บทที่ 4

ผลของการวิจัย

ในบทที่ 4 นำเสนอผลการทดลองระบบ DSDSE โดยแบ่งการทดลองออกเป็น 2 ส่วน ได้แก่ การตรวจสอบบอตเน็ต และการควบคุมอัตราการส่งจดหมายอิเล็กทรอนิกส์สแปมจากเครื่องบอตเน็ต สำหรับวิธีในการตรวจสอบบอตเน็ตสามารถแบ่งการตรวจสอบออกเป็น 2 กรณี คือ การตรวจสอบเครื่องคอมพิวเตอร์จากพฤติกรรมการสอบถามชื่อ Domain Name จากระบบ DNS และการตรวจสอบปริมาณการส่งจดหมายอิเล็กทรอนิกส์ที่ผิดปรกติโดยการกำหนดค่า Threshold มีขั้นตอนรายละเอียด ดังนี้

รายละเอียดในการตรวจสอบบอตเน็ตแบ่งเป็น 2 กรณี มีรายละเอียดดังนี้

1) การตรวจสอบบอตเน็ต โดยพิจารณากลุ่มเครื่องคอมพิวเตอร์ที่สอบถาม Domain Name เดียวกันในเวลาเดียวกัน ผ่านระบบ DNS โดยอาศัยค่า Thershold และการคำนวณค่า Similarity เข้ามาช่วยวิเคราะห์ว่า Domain Name ใดที่เครื่องบอตเน็ตต้องการติดต่อเพื่อรอรับคำสั่งให้ทำงานตามที่ผู้ควบคุมบอตเน็ตต้องการ นอกจากนั้นยังสามารถสอบกลับได้ว่า เครื่องใดบ้างที่พยายามติดต่อกับ Domain Name นี้ ซึ่งเป็นเครื่องที่น่าสงสัยว่า เป็นเครื่องบอตเน็ต

2) การตรวจสอบปริมาณการส่งจดหมายอิเล็กทรอนิกส์ที่ผิดปรกติ โดยพิจาณาหาค่ากลางหรือค่า Threshold ของการส่งจดหมายอิเล็กทรอนิกส์ จำนวน 7 วันเมื่อได้ค่า Threshold มาแล้ว นำค่า Threshold เปรียบเทียบกับวันที่เราต้องการหาผลลัพธ์ว่าเป็นบอตเน็ตส่งจดหมายอิเล็กทรอนิกส์สแปมหรือไม่ โดยการพิจารณาจำนวนการส่งจดหมายทั้งหมดในวันนั้น ถ้าจำนวนจดหมายอิเล็กทรอนิกส์มีปริมาณการส่งมากกว่า หรือ เท่ากับ ค่า Thresold แสดงว่าวันนั้นมีการส่งจดหมายอิเล็กทรอนิกส์

รายละเอียดการควบคุมอัตราการส่งจดหมายอิเล็กทรอนิกส์

นำเสนอวิธีและผลการทดลอง การลดอัตราการส่งจดหมายอิเล็กทรอนิกส์สแปมจากเครื่องบอตเน็ต โดยผ่านระบบ DSDSE 3 รูปแบบ

4.1 การตรวจสอบเครื่องบอตเน็ตจากพฤติกรรมการร้องขอตรวจสอบ Domain Name จากระบบ DNS

แบ่งการวิเคราะห์ 2 ขั้นตอน ได้แก่ การวิเคราะห์หาค่า Threshold ที่เหมาะสม และการวิเคราะห์หาค่า Similarity ของแต่ละ Domain Name ที่บอตเน็ตพยายามติดต่อด้วย

ตารางที่ 4.1 แสดงข้อมูลการหาค่า Threshold

Domain Name	Total
1 Count	134
ad.tarad.com	2
ad.yieldmanager.com	2
ads.adbrite.com	2
ads.admaxasia.com	2
ads.manager.co.th	2
ads.manager.co.th	2
ads.silkspan.com	2
amp.sageanalyst.net	2
content.yieldmanager.edgesuite.net	2
eas.apm.emediate.eu	2
imgads.manager.co.th	2
m1.webstats.motigo.com	2
m1.webstats4u.com	2
mail.lmginsurance.co.th	2
mpics.manager.co.th	2
pagead2.googlesyndication.com	2
search.tarad.com	2
spe.atdmt.com	2
st.sageanalyst.net	2
truehits1.gits.net.th	2
view.atdmt.com	2
www.google.co.th	2
www.hp.com	2
www.manager.co.th	2

Domain Name	Total
www.positioningmag.com	2
www.sib.co.th	2
www.tarad.com	2
www.thaiasiatoday.com	2
www.thaisecondhand.com	2
www.weareit.net	2
2 Count	29
www.settrade.com	3
3 Count	1
www.google-analytics.com	4
4 Count	1
hits.truehits.in.th	6
lvs.truehits.in.th	6
6 Count	2
irc.bot.com	10
10 Count	1
Grand Count	168

จากตารางที่ 4.1 ข้อมูลการหาค่า Threshold ผลลัพธ์ที่ได้แบ่งเป็น 6 กลุ่ม มีรายละเอียดดังนี้

1) **1 count** หมายถึง มีจำนวน Domain Name ทั้งหมด 134 Domain Names และแต่ละ Domain Name มี IP Address เดียว หรือเครื่องเพียง 1 เครื่องเท่านั้นที่ทำการสอบถามชื่อ Domain Name นี้ในแต่ละชั่วโมง ซึ่งถือว่าเป็นเครื่องที่ปรกติ (ไม่เป็นกลุ่มบอตเน็ต)

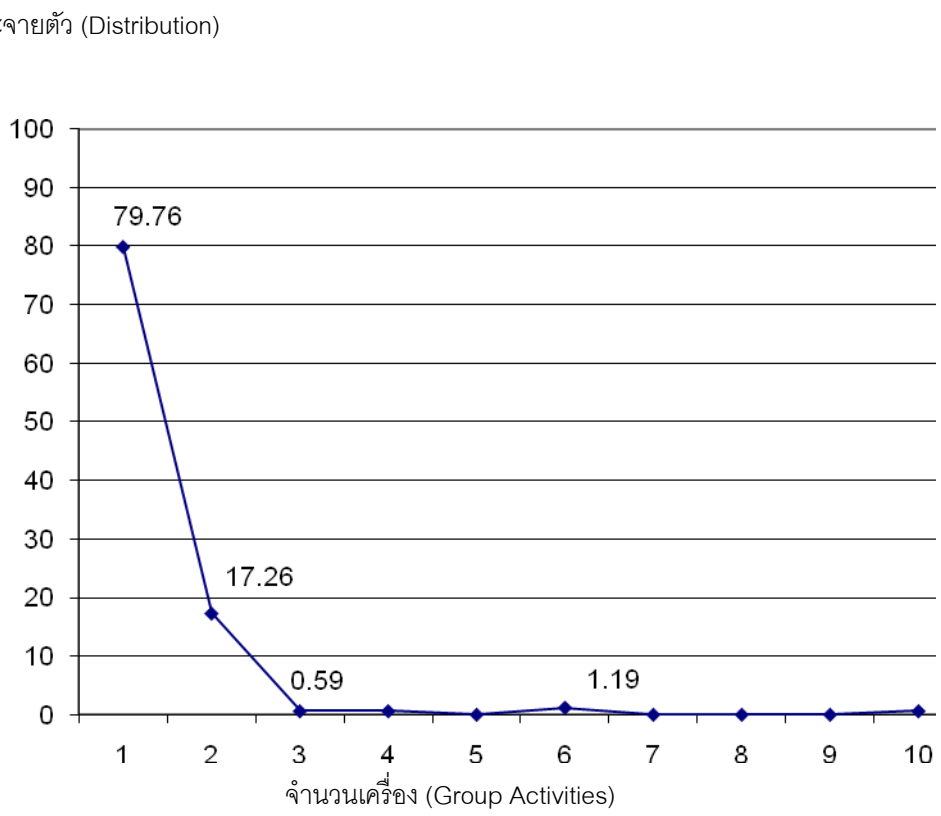
2) **2 count** หมายถึง มีจำนวน Domain Name ทั้งหมด 29 Domain Names ที่แต่ละ Domain Name มี 2 IP Address ที่ติดต่อกับ Domain Name นี้พร้อมกัน ซึ่งถือว่าปรกติ

3) **3 count** หมายถึง มีจำนวน Domain Name เพียง 1 Domain Name เท่านั้นที่มี 3 IP Addresses ติดต่อกับ Domain Name นี้พร้อมกัน ซึ่งยังถือว่าปรกติ

4) **7, 8, 9 count** ไม่พบ หมายถึง ไม่มี Domain Name ใดที่มี IP Address ติดต่อพร้อมกันเป็นจำนวน 7, 8, 9 IP Address

5) **10 count** หมายถึง มีเพียง 1 Domain Name เท่านั้นที่มีกลุ่มเครื่องคอมพิวเตอร์เรียกสอบถาม Domain Name นี้พร้อมๆกันในเวลาเดียวกัน เป็นจำนวน 10 เครื่องหรือ 10 IP Address ซึ่งถือว่าเป็น Domain Name ที่ผิดปรกติ ซึ่งเป็นกลุ่มกิจกรรมหรือพฤติกรรมของเครื่องที่เป็นบอตเน็ต

ภาพที่ 4.1 แสดงการหาค่า Threshold ของกลุ่มเครื่องที่สอบถามระบบ DNS



จากภาพที่ 4.1 นำผลการวิเคราะห์แสดงออกมาเป็นกราฟผลลัพธ์ของ Threshold ที่เหมาะสม คือ 80% หรือคิดเป็น 0.8 ค่า Threshold นี้จะถูกนำมาเป็นมาตรฐานในการเปรียบเทียบต่อไป

การพิจารณากราฟ จำนวนเครื่อง 1 เครื่องมีการกระจายตัวที่ 80% หรือ 0.8 หมายความว่าเป็นการทำงานของแต่ละเครื่องที่ปรกติ โดยมีเพียง 1 เครื่องที่ทำการตรวจสอบรายชื่อ Domain Name แตกต่างกัน การพิจารณาจะเริ่มพิจารณาจากกลุ่มของเครื่องตั้งแต่

2 เครื่องขึ้นไปที่มีการร้องขอให้ตรวจสอบรายชื่อ Domain Name เดียวกันพร้อมๆกัน 2 เครื่องขึ้นไป เรียกว่าเป็นการพิจารณากลุ่มของกิจกรรม (Group Activities)

ตารางที่ 4.2 แสดงการหาค่า Similarity จาก Domain Name ทั้งหมดในวันที่ 29 เมษายน 2551

DomainName/Hrs.	Similarity (0.5 * ((C/A) + (C/B))										Average
	1	2	3	4	5	6	7	8	9	10	10 HRs.
10.132.194.80	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
a.rad.live.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.1
ad.doubleclick.net	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ad.tarad.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ad.yieldmanager.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.1
ads.adapterproject.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ads.adbrite.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ads.admaxasia.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ads.manager.co.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ads.mweb.co.th	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.2
ads.silkspan.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ads1.msn.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
amp.sageanalyst.net	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
bs.serving-sys.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
c.msn.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.1
config.office.microsoft.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
content.yieldmanager.edgesuite.net	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
d.as.sanook.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
eas.apm.emediate.eu	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
edge.quantserve.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
gfx1.hotmail.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.1
gfx2.hotmail.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
go.microsoft.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
h.live.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.1
hits.truehits.in.th	0.0	0.7	0.0	0.5	0.8	0.3	0.3	0.7	0.7	0.6	0.5
i.ytimg.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
img.as.sanook.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
imgads.manager.co.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
irc.bot.com	0.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9
j.clickdensity.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
java.sun.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
javadl-esd.sun.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ll.atdmt.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
login.yahoo.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
lvs.truehits.in.th	0.0	0.7	0.3	0.7	0.8	0.3	0.3	0.7	0.7	0.7	0.5
m1.2mdn.net	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

DomainName/Hrs.	Similarity (0.5 * ((C/A) + (C/B))										Average
	1	2	3	4	5	6	7	8	9	10	10 HRs.
m1.webstats.motigo.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
m1.webstats4u.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mail.lmginsurance.co.th	0.0	0.0	0.0	0.0	1.0	1.0	0.5	0.8	0.8	1.0	0.5
media.nipa.co.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mpics.manager.co.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
msnportal.112.2o7.net	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
office.microsoft.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
pagead2.googlesyndication.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
pixel.quantserve.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
r.clickdensity.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
r.office.microsoft.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
rad.live.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.1
rad.msn.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
rmd.atdmt.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sb.google.com	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
search.tarad.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
spe.atdmt.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.4	0.0	0.1
st.sageanalyst.net	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
static.sanook.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
th.msn.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
tns.nipa.co.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
truehits1.gits.net.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
us.bc.yahoo.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
view.atdmt.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.8	0.8	0.2
www.google.co.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	1.0	0.7	0.3
www.google.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.google-analytics.com	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.8	0.8	0.8	0.3
www.hp.com	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	1.0	1.0	0.6
www.jobsdb.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.manager.co.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.microsoft.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.one2car.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.positioningmag.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.sanook.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.settrade.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.8	0.4	0.2
www.sib.co.th	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.tarad.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.thaiasiatoday.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.thaisecondhand.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.weareit.net	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
www.youtube.com	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

หมายเหตุ ค่า A,B,C ของแต่ละ Domain Name สามารถดูได้จากภาคผนวก

สูตรการคำนวณค่า Similarity = (0.5 * ((C/A) + (C/B)) โดยที่ (A และ B <> 0)

