บทความวิจัย

ระบบแนะนำเพื่อแก้ไขความผิดปกติของกฎไฟร์วอลล์

Recommendation System to Resolve the Firewall Rule Anomalies

ขวัญฤดี โพธิ์ล้อม

Kwanruedee Pholom

บทคัดย่อ

ระบบแนะนำเพื่อแก้ไขความผิดปกติของกฎไฟร์วอลล์ เป็นระบบที่จะให้คำแนะนำในการตัดสินแก้ไขกฎไฟร์วอลล์ พัฒนาด้วยภาษา java โดยผู้ใช้งานสามารถตัดสินใจ ว่าจะทำตามคำแนะนำที่ระบบวิเคาระห์หรือไม่

ระบบแนะนำเพื่อแก้ไขความผิดปกติของกฎไฟร์วอลล์ จะคำนวณความผิดปกติของกฎไฟร์วอลล์ ที่มีอยู่ในระบบ และคำนวณความน่าจะเป็นที่กฎนั้นจะเกิดผลกระทบ หากมีการแก้ไขกฎ โดยจะใช้เบย์ เข้ามาในการหาความน่าจะเป็น

คำสำคัญ : กฎไฟร์วอลล์, ความผิดปกติของกฎไฟร์วอลล์

1. บทนำ

ไฟร์วอลล์เป็นเครื่องมือหรือระบบของระบบคอมพิวเตอร์ที่ทำหน้าที่ป้องกันเครือข่ายจากผู้ที่ไม่ได้รับอนุญาต ไม่ให้สามารถมาใช้หรือมองเห็นข้อมูลหรือเครือข่ายคอมพิวเตอร์ได้ ถือว่าเป็นระบบที่ ทำหน้าที่รักษาความปลอดภัยสำหรับเครือข่ายคอมพิวเตอร์ที่มีความสำคัญยิ่ง เพราะไฟร์วอลล์สามารถตรวจสอบและป้องกันการบุกรุกหรือการโจมตีที่เป็นอันตรายต่อระบบเครือข่าย โดยขั้นตอนในการทำงานของระบบไฟร์วอลล์จะทำหน้าที่ตรวจสอบการบุกรุกหรือการโจมตีโดยการค้นหาความผิดปกติ โดยการเปรียบเทียบกฎที่ถูกกำหนดไว้ เมื่อตรวจสอบพบความผิดปกติของข้อมูลไฟร์วอลล์ก็จะทำการปิดกั้นไม่ให้ข้อมูลเหล่านั้นผ่านไปได้ ในทาง กลับกัน ถ้าข้อมูลที่ถูกตรวจสอบไม่มีสิ่งผิดปกติใดๆ ไฟร์วอลล์ก็จะอนุญาตให้ข้อมูลเหล่านั้นผ่านไปทำงานที่ต้องการได้

ประสิทธิภาพของไฟร์วอลล์ขึ้นอยู่กับกฎ โดยปกติกฎจะเกิดความผิดพลาดหรือความขัดแย้งอยู่เสมอ ยิ่งกฎมีจำนวนมากขึ้นเท่าไหร่ ความผิดพลาดหรือความขัดแย้งก็จะเกิดขึ้นมากตามไปด้วย ความผิดพลาดที่ ผู้เชี่ยวชาญไฟร์วอลล์แบ่งออกได้เป็น 5 แบบ คือ 1.Shadowing Anomaly 2.Correlation Anomaly 3.Generalization Anomaly 4.Redundancy Anomaly 5.Semantic Loss

โดยปัจจุบันวิธีการแก้ไข ขึ้นอยู่กับผู้ดูแลระบบ จะโยนภาระให้กับผู้ดูแลระบบเป็นคนแก้ไข ซึ่งจะส่งผลให้กฎมีประสิทธิภาพก็อยู่กับผู้ดูแลระบบที่เชี่ยวชาญ

