บทที่ 2

แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

การศึกษาและวิจัยในครั้งนี้เป็นการวิเคราะห์หาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์ ศึกษาการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ พัฒนาตัวแบบการสืบค้นผู้กระทำปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ พร้อมทั้งพัฒนาตัวแบบการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า เพื่อที่จะค้นหากลยุทธ์การติดตามภัยจากการใช้เครือข่ายสังคมออนไลน์จากตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ ดังนั้นงานวิจัยชิ้นนี้จึงได้รวบรวมแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ไว้ดังต่อไปนี้

2.1 เครือข่ายสังคมออนไลน์ (Online Social Networks)

2.2 ความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ (Anomaly and Attack Patterns in Online Social Networks)

2.3 การวิเคราะห์ความมั่นคงปลอดภัย (Security Analytics)

2.4 การรักษาความมั่นคงปลอดภัยไซเบอร์ (Cybersecurity)

2.5 อันตรายของสแปม (Spam)

2.6 เทคนิคเหมืองข้อมูล (Data Mining)

2.7 โครงสร้างกราฟ (Graph-based Structure)

2.8 สถิติวิเคราะห์ (Statistic Analysis)

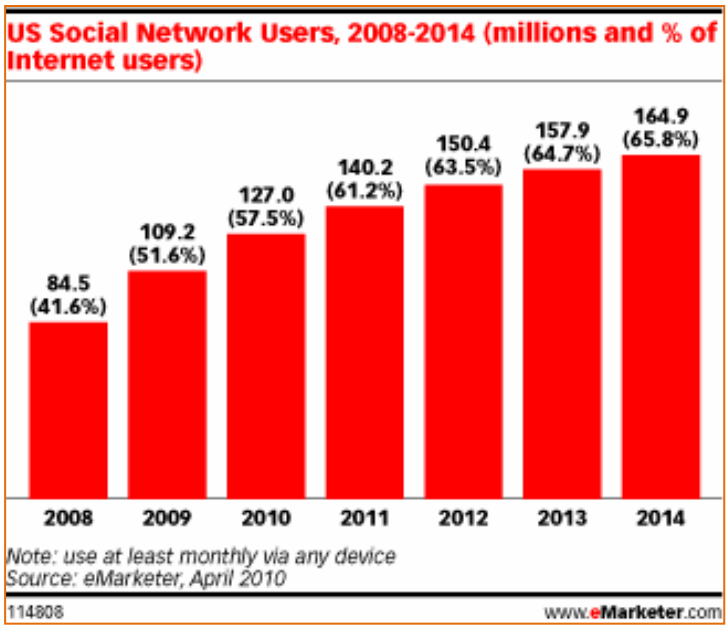
2.9 การแสดงภาพวิชวลไลเซชัน (Visualization)

2.10 งานวิจัยที่เกี่ยวข้อง

2.11 สรุป

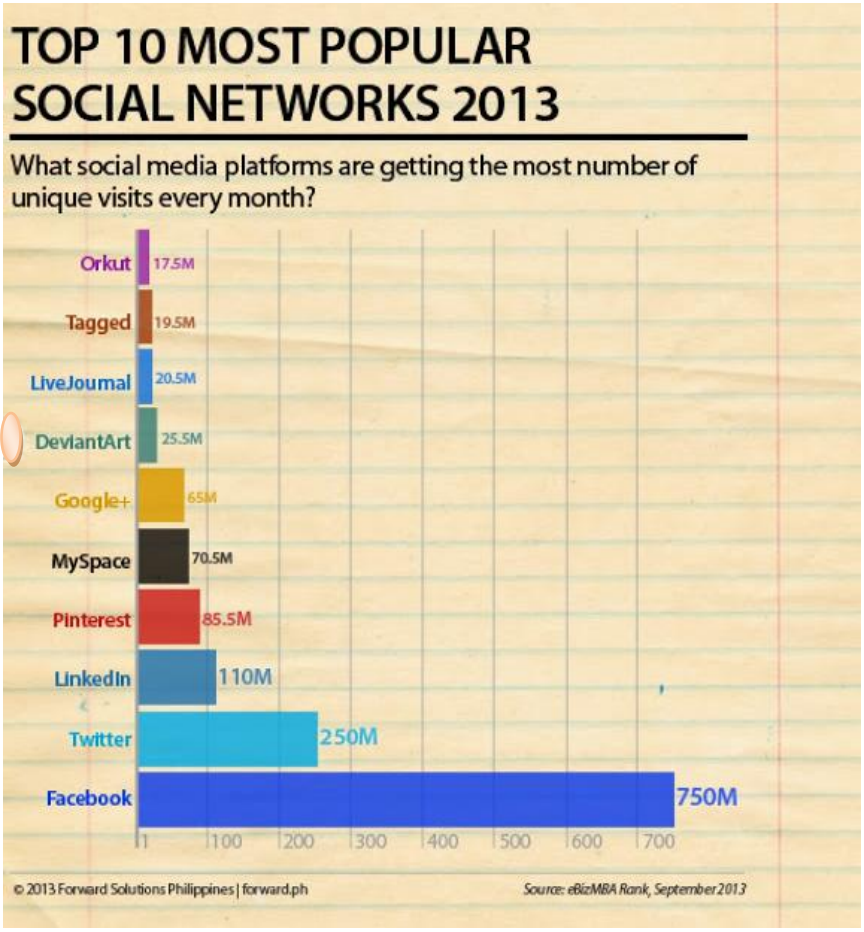
2.1 เครือข่ายสังคมออนไลน์ (Online Social Networks)

