# บทที่ 3

# รูปแบบความสัมพันธ์และการแทนกฎของไฟร์วอลล์

## 3.1 รูปแบบความสัมพันธ์ของกฎ

รูปแบบความสัมพันธ์ของกฎในไฟร์วอลล์ (Firewall rule relation format) ถูกจัดเตรียมไว้สำหรับการวิเคราะห์ความผิดปกติของกฎในไฟร์วอลล์ทั้งสภาพแวดล้อมแบบไฟร์วอลล์เดี่ยวและไฟร์วอลล์แบบกระจาย โดย Al-Shaer and Hamed (2002) ได้ทำการพิสูจน์ว่ารูปแบบความสัมพันธ์ที่มีอยู่ระหว่างกฎมี 5 รูปแบบ รูปแบบความสัมพันธ์ถูกพิจารณาโดยการเปรียบเทียบกับค่าเซตข้อมูลเครือข่าย (Network field) ต่างๆ ของกฎ ได้แก่ โพรโทคอล (Protocol) เลขที่อยู่ต้นทาง (Source IP address) ช่องทางต้นทาง (Source Port) เลขที่อยู่ปลายทาง (Destination IP address) และช่องทางปลายทาง (Destination Port)

ค่าเซตข้อมูลเครือข่ายเดียวกันของกฎที่ต่างกัน 2 กฎใดๆ อาจจะเท่ากัน (Equal) หรือค่าหนึ่งเป็นส่วนหนึ่งของอีกค่าหนึ่ง (Inclusive) หรือแตกต่าง (Distinct) สำหรับทุกๆ ค่าเซตข้อมูลเครือข่าย ถ้าค่าเซตข้อมูลเครือข่าย 2 ค่าเท่ากัน เราเรียกว่า "exactly match" ตัวอย่างเช่น กฎลำดับที่ 1 คือ 129.162.37.3, *, 184.140.18.*, 80, deny และกฎลำดับที่ 2 คือ 129.162.37.3, *, 184.140.18.*. 80. deny กฎทั้งสองลำดับเป็น exactly match กัน เป็นต้น ถ้ามีหนึ่งค่าเซตข้อมูลเครือข่ายเป็นเซตย่อยที่ซึ่งไม่ใช่เซตย่อยโดยแท้จริง (Proper subset) ของค่าเซตข้อมูลเครือข่ายอีกค่าหนึ่ง เราเรียกว่า "inclusive match" ตัวอย่างเช่น กฎลำดับที่ 1 คือ 129.162.37.3, *, 184.140.18.*, 80, deny และกฎลำดับที่ 2 คือ 129.162.37.*, *, 184.140.18.*. 80, allow จะได้ว่ากฎลำดับที่ 1 เป็น inclusive match กับกฎลำดับที่ 2 เป็นต้น และถ้าค่าเซตข้อมูลเครือข่าย 2 ค่าแตกต่างกัน เราเรียกว่า "distinct" ตัวอย่างเช่น กฎลำดับที่ 1 คือ 129.162.37.3, *, 184.140.18.*, 80, deny และกฎลำดับที่ 2 คือ 129.162.37.1, *, 184.140.36.*, 80, deny กฎทั้งสองลำดับเป็น distinct กัน เป็นต้น ถ้าค่าเซตข้อมูลเครือข่าย 2 ค่าจับคู่กันได้ (Match) ค่าทั้ง 2 จะเท่ากันหรือค่าหนึ่งเป็นส่วนหนึ่งของอีกค่าหนึ่ง ตัวอย่างเช่น เลขที่อยู่ไอพี 129.162.37.3 จับคู่กับ 129.162.37.* ได้ แต่ไม่จับคู่กับ 129.162.37.1 หรือช่องทาง 80 จับคู่กับ * ได้แต่ไม่จับคู่กับ 25 เป็นต้น

กำหนดให้ $R_x$ และ $R_y$ เป็นกราฟย่อยของกราฟ $G$ ให้ $\bowtie \in \{\subset, \supset, =\}$ และ $i, j \in$ $\{protocol, src\_ip, src\_port, dst\_ip, dst\_port\}$ โดยที่ $i \neq j$