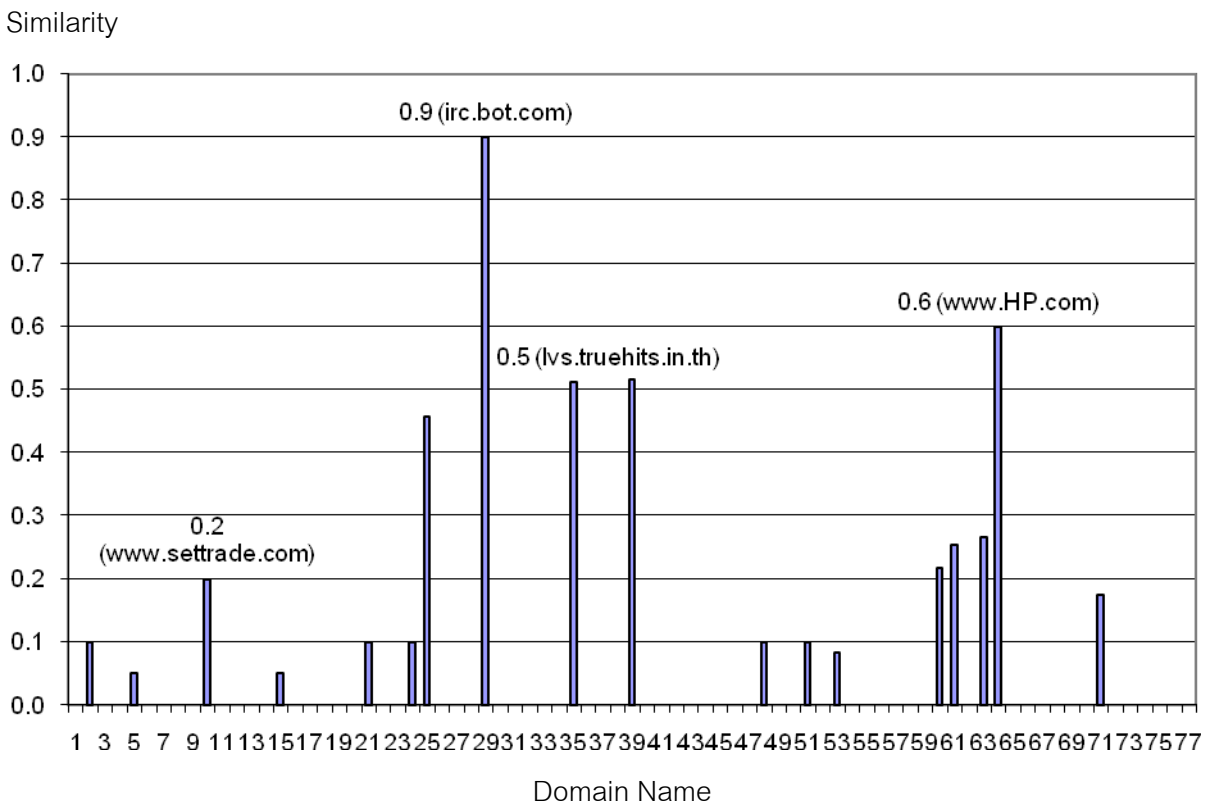
A = จำนวน IP ณ ช่วงเวลาที่ 1 หรือในชั่วโมงที่ 1

B = จำนวน IP ณ ช่วงเวลาที่ 2 หรือในชั่วโมงที่ 2 ชั่วโมงถัดไป

C = จำนวน IP ที่เหมือนกัน (Intersection) ของ IP ระหว่าง A กับ B

จากตารางที่ 4.2 แสดงให้เห็นค่า Semilarity ของแต่ละ Domain Name ย้อนหลัง 10 ชั่วโมงใน 1 วัน พิจารณาค่า Similarity ของแต่ละ Domain Name เปรียบเทียบกับค่า Threshold ที่ได้คำนวณคือ 0.8 ถ้า Domain Name ใดที่มีค่าคำนวณ Similarity เกินกว่า 0.8 แสดงว่า Domain Name นั้นเป็น Domain Name ที่มีการติดต่อกับเครื่องบอตเน็ต และเราสามารถสรุปได้ว่าเครื่องคอมพิวเตอร์เครื่องใด IP Address ใดที่มีการติดต่อกับ Domain Name นี้ถือว่าเป็นเครื่องบอตเน็ต เช่นกัน

ภาพที่ 4.2 แสดงค่า Similarity ของแต่ละ Domain Name



จากภาพที่ 4.2 แสดงให้เห็นว่าหลังจากการคำนวณค่า Similarity ของแต่ละ Domain Name แสดงให้เห็นว่า Domain Name ของ irc.bot.com มีค่า Similarity เท่ากับ 0.9 ซึ่งสูงเกินกว่าค่า Threshold ที่กำหนดไว้ คือ 0.8 ดังนั้น สามารถสรุปได้ว่า irc.bot.com เป็น Domain ที่มีกลุ่มของบอตเน็ตพยายามติดต่อสื่อสารด้วย

4.2 การตรวจสอบเครื่องบอตเน็ตจากการตรวจสอบการส่งจดหมายอิเล็กทรอนิกส์สแปมโดยวิธีกำหนดค่า Threshold

การตรวจสอบเครื่องบอตเน็ตจากการตรวจสอบการส่งจดหมายอิเล็กทรอนิกส์สแปม มีขั้นตอนในการตรวจสอบ ดังนี้

4.2.1 ขั้นตอนการจำลองข้อมูลจดหมายอิเล็กทรอนิกส์ปรกติและจดหมายอิเล็กทรอนิกส์สแปม

1) เริ่มการจำลองข้อมูลตั้งแต่ วันที่ 1 ถึงวันที่ 7 โดยกำหนดให้แต่ละวันมีการส่งจดหมายอิเล็กทรอนิกส์แบบปกติของผู้ใช้งานทั่วไปสูงสุดไม่เกิน 120 ฉบับต่อวัน มีหลักการคิดคำนวณดังนี้

1.1) การส่งจดหมายอิเล็กทรอนิกส์ได้ออกแบบให้มีการส่งจดหมายอิเล็กทรอนิกส์ทุกๆ 12 นาทีทำการส่ง 1 ครั้ง ดังนั้นจะมีการส่งสูงสุดได้เพียง 5 ครั้งใน 1 ชั่วโมง

1.2) การกำหนดให้แต่ละครั้งใน 1 ชั่วโมง มีโอกาสที่จะส่งหรือไม่ส่งก็ได้ โดยแต่ละครั้งจะสุ่มค่า 2 ค่า คือ ส่ง หรือไม่ส่ง จดหมายอิเล็กทรอนิกส์

1.3) นับรวมทุกๆ ชั่วโมงรวม 24 ชั่วโมง ผลลัพธ์ที่ได้จะเป็นข้อมูลการส่งจดหมายอิเล็กทรอนิกส์ปรกติของทุกๆเครื่อง (IP Address) ในเวลา 1 วัน จำนวนจดหมายอิเล็กทรอนิกส์สูงสุดไม่เกิน 120 ฉบับต่อ 1 วัน

1.4) จำลองข้อมูลในลักษณะนี้รวมทั้งหมด 7 วัน

2) ข้อมูลในวันที่ 8 เป็นการจำลองข้อมูลส่วนที่มีเครื่องบอตเน็ตเริ่มส่งจดหมายอิเล็กทรอนิกส์สแปมแบ่งออกเป็น 2 ส่วนย่อยคือ

2.1) ส่วนเครื่องที่เป็นบอตเน็ต การส่งจดหมายอิเล็กทรอนิกส์สแปมจะเพิ่มขึ้นตามจำนวนเครื่องที่เป็นบอตเน็ต โดยเริ่มตั้งแต่บอตเน็ต 10 20 30 40 50 60 70 80 90 และบอต

เน็ต 100 เปอร์เซ็นต์ เช่น บอตเน็ต 10 เปอร์เซ็นต์ หมายถึง มีจำนวนเครื่องที่เป็นบอตเน็ตจำนวน 10 เครื่อง อีก 90 เครื่องเป็นเครื่องปรกติ

2.2) เครื่องที่เป็นบอตเน็ตจะมีการสุ่มจำนวนจดหมายอิเล็กทรอนิกส์สแปมที่แตกต่างออกไป โดยจะใช้วิธีสุ่มส่งจดหมายอิเล็กทรอนิกส์ทุก ๆ 3 นาที (ตามพฤติกรรมของบอตเน็ต) ดังนั้นจะส่งทั้งหมด 20 ครั้งใน 1 ชั่วโมง รวม 24 ชั่วโมง ในการส่งแต่ละครั้งจะถูกกำหนดค่าความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปม เริ่มตั้งแต่ 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 และ 1 ยกตัวอย่างเช่น เครื่องบอตเน็ต 10 เครื่อง มีโอกาสความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปม 0.1 หมายถึง บอตเน็ต จำนวน 10 เครื่อง มีโอกาสในการส่งจดหมายอิเล็กทรอนิกส์สแปมแต่ละครั้งของการส่งเริ่ม ครั้งที่ 1 จนถึงครั้งที่ 20 ในช่วง 1 ชั่วโมง มีโอกาสแต่ละครั้งที่จะส่งจดหมายอิเล็กทรอนิกส์สแปม ที่ 0.1

2.3) ส่วนของเครื่องที่ไม่ได้เป็นบอตเน็ตที่เหลืออีก 90 เครื่อง จะใช้เงื่อนไขการสุ่มตัวเลขเช่นเดียวกันกับวิธีการส่งจดหมายอิเล็กทรอนิกส์แบบปรกติ

4.2.2 ขั้นตอนการตรวจสอบข้อมูลจดหมายเล็กทรอนิกส์ปรกติและจดหมายอิเล็กทรอนิกส์สแปมโดยวิธีการกำหนดค่า Threshold

ในการทำงาน ระบบจะพิจารณาจดหมายอิเล็กทรอนิกส์เพื่อระบุว่าเป็นจดหมายอิเล็กทรอนิกส์สแปมหรือไม่ อัลกอริธึมที่นำมาใช้ คือ การหาค่า Threshold ของการส่งจดหมายฉบับก่อนๆ ย้อนหลังไป 7 วัน แล้วนำค่า Threshold ที่ได้มาเปรียบเทียบกับข้อมูลการส่งสูงสุดรวมทั้งวันของวันที่ 8 หรือวันที่ต้องการเปรียบเทียบ ถ้าอัตราการส่งสูงสุดของวันที่ต้องการเปรียบเทียบ มากกว่าหรือเท่ากับค่าของ Threshold แสดงว่าวันนั้นมีการส่งจดหมายอิเล็กทรอนิกส์สแปมจากเครื่องบอตเน็ต

ในการกำหนดค่า Threshold ที่นำมาใช้ในการตรวจสอบเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปม ผู้วิจัยได้ทดลองกำหนดค่า Threshold ที่แตกต่างกันไป วัตถุประสงค์เพื่อหาค่า Threshold ที่เหมาะสมและดีที่สุดในทุกๆกรณี โดยเมื่อระบบคำนวณหาค่า Thresold เรียบร้อยแล้ว ผู้วิจัยจะทำการทดลองเพิ่มเติมโดยการกำหนดค่า Threshold ที่แตกต่างกัน เริ่มจากทดสอบค่า Threshold ที่ ¼ Threshold ½ Threshold ¾ Threshold และ 1 Threshold

4.2.3 การทดลองการส่งจดหมายอิเล็กทรอนิกส์สแปมโดยวิธีการกำหนดค่า Threshold

การทดลองเริ่มจากกำหนดจำนวนบอตเน็ตที่ทำการส่งจดหมายอิเล็กทรอนิกส์สแปมตั้งแต่ 10 เปอร์เซ็นต์ จนถึง 100 เปอร์เซ็นต์ จากจำนวนเครื่องคอมพิวเตอร์ทั้งหมด 100 เครื่อง (จำนวนบอตเน็ตเริ่มที่10 20 30 40 ถึง 100 เครื่อง) แต่ละเปอร์เซ็นต์จะมีโอกาสส่งจดหมายอิเล็กทรอนิกส์สแปมที่ความน่าจะเป็นแตกต่างกัน 10 กรณี คือ 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 และ 1 ยกตัวอย่างเช่น ความน่าจะเป็นในการส่งจดหมาย 0.1 หมายถึง ทุกครั้งที่บอตเน็ตติดต่อกับผู้ควบคุมบอต เมื่อได้รับการติดต่อให้บอตเน็ตส่งจดหมายอิเล็กทรอนิกส์สแปมจะมีโอกาสในการส่งจดหมายอิเล็กทรอนิกส์สแปมเพียง 0.1 ซึ่งหมายถึงโอกาสที่จะส่งจดหมายอิเล็กทรอนิกส์สแปมน้อยกว่าความน่าจะเป็นที่ 1

ตัวอย่างข้อมูลการทดลองการส่งจดหมายอิเล็กทรอนิกส์สแปมจากเครื่องบอตเน็ต และการส่งจดหมายอิเล็กทรอนิกส์ธรรมดาจากเครื่องปรกติ แสดงการคิดคำนวณหาค่า MAX และการเปรียบเทียบค่า Threshold เพื่อสามารระบุได้ว่าเป็นจดหมายอิเล็กทรอนิสก์สแปม ดังแสดงข้อมูลตารางที่ 4.3

ตารางที่ 4.3 แสดงข้อมูลตัวอย่างการส่งจดหมายอิเล็กทรอนิกส์ของเครื่อง 100 เครื่องโดยมีเครื่องบอตเน็ต 10 % ความน่าจะเป็น 0.2 และ Threshold = ½ Threshold

IP	ปริมาณการส่งจดหมายอิเล็กทรอนิกส์ วันที่ 1-7 ส่งแบบปรกติ วันที่ 8 ส่งแบบสแปม								ผลลัพธ์		
	วันที่ 1	วันที่ 2	วันที่ 3	วันที่ 4	วันที่ 5	วันที่ 6	วันที่ 7	วันที่ 8	MAX	½ TH	Spam
1	67	67	54	61	63	66	68	97	68	60.5	1
2	61	61	57	65	62	53	61	103	65	60.5	1
3	56	48	61	59	61	56	58	91	61	60.5	1
4	62	62	56	56	60	65	67	96	67	60.5	1
5	56	61	69	60	54	51	63	98	69	60.5	1
6	63	67	68	47	58	62	64	109	68	60.5	1
7	67	61	57	60	63	59	62	101	67	60.5	1
8	60	61	55	49	64	64	63	113	64	60.5	1
9	58	66	63	60	54	62	59	94	66	60.5	1
10	66	60	59	53	61	66	59	98	66	60.5	1
11	55	56	63	74	50	64	62	61	74	60.5	1
12	59	62	61	53	61	58	58	67	62	60.5	1
13	60	56	72	65	67	52	60	60	72	60.5	0

IP	ปริมาณการส่งจดหมายอิเล็กทรอนิกส์ วันที่ 1-7 ส่งแบบปรกติ วันที่ 8 ส่งแบบสแปม								ผลลัพธ์		
	วันที่ 1	วันที่ 2	วันที่ 3	วันที่ 4	วันที่ 5	วันที่ 6	วันที่ 7	วันที่ 8	MAX	½ TH	Spam
14	58	60	65	59	73	60	65	56	73	60.5	0
15	52	61	70	60	53	58	66	57	70	60.5	0
16	62	63	62	60	58	59	56	51	63	60.5	0
17	64	57	64	60	56	61	55	64	64	60.5	1
18	60	57	57	55	50	57	57	63	60	60.5	1
19	53	65	63	66	61	63	65	66	66	60.5	1
20	57	52	66	60	53	59	70	64	70	60.5	1
21	56	54	60	56	56	62	59	48	62	60.5	0
22	60	62	64	43	54	62	63	65	64	60.5	1
23	66	57	63	67	59	56	57	72	67	60.5	1
24	55	55	57	54	64	55	68	57	68	60.5	0
25	52	57	62	64	61	58	64	55	64	60.5	0
26	71	55	45	63	67	65	64	57	71	60.5	0
27	68	58	51	62	60	68	57	63	68	60.5	1
28	65	56	54	63	49	57	52	72	65	60.5	1
29	54	65	52	56	56	54	55	69	65	60.5	1
30	65	59	65	58	68	63	50	61	68	60.5	1
31	56	57	59	60	59	57	54	60	60	60.5	0
32	56	61	53	63	65	51	51	60	65	60.5	0
33	64	57	59	59	69	59	58	64	69	60.5	1
34	75	65	64	60	71	63	59	63	75	60.5	1
35	63	66	71	64	64	60	60	55	71	60.5	0
36	62	54	61	61	62	63	67	67	67	60.5	1
37	57	67	66	61	50	67	62	59	67	60.5	0
38	57	67	62	57	46	69	60	56	69	60.5	0
39	47	60	67	65	71	60	61	66	71	60.5	1
40	60	58	59	52	50	59	54	62	60	60.5	1
41	62	51	62	55	74	58	61	58	74	60.5	0
42	58	58	63	54	65	49	64	63	65	60.5	1
43	65	64	61	54	59	65	58	71	65	60.5	1
44	57	61	60	52	53	54	66	58	66	60.5	0
45	54	63	55	57	60	67	63	61	67	60.5	1
46	56	67	60	60	65	69	52	50	69	60.5	0
47	62	60	58	68	67	56	62	61	68	60.5	1
48	53	60	54	62	63	57	41	60	63	60.5	0
49	53	61	62	60	56	59	63	66	63	60.5	1

IP	ปริมาณการส่งจดหมายอิเล็กทรอนิกส์ วันที่ 1-7 ส่งแบบปรกติ วันที่ 8 ส่งแบบสแปม								ผลลัพธ์		
	วันที่ 1	วันที่ 2	วันที่ 3	วันที่ 4	วันที่ 5	วันที่ 6	วันที่ 7	วันที่ 8	MAX	½ TH	Spam
50	60	53	55	62	49	64	58	57	64	60.5	0
51	59	58	56	55	66	59	61	64	66	60.5	1
52	62	67	62	56	60	48	53	65	67	60.5	1
53	53	63	58	60	63	51	64	67	64	60.5	1
54	59	45	52	63	61	61	55	60	63	60.5	0
55	59	65	72	61	61	54	65	55	72	60.5	0
56	60	61	61	48	63	57	61	60	63	60.5	0
57	59	59	63	64	62	65	64	52	65	60.5	0
58	55	58	58	62	57	53	55	58	62	60.5	0
59	57	63	57	54	62	53	72	56	72	60.5	0
60	51	56	53	48	59	61	55	53	61	60.5	0
61	61	68	60	64	61	55	61	49	68	60.5	0
62	63	44	61	50	55	53	63	57	63	60.5	0
63	61	53	56	63	62	60	59	58	63	60.5	0
64	58	68	56	64	64	51	59	56	68	60.5	0
65	45	62	60	67	62	74	68	63	74	60.5	1
66	62	74	58	60	59	56	60	61	74	60.5	1
67	63	63	60	57	62	57	59	71	63	60.5	1
68	61	60	46	64	58	59	55	58	64	60.5	0
69	58	58	59	56	58	57	60	48	60	60.5	0
70	61	60	59	58	51	74	62	62	74	60.5	1
71	63	56	58	61	56	56	59	63	63	60.5	1
72	61	60	52	52	66	55	61	66	66	60.5	1
73	55	63	58	56	58	53	61	60	63	60.5	0
74	54	71	67	62	58	66	59	66	71	60.5	1
75	52	52	56	50	54	59	58	64	59	60.5	1
76	64	61	49	63	58	53	65	48	65	60.5	0
77	61	57	63	48	65	56	54	52	65	60.5	0
78	70	59	53	62	52	64	61	66	70	60.5	1
79	53	55	64	67	55	52	59	65	67	60.5	1
80	54	60	65	59	55	55	53	62	65	60.5	1
81	63	59	63	60	59	51	56	53	63	60.5	0
82	48	62	65	55	58	53	59	63	65	60.5	1
83	59	60	62	59	52	56	61	53	62	60.5	0
84	48	58	66	60	50	59	66	62	66	60.5	1
85	50	47	64	54	58	51	63	61	64	60.5	1

IP	ปริมาณการส่งจดหมายอิเล็กทรอนิกส์ วันที่ 1-7 ส่งแบบปรกติ วันที่ 8 ส่งแบบสแปม								ผลลัพธ์		
	วันที่ 1	วันที่ 2	วันที่ 3	วันที่ 4	วันที่ 5	วันที่ 6	วันที่ 7	วันที่ 8	MAX	½ TH	Spam
86	60	50	69	65	57	65	63	56	69	60.5	0
87	62	61	55	58	66	58	63	58	66	60.5	0
88	64	48	61	61	56	76	55	60	76	60.5	0
89	60	56	55	74	57	67	58	68	74	60.5	1
90	59	67	61	64	57	62	59	63	67	60.5	1
91	71	58	66	66	60	55	64	62	71	60.5	1
92	60	59	69	73	57	56	52	64	73	60.5	1
93	67	58	62	61	63	80	60	64	80	60.5	1
94	67	58	58	70	57	54	62	59	70	60.5	0
95	67	64	69	62	61	55	62	69	69	60.5	1
96	60	53	57	60	61	60	67	58	67	60.5	0
97	63	54	62	52	66	50	59	68	66	60.5	1
98	59	68	58	55	66	62	70	60	70	60.5	0
99	62	52	62	55	56	59	63	61	63	60.5	1
100	56	63	57	59	66	56	65	61	66	60.5	1

จากตารางที่ 4.3 แสดงจำนวนการส่งจดหมายอิเล็กทรอนิกส์ของเครื่องคอมพิวเตอร์จำนวน 100 เครื่องโดยมีเครื่องคอมพิวเตอร์ที่เป็นบอตเน็ต 10 เปอร์เซ็นต์ (IP ที่ 1 -10) เป็นเวลา 8 วัน โดยที่ข้อมูลดังกล่าวถูกสร้างขึ้นจากการสุ่มข้อมูลแบบมีเงื่อนไขคือ ความน่าจะเป็น 0.2และกำหนดค่า Threshold = ½ Threshold ซึ่งจำแนกเป็นข้อมูลการส่งจดหมายอิเล็กทรอนิกส์ปรกติตั้งแต่วันที่ 1-7 และข้อมูลการส่งจดหมายอิเล็กทรอนิกส์สแปมในวันที่ 8 มีรายละเอียดดังนี้

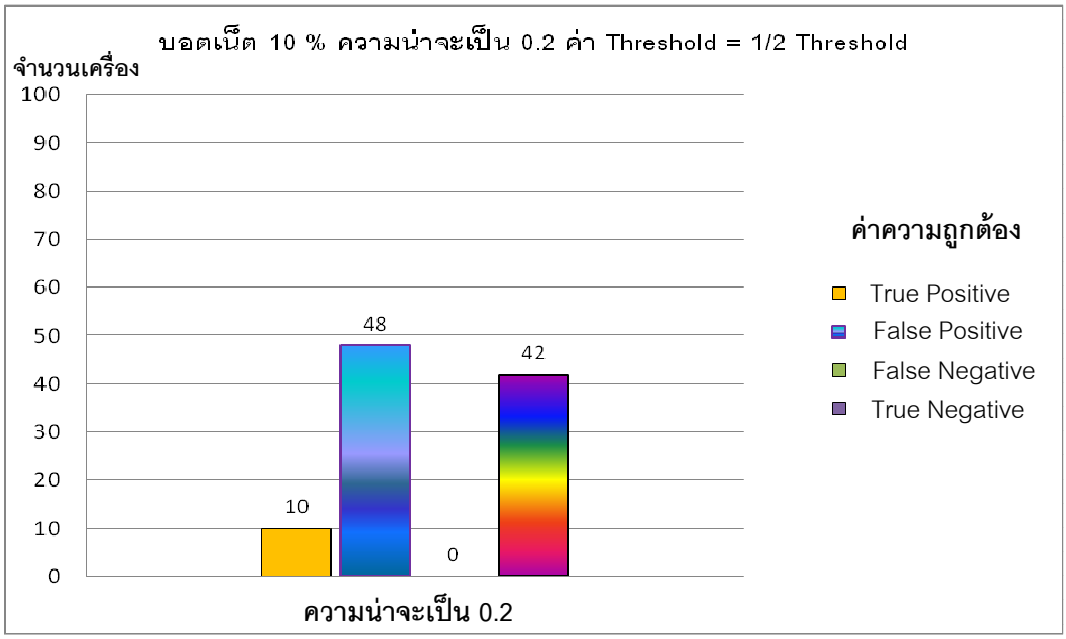
วิธีการคำนวนหาค่า Threshold จากสูตรการหาค่า Thereshld = MAX + 1

จากตารางที่ 4.3 ที่ IP Address ที่ 1 มีการคำนวณหาค่า ½ Threshold โดยคิดคำนวณหาค่า MAX จากค่าส่งสูงที่สุดของ IP Addressที่ 1 ในรอบ 7 วัน เป็น 68 ฉบับ ต้องการคำนวณหาค่า ½ Threshold โดยทำการหาค่า 1 Threshold ก่อน ดังนั้น จากจำนวนการส่งจดหมายสูงสุดใน 1 วันมีการส่งจดหมายประมาณการ 120 ฉบับ จากสูตรหาค่า MAX = MAX (68, 120) ดังนั้น ค่า MAX จะได้ 120 แล้วนำค่า MAX ดังกล่าวมาหาค่า Threshold โดยกำหนดให้เป็น Threshold = MAX + 1 จะได้ค่า 1 Threshold = 121 ฉบับ ต้องการคิด ½ Threshold จะได้ (0.5*121) = 60.5 เป็นค่า Threshold ที่เกิดจากการคำนวณการส่งจดหมายอิเล็กทรอนิกส์ปรกติรวม 7 วัน

คิดคำนวณลักษณะนี้กับทุกๆ IP Address เมื่อได้ค่า Threshold แล้วจะนำค่าดังกล่าวมาเปรียบเทียบกับปริมาณการส่งจดหมายอิเล็กทรอนิกส์ของเครื่องแต่ละเครื่องในวันที่ 8 ถ้าเครื่องใดมีค่าการส่งจดหมายอิเล็กทรอนิกส์ในวันที่ 8 เกินกว่าค่า Threshold จะถือว่าเป็นการส่งจดหมายอิเล็กทรอนิกส์สแปม

จากตารางที่ 4.3 สามารถนำมาโดยการสร้างกราฟความน่าจะเป็น 0.2 และ Threshold = ½ Threshold เพื่อวิเคราะห์ผลการทำงานของระบบ DSDSE ได้ดังภาพที่ 4.3

ภาพที่ 4.3 แสดงข้อมูลค่าความถูกต้องของ บอตเน็ต 10 เปอร์เซ็นต์ ความน่าจะเป็น 0.2 และการกำหนดค่า Threshold = ½ Threshold



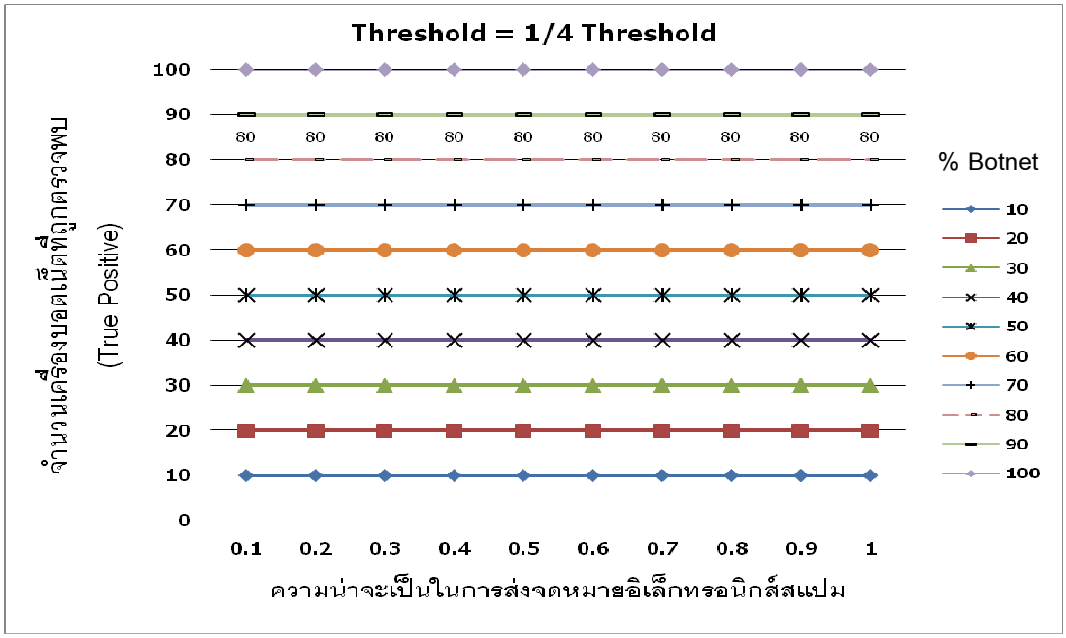
จากภาพที่ 4.3 สามารถแสดงค่าความถูกต้องจากตัวอย่างข้อมูลตารางที่ 4.3 ได้ดังนี้