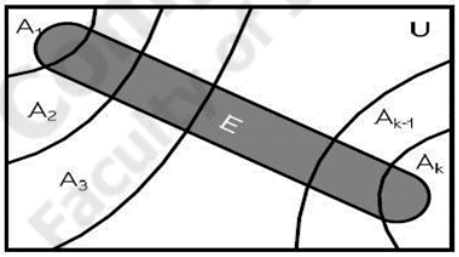
2. ทฤษฎีที่เกี่ยวข้อง

ไฟร์วอลล์ เป็นเครื่องมือหรือระบบของระบบคอมพิวเตอร์ที่ทำหน้าที่ป้องกันเครือข่ายจากผู้ที่ไม่ได้รับอนุญาตไม่ให้สามารถมาใช้หรือมองเห็นข้อมูลหรือเครือข่ายคอมพิวเตอร์ได้ ถือว่าเป็นระบบที่ทำหน้าที่รักษา ความปลอดภัยสำหรับเครือข่ายคอมพิวเตอร์ที่มีความสำคัญยิ่ง

IPTables Linuxสามารถใช้งานเป็นไฟร์วอลล์ได้ตั้งแต่เคอร์เนล 1.1 ซึ่งเป็นเวอร์ชันแรกโดย Alan Cox ใช้ชื่อ ว่า ipfw (จาก BSD) ต่อมา Linux 2.0 ได้ถูกพัฒนาและปรับปรุงได้เครื่องมือที่มีชื่อว่า ipfwadm โดยเครื่องมือ ชิ้นนี้อนุญาตให้ผู้ใช้สามารถควบคุม filtering rule ได้ และต่อมา Linux 2.2 ก็ได้สร้างเครื่องมือตัวใหม่ที่มีชื่อ ว่า ipchains ซึ่งเผยแพร่ในปี ค.ศ.1998 โดย Rusty Russell และทีมงาน ทั้งนี้ ipchains นี้ถือได้ว่าเป็นพัฒนาการขั้นที่สามของ Linux Firewall จวบจนกระทั่งในปัจจุบัน ก็มีการพัฒนา Netfilter และ IPTables ซึ่งถือได้ว่าเป็นพัฒนาการชั้นที่สี่ของ Linux Firewal

ทฤษฎีของเบย์ (Bayes' Theorem)

ถ้าให้เอกภพสัมพัทธ์ U ประกอบด้วยเหตุการณ์ที่ไม่สามารถเกิดขึ้นได้พร้อมกันk จำนวนเหตุการณ์คือ A1, A2, A3, ..., Ak ดังรูปที่ 7.1 และให้ E เป็นเหตุการณ์หนึ่งในปริภูมิตัวอย่างที่เกิดจาก การ ทดลองเดียวกันนี้และต้องเป็นส่วนหนึ่งของ Ai(i=1, 2, 3, ..., k) จะสามารถคำนวณความน่าจะ เป็นแบบมี เงื่อนไขของเหตุการณ์หนึ่งใน A i เมื่อเหตุการณ์ E เกิดขึ้นแล้ว



ภาพที่ 1 เหตุการณ์ E บนเหตุการณ์ k เหตุการณ์ที่เกิดพร้อมกันไม่ได้

การเรียนรู้แบบเบย์อาศัยหลักการของการคำนวณความน่าจะเป็นของแต่ละสมมติฐาน(ในที่นี่คือคลาสเป้าหมายหรือผลลัพธ์การทำนาย)โดยการเรียนรู้แบบเบย์เป็นการเรียนรู้เพิ่มได้เนื่องจากตัวอย่าง ใหม่ที่ได้มาถูกนำมาปรับเปลี่ยนการแจกแจงซึ่งมีผลต่อการเพิ่มหรือลดความน่าจะเป็น ทำให้มีการเรียนรู้ที่ เปลี่ยนไป

การปรับคุณสมบัติ Min-Max

การปรับคุณสมบัติ Min-Max (หรือเรียกว่าการปรับสภาพข้อมูล) เป็นวิธีมาตรฐานที่ใช้ในการปรับช่วง ของข้อมูล เนื่องจากช่วงของค่าข้อมูลอาจแตกต่างกันมาก ดังนั้นจึงเป็นขั้นตอนที่จำเป็นในการประมวลผล ข้อมูลล่วงหน้าก่อนประมวลผลในขั้นตอนถัดไป โดยปกติจะใช้เพื่อปรับขนาดช่วงข้อมูลใด ๆ ให้อยู่ในช่วง [0, 1] เรียกว่าการทำให้เป็น unity-based normalization