*บทนิยาม 1*: $R_x$ exactly matches $R_y$ นั่นคือ $\Re_{EM}$ ก็ต่อเมื่อ $\forall i: R_x(i) = R_y(i)$

*บทนิยาม 2*: $R_x$ inclusively matches $R_y$ นั่นคือ $\Re_{IM}$ ก็ต่อเมื่อ

$\forall i: R_x(i) \subseteq R_y(i)$ และ $\exists j: R_x(j) \neq R_y(j)$

*บทนิยาม 3*: $R_x$ และ $R_y$ เป็น completely disjoint นั่นคือ $\Re_{CD}$ ก็ต่อเมื่อ $\forall i: R_x(i) \not\bowtie R_y(i)$

*บทนิยาม 4*: $R_x$ และ $R_y$ เป็น partially disjoint นั่นคือ $\Re_{PD}$ ก็ต่อเมื่อ

$\exists i, j: R_x(i) \bowtie R_y(i)$ และ $R_x(j) \not\bowtie R_y(j)$

*บทนิยาม 5*: $R_x$ และ $R_y$ เป็น correlated นั่นคือ $\Re_C$ ก็ต่อเมื่อ

$\forall i: R_x(i) \bowtie R_y(i)$ และ $\exists i, j: R_x(i) \subset R_y(i)$ และ $R_x(j) \supset R_y(j)$

เซตเอกภพ (Universal set) ของความสัมพันธ์กฎ นั่นคือ $\Re$ ถูกนิยามได้ว่า

$\Re = \{\Re_{CD}, \Re_{PD}, \Re_{EM}, \Re_{IM}, \Re_C\}$ โดยที่ $\Re_{CD}$ คือ completely disjoint $\Re_{PD}$ คือ partially disjoint $\Re_{EM}$ คือ exactly match $\Re_{IM}$ คือ inclusively match และ $\Re_C$ คือ correlated