True Positive	10 เครื่อง	จำนวนเครื่องที่ส่งจดหมายอิเล็กทรอนิกส์สแปมและระบบสามารถระบุถูก
False Positive	48 เครื่อง	จำนวนเครื่องที่ไม่ส่งจดหมายอิเล็กทรอนิกส์สแปม แต่ระบบระบุผิดว่าส่งจดหมายอิเล็กทรอนิกส์สแปม
False Negative	0 เครื่อง	จำนวนเครื่องที่ส่งจดหมายอิเล็กทรอนิกส์สแปม แต่ระบบระบุผิดว่าไม่ส่งจดหมายอิเล็กทรอนิกส์สแปม
True Negative	42 เครื่อง	จำนวนเครื่องที่ไม่ส่งจดหมายอิเล็กทรอนิกส์สแปมและระบบระบุถูกว่าไม่ส่งจดหมายอิเล็กทรอนิกส์สแปม

4.2.4 ผลการวิเคราะห์ข้อมูลการตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมโดยวิธีกำหนดค่า Threshold

ผลการวิเคราะห์นี้จะแสดงให้เห็นถึงผลลัพธ์ของการกำหนดค่า Threshold ที่แตกต่างกัน และการกำหนดค่าความน่าจะเป็นของการส่งจดหมายอิเล็กทรอนิกส์สแปมที่แตกต่างกัน

ภาพที่ 4.4 แสดงจำนวนบอตเน็ตที่ถูกตรวจพบเมื่อกำหนด Threshold = ¼ Threshold



ภาพที่ 4.4 แสดงให้เห็นถึงจำนวนเครื่องบอตเน็ตที่ถูกตรวจพบ เมื่อกำหนดค่า Threshold เป็น ¼ Threshold จะเห็นว่าระบบสามารถตรวจพบบ็อตที่ทำการส่งสแปมได้ครบทั้งหมด ไม่ว่าจะมีความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปมเท่าใดก็ตาม แสดงข้อมูลตามตารางที่ 4.4

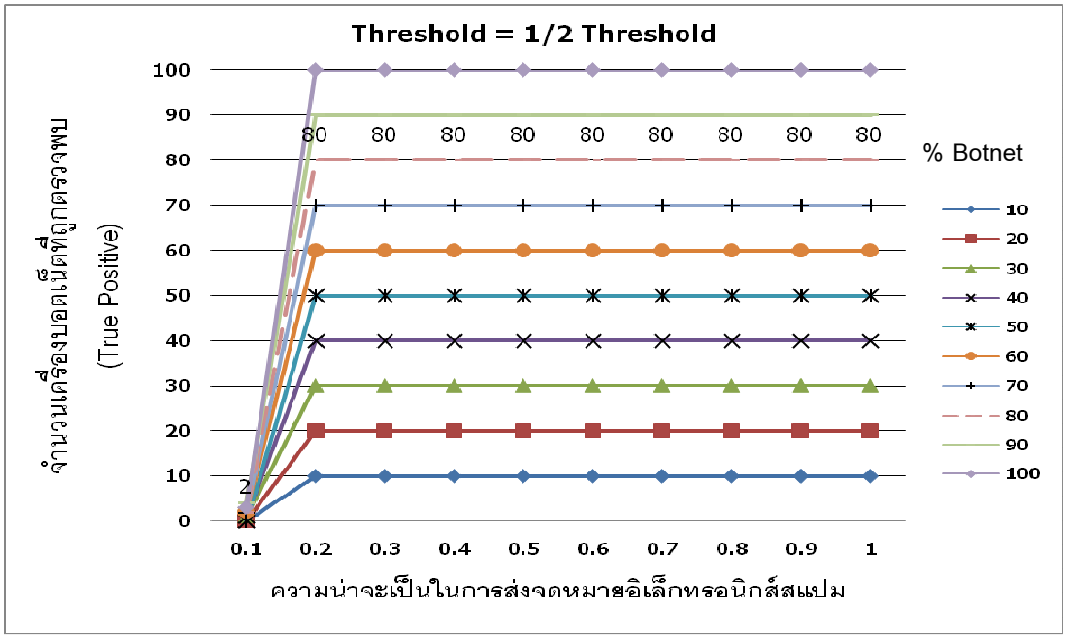
ตารางที่ 4.4 แสดงข้อมูลจำนวนบอตเน็ตที่ตรวจพบ ณ Threshold = ¼ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	10	20	30	40	50	60	70	80	90	100
0.2	10	20	30	40	50	60	70	80	90	100
0.3	10	20	30	40	50	60	70	80	90	100
0.4	10	20	30	40	50	60	70	80	90	100

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.5	10	20	30	40	50	60	70	80	90	100
0.6	10	20	30	40	50	60	70	80	90	100
0.7	10	20	30	40	50	60	70	80	90	100
0.8	10	20	30	40	50	60	70	80	90	100
0.9	10	20	30	40	50	60	70	80	90	100
1	10	20	30	40	50	60	70	80	90	100

จากตารางที่ 4.4 ข้อมูลในแต่ละแถวแสดงจำนวนบ็อตเน็ตที่ถูกตรวจจับได้เมื่อค่าความน่าจะเป็น ในการส่งข้อมูลในแต่ละช่วงเวลานั้นแตกต่างกันไป ถ้าค่าความน่าจะเป็นสูง หมายถึง เครื่องบ็อตเน็ตแต่ละเครื่องส่งจดอิเล็กทรอนิกส์สแปมออกมามาก แต่ถ้าค่าความน่าจะเป็นต่ำ หมายถึง บอตเน็ตแต่ละเครื่องส่งจดหมายอิเล็กทรอนิกส์สแปมออกมาน้อย สำหรับในแต่ละคอลัมภ์เป็นการแสดงเปอร์เซ็นต์ของบ็อตเน็ตจากจำนวนเครื่องคอมพิวเตอร์ทั้งหมด จากตารางแสดงให้เห็นว่าทุกๆ ความน่าจะเป็น ตั้งแต่ 0.1 - 1 ระบบสามารถระบุได้ว่าเป็นจดหมายอิเล็กทรอนิกส์สแปมถูกส่งออกมาจากเครื่องใดบ้างที่เป็นบอตเน็ต ซึ่งค่าเริ่มต้นของบอตเน็ตกำหนดไว้ตามเปอร์เซ็นต์บอตเน็ต จากข้อมูลที่แสดงระบบสามารถตรวจพบครบทุกเครื่องเมื่อกำหนดค่า Threshold = ¼ Threshold

ภาพที่ 4.5 แสดงจำนวนบอตเน็ตที่ถูกตรวจพบเมื่อกำหนด Threshold = ½



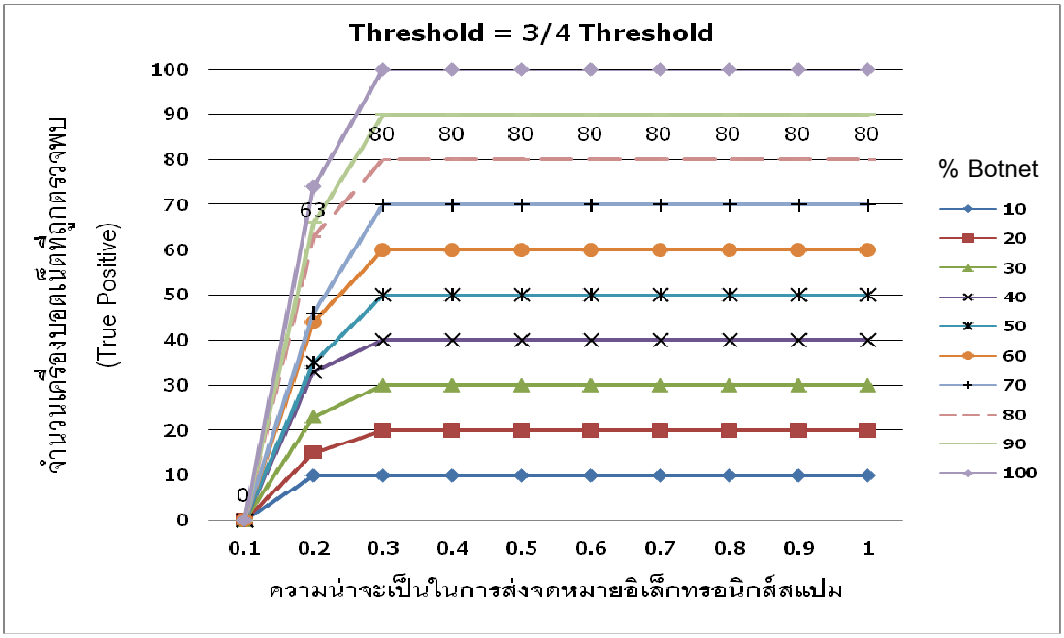
ภาพที่ 4.5 แสดงให้เห็นถึงจำนวนเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปมถูกตรวจพบ เมื่อกำหนดค่า Threshold เป็น ½ Threshold จะพบว่าระบบสามารถตรวจพบเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปมได้ครบทั้งหมด เมื่อมีค่าความน่าจะเป็นของการส่งจดหมายอิเล็กทรอนิกส์สแปมตั้งแต่ 0.2 ฉบับต่อชั่วโมง โดยระบบจะเริ่มตรวจสอบเครื่องที่เป็นบอตเน็ตไม่ได้ทั้งหมดเมื่อมีความน่าจะเป็นของการส่งจดหมายอิเล็กทรอนิกส์สแปมที่ 0.1 แสดงข้อมูลตามตารางที่ 4.5

ตารางที่ 4.5 แสดงข้อมูลจำนวนบอตเน็ตที่ตรวจพบ ณ Threshold = ½ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	0	0	1	1	0	2	2	2	4	3
0.2	10	20	30	40	50	60	70	80	90	100
0.3	10	20	30	40	50	60	70	80	90	100
0.4	10	20	30	40	50	60	70	80	90	100
0.5	10	20	30	40	50	60	70	80	90	100
0.6	10	20	30	40	50	60	70	80	90	100
0.7	10	20	30	40	50	60	70	80	90	100
0.8	10	20	30	40	50	60	70	80	90	100
0.9	10	20	30	40	50	60	70	80	90	100
1	10	20	30	40	50	60	70	80	90	100

จากตารางที่ 4.5 แสดงให้เห็นว่า Probability ที่ 0.1 มีบางเครื่องที่เป็นบอตเน็ตส่งสแปมออกมาและสามารถถูกตรวจพบ ซึ่งค่าเริ่มต้นตั้งไว้เครื่องที่ 1 ถึง10 เป็นบอตเน็ตที่ทำการส่งจดหมายอิเล็กทรอนิกส์สแปม ระบบจะเริ่มตรวจพบจดหมายอิเล็กทรอนิกส์สแปมได้เมื่อกำหนดมีค่า Probability ที่ 0.2 เช่นเดียวกับ ¼ Threshold

ภาพที่ 4.6 แสดงจำนวนบอตเน็ตที่ถูกตรวจพบเมื่อกำหนด Threshold = ¾ Threshold



ภาพที่ 4.6 แสดงให้เห็นถึงจำนวนเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปมถูกตรวจพบ เมื่อกำหนดค่า Threshold เป็น ¾ Threshold จะพบว่าระบบสามารถตรวจพบจดหมายอิเล็กทรอนิกส์จากเครื่องบอกเน็ตได้ครบทั้งหมด เมื่อมีค่าความน่าจะเป็นที่จะส่งจดหมายอิเล็กทรอนิกส์สแปมตั้งแต่ 0.3 ฉบับต่อชั่วโมงขึ้นไป แนวโน้มระบบจะตรวจพบจดหมายอิเล็กทรอนิกส์สแปมได้มากขึ้นเมื่อกำหนดค่า Threshold ที่สูงขึ้น และจะตรวจสอบได้ถูกต้องมากขึ้นเมื่อมีความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปมสูงๆ แสดงข้อมูลตามตารางที่ 4.6

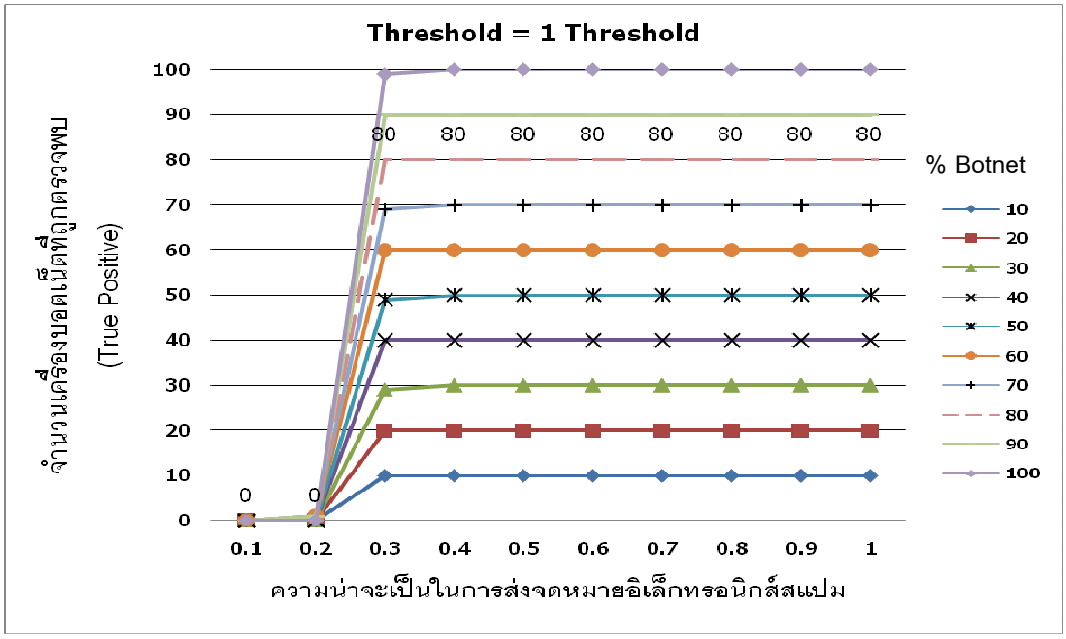
ตารางที่ 4.6 แสดงข้อมูลจำนวนบอตเน็ตที่ตรวจพบ ณ Threshold = ¾ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	0	0	0	0	0	0	0	0	0	0
0.2	10	15	23	33	35	44	46	63	66	74
0.3	10	20	30	40	50	60	70	80	90	100
0.4	10	20	30	40	50	60	70	80	90	100
0.5	10	20	30	40	50	60	70	80	90	100
0.6	10	20	30	40	50	60	70	80	90	100

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.7	10	20	30	40	50	60	70	80	90	100
0.8	10	20	30	40	50	60	70	80	90	100
0.9	10	20	30	40	50	60	70	80	90	100
1	10	20	30	40	50	60	70	80	90	100

จากตารางที่ 4.6 แสดงให้เห็นว่า Probability ที่ 0.1 ไม่มีเครื่องไหนที่ถูกระบุว่าเป็นบอตเน็ตส่งจดหมายอิเล็กทรอนิกส์สแปม ระบบจะเริ่มตรวจสอบได้ว่าเป็นจดหมายอิเล็กทรอนิกส์สแปมได้ เมื่อความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปมสูงขึ้น จากตารางที่ 4.6 จะพบว่าค่าความน่าจะเป็นอยู่ที่ 0.3 จะตรวจพบเครื่องที่ส่งจดหมายอิเล็กทรอนิกส์สแปมทั้งหมด ที่ Threshold = ¾ Threshold

ภาพที่ 4.7 แสดงจำนวนบอตเน็ตที่ถูกตรวจสอบพบ เมื่อกำหนด Threshold = 1 Threshold



ภาพที่ 4.7 แสดงให้เห็นถึงจำนวนเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปมที่ถูกตรวจพบ เมื่อกำหนดค่า Threshold เป็น 1 Threshold จะเห็นว่าระบบสามารถตรวจสอบได้ว่าเป็นจดหมายอิเล็กทรอนิกส์สแปมได้ครบ เมื่อมีค่าความน่าจะเป็นที่จะส่งจดหมายอิเล็กทรอนิกส์

สแปมตั้งแต่ 0.3 ฉบับต่อชั่วโมงขึ้นไปเช่นเดียวกับการกำหนดค่า Threshold = ¾ Threshold แสดงข้อมูลตามตารางที่ 4.7

ตารางที่ 4.7 แสดงข้อมูลจำนวนบอตเน็ตที่ตรวจพบ ณ Threshold = 1 Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	0	0	0	0	0	0	0	0	0	0
0.2	0	0	0	0	0	1	1	0	1	0
0.3	10	20	29	40	49	60	69	80	90	99
0.4	10	20	30	40	50	60	70	80	90	100
0.5	10	20	30	40	50	60	70	80	90	100
0.6	10	20	30	40	50	60	70	80	90	100
0.7	10	20	30	40	50	60	70	80	90	100
0.8	10	20	30	40	50	60	70	80	90	100
0.9	10	20	30	40	50	60	70	80	90	100
1	10	20	30	40	50	60	70	80	90	100

จากตารางที่ 4.7 แสดงให้เห็นว่าระบบสามารถตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมได้ถูกต้องมากขึ้นเมื่อความน่าจะเป็นมากขึ้นคือ ความน่าจะเป็นที่ 0.4 และระบบไม่มีสามารถตรวจสอบเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์ออกมาได้เลยเมื่อมีความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปมที่ 0.1 เมื่อความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปมน้อยๆ และมีการกำหนดค่า Threshold ที่สูงขึ้น จะทำให้ระบบตรวจสอบจดหมายอิเล็กทรอนิกส์ได้น้อยลง

สรุปจากข้อมูลภาพที่ 4.3-4.6 แสดงให้เห็นว่า เมื่อกำหนดค่า Threshold ต่ำๆ จะทำให้ระบบสามารถระบุได้ว่าเป็นจดหมายอิเล็กทรอนิกส์สแปมได้มากขึ้น แม้ว่าจะมีความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปมต่ำ ระบบก็สามารถตรวจสอบได้เช่นกัน

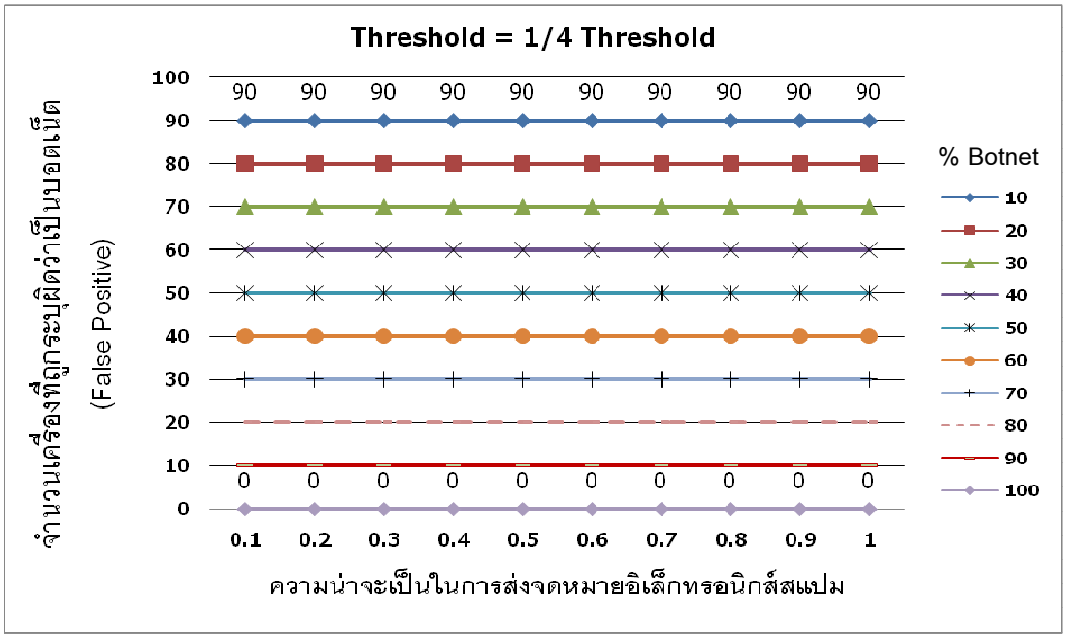
เมื่อพิจารณาจากข้อมูลตารางที่ 4.3 ข้อมูลตัวอย่าง การส่งจดหมายอิเล็กทรอนิกส์ของเครื่อง 100 เครื่องโดยมีเครื่องบอตเน็ต 10 เปอร์เซ็นต์ ความน่าจะเป็น 0.2 และ Threshold = ½ Threshold พบว่าระบบตรวจสอบได้ว่า เป็นจดหมายอิเล็กทรอนิกส์สแปมจากเครื่องบอตเน็ต 10 เครื่องตามที่กำหนดไว้ แต่ระบบมีการตรวจระบุผิดโดยระบุจดหมายอิเล็กทรอนิกส์ปรกติที่ถูกส่งออกมาจากเครื่องที่ไม่เป็นบอตเน็ต (เครื่องปรกติ) เป็นจดหมายอิเล็กทรอนิกส์สแปม ซึ่งถือว่า

เป็นข้อผิดพลาดของระบบเพราะเกิดค่า False Positive มาก (False Positive = 48 เครื่อง) ดังนั้น การระบุว่าค่า Threshold น้อยจะดีในการระบุว่าเป็นจดหมายอิเล็กทรอนิกส์สแปมได้มากขึ้น แต่ไม่ได้หมายความว่าจะระบุได้ถูกต้องว่าเป็นจดหมายอิเล็กทรอนิกส์สแปม จริงๆ

4.2.5 ผลการวิเคราะห์ค่า False Positive ของการกำหนดค่า Threshold

ผลการวิเคราะห์ดังต่อไปนี้แสดงให้เห็นถึงระบบที่เกิดความผิดพลาดในตรวจสอบจดหมายอิเล็กทรอนิกส์สแปม โดยระบบระบุผิดว่าเครื่องปรกติเป็นเครื่องบอตเน็ตที่มีการส่งจดหมายอิเล็กทรอนิกส์สแปม เมื่อมีการกำหนดค่า Threhold ที่แตกต่างกัน

ภาพที่ 4.8 แสดงจำนวนค่า False Positive ของการกำหนดค่า Threshold = ¼ Threshold



จากภาพที่ 4.8 แสดงจำนวน False Positive (ระบบตรวจสอบผิดเครื่อง) เมื่อกำหนดค่า Threshold ของระบบ เป็น ¼ Threshold แล้วจะเกิดกรณีที่ระบบระบุว่าเครื่องธรรมดาเป็นเครื่องบอตเน็ตที่กำลังส่งจดหมายอิเล็กทรอนิกส์สแปม จากภาพเปอร์เซ็นต์บอตเน็ตที่ 100 ระบบไม่ระบุผิด เพราะว่าทั้ง 100 เครื่องเป็นบอตเน็ตและส่งจดหมายอิเล็กทรอนิกส์สแปมทั้งหมด แต่เปอร์เซ็นต์บอตเน็ตที่ต่ำกว่า 90 ลงไป จะเกิดค่า False Positive ทุกเปอร์เซ็นต์บอตเน็ต ดังนั้นค่า Threshold = ¼ Threshold นี้ไม่ใช่ค่า Threshold ที่ดี ถึงแม้จะสามารถระบุถูกว่าเป็นจดหมาย

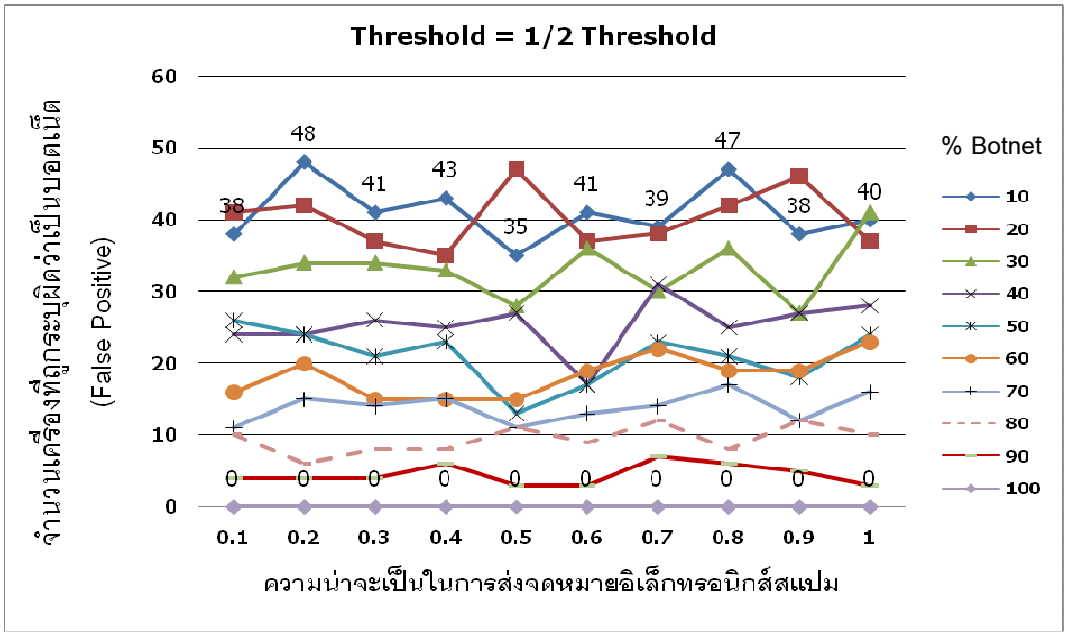
อิเล็กทรอนิกส์สแปม ค่า True Positive สูง แต่ก็มีค่า False Positive ที่สูงมากเช่นกัน แสดงข้อมูลตามตารางที่ 4.8

ตารางที่ 4.8 แสดงข้อมูล False Positive ของการกำหนดค่า Threshold = ¼ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	90	80	70	60	50	40	30	20	10	0
0.2	90	80	70	60	50	40	30	20	10	0
0.3	90	80	70	60	50	40	30	20	10	0
0.4	90	80	70	60	50	40	30	20	10	0
0.5	90	80	70	60	50	40	30	20	10	0
0.6	90	80	70	60	50	40	30	20	10	0
0.7	90	80	70	60	50	40	30	20	10	0
0.8	90	80	70	60	50	40	30	20	10	0
0.9	90	80	70	60	50	40	30	20	10	0
1	90	80	70	60	50	40	30	20	10	0

จากตารางที่ 4.8 แสดงให้เห็นว่าระบบมีการตรวจสอบที่เกิดความผิดพลาดขึ้นเมื่อกำหนดค่า Threshold = ¼ Threshold จากข้อมูลนี้แสดงว่าจำนวนของเครื่องปรกติที่ส่งจดหมายอิเล็กทรอนิกส์ปรกติ ถูกระบุว่าเป็นเครื่องบอตเน็ตส่งจดหมายอิเล็กทรอนิกส์สแปม

ภาพที่ 4.9 แสดงจำนวนค่า False Positive ของการกำหนดค่า Threshold = ½ Threshold



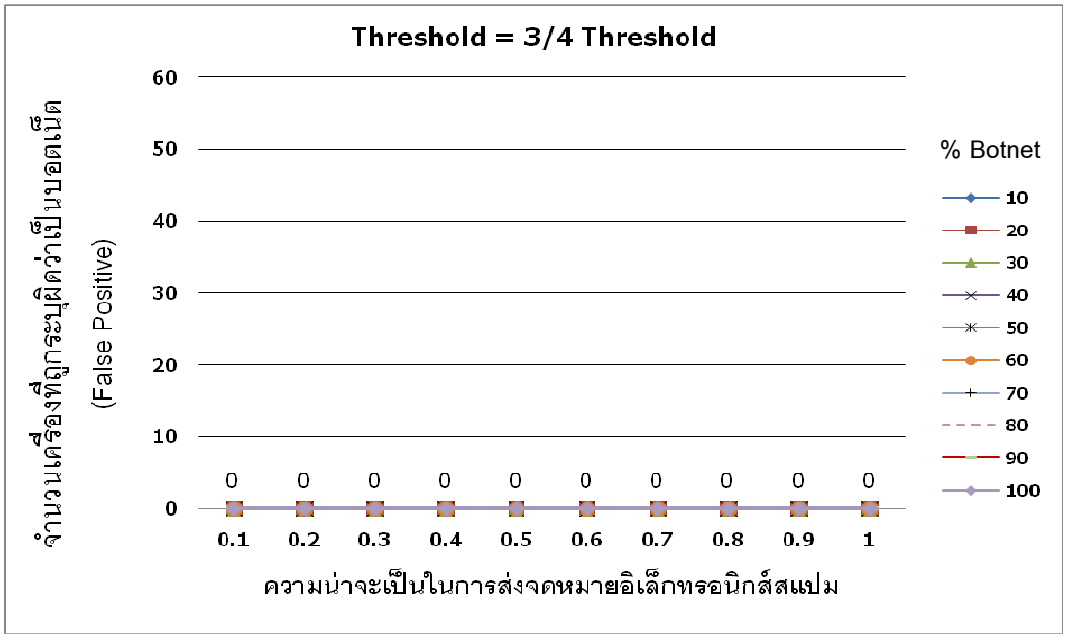
จากภาพที่ 4.9 แสดงจำนวน False Positive เมื่อกำหนดค่า Threshold ของระบบเป็น ½ Threshold พบว่าจดหมายอิเล็กทรอนิกส์ปรกติที่ถูกส่งออกมาจากเครื่องปรกติกถูกระบบระบุว่าเป็นจดหมายอิเล็กทรอนิกส์สแปมน้อยลงแต่ก็ยังปรากฏว่ามีค่า False Positive ทุกๆกรณี

ตารางที่ 4.9 แสดงข้อมูล False Positive ของการกำหนดค่า Threshold = ½ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	38	41	32	24	26	16	11	10	4	0
0.2	48	42	34	24	24	20	15	6	4	0
0.3	41	37	34	26	21	15	14	8	4	0
0.4	43	35	33	25	23	15	15	8	6	0
0.5	35	47	28	27	13	15	11	11	3	0
0.6	41	37	36	17	17	19	13	9	3	0
0.7	39	38	30	31	23	22	14	12	7	0
0.8	47	42	36	25	21	19	17	8	6	0
0.9	38	46	27	27	18	19	12	12	5	0
1	40	37	41	28	24	23	16	10	3	0

จากตารางที่ 4.9 แสดงให้เห็นว่าระบบมีการตรวจสอบที่เกิดความผิดพลาดขึ้นเมื่อกำหนดค่า Threshold = ½ Threshold และมีการตรวจสอบผิดพลาดน้อยลงเมื่อค่า Threshold สูงขึ้นแต่ยังมีค่าที่ผิดพลาดสูง

ภาพที่ 4.10 แสดงจำนวนค่า False Positive ของการกำหนดค่า Threshold = ¾ Threshold



จากภาพที่ 4.10 แสดงจำนวนเครื่องที่ระบบระบุไม่ผิดว่าเป็นเครื่องบอตเน็ตส่งจดหมายอิเล็กทรอนิกส์สแปม ระบบตรวจสอบถูกต้องสูงมาก ดังแสดงข้อมูลตามตารางที่ 4.10

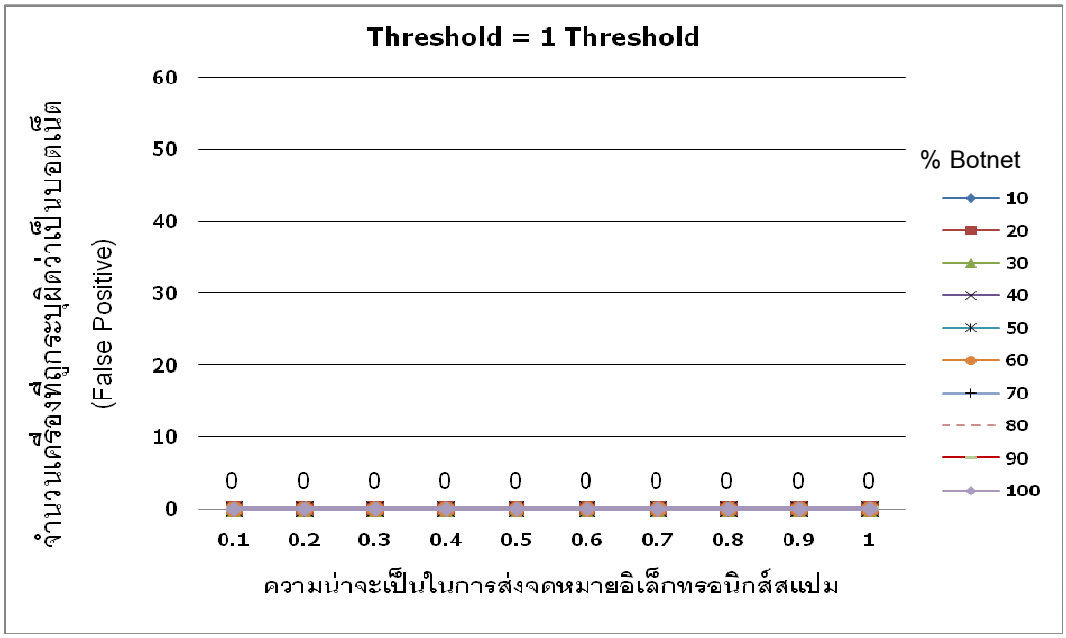
ตารางที่ 4.10 แสดงข้อมูล False Positive ของการกำหนดค่า Threshold = ¾ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	0	0	0	0	0	0	0	0	0	0
0.2	0	0	0	0	0	0	0	0	0	0
0.3	0	0	0	0	0	0	0	0	0	0
0.4	0	0	0	0	0	0	0	0	0	0
0.5	0	0	0	0	0	0	0	0	0	0
0.6	0	0	0	0	0	0	0	0	0	0
0.7	0	0	0	0	0	0	0	0	0	0

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.8	0	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

จากตารางที่ 4.10 ข้อมูลการระบุผิดว่าเครื่องปรกติที่ส่งจดหมายอิเล็กทรอนิกส์ปรกติเป็นเครื่องบอตเน็ตมีค่าเป็น 0 ทั้งหมดแสดงว่าไม่มีค่า False Positive เกิดขึ้นเลยเมื่อกำหนดค่า Threshold = ¾ Threshold

ภาพที่ 4.11 แสดงจำนวนค่า False Positive ของการกำหนดค่า Threshold = 1 Threshold



จากภาพที่ 4.11 แสดงจำนวนเครื่องที่ระบบระบุไม่ผิดว่าเป็นเครื่องบอตเน็ตส่งจดหมายอิเล็กทรอนิกส์สแปม ซึ่งแสดงว่าระบบมีค่าความถูกต้องในการตรวจสอบสูงมาก เช่นเดียวกับการกำหนดค่า Threshold = ¾ Threshold แสดงข้อมูลตามตารางที่ 4.11

ตารางที่ 4.11 แสดงข้อมูล False Positive ของการกำหนดค่า Threshold = 1 Threshold

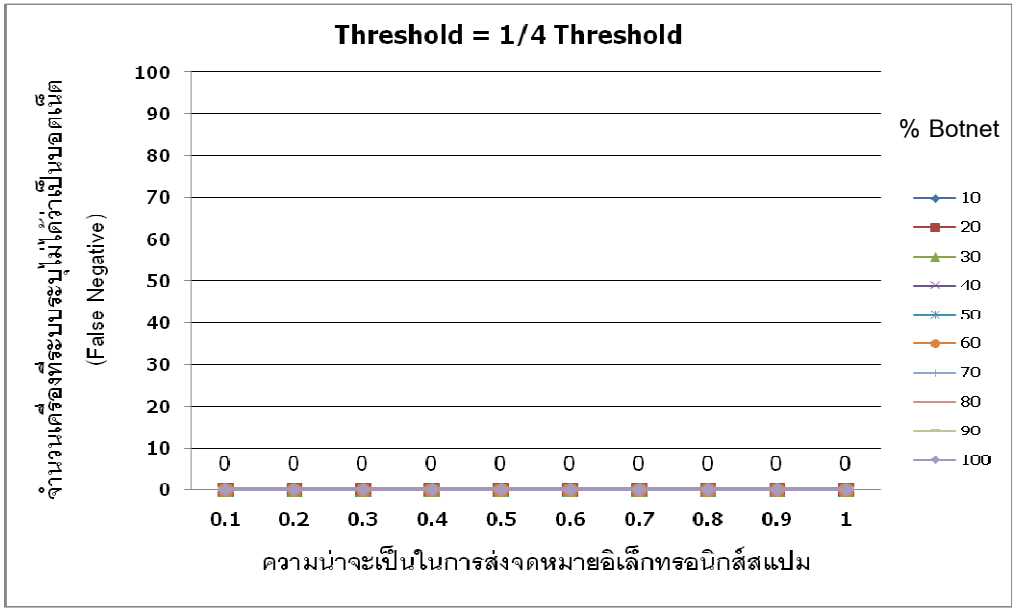
Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	0	0	0	0	0	0	0	0	0	0
0.2	0	0	0	0	0	0	0	0	0	0
0.3	0	0	0	0	0	0	0	0	0	0
0.4	0	0	0	0	0	0	0	0	0	0
0.5	0	0	0	0	0	0	0	0	0	0
0.6	0	0	0	0	0	0	0	0	0	0
0.7	0	0	0	0	0	0	0	0	0	0
0.8	0	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

จากตารางที่ 4.11 ข้อมูลการระบุผิดว่าเครื่องปรกติที่ส่งจดหมายอิเล็กทรอนิกส์ปรกติเป็นเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปม มีค่าเป็น 0 ทั้งหมดแสดงว่าไม่มีค่า False Positive เกิดขึ้นเลยเมื่อกำหนดค่า Threshold = 1 Threshold

4.2.6 **ผลการวิเคราะห์ค่า** False Negative **ของการกำหนดค่า** Threshold

ผลการวิเคราะห์ดังต่อไปนี้แสดงให้เห็นถึงระบบที่เกิดความผิดพลาดในตรวจสอบจดหมายอิเล็กทรอนิกส์สแปม โดยในระบบเครือข่ายมีเครื่องคอมพิวเตอร์ที่เป็นบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปม แต่ระบบ DSDSE ระบุผิดว่าจดหมายดังกล่าวไม่เป็นจดหมายอิเล็กทรอนิกส์สแปม พิจารณาโดยการกำหนดค่า Threhold ที่แตกต่างกัน

ภาพที่ 4.12 แสดงจำนวนค่า False Negative ของการกำหนดค่า Threshold = ¼ Threshold



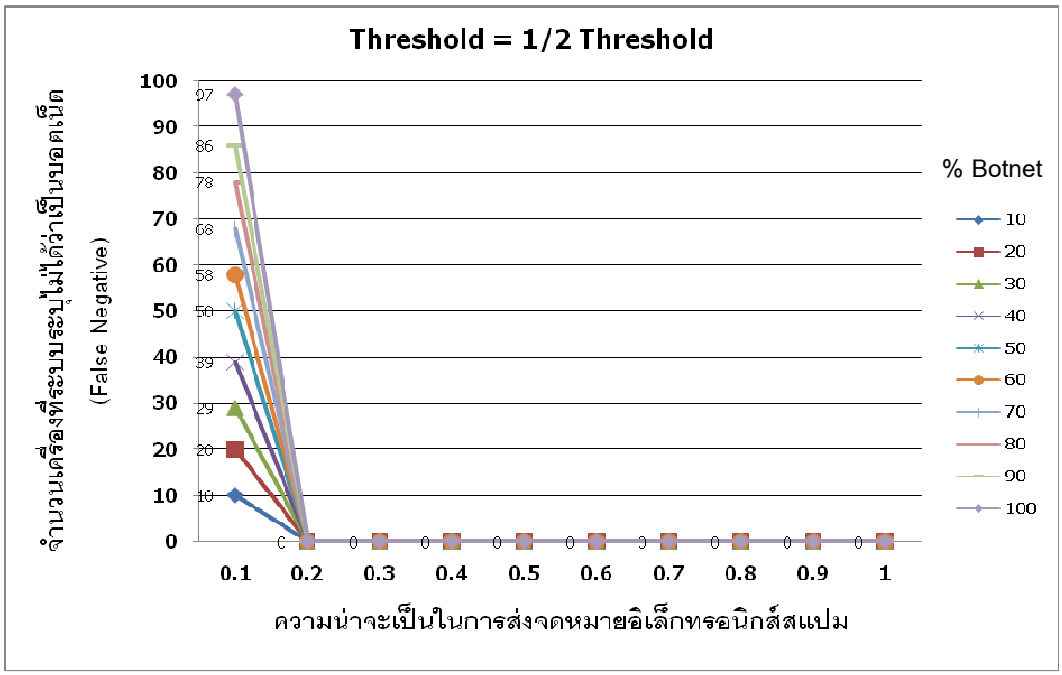
จากภาพที่ 4.1.2 แสดงค่า False Negative ของ Threshold = 1/4 Threshold แสดงให้เห็นว่าทุกความน่าจะเป็น ระบบสามารถระบุได้ว่ามีจดหมายอิเล็กทรอนิกส์ถูกส่งออกมาจากเครื่องบอตเน็ต

ตารางที่ 4.12 แสดงข้อมูล False Negative ของการกำหนดค่า Threshold = ¼ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	0	0	0	0	0	0	0	0	0	0
0.2	0	0	0	0	0	0	0	0	0	0
0.3	0	0	0	0	0	0	0	0	0	0
0.4	0	0	0	0	0	0	0	0	0	0
0.5	0	0	0	0	0	0	0	0	0	0
0.6	0	0	0	0	0	0	0	0	0	0
0.7	0	0	0	0	0	0	0	0	0	0
0.8	0	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

จากตารางที่ 4.12 ระบบสามารถระบุได้ไม่ผิดว่าเป็นบอตเน็ตส่งจดหมายอิเล็กทรอนิกส์สแปม แม้ว่าจะมีความน่าจะเป็นในการส่งน้อยก็ตาม

ภาพที่ 4.13 แสดงจำนวนค่า False Negative ของการกำหนดค่า Threshold = ½ Threshold



ภาพที่ 4.13 แสดงให้เห็นจำนวนเครื่องคอมพิวเตอร์ที่เป็นบอตเน็ตไม่สามารถตรวจสอบได้ว่าเป็นเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปม เมื่อมีความน่าจะเป็น ที่ 0.1

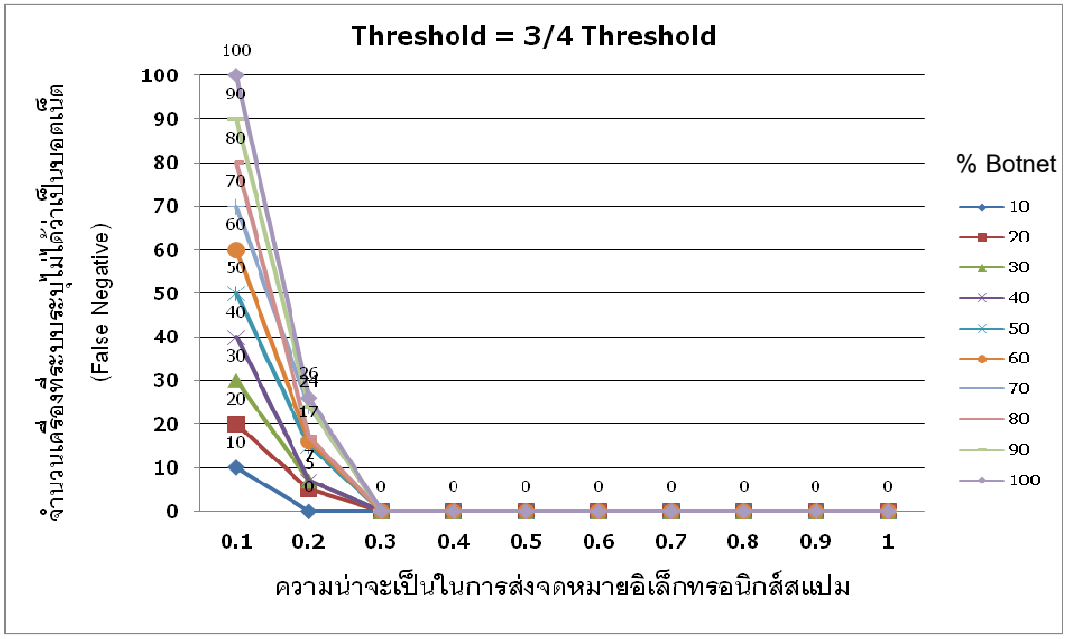
ตารางที่ 4.13 แสดงข้อมูล False Negative ของการกำหนดค่า Threshold = ½ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	10	20	29	39	50	58	68	78	86	97
0.2	0	0	0	0	0	0	0	0	0	0
0.3	0	0	0	0	0	0	0	0	0	0
0.4	0	0	0	0	0	0	0	0	0	0
0.5	0	0	0	0	0	0	0	0	0	0
0.6	0	0	0	0	0	0	0	0	0	0
0.7	0	0	0	0	0	0	0	0	0	0

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.8	0	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

จากตารางที่ 4.13 ที่ความน่าจะเป็นที่ 0.1 มีจดหมายอิเล็กทรอนิกส์จากเครื่องบอตเน็ตส่วนมาก ระบบ ไม่สามารถถูกระบุได้ว่าเป็นบอตเน็ต

ภาพที่ 4.14 แสดงจำนวนค่า False Negative ของการกำหนดค่า Threshold = ¾ Threshold



จากภาพที่ 4.14 แสดงให้เห็นจำนวนเครื่องบอตเน็ตที่ส่งจดหมายอิเล็กทรอนิกส์สแปมไม่สามารถถูกตรวจสอบ ในช่วงความน่าจะเป็น 0.1 และ 0.2

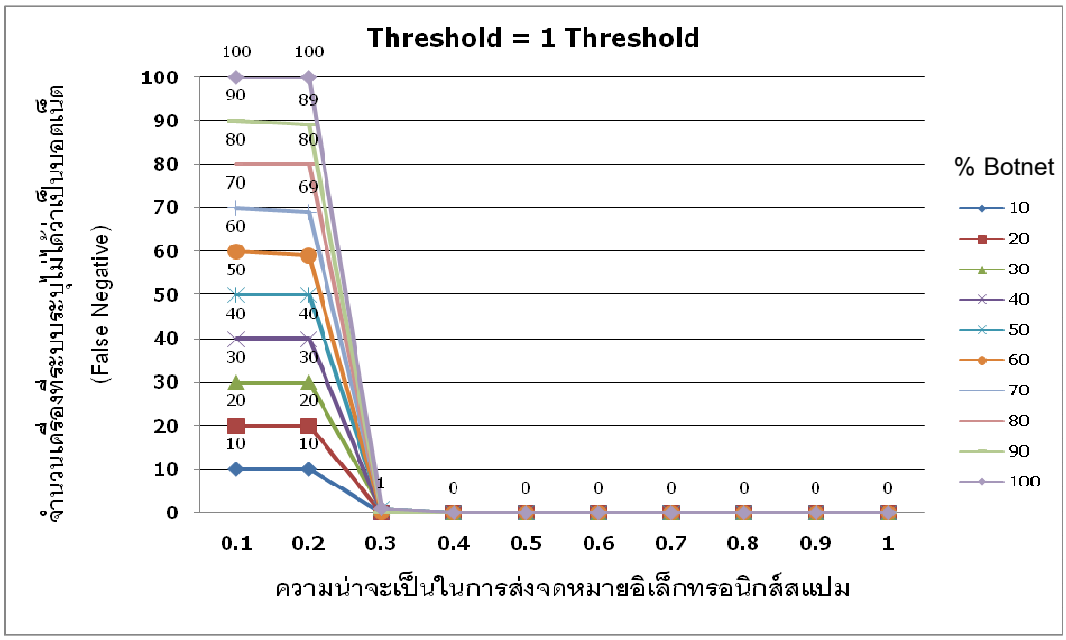
ตารางที่ 4.14 แสดงข้อมูล False Negative ของการกำหนดค่า Threshold = ¾ Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	10	20	30	40	50	60	70	80	90	100
0.2	0	5	7	7	15	16	24	17	24	26
0.3	0	0	0	0	0	0	0	0	0	0

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.4	0	0	0	0	0	0	0	0	0	0
0.5	0	0	0	0	0	0	0	0	0	0
0.6	0	0	0	0	0	0	0	0	0	0
0.7	0	0	0	0	0	0	0	0	0	0
0.8	0	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

จากตารางที่ 4.14 จะแสดงให้เห็นว่าเมื่อค่า Threshold ที่เพิ่มขึ้นระบบไม่สามารถตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมจากเครื่องที่เป็นบอตเน็ตได้ โดยเฉพาะถ้ามีความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปมน้อยๆ

ภาพที่ 4.15 แสดงจำนวนค่า False Negative ของการกำหนดค่า Threshold = 1 Threshold



ภาพที่ 4.15 แสดงจำนวนเครื่องคอมพิวเตอร์บอตเน็ตที่ระบบไม่สามารถระบุได้ว่าส่งจดหมายอิเล็กทรอนิกส์สแปม แสดงให้เห็นว่าค่า Threshold สูง คือ Threshold = 1 Threshold ระบบจะทำงานมีประสิทธิภาพดีที่สุดเมื่อมีความน่าจะเป็นในการส่งจดหมายอิเล็กทรอนิกส์สแปม

ที่ 0.4 ขึ้นไป หมายถึง ถ้าค่า Threshold สูงๆ ระบบจะไม่สามารถตรวจจดหมายอิเล็กทรอนิกส์สแปมจำนวนน้อยๆได้

ตารางที่ 4.15 แสดงข้อมูล False Negative ของการกำหนดค่า Threshold = 1 Threshold

Probability/Botnet (%)	10	20	30	40	50	60	70	80	90	100
0.1	10	20	30	40	50	60	70	80	90	100
0.2	10	20	30	40	50	59	69	80	89	100
0.3	0	0	1	0	1	0	1	0	0	1
0.4	0	0	0	0	0	0	0	0	0	0
0.5	0	0	0	0	0	0	0	0	0	0
0.6	0	0	0	0	0	0	0	0	0	0
0.7	0	0	0	0	0	0	0	0	0	0
0.8	0	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

จากตารางที่ 4.15 แสดงให้เห็นว่าจำนวนเครื่องที่เป็นบอตเน็ตส่งจดหมายอิเล็กทรอนิกส์สแปม แต่ระบบระบุว่าไม่เป็นจดหมายอิเล็กทรอนิกส์สแปมมีจำนวนมาก เมื่อมีการใช้งานค่า Threshold สูงๆ และโอกาสในการส่งจดหมายอิเล็กทรอนิกส์สแปมจำนวนน้อยๆ

ตารางที่ 4.16 สรุปข้อเปรียบเทียบของการกำหนดค่า Threshold

ค่า Threshold	ข้อดี	ข้อเสีย
Threshold = ¼ Threshold	ตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมได้ครบแม้ว่าจะมีปริมาณน้อย	เกิดค่า False Positive สูงที่สุด
Threshold = ½ Threshold	ตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมได้ครบเมื่อมีปริมาณน้อยที่ความน่าจะเป็น 0.2 ขึ้นไป	ยังเกิดค่า False Postive สูง
Threshold = ¾ Threshold	ไม่เกิดค่า False Positive	ตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมได้ครบที่ความน่าจะเป็น 0.3 ขึ้นไป
Threshold = 1 Threshold	ไม่เกิดค่า False Positive	ตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมได้ครบเมื่อที่ความน่าจะเป็น 0.4 ขึ้นไป

จากตารางที่ 4.16 แสดงผลการวิเคราะห์ข้อมูลการส่งจดหมายอิเล็กทรอนิกส์สแปมโดยอาศัยค่า Threshold = ¼ Threshold ½ Threshold ¾ Threshold และ 1 Threshold สามารถสรุปรายละเอียดได้ดังนี้

1) ค่า Threshold = ¼ Threshold ข้อดีคือ ไม่เกิดค่า False Negative ข้อเสียคือ ระบบเกิดค่า False Positive มาก

2) ค่า Threshold = ½ Threshold ข้อดีคือ เกิดค่า False Positive น้อยกว่า Threshold = ¼ Threshold แต่ก็ยังมีอยู่ ถึงแม้ระบบจะสามารถตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมได้ครบก็ตาม

3) ค่า Threshold = ¾ Threshold ข้อดีคือ ไม่เกิดค่า False Positive และสามารถตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมจากเครื่องบอตเน็ตได้ครบ ซึ่งโอกาสในการส่งจดหมายอิเล็กทรอนิกส์สแปมที่ตรวจครบจะต้องมีมากพอสมควร ควรจะมีความน่าจะเป็น 0.3 ขึ้นไป

4) ค่า Threshold = 1 Threshold ข้อดีคือ ไม่เกิดค่า False Positive แต่การที่จะตรวจสอบจดหมายอิเล็กทรอนิกส์สแปมจากเครื่องบอตเน็ตได้ครบ จะต้องมีปริมาณจดหมายอิเล็กทรอนิกส์สแปมที่มากพอสมควร ถึงจะตรวจสอบพบ จากการทดลองพบที่ความน่าจะเป็น 0.4 ขึ้นไป

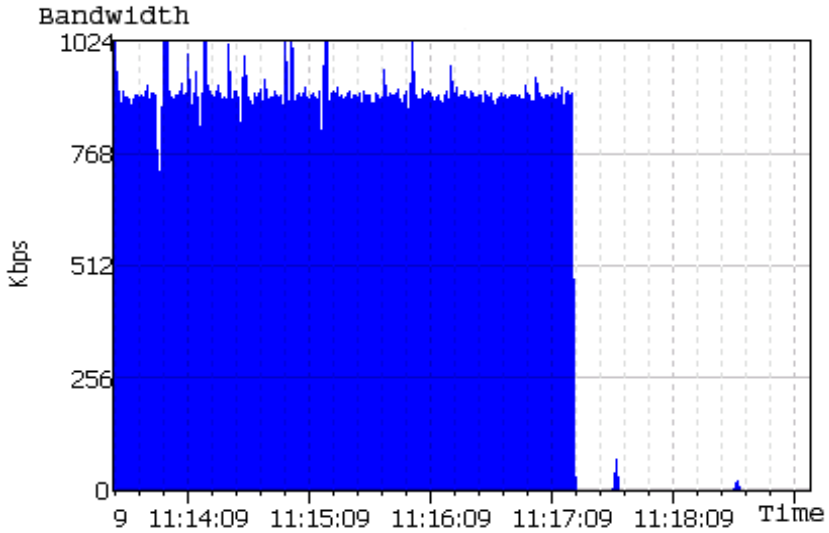
ปัจจัยที่สำคัญ ที่ทำให้ระบบ DSDSE สามารถพิจารณาว่าจดหมายอิเล็กทรอนิกส์ที่ถูกส่งออกมาเป็นจดหมายอิเล็กทรอนิกส์ปรกติ หรือเป็นจดหมายอิเล็กทรอนิกส์สแปม ก็คือ ค่า Threshold ที่เหมาะสม จากผลการทดลองในบทนี้ถ้าพิจารณากรณีของ ค่าความถูกต้อง (False Positive) โดยค่า Threshold ที่เหมาะสมและไม่ทำให้ระบบเกิดค่า False Positive ได้แก่ ค่า Threshold ที่ ¾ Threshold และ 1 Threshold

สามารถดูรายละเอียดเพิ่มเติมของค่าความถูกต้อง True Negative ได้ที่ภาคผนวก ข

4.3 การลดอัตราการส่งจดหมายอิเล็กทรอนิกส์สแปม (Traffic Control)

เมื่อระบบ DSDSE สามารถระบุได้ว่าเครื่องคอมพิวเตอร์ใดเป็นเครื่องที่คาดว่าเป็นบอตเน็ต และทำการส่งจดหมายอิเล็กทรอนิกส์มากผิดปรกติ ระบบจะทำการลดอัตราการส่งจดหมายอิเล็กทรอนิกส์โดยวิธีการลดขนาด Bandwidth บอตเน็ตจะทำการส่งจดหมายอิเล็กทรอนิกส์จำนวน 2,000 ฉบับ และทำการส่งพร้อมกันรวมเป็นบอตเน็ต 10 เครื่อง รวมจดหมายทั้งสิ้น 20,000 ฉบับ ใช้เวลาโดยประมาณ 4 นาทีในการส่งทั้งหมด

ภาพที่ 4.16 แสดงปริมาณ Bandwidth ที่ถูกใช้ไปในขณะที่บอตเน็ตส่งจดหมายอิเล็กทรอนิกส์



จากภาพที่ 4.16 เป็นการทดลองส่งจดหมายอิเล็กทรอนิกส์จำนวน 20,000 ฉบับจากเครื่องบอตเน็ตจำนวน 10 เครื่อง ส่งเครื่องละ 2,000 ฉบับ จากกราฟที่ได้จะพบว่าบอตเน็ตเริ่มส่ง

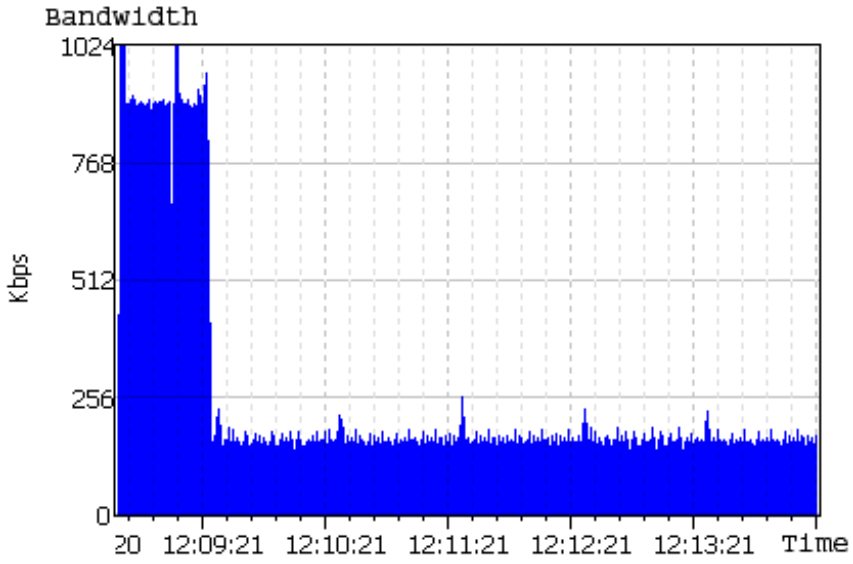
จดหมายอิเล็กทรอนิกส์เมื่อเวลา 11.13.33 น. เสร็จสิ้นที่เวลา 11.17.21 น. รวมเวลาที่ใช้ส่งทั้งสิ้น 3 นาที 48 วินาที จากภาพกราฟแสดงให้เห็นถึงการแพร่กระจายของจดหมายอิเล็กทรอนิกส์ปริมาณมากในช่วงเวลาสั้น ๆ

ระบบ DSDSE จะทำการลดขนาด Bandwidth โดยผู้วิจัยนำเสนอรูปแบบ 3 รูปแบบ ดังนี้

4.3.1 การลดอัตราการส่งจดหมายอิเล็กทรอนิกส์สแปมโดยวิธีกำหนดค่าเริ่มต้น

กำหนดปริมาณ Bandwidth แบบคงที่มีหน่วยเป็น Kbps. โดยปริมาณ Bandwidth สามารถกำหนดโดยผู้ดูแลระบบ

ภาพที่ 4.17 แสดงปริมาณ Bandwidth ที่ถูกควบคุมในการส่งจดหมายอิเล็กทรอนิกส์ แบบที่ 1 โดยการกำหนดค่าเริ่มต้น



จากภาพที่ 4.17 โดยการกำหนดค่าปริมาณการส่งอิเล็กทรอนิกส์ ประมาณ 100 KB ซึ่งใช้เวลาในการส่งทั้งหมดอยู่ที่ประมาณ 25 นาที

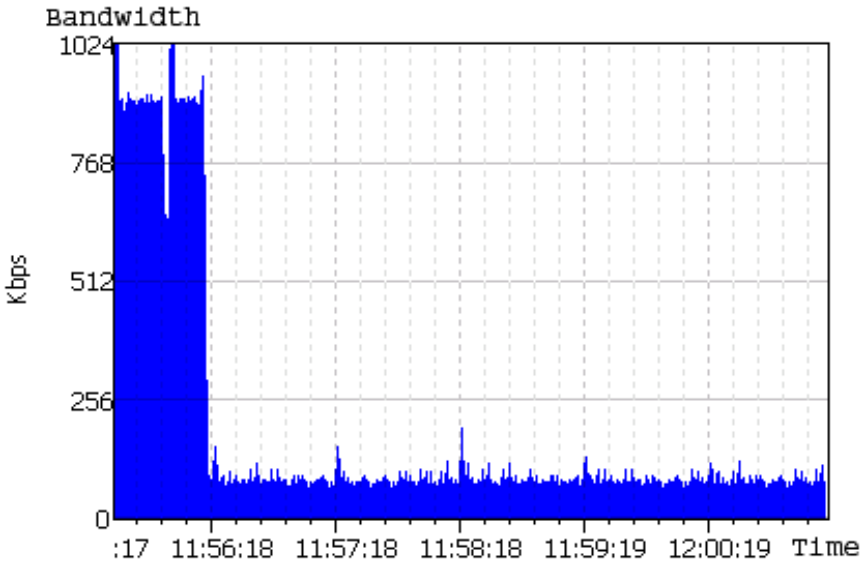
หมายเหตุ ยอดกราฟ แสดงถึงบอตเน็ตกำลังเข้าไปใน IRC Server ซึ่งระบบ DSDSE จะทำการควบคุม Bandwidth เฉพาะจดหมายอิเล็กทรอนิกส์ ไม่ได้ทำการควบคุมข้อมูลอื่นๆด้วย โดยยอดกราฟแหลมๆ บ่งบอกว่าบอตเน็ตจะเข้าไปที่ IRC Server เพื่อเข้าไปรอรับคำสั่งครั้งละ

3 นาที หลังจากนั้นจะหยุดการติดต่อ และจะทำการติดต่อใหม่ทันที โดยในการเชื่อมต่อแต่ละครั้งจะเปลี่ยนชื่อ login ใหม่ทุกครั้ง ในกรณีที่ไม่พบเจอ IRC server บอตเน็ตจะหยุดการเชื่อมต่อเป็นเวลา 30 นาที แล้วทำการติดต่อใหม่

4.3.2 การลดอัตราการส่งจดหมายอิเล็กทรอนิกส์สแปมจากประวัติการใช้งาน

กำหนดปริมาณ Bandwidth โดยพิจารณาปริมาณการส่งข้อมูลจดหมายอิเล็กทรอนิกส์ของเครื่องนั้น ช่วงเวลา 1 ชั่วโมง ย้อนหลังไป 7 วัน โดยเทียบกับเวลาปัจจุบัน (ผู้วิจัยใช้วิธีนี้)

ภาพที่ 4.18 แสดงปริมาณ Bandwidth ที่ถูกควบคุมในการส่งจดหมายอิเล็กทรอนิกส์ แบบที่ 2 โดยการพิจารณาประวัติการใช้งาน



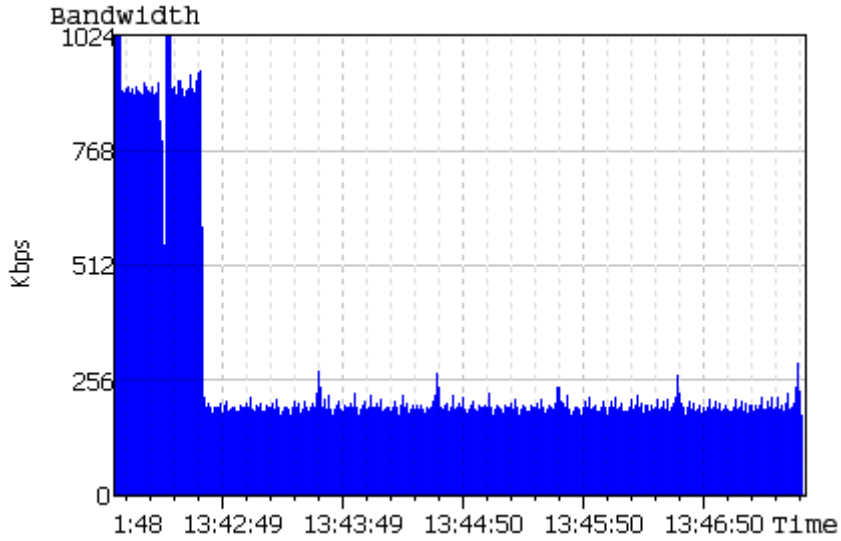
จากภาพที่ 4.18 การควบคุม Bandwidth ซึ่งผลการทดลองที่ได้ ในช่วงระยะเวลาแรกปริมาณข้อมูลที่เกิดจากการส่งจดหมายอิเล็กทรอนิกส์ของบอตเน็ตจะมีขนาดประมาณ 896 Kbps. และต่อจากนั้นในช่วงระยะเวลาที่สอง ระบบ DSDSE จะทำการควบคุม Bandwidth ให้สามารถส่งจดหมายอิเล็กทรอนิกส์สแปมได้ช้าลงจนทำให้ขนาดของข้อมูลอยู่ที่ 100 Kbps ซึ่งสามารถลดขนาดการส่งข้อมูลได้ถึง 896 – 100 = 796 Kbps. คิดเป็นอัตราส่วนประมาณ 88.8 เปอร์เซ็นต์ ผลการทดลองพบว่าระบบ DSDES จะเริ่มลดอัตรา Bandwidth ตั้งแต่นาทีที่ 1 และ

บอตเน็ตยังใช้เวลาส่งจดหมายอิเล็กทรอนิกส์สแปมต่อเนื่องแต่จะใช้เวลานานขึ้นจากเดิมที่ใช้เวลาส่ง 4 นาที เมื่อใช้ระบบ DSDSE แล้วทำให้บอตเน็ตใช้เวลาในการส่งรวม 31 นาที

4.3.3 การลดอัตราการส่งจดหมายอิเล็กทรอนิกส์สแปมแบบกำหนดค่ากลาง

กำหนดปริมาณ Bandwidth โดยกำหนดค่ากลางไว้ให้ทุกๆเครื่องใช้งานร่วมกัน เมื่อระบบติดบอตเน็ต ระบบ Detection จะวิเคราะห์จากปริมาณค่ากลางที่ตั้งไว้หารด้วยจำนวนบอตเน็ตทั้งหมด ซึ่งจะได้ค่าผลลัพธ์เป็นปริมาณ Bandwidth ที่ให้ใช้ในแต่ละเครื่องที่ติดบอตเน็ต

ภาพที่ 4.19 แสดงปริมาณ Bandwidth ที่ถูกควบคุมในการส่งจดหมายอิเล็กทรอนิกส์ แบบที่ 3 โดยการกำหนดค่ากลาง



จากภาพที่ 4.19 การควบคุมแบบนี้จะขึ้นอยู่กับจำนวนเครื่องที่ติดบอตเน็ตหารด้วยปริมาณค่ากลางซึ่งในการทดลอง กำหนดค่ากลางอยู่ที่ 2,000 Kbps. จากภาพมีการติดบอตทั้งหมด 10 เครื่อง ดั้งนั้นค่า Bandwidth ที่ให้แต่ละเครื่องใช้นั้นจะได้ประมาณ 200 Kbps.

การลดอัตราการส่งจดหมายอิเล็กทรอนิกส์สแปม โดยวิธีการควบคุมอัตราการ Bandwidth ทั้ง 3 วิธี ผู้วิจัยได้เลือกแบบที่ 2 ทั้งนี้ผู้วิจัยไม่ได้กำหนดว่าแบบไหนดีที่สุด เพราะยังคงต้องทำการวิจัยศึกษาในเรื่องดังกล่าว เพื่อให้ได้ค่าที่เหมาะสมที่สุดในการควบคุม Bandwidth ควรจะอยู่ที่ค่าใด เนื่องจากมีประเด็นที่ต้องพิจารณาหลายประเด็น เช่น ความพึงพอใจของผู้ใช้งาน ความคุ้มทุนในการใช้ พฤติกรรมของระบบเครือข่ายของแต่ละองค์กร เป็นต้น