การแปลงข้อมูลให้อยู่ในรูปแบบจำนวนเต็มบวก ที่อยู่อินเทอร์เน็ตโพรโทคอล (ที่รู้จักกันในชื่อที่อยู่ IP) เป็นที่อยู่เฉพาะที่อุปกรณ์เครือข่าย เช่น เรา เตอร์, สวิตช์ และ คอมพิวเตอร์ ใช้เพื่อระบุตัวเองและสื่อสารผ่านอุปกรณ์อื่น ๆ ในเครือข่ายคอมพิวเตอร์ ที่อยู่ IPv4 (IP รุ่น 4) มีค่าเท่ากับ 32 บิตตั้งแต่ 0 ถึง 232 – 1 โดยปกติจะแบ่งออกเป็น 4 ส่วนแต่ละส่วน (8 บิต = ออคเต็ต) คั่นด้วยจุดเช่น $A1.\ A2.\ A3.\ A4$ โดยที่ A [0, 255] ที่อยู่ IPv4 สามารถแปลงเป็น จำนวนเต็มใด ๆ ที่ไม่ติดลบด้วยสมการต่อไปนี้:

Moving Average (MA)

ค่าเฉลี่ยเคลื่อนที่ (MA) เป็นตัวบ่งชี้ที่ใช้กันอย่างแพร่หลายสำหรับการวิเคราะห์แนวโน้ม

ข้อมูล ช่วย ให้การดำเนินการข้อมูลราบรื่นขึ้นโดยกรองการรบกวนจากความผันผวนของข้อมูลระยะสั้น ค่าเฉลี่ยเคลื่อนที่มี อยู่ 2 ประเภทซึ่งเป็นที่นิยมและใช้กันอย่างแพร่หลายคือ Simple Moving Average (SMA) และ Exponential Moving Average (EMA) SMA คำนวณค่าเฉลี่ยของข้อมูลล่าสุด เมื่อ n แสดงถึงจำนวนของช่วงเวลาที่เราต้องการเค่าเฉลี่ย

งานวิจัยที่เกี่ยวข้อง ไฟร์วอลล์เป็นอุปกรณ์เน็ตเวิร์คที่มีความสำคัญซึ่งจะเตรียมหรือสนับสนุนเรื่องของภัยคุกคามบน เครือข่าย มาตราการของการป้องกันขึ้นอยู่กับกฎไฟร์วอลล์ ไฟร์วอลล์แบบดั้งเดิมได้แก่ i.n .Cisco ACL,IPTABLES, Checkpoint และ Juniper NetScreen firewall จะตรวจสอบแพ็คเกจจากกฎ และกฎก็จะ ถูกออกแบบในลักษณะของลิสต์

3. แผนการดำเนินงาน

การออกแบบระบบ มี 4 ขั้นตอน

ขั้นตอนที่ 1 การเตรียมกฎไฟร์วอลล์ จะแบ่งออกเป็น 2 ส่วนคือ 1. เงื่อนไข และ การตัดสินใจ 2. ส่วนการความน่าจะเป็น กล่าวคือส่วนที่ 1 จะมีการปรับกฎไฟร์วอลล์ให้อยู่ในรูปแบบจำนวนเต็มบวก โดยสมการ

$(A_1 \times 2^{24}) + (A_2 \times 2^{16}) + (A_3 \times 2^8) + (A_4 \times 2^0)$

ใช้ในการ แปลงข้อมูลของ sip และ dip ส่วน sp และ dp จะมีค่าเท่ากับ $2^{16} - 1$ เมื่อค่าในช่องนั้นเป็น *

ขั้นตอนที่ 2 วิเคราะห์ความผิดปกติ หลังจากที่มีการปรับค่า ip ต่างๆ ให้อยู่ในรูปแบบเลขจำนวนเต็มบวกแล้วนั้นเราจะนำค่า ip ทั้งหมดไปวิเคราะห์ ความผิดปกติ