เครือข่ายสังคมออนไลน์ช่วยให้ผู้คนจำนวนมากได้มามีปฏิสัมพันธ์กัน ซึ่งวัตถุประสงค์ที่แท้จริงของคำว่าเครือข่ายสังคมออนไลน์ก็คือ ผู้ใช้มีส่วนสร้างเนื้อหาบนอินเทอร์เน็ต และ ผู้ใช้ยังเป็นผู้ร่วมกำหนดคุณค่าของเนื้อหา ทำให้สังคมพิจารณาได้ว่าเนื้อหาใดมีคุณภาพ รวมทั้งการให้ความสำคัญกับทัศนคติของคนที่มีต่อการเรียนรู้เรื่องราวต่างๆ มารวมกันให้เป็นองค์กรการเรียนรู้ขนาดใหญ่ของทุกคนที่ใช้เว็บ ทำให้เกิดการต่อยอดความรู้และความคิด นำมาซึ่งองค์ความรู้ ทำให้เราเห็นปรากฏการณ์การเกิดชุมชนออนไลน์ ทั้งนี้เครือข่ายสังคมออนไลน์ที่เป็นที่นิยมอย่างมากในกลุ่มคนที่ใช้อินเทอร์เน็ต ทำให้เครือข่ายขยายวงกว้างออกไปเรื่อยๆ และจะยังคงเป็นเช่นนี้ต่อไปอีกในอนาคต



รูปที่ 2.1 แนวโน้มความนิยมใช้งานเครือข่ายสังคมออนไลน์ของประเทศสหรัฐอเมริกา
(ที่มา: www.eMarketer.com, ออนไลน์, 2553)

จากรูปที่ 2.1 แสดงผลการสำรวจจากประเทศสหรัฐอเมริกายืนยันการใช้บริการเครือข่ายสังคมออนไลน์เพิ่มสูงขึ้นอย่างต่อเนื่องทุกปี และมาแรงเป็นอันดับต้นๆ ของโลกออนไลน์ โดยการคาดการณ์การใช้งานเครือข่ายสังคมออนไลน์ของประเทศสหรัฐอเมริกาในปี 2557 จะสูงถึง 65.8% ของผู้ใช้งานอินเทอร์เน็ตทั้งหมด



รูปที่ 2.2 ความนิยมใช้งานเครือข่ายสังคมออนไลน์

(ที่มา: www.ebizmba.com/, ออนไลน์, 2556)

จากรูปที่ 2.2 แสดงถึงความนิยมใช้งานเครือข่ายสังคมออนไลน์ โดยที่เครือข่ายสังคมออนไลน์ที่มีผู้นิยมใช้งานมากที่สุดคือ Facebook, Twitter และ Linkedln และเว็บไซต์เครือข่ายสังคมออนไลน์ที่มีเปอร์เซ็นต์เติบโตสูงมากขึ้นเป็นลำดับได้แก่ Facebook และ Twitter ซึ่งเป็นที่นิยมมากๆ ในประเทศไทย [3, 34, 59, 81]