## 3.2 การแทนกฎ

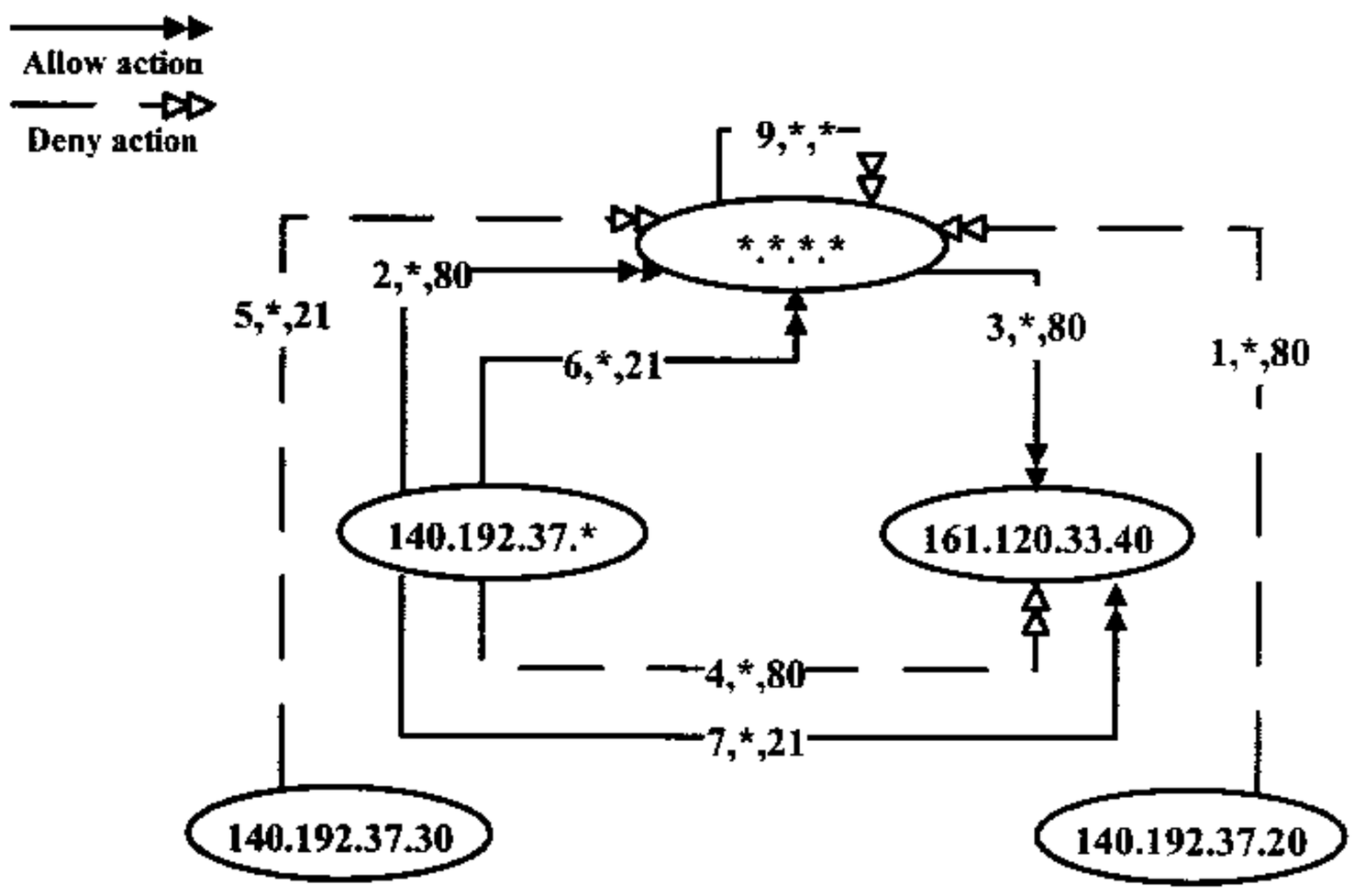
ในงานวิจัยนี้ได้กำหนดให้กฎของไฟร์วอลล์ (Firewall rule) หรือเรียกว่ากฎการกรองถูกแทนด้วยกราฟของกฎ (Policy graph) โดยกราฟนี้จะเป็นตัวแทนของกฎต่างๆ ในไฟร์วอลล์ และในขณะเดียวกันจะใช้กราฟนี้เพื่อการวิเคราะห์และค้นหาความสัมพันธ์และข้อบกพร่องของกฎต่างๆ นั้น ถ้ากลุ่มข้อมูลสามารถจับคู่กันได้ (Match) กับกฎใดกฎหนึ่งในกราฟนั้นได้ ก็ทำให้เราสามารถรู้ได้ว่ากลุ่มข้อมูลจะถูกยอมรับหรือกลุ่มข้อมูลจะถูกปฏิเสธ

บัพ (Node) หนึ่งบัพในกราฟจะแทนเขตข้อมูลหนึ่งเขตข้อมูลของกฎที่ระบุค่าไอพี และแต่ละเส้นเชื่อม (Edge) เชื่อมระหว่างบัพ 2 บัพ จะกำกับด้วย ลำดับ ช่องทางต้นทาง ช่องทางปลายทาง หรือ ชื่อไฟร์วอลล์ และกฎที่มีแอ็คชันเป็น "allow" เราจะแทนเส้นเชื่อมเส้นทึบด้วยลูกศรหัวทึบสีน้ำเงิน " ⟶ " และกฎที่มีแอ็คชันเป็น "deny" เราจะแทนเส้นเชื่อมเส้นประด้วยลูกศรหัวโปร่งสีแดง " ⇢ " โดยปลายลูกศรจะอยู่ที่บัพต้นทาง (Source node) และหัวลูกศรจะชี้ไปที่บัพปลายทาง (Destination node) ดังภาพที่ 3.1 และภาพที่ 3.3

เส้นเชื่อม 1 เส้นเชื่อม เป็นตัวแทนของกฎ 1 กฎในไฟร์วอลล์ ที่ซึ่งเริ่มต้นที่บัพๆ หนึ่งเป็นไอพีต้นทาง (src_ip) และไปสิ้นสุดที่อีกบัพหนึ่งเป็นไอพีปลายทาง (dst_ip)