ขั้นตอนที่ 3 คำนวณความน่าจะเป็น การคำนวณความน่าจะเป็นจะถูกนำมาใช้ในการหาความน่าจะเป็น ในการปรับปรุงกฎ โดยจะมีการนำคุณสมบัติพิเศษเข้ามาช่วนในการคำนวณ เพื่อให้มีความน่าเชื่อถือของกฎไฟร์วอลล์มากยิ่งขึ้น

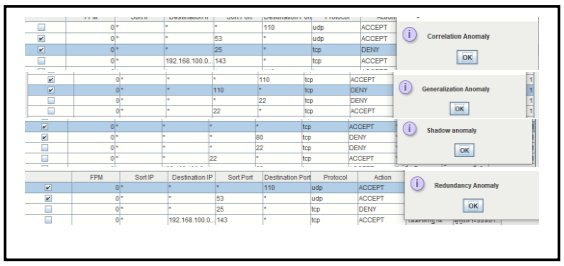
ขั้นตอนที่ 4 การปรับปรุงกฎ ผู้ใช้ระบบสามารถปรับปรุงกฎตามคำแนะนำที่ระบบแจ้งเตรียน ในกรณีที่ผู้ใช้ต้องการปรับปรุ่งกฎไฟร์วอลล์ โปรแกรมของระบบสามารถ ลบกฎ ย้ายกฎ และ รวมกฎ ให้ได้

4. การทดสอบระบบ

ทดสอบการนำเข้าข้อมูล ทดสอบการอ่านไฟล์ txt และการสร้างไฟล์ขึ้นมาใหม่ ผลการทดสอบพบว่าสามารถอ่านไฟล์จากภายในที่เป็นข้อมูล txt ได้ และสามารถสร้างไฟล์ขึ้นมาใหม่ภายในระบบได้ ถูกต้อง

ทดสอบการวิเคราะห์ความผิดปกติ ระบบจะทำการตรวจสอบความผิดปกติของกฎไฟร์วอลล์ โดยการเช็คความผิดปกติของกฎที่ 1 และกฎที่ 2 ว่ากฎนั้นมีความขัดแย้งในรูปแบบใด

หลังจากนั้นจะทำการแจ้งเตื่อน ข้อมูลในรูปแบบข้อความผิดปกติ



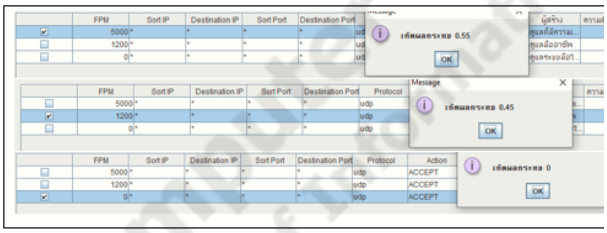
ภาพที่ 2 การเกิดความขัดแย้งการกฎไฟร์วอลล์

ทดสอบความน่าจะเป็น จะทำการคำนวณหาค่าพิเศษ จากนั้นจึงทำการตรวจสอบค่าความน่าจะเป็น

ตารางที่ 1 แปลงค่า max-min

กฎที่	ความถี่การตกกระทบ	หลักฐานการสร้าง	ผู้สร้างกฎ	ความสำคัญ Protocol
R_1	5000	1	1	4
	1	0.33	0.33	0.625
R_2	1200	3	2	6
	0.24	1	0.67	0.25
R_3	0	0	0	9
	0	0	0	0

โดยการแปลงคุณสมบัติ ค่า max-min ดังตารางที่ 1 ค่าที่ถูกแปลงคุณสมบัติแล้วจะถูกนำมาคำนวณความน่าจะเป็น



ภาพที่ 3 การตรวจสอบความน่าจะเป็น

ทดสอบการแก้ไขความผิดปกติของกฎ

ระบบแนะนำเพื่อการแก้ไขความผิดปกติของกฎไฟร์วอลล์ การแก้ไขกฎไฟร์วอลล์ก่อให้เกิดผลกระทบระหว่างกฎด้วยกันอย่างมาก ซึ่งในระบบนี้สามารถดำเนินการแก้ไขกฎ โดยที่ผู้ใช้สามารถทำการแก้ไขกฎที่ขัดแย้งกันได้

ตารางที่ 1 การทดสอบระบบ

การแก้ไขกฎฟร์วอลล์	ความสามารถของระบบ
การเพิ่มกฎ	ได้
การรวมกฎ	ไม่ได้
การลบกฎ	ได้
การย้ายกฎ	ได้

5 สรุป

ในทางปฏิบัติการแก้ไขความผิดปกติของกฎไฟร์วอลล์ค่อนข้างซับซ้อน ขึ้นอยู่กับมุมมองและประสบการณ์ของผู้ดูแลระบบ การแก้ไขข้อผิดพลาดอาจนำไปสู่ความผิดปกติอื่น ๆ ตัวอย่างเช่นเมื่อแก้ไขความ ผิดปกติซ้ำซ้อนมันอาจกลายเป็นการสูญเสียความหมายของกฎ เพื่อลดผลกระทบของข้อผิดพลาดในการแก้ไขความผิดปกติของผู้ดูแลระบบ ดังนั้นบทความนี้ได้ออกแบบและพัฒนาระบบเพื่อช่วยในการตัดสินใจของผู้ดูแล ระบบโดยใช้ความน่าจะเป็นพร้อมกับคุณสมบัติเพิ่มเติม 4 ประการของกฎ คือ ความถี่ของการจับคู่ระหว่าง แพ็กเก็ต, หลักฐานของการสร้างกฎ, ความเชี่ยวชาญของผู้สร้างกฎ แต่ละกฎคำนวณความน่าจะเป็นตามคุณลักษณะทั้ง 4 นี้ หากความน่าจะเป็นของกฎใด ๆ สูงแสดงว่ากฎมีลำดับความสำคัญสูง ในขณะที่กฎใด ๆ ใน ไฟร์วอลล์มีข้อขัดแย้งกฎที่มีค่าความน่าจะเป็นสูงจะถือเป็นอันดับแรกเสมอ จากการทดสอบระบบผู้ดูแลระบบ สามารถตัดสินใจได้อย่างแม่นยำมากขึ้นเกี่ยวกับกฎข้อขัดแย้งในไฟร์วอลล์ สำหรับประสิทธิภาพโดยรวมของ ระบบความซับซ้อนของเวลาในการสร้างระบบ (PST) เท่ากับ o(n) เวลาค้นหาผ่าน PST คือ 0(logmn) และความซับซ้อนคือ O(m*n) อย่างไรก็ตามระบบยังมีข้อ จำกัด ใน การสร้างโครงสร้างต้นไม้ใหม่ ในขณะที่ การ

แก้ไขความผิดปกติใดๆ ของกฎในแต่ละช่วงเวลา นั้นต้องการโครงสร้างตันไม้ PST ทั้งชุด

6 เอกสารอ้างอิง

1 “การเรียนรู้แบบเบย์ (Bayesain Learning)”. สืบค้น 24 กุมภาพันธ์ 2019. http://webcache.googleusercontent.com/search?q=cache:OcIoLVFbqToJ:mis.csit.sci.tsu.ac.th/noppamas/download/DataMining/DataMiningCh7V1.pdf+&cd=11&hl=th&ct=cl nk&gl=th.

2 Alfred V. Aho, J.D.U., John E. Hopcroft The Design and Analysis of Computer Algorithms (Addison-Wesley Series in Computer Science and Information Processing) by January 11, 1983: Addison Wesley in Computer Science and Information Processing.

3 Chomsiri, T., X. He, P. Nanda, และ Z. Tan. “An Improvement of Tree-Rule Firewall for a Large Network: Supporting Large Rule Size and Low Delay”. ใน 2016 IEEE Trustcom/BigDataSE/ISPA, 178–84, 2016.https://doi.org/10.1109/TrustCom.2016.0061.