในช่วงหนึ่งของแกนทั้งสาม ในขณะที่แกนที่เหลือ เส้นจะตกลงบนจุดเพียงจุดเดียว

(e) Flow ของ Trudy, Trudy จะใช้ Source IP address 3 หมายเลข ทำการโจมตีแบบ Host scan, Port scan และ Single source DoS ตามลำดับ ทำให้เรามองเห็นเส้นที่ลากไปยังแกน Source IP address มีสามจุด และแกน Time กระจายในสามช่วงส่วนแกน Destination IP address และแกน Destination service จะมีลักษณะของเส้นที่ผสมผสานกันระหว่าง Host scan, Port scan, และ Single source Dos ในขณะที่แกนที่เหลือ เส้นจะตกลงบนจุดเพียงจุดเดียว

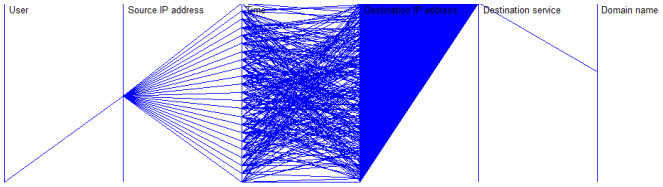
จะเห็นว่า Trudy มีความพยายามทำการโจมตี 3 แบบต่อเนื่องกัน โดยการทำ Host scan ในครั้งแรกจะทำให้สามารถทราบข้อมูล Host ที่ให้บริการในเครือข่าย การทำ Port scan ในครั้งถัดมาจะทำให้สามารถทราบข้อมูล Port ที่เปิดให้บริการ และสุดท้ายการทำ Single source DoS จะทำให้ Host เป้าหมายหยุดการให้บริการ ซึ่งโดยปกติการตรวจจับการโจมตีแบบ Single source DoS เป็นเรื่องที่ยากเนื่องจากลักษณะของ Flow จะมีความใกล้เคียงกับ Flow ของการติดต่อสื่อสารแบบปกติ ดังนั้นเมื่อพบว่ามีความพยายามทำ Host scan ในครั้งแรก และ Port scan ในเวลาถัดมา Flow ที่เกี่ยวข้องต่อไปอาจจะเป็น Flow ที่น่าสงสัยซึ่งสามารถทำการวิเคราะห์เพิ่มเติมต่อไป

4. ผลการดำเนินงาน

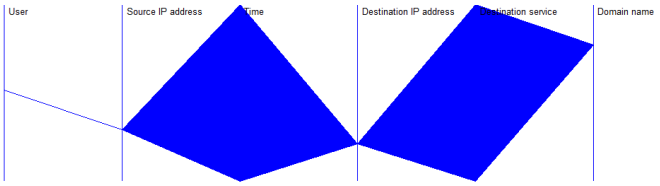
การทดลองกับสถานการณ์การโจมตี ทั้งหมด 8 กรณี สามารถสรุปเป็น Attack signatures ได้ดังรูปที่ 4.

Attack signatures ของ Host scan, Port scan, Single source DoS และ Back scatter ถูกแสดงด้วยเทคนิค In->Out Parallel coordinates, Out->In Parallel coordinates และ PCAV [4] โดยมีลำดับของตัวแปร ดังตารางที่ 1.

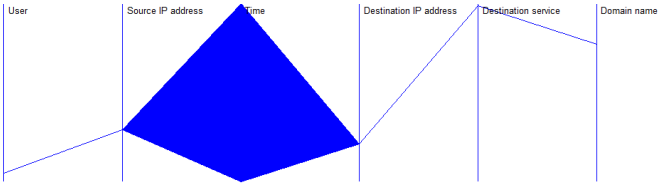
(a) Host scan



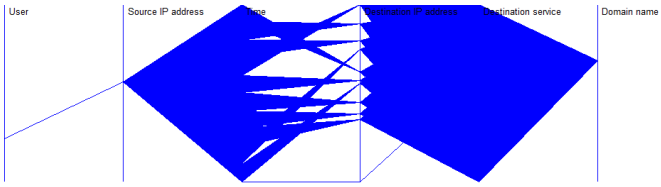
(b) Port scan



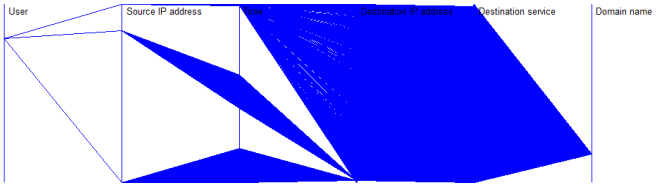
(c) Single source DoS



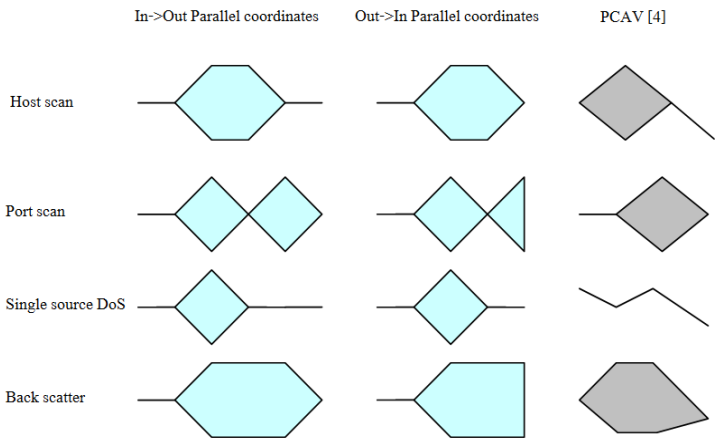
(d) Backscatter



(e) Host scan, Port scan and Single source DoS by Trudy



รูปที่ 3. In->Out Parallel coordinates



รูปที่ 4. Attack signatures โดยใช้เทคนิค In->Out Parallel coordinates, Out->In Parallel coordinates และ PCAV [4]

ตารางที่ 1. ลำดับของตัวแปรในเทคนิค Parallel coordinates

	User	Source IP address	Time	Destination IP address	Destination service	Domain name	Packet length
In->Out	1	2	3	4	5	6	
Out->In		2	3	4	5	1	
PCAV [4]		1		2	3		4

Host scan ผู้บุกรุกจะส่ง Traffic ไปยัง Host ปลายทางโดยกระจาย Destination IP address ภายในช่วงของ Subnet ทำให้รูปจากเทคนิค In->Out Parallel coordinates และ Out->In Parallel coordinates มีช่วงกว้างที่ Time และ Destination IP address ในขณะที่รูปจากเทคนิค PCAV [4] จะมีช่วงกว้างที่ Destination IP address

Port scan ผู้บุกรุกจะส่ง Traffic กระจายไปยัง Port ต่างๆ ของ Host เป้าหมาย ทำให้รูปจากเทคนิค In->Out Parallel coordinates และ Out->In Parallel coordinates มีช่วงกว้างที่ Time และ Destination service ในขณะที่รูปจากเทคนิค PCAV [4] จะมีช่วงกว้างที่ Destination service

Single source DoS ผู้บุกรุกจะส่ง Traffic จำนวนมากไปยัง Host เป้าหมาย ทำให้รูปจากเทคนิค In->Out Parallel coordinates และ Out->In Parallel coordinates มีช่วงกว้างที่ Time ในขณะที่รูปจากเทคนิค PCAV [4] จะไม่มีช่วงที่กว้าง

Backscatter ผู้บุกรุกจะทำ Host scan และ Port scan พร้อมกัน ดังนั้นรูปจากเทคนิค In->Out Parallel coordinates และ Out->In Parallel coordinates จะมีช่วงกว้างที่ Time, Destination IP address และ Destination service ในขณะที่รูปจากเทคนิค PCAV [4] จะมีช่วงกว้างที่ Destination IP address และ Destination service

ดังนั้นจากรูปที่ 4. จะเห็นได้ว่าการเพิ่มตัวแปร Time ในเทคนิค In->Out Parallel coordinates และ Out->In Parallel coordinates จะทำให้เราสามารถมองเห็นภาพ Attack signatures ในมิติของเวลาเพิ่มขึ้น และจากตัวอย่าง Flow ของ Trudy, การเพิ่มตัวแปร User จะทำให้เราสามารถมองเห็นลำดับเหตุการณ์ของการโจมตี ซึ่งผู้วิเคราะห์สามารถตรวจสอบพฤติกรรมการกระทำความผิดบนเครือข่ายเป็นรายๆ ไป ในขณะที่เทคนิค PCAV[4] ไม่สามารถตอบคำถามเหล่านี้ได้

4. บทสรุป

งานวิจัยนี้ทำการศึกษาพฤติกรรมของผู้ใช้บริการเครือข่ายและ Domain name จาก Flow data โดยใช้เทคนิค Parallel coordinates แบบ In -> Out Parallel coordinates และ

Out -> In Parallel coordinates ซึ่งผู้วิจัยคาดว่าการศึกษาในมิติของเวลาจะสามารถบ่งชี้ถึงพฤติกรรมการใช้งานเครือข่าย จึงกำหนดให้ User, Time และ Domain name เป็นตัวแปรหนึ่งใน Parallel coordinates เพื่อทดสอบสมติฐานดังกล่าว ในเบื้องต้นจึงจำลองสถานการณ์การโจมตีแบบ Host scan, Port scan, Single source DoS และ Back scatter และแสดง Flow ของการโจมตีในแบบต่างๆ โดยสามารถสรุปเป็น Attack signatures ได้ทั้งหมด 8 Signatures ซึ่งจะเห็นได้ว่าการเพิ่มตัวแปร Time และ User (และเช่นเดียวกันกับตัวแปร Domain name) จะทำให้เราสามารถมองเห็นภาพ Attack signatures ในมิติของเวลาและลำดับเหตุการณ์ของการโจมตีของผู้ใช้เป็นรายๆ ไป

ในอนาคตผู้วิจัยจะทำการศึกษาพฤติกรรมการใช้งานเครือข่ายบนสถานที่จริง และตรวจจับพฤติกรรมการใช้งานที่ละเมิดนโยบายความปลอดภัยของเครือข่ายด้วยเทคนิค Anomaly detection และ Misuse detection

เอกสารอ้างอิง

[1] J. Haggerty, D. Llewellyn-Jones, and M. Taylor, “Forweb: file fingerprinting for automated network forensics investigations,” in e-Forensics ’08: Proceedings of the 1st international conference on Forensic applications and techniques in telecommunications, information, and multimedia and workshop, January 2008, pp. 1-6.

[2] M. I. Cohen, “Source attribution for network address translated forensic captures,” Digital Investigation, vol. 5, pp. 138-145, 2009.

[3] R. Hadjidj, M. Debbabi, H. Lounis, F. Iqbal, A. Szporer, and D. Benredjem, “Towards an integrated e-mail forensic analysis framework,” Digital Investigation, vol. 5, pp.124-137, 2009.

[4] H. Choi, H. Lee, and H. Kim, “Fast detection and visualization of network attacks on parallel coordinates,” Computers and Security, vol. 28, pp. 276-288, 2009.

[5] R. Ball, G. A. Fink, and C. North, “Home-centric visualization of network traffic for security administration,” in VizSEC/DMSEC ’04: Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security, October 2004, pp. 55-64.

[6] D. Phan, A. Paepcke, , and T. Winograd, “Progressive multiples for communication-minded visualization,” in Graphics Interface Conference, May 2007, pp. 225-232.

[7] X. Yin, W. Yurcik, M. Treaster, Y. Li, and K. Lakkaraju, “Visflowconnect: netflow visualizations of link relationships for security situational awareness,” in VizSEC/DMSEC ’04: Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security, October 2004, pp. 26-34.

[8] G. Conti and K. Abdullah, “Passive visual fingerprinting of network attack tools,” in VizSEC/DMSEC ’04: Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security, October 2004, pp. 45-54.

[9] S. Krasser, G. Conti, J. Grizzard, J. Gribschaw, and H. Owen, “Real-time and forensic network data analysis using animated and coordinated visualization,” in Information Assurance Workshop, 2005. IAW ’05. Proceedings from the Sixth Annual IEEE SMC, June 2005, pp. 42-49.

[10] G. Conti, K. Abdullah, J. Grizzard, J. Stasko, J. A. Copeland, M. Ahamad, H. L. Owen, and C. Lee, “Countering security information overload through alert and packet visualization,” IEEE Comput. Graph., vol. 26, pp. 60-70, March/April 2006.

[11] A. Inselberg, “Multidimensional detective,” in INFOVIS ’97: Proceedings of the 1997 IEEE Symposium on Information Visualization (InfoVis ’97), 1997, pp. 100–107.

[12] N. Promrit and A. Mingkhwan, “User investigations with visualization time machine for network forensic,” Information Technology Journal, vol. 11, pp. 31–36, 201