## 3.3 การแทนกฎสำหรับไฟร์วอลล์เดี่ยว

กฎต่างๆ ในไฟร์วอลล์เดี่ยว (Single firewall) ที่ถูกแทนด้วยกราฟ ประกอบด้วยบัพและเส้นเชื่อมจำนวนมาก ทั้งนี้เนื่องมาจากในไฟร์วอลล์ประกอบด้วยกฎหลายๆ กฎ โดยบัพแต่ละบัพ แทนไอพีต้นทาง หรือไอพีปลายทาง และเส้นเชื่อมแต่ละเส้นเชื่อมแทนส่วนผสม หรือ Combination ของ 3 ส่วน (3-tuples) คือ <ลำดับ (Order) ช่องทางต้นทาง (Source port หรือ src_port) และช่องทางปลายทาง (Destination port หรือ dst_port)> ส่วนแอ็คชัน (Action) จะแทนด้วยสีของเส้นเชื่อมและชนิดของหัวลูกศรที่แตกต่างกันซึ่งอธิบายไว้ตามหัวข้อ 3.1 ตัวอย่างดังภาพที่ 3.1

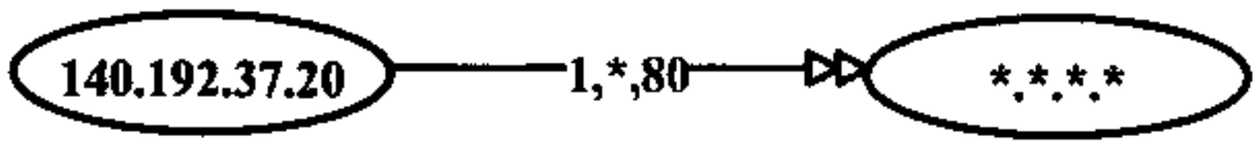



**ภาพที่ 3.1 กราฟแสดงความสัมพันธ์ของกฎต่างๆ ในไฟร์วอลล์เดี่ยว**

ตามภาพที่ 3.1 แสดงให้เห็นว่ากฎลำดับที่ 2 4 6 และ 7 ใช้บัพต้นทางร่วมกัน คือบัพ 140.192.37.* ส่วนกฎลำดับที่ 1 2 5 6 และ 9 รวมถึงกฎลำดับที่ 3 4 และ 7 ใช้บัพปลายทางร่วมกัน คือบัพ *.*.*.* (ไอพีใดๆ) และบัพ 161.120.33.40 ตามลำดับ สำหรับกฎลำดับที่ 9 ซึ่งเป็นกฎโดยปริยาย (Default filtering rule) ใช้บัพเดียวแทนทั้งไอพีต้นทางและไอพีปลายทาง เป็นที่น่าสังเกตว่ากฎลำดับที่ 8 ไม่ได้แสดงไว้ในกราฟ ทั้งนี้เป็นเพราะเกิดข้อผิดพลาดที่เรียกว่า Irrelevance Anomaly ขึ้นดังคำอธิบายในหัวข้อ 4.1.1

ตัวอย่างเช่น กฎคือ

1: 140.192.37.20, any, *.*.*.*, 80, deny



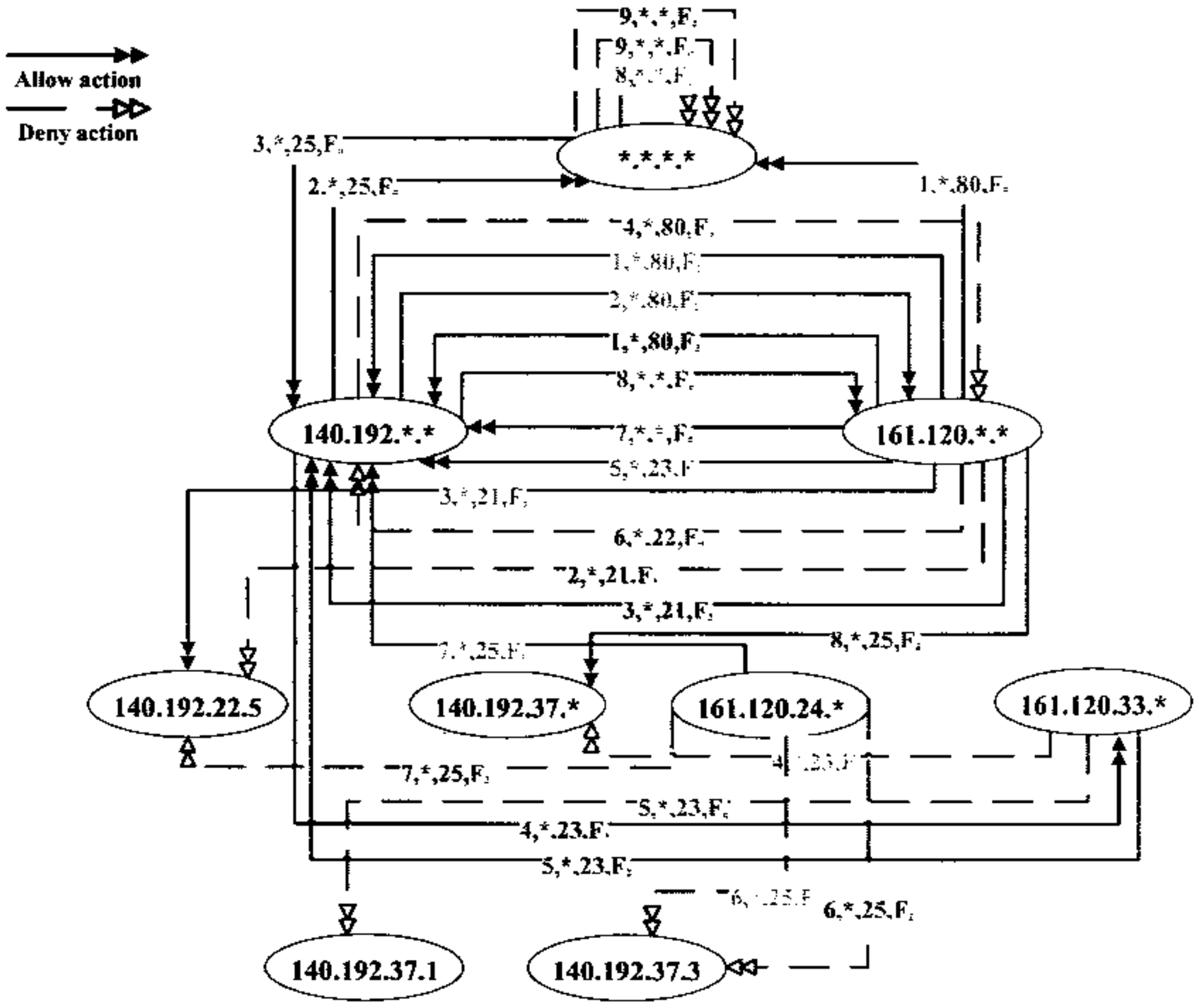

**ภาพที่ 3.2 ตัวอย่างเส้นเชื่อมที่ใช้แทนกฎ 1 กฎในไฟร์วอลล์เดี่ยว**

**สำหรับกฎดังภาพที่ 3.2 นี้หมายถึง บัพต้นทาง** 140.192.37.20 **และบัพปลายทางใดๆ มีค่าบนเส้นเชื่อมคือ** 1, *, 80 **หมายถึง กฎลำดับที่** 1 **และช่องทางต้นทางใดๆ และช่องทางปลายทางเท่ากับ** 80 **และกฎมีแอ็คชันเป็น** deny (หัวลูกศรโปร่ง)

**ในความหมายสำหรับกลุ่มข้อมูล หมายถึง กฎลำดับที่** 1 **นี้ได้ระบุไว้ว่าจะไม่อนุญาตให้กลุ่มข้อมูลที่มีไอพีต้นทาง** 140.192.37.20 **และหมายเลขช่องทางต้นทางใดๆ และไอพีปลายทางใดๆ และหมายเลขช่องทางปลายทางเท่ากับ** 80 **ผ่านไฟร์วอลล์ไปได้**

## 3.4 การแทนกฎสำหรับไฟร์วอลล์แบบกระจาย

**สำหรับในไฟร์วอลล์แบบกระจาย** (Distributed firewall) **แต่ละไฟร์วอลล์ประกอบด้วยกฎมากมาย จึงทำให้กราฟกฎของไฟร์วอลล์โดยรวมมีเส้นเชื่อมจำนวนมากที่ซ้อนทับกันอยู่ ดังภาพที่** 3.3 **แต่ละบัพจะแทนบัพต้นทาง** (src_ip) **หรือบัพปลายทาง** (dst_ip) **เช่นเดียวกับการแทนบัพในไฟร์วอลล์เดี่ยว และเส้นเชื่อมแทน** Combination 4 **ส่วน** (4-tuples) **คือ <ลำดับ** (Order) **ช่องทางต้นทาง** (src_port) **ช่องทางปลายทาง** (dst_port) **และชื่อไฟร์วอลล์** (Firewall identification ($F_n$)> **เมื่อ** n **บ่งบอกว่าเป็นไฟร์วอลล์เดี่ยวตัวใด)**

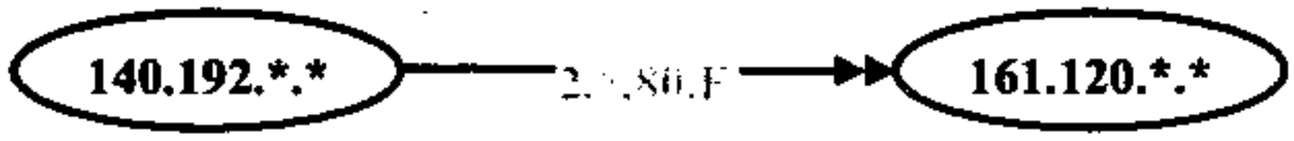



**ภาพที่ 3.3** กราฟแสดงความสัมพันธ์ของกฎโดยรวมของไฟร์วอลล์แบบกระจาย

ตามภาพที่ 3.3 แสดงให้เห็นว่าเส้นเชื่อมต่างๆ ที่แทนกฎเหล่านี้มีอยู่เป็นจำนวนมาก บัพแต่ละบัพจะถูกใช้แทนไอพีต้นทางหรือไอพีปลายทางหรือแทนค่าของทั้งสองไอพีก็เป็นได้ ตัวอย่างเช่น กฎลำดับที่ $2/F_0$ $4/F_0$ $8/F_0$ $2/F_1$ และ $4/F_2$ ใช้บัพต้นทางร่วมกันคือบัพ 140.192.*.* ส่วนกฎลำดับที่ $3/F_1$ $2/F_2$ และ $7/F_2$ ใช้บัพปลายทางร่วมกันคือบัพ 140.192.22.5 สำหรับกฎลำดับที่ $9/F_0$ $8/F_1$ และ $9/F_2$ ซึ่งเป็นกฎโดยปริยายใช้บัพเดียวแทนทั้งไอพีต้นทางและไอพีปลายทาง

ตัวอย่างเช่นไฟร์วอลล์ตัวที่ 1 มีกฎคือ

2: 140.192.*.*, any, 161.120.*.*, 80, allow




**ภาพที่ 3.4 ตัวอย่างเส้นเชื่อมที่ใช้แทนกฎ 1 กฎในไฟร์วอลล์แบบกระจาย**

**สำหรับกฎดังภาพที่ 3.4 นี้หมายถึง บัพต้นทาง 140.192.*.* และบัพปลายทาง 161.120.*.* มีค่าบนเส้นเชื่อมคือ 2, *, 80, $F_1$ หมายถึง กฎลำดับที่ 2 และช่องทางต้นทางใดๆ และช่องทางปลายทางเท่ากับ 80 และกฎนี้อยู่ในไฟร์วอลล์ตัวที่ 1 และกฎมีแอ็คชันเป็น allow (หัวลูกศรทึบ)**

**ในความหมายสำหรับกลุ่มข้อมูล หมายถึง กฎลำดับที่ 2 นี้ได้ระบุไว้ว่าจะอนุญาตให้กลุ่มข้อมูลที่มีไอพีต้นทาง 140.192.*.* และหมายเลขช่องทางต้นทางใดๆ และปลายทางไปที่ไอพี 161.120.*.* และหมายเลขช่องทางปลายทางเท่ากับ 80 ผ่านไฟร์วอลล์ตัวที่ 1 ไปได้**