เครือข่ายสังคมออนไลน์ เป็นปรากฏการณ์ของการเชื่อมต่อระหว่างบุคคลในโลกอินเทอร์เน็ต และ รวมถึงการเชื่อมต่อระหว่างเครือข่ายกับเครือข่ายสังคมออนไลน์เข้าด้วยกัน เป็นการเน้นไปที่การสร้างชุมชนออนไลน์ซึ่งผู้คนสามารถที่จะแลกเปลี่ยน แบ่งปันตามผลประโยชน์ กิจกรรม หรือความสนใจเฉพาะเรื่อง ซึ่งอาศัยระบบพื้นฐานของเว็บไซต์ที่ทำให้มีการโต้ตอบกันระหว่างผู้คนโดยแต่ละเว็บไซต์นั้นอาจมีการให้บริการที่ต่างกัน เช่น อีเมล กระดานข่าว และในต่อมาเป็นการแบ่งปันพื้นที่ให้สมาชิกเป็นเจ้าของพื้นที่ร่วมกันและแบ่งปันข้อมูล

ประเภทของเครือข่ายสังคมออนไลน์ในอินเทอร์เน็ต โดยพิจารณาจากเป้าหมายของการเข้าเป็นสมาชิกในเครือข่ายสังคมออนไลน์ สามารถแบ่งได้เป็น 5 กลุ่มใหญ่ๆ คือ [6]

1. การแสดงตัวตนและภาพลักษณ์ของตน (Identity Network) เช่น http://www.hi5.com/ http://www.facebook.com/

2. กลุ่มเครือข่ายที่มีความสนใจร่วมกัน (Interested Network) เป็นการรวมตัวกันโดยอาศัย “ความสนใจ” ตรงกัน เช่น Digg.com , del.icio.us

3. กลุ่มเครือข่ายที่ทำงานร่วมกัน (Collaboration Network) เป็นกลุ่มเครือข่ายที่ร่วมกัน “ทำงาน” ยกตัวอย่างเช่น http://www.wikipedia.org/

4. กลุ่มจำลองตัวเองในโลกเสมือน (Gaming/Virtual Reality) หรือ โลกเสมือน ในบางครั้งเราเจอคำว่า Second Life ซึ่งเป็นลักษณะของเครือข่ายสังคมออนไลน์ที่มีลักษณะเป็นการสวมบทบาทของผู้เล่นในชีวิตจริงกับตัวละครในเกม เช่น Ragnarok

5. กลุ่มที่ใช้ในงานอาชีพ (Professional Network) กลุ่มนี้มีวัตถุประสงค์เพื่อใช้ประโยชน์ในงานอาชีพ เช่น LinkedIn

2.2 ความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ (Anomaly and Attack Patterns in Online Social Networks)

เนื่องจากการใช้อินเทอร์เน็ตและการใช้เวลาอยู่บนโลกออนไลน์ในปัจจุบันนี้มีอัตราการเติบโตสูงขึ้นอย่างรวดเร็ว ที่เห็นได้อย่างชัดคือ มีการใช้อินเทอร์เน็ตในการสืบค้นข้อมูล การอ่านข่าวสารผ่านเว็บไซต์ หรือแม้แต่การใช้โทรศัพท์เคลื่อนที่ในการเชื่อมต่ออินเทอร์เน็ต ส่งผลให้พบปัญหาที่ตามมาจากอัตราการใช้อินเทอร์เน็ตได้เพิ่มมากขึ้นเรื่อยๆ คือ ความมั่นคงปลอดภัยของผู้ใช้อินเทอร์เน็ต ซึ่งภัยคุกคามทางอินเทอร์เน็ตได้ลุกลามมากขึ้น และยังเป็นเป้าหมายใหม่ของเหล่ามิจฉาชีพ และมิจฉาชีพก็ได้หาวิธีที่แยบยลมากขึ้น เพื่อใช้เป็นกลลวงและหลอกล่อเหยื่อ โดยวิธีที่ใช้งานกันอย่างแพร่หลายสำหรับการเข้าถึงคอมพิวเตอร์ก็คือ สแปม ซึ่งยังคงเป็นสื่อหลักสำหรับการแพร่กระจายเวิร์มและมัลแวร์ รวมถึงการปิดกั้นข้อมูลจราจรทางอินเทอร์เน็ต

ปัจจุบันมีข้อความสแปมที่ถูกส่งออกไปถึง 180,000 ล้านฉบับในแต่ละวัน โดยคิดเป็นสัดส่วนราว 90 เปอร์เซ็นต์ของข้อมูลจราจรของอีเมลทั่วโลก ซึ่งในช่วงที่ผ่านมาตัวอย่างของภัยร้ายนี้ก็คือ ผู้ใช้ Facebook จะเห็นได้ว่าสแปมในปัจจุบันไม่ได้มีจำกัดอยู่แต่ในอีเมลเหมือนที่ผ่านมาเท่านั้น หากแต่ขยายไปถึงเครือข่ายสังคมออนไลน์อื่นๆ ด้วย [96] ภัยอินเทอร์เน็ตสมัยใหม่จึงมุ่งไปยังกลุ่มผู้ใช้งานตามบ้านผ่านทางการเข้าเว็บไซต์เครือข่ายสังคมออนไลน์เหล่านั้น โดยพบการหลอกหลวงผ่านทางเว็บไซต์ยอดนิยมต่างๆ โดยใช้เทคนิคการวิศวกรรมทางสังคม (Social Engineering) และภัย

จากโปรแกรมมุ่งร้ายหรือโปรแกรมมัลแวร์ (Threat from Malware) ที่แพร่กระจายอยู่ที่เครือข่ายสังคมออนไลน์ เช่นมีโปรแกรมม้าโทรจันแพร่กระจายผ่านทางเว็บไซต์เครือข่ายสังคมออนไลน์ หรือผ่านทาง Instant Messaging เช่น MSN เป็นต้น

จากรายงานของบริษัท Symantec พบว่า Facebook เป็นเว็บไซต์เครือข่ายสังคมออนไลน์ ที่มีการระบาดของสแปมสูงมากเป็นอันดับ 1 นั่นเป็นเพราะว่า Facebook เป็นเว็บไซต์เครือข่ายสังคมออนไลน์ ที่เป็นที่นิยมมาก และสามารถเข้าไปเพื่อติดต่อพูดคุย เล่นเกม หรือใช้แอพพลิเคชั่นต่างๆ

อีกทั้งทางบริษัท BitDefender ได้มีการศึกษาถึงความรวดเร็วในการแพร่กระจายไวรัสที่มาจากข้อความสแปม โดยใช้ระบบรายงานการตรวจจับการแพร่กระจายของไวรัสแบบเชิงเวลาจริง (BitDefender Real-Time Virus Reporting System) โดยระบบดังกล่าวสามารถตรวจพบการแพร่กระจายของไวรัสที่ชื่อ “Trojan.Dropper.Oficla.G” จากข้อความสแปม และมีแนวโน้มเห็นการเริ่มต้นการแพร่กระจายไวรัส Trojan.Dropper.Oficla.G ที่มากขึ้น [5]

2.3 การวิเคราะห์ความมั่นคงปลอดภัย (Security Analytics) [4, 13, 43, 92, 101]

กระบวนการระบบบริหารจัดการด้านความมั่นคงปลอดภัยประกอบด้วย 4 ขั้นตอนหลัก กล่าวคือ

1. การวางแผน (Plan)
2. การดำเนินการตามแผน (Do)
3. การเฝ้าระวังและติดตามการดำเนินการตามแผน (Check)
4. การดำเนินการเพิ่มเติมตามที่เห็นสมควร (Act)

เมื่อพิจารณากระบวนการระบบบริหารจัดการด้านความมั่นคงปลอดภัยจะพบว่าหลักของทั้ง 4 ข้อนั้นคือการประเมินความเสี่ยงและการจัดทำแผนการลดความเสี่ยง (ขั้นตอนการวางแผน หรือ Plan) การดำเนินการตามแผนการลดความเสี่ยง (ขั้นตอนการดำเนินการตามแผน หรือ Do) การเฝ้าระวังและติดตามการดำเนินการตามแผน (ขั้นตอนการเฝ้าระวังและติดตามการดำเนินการตามแผน หรือ Check) และการดำเนินการเพิ่มเติมตามที่เห็นสมควร (ขั้นตอนการดำเนินการเพิ่มเติมตามที่เห็นสมควร หรือ Act)

2.3.1 การวางแผน (Plan)

ขั้นตอนการวางแผน เป็นการประเมินความเสี่ยงที่มีต่อทรัพย์สินสารสนเทศ ซึ่งส่วนใหญ่จะหมายถึงทรัพย์สินสารสนเทศใหม่ที่กำลังนำเข้ามาสู่การใช้งาน ที่จะต้องมีการประเมินความเสี่ยงเพื่อเตรียมการป้องกันก่อนเริ่มต้นใช้งานทรัพย์สินใหม่เหล่านั้น

2.3.2 การดำเนินการตามแผน (Do)

ขั้นตอนการดำเนินการตามแผนที่ได้กำหนดไว้ทั้งหมดในข้างต้น กล่าวคือการพัฒนาระบบงานตามแผนการที่เตรียมไว้ในข้างต้นทั้งหมด

2.3.3 การเฝ้าระวังและติดตามการดำเนินการตามแผน (Check)

ขั้นตอนการเฝ้าระวังและติดตามการดำเนินการตามแผน คือการติดตามเพื่อดูว่าการพัฒนาระบบงาน ได้ปฏิบัติตามแผนครบทั้งหมดและได้ผลลัพธ์ที่ถูกต้องหรือไม่

2.3.4 การดำเนินการเพิ่มเติมตามที่เห็นสมควร (Act)

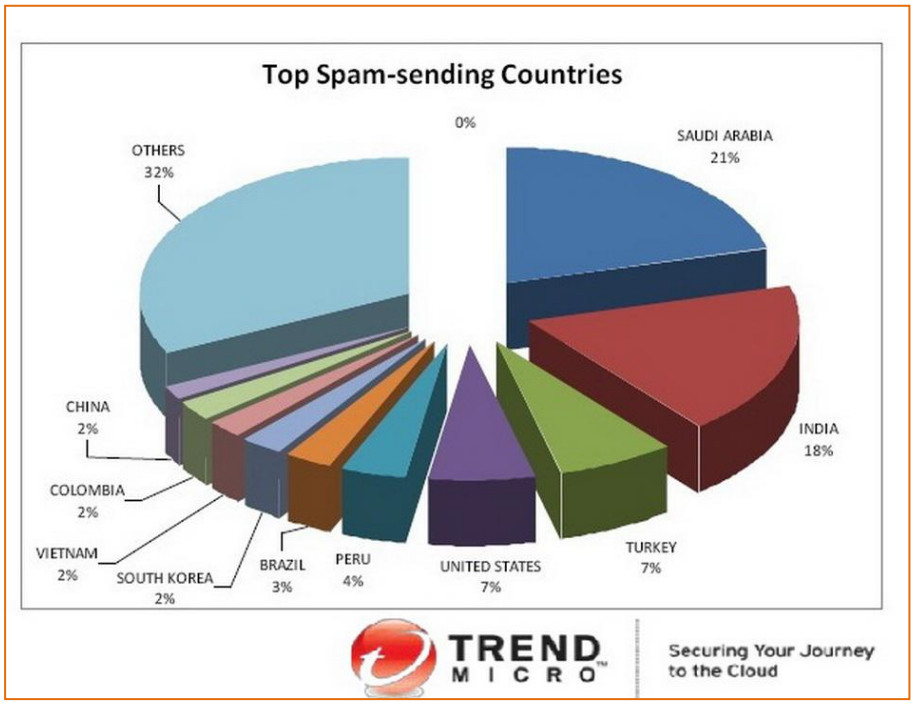
หากเมื่อการเฝ้าระวังและติดตามการดำเนินการตามแผนในทุกรายการแล้ว และพบปัญหาที่ต้องทำการแก้ไขให้ดำเนินการเพิ่มเติมตามที่เห็นสมควร (Act) เช่น กรณีพบว่าระบบงานยังไม่รองรับปัญหาการโจมตีระบบแบบเอส คิว แอล อินเจคชัน (SQL Injection) ให้ดำเนินการแก้ไข (Take Action) เพื่อให้ระบบสามารถจัดการกับปัญหานี้ได้

2.3.5 การเฝ้าระวังและติดตามการดำเนินการตามแผนและการดำเนินการเพิ่มเติมตามที่เห็นสมควร

โดยสรุปวงจรชีวิตของกระบวนการระบบบริหารจัดการด้านความมั่นคงปลอดภัย (Plan, Do, Check, Act) นี้ เริ่มต้นตั้งแต่เริ่มมีทรัพย์สินสารสนเทศใหม่ที่กำลังถูกนำเข้ามาใช้งานกับองค์กร จากนั้นแม้ว่าจะมีการติดตั้งและใช้งานทรัพย์สินเหล่านั้นไปแล้วก็ตาม ก็ยังคงการเฝ้าระวังและติดตามการดำเนินการตามแผน (Check) และ การดำเนินการเพิ่มเติมตามที่เห็นสมควร (Act) อย่างต่อเนื่อง จนกระทั่งทรัพย์สินสารสนเทศเหล่านั้นหมดอายุการใช้งานนั่นเอง จึงยุติการประเมินความเสี่ยงสำหรับทรัพย์สินเหล่านั้น

2.4 อันตรายของสแปม (Spam) [5]

ผลกระทบและความเสียหายที่เกิดจากสแปม (Spam) นั้น อาจสร้างความเดือนร้อนไม่มาก รวมไปถึงการทำให้พื้นที่ในระบบที่ใช้งานมีไม่เพียงพอ หรือระบบมีข้อมูลมากจนไม่สามารถทำงานได้อย่างปกติ หรืออาจเกิดความเสียหายร้ายแรงอื่น ๆ เกิดขึ้นตามมา นอกจากนั้นยังทำให้ผู้รับและผู้ดูแลระบบต้องเสียเวลาหรือต้องเสียค่าใช้จ่ายในการลบหรือทำลายข้อมูลในลักษณะที่เป็น สแปม



รูปที่ 2.3 สถิติประเทศที่มีการปล่อยสแปมระบาดผ่านลงบนเว็บไซต์เครือข่ายสังคมออนไลน์ (ที่มา: http://www.trendmicro.co.th/, ออนไลน์, 2555)

จากรูปที่ 2.3 แสดงสถิติประเทศที่มีการปล่อยสแปมระบาดผ่านลงเว็บไซต์เครือข่ายสังคมออนไลน์ โดยในปี 2555 พบว่า ประเทศซาอุดิอาระเบียมีการปล่อยสแปมระบาดมากที่สุด โดยมีปริมาณมากถึง 21% ของทั้งหมด รองลงมาคือประเทศอินเดียโดยมีปริมาณมากถึง 18% ของทั้งหมด โดยที่ประเทศสหรัฐอเมริกาและประเทศตุรกีมีปริมาณ 7% ของทั้งหมด

นอกจากนี้ เครือข่ายสังคมออนไลน์ก็ยังเป็นสถานที่ที่เหมาะสมอย่างยิ่งในการขโมยข้อมูลที่ระบุตัวบุคคลจากการโพสต์ข้อมูลกันอย่างเปิดเผยโดยเฉพาะในส่วนของข้อมูลส่วนตัว (Profile) ของผู้ใช้งานก็ดูน่าเชื่อถืออย่างมาก และเมื่อรวมเข้ากับข้อมูลการโต้ตอบบนเครือข่ายสังคมออนไลน์

ก็สามารถให้อาชญากรที่จะขโมยข้อมูลส่วนบุคคลและถูกโจมตีด้วยเทคนิคกลลวงทางสังคมแบบมีเป้าหมาย โดยนำไปใช้ในการปลอมตัวเป็นเหยื่อหรือใช้ในการดำเนินการต่างๆ ได้ [10, 11, 14]

ทั้งนี้จากรายงานด้านความมั่นคงปลอดภัยประจำไตรมาสที่ 3 ของปี 2555 (Trend Micro's Top Security Trends for 2012) โดยบริษัทเทรนด์ ไมโคร (ประเทศไทย) จำกัด จัดทำสรุปไว้ดังนี้

1. มัลแวร์ ZeroAccess เป็นตัวแพร่ไวรัสอันดับหนึ่งที่มีอยู่ในระบบคอมพิวเตอร์ของไตรมาสนี้ ขณะที่หนอน DOWNAD/Conficker ตามมาเป็นอันดับที่สอง ซึ่งมัลแวร์เหล่านี้ยังคงแพร่ระบาดอย่างหนักในอินเดีย อินโดนีเซีย และฟิลิปปินส์ที่ซึ่งผู้คนยังคงนิยมใช้บริการอินเทอร์เน็ตสาธารณะอยู่เป็นประจำ

2. ประเทศเกาหลีใต้ก้าวขึ้นแซงหน้าประเทศญี่ปุ่นในฐานะประเทศที่มีเว็บไซต์ซึ่งโฮสต์มัลแวร์ไว้เป็นจำนวนสูงสุดในภูมิภาคเอเชียแปซิฟิก โดยเกาหลีใต้ยังโฮสต์โดเมนที่เป็นอันตรายซึ่งถูกบล็อกไว้เป็นอับดับที่สองในโลกอีกด้วย จะเห็นได้ว่าญี่ปุ่น, เกาหลีใต้, ไต้หวัน, อินเดีย และออสเตรเลีย ติดอับดับสิบประเทศแรกที่มีการเรียกใช้เว็บไซต์ที่เป็นอันตรายมากที่สุด

3. บริษัทและหน่วยงานภาครัฐยังคงตกเป็นเป้าหมายของภัยคุกคามต่อเนื่องขั้นสูง (APT) ซึ่งตรวจพบว่า APT ที่ชื่อ Lurid and Nitro ได้รับการปรับปรุงให้มีประสิทธิภาพมากขึ้น และจากการติดตามตรวจสอบการโจมตีแบบมีเป้าหมายในระบบของลูกค้าของเรา พบว่าประเทศญี่ปุ่นและภูมิภาคเอเชียแปซิฟิกยังคงตกเป็นเป้าหมายอันดับหนึ่ง

4. PayPal ยังคงดึงดูดบรรดาผู้หลอกลวงทางอินเทอร์เน็ตในระดับสูงสุด ขณะที่ Linkedin ติดอับดับเป้าหมายที่ถูกเลือกของ Blackhole Exploit Kit

5. จากสถิติสามารถระบุได้ว่าสแปมน่าจะมีที่มาจากประเทศซาอุดิอาระเบีย (หรืออินเดีย)

6. ภัยคุกคามสื่อสังคมออนไลน์และความกังวลด้านสิทธิส่วนบุคคลยังคงเกิดขึ้นอย่างต่อเนื่อง

โดยการคาดการณ์ที่สำคัญเพิ่มเติมในอนาคต อาจเป็นดังนี้

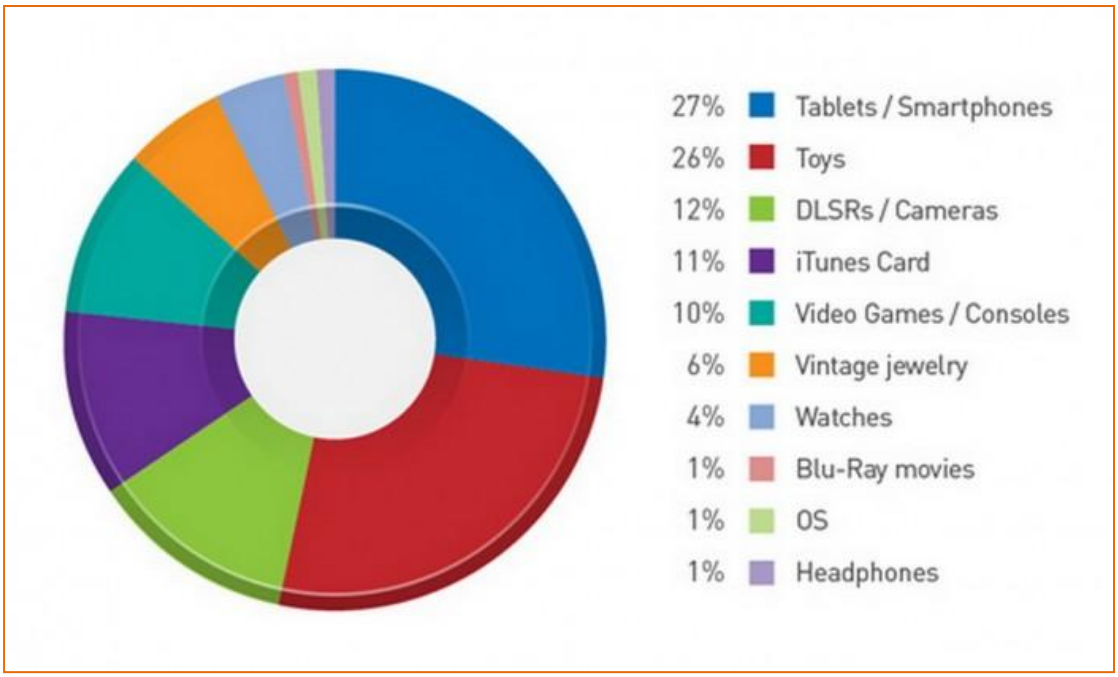
1. ด้านเกี่ยวกับธุรกรรมทางการเงิน จะยังคงมีอาชญากรรมไซเบอร์เคียงข้างอยู่ด้วย การลดความเสี่ยงไม่ใช่ตัวเลือกที่ใช้ได้อีกต่อไป แม้แต่การใช้เบราว์เซอร์อื่น หรือระบบปฏิบัติการทางเลือก

2. มัลแวร์ (Malware) จะเปลี่ยนรูปแบบไปในเวลาเพียงไม่กี่ชั่วโมง และการติดมัลแวร์จะเป็นเรื่องปกติ

3. การโจมตีแบบใหม่ๆ จะเพิ่มมากขึ้นในสภาพแวดล้อมแบบเสมือนจริงและระบบการประมวลผลแบบคลาวด์ (Cloud Computing)

4. การกระจายของโปรแกรมที่ทำงานโดยอัตโนมัติ (Bot) จะขยายวงเพิ่มขึ้น

5. เครือข่ายบริษัทหรือเครือข่ายทางสังคมก็อาจจะยังคงถูกก่อกวนจากอาชญากรที่จะขโมยข้อมูลไปใช้หาประโยชน์



รูปที่ 2.4 ประเภทของเว็บไซต์เครือข่ายสังคมออนไลน์ที่ถูกเลียนแบบในช่วงเทศกาลปีใหม่ (ที่มา: http://www.trendmicro.co.th/, ออนไลน์, 2556)

ทั้งนี้ นายไรอัน เซอเทซา ฝ่ายสื่อสารด้านเทคนิค ศูนย์วิจัยเทรนด์แล็บส์ บริษัท เทรนด์ ไมโคร อิงค์ ผู้นำระดับโลกด้านโซลูชั่นรักษาความปลอดภัยข้อมูลสำหรับระบบคลาวด์ เปิดเผยว่า “สำหรับหลายคนแล้ว เทศกาลวันหยุดถือเป็นฤดูกาลสำคัญแห่งการจับจ่ายซื้อสินค้าให้คนสำคัญ แต่สำหรับอาชญากรไซเบอร์กลับมองเห็นช่องทางที่ต่างออกไป นั่นคือโอกาสครั้งสำคัญในการโจรกรรมข้อมูล”

ตัวอย่างเช่น การจับจ่ายซื้อสินค้าออนไลน์ เว็บไซต์ที่เป็นอันตรายจะพยายามล่อลวงให้ผู้ซื้อสินค้าทางออนไลน์ชำระเงินให้พวกเขาแทนที่จะจ่ายให้กับเว็บไซต์ที่แท้จริง เว็บไซต์เหล่านี้มักจะถูกสร้างขึ้นมาให้มีลักษณะเหมือนกับเว็บไซต์ขายสินค้าที่ถูกเลียนแบบ รวมทั้งมีหน้าจอล็อกอินที่จะขอให้ผู้ใช้ใส่ข้อมูลส่วนตัว เว็บไซต์ลวงเหล่านี้ให้ความสนใจในข้อมูลล็อกอินทุกประเภท โดยล่าสุดเราพบไซต์ฟิชชิ่งที่ขโมย Apple ID ของผู้ใช้ เป็นต้น

2.5 การรักษาความมั่นคงปลอดภัยไซเบอร์ (Cybersecurity)

การรักษาความมั่นคงปลอดภัยไซเบอร์ (Cybersecurity) [1, 37, 38, 43, 56, 59, 60, 61, 67, 80, 81, 85, 91, 94, 107, 110, 112, 115, 119] เป็นมาตรการที่ใช้สำหรับป้องกันผู้ที่ไม่ได้รับอนุญาตให้เข้าใช้เครือข่ายคอมพิวเตอร์และอินเทอร์เน็ตเพื่อทำลายข้อมูลหรือทำในสิ่งที่อาจก่อให้เกิดความเสียหายต่อคุณสมบัติของข้อมูล หมายถึง การปกป้องข้อมูลโดยการป้องกัน การตรวจหา และการตอบสนองต่อการถูกโจมตี

การรักษาความมั่นคงปลอดภัย หมายถึง มาตรการที่ใช้ในการป้องกันทรัพยากรจากภัยคุกคามทั้งโดยเจตนาและไม่เจตนา ซึ่งเป็นหนึ่งในวิธีช่วยลดความเสี่ยงด้านความมั่นคงปลอดภัย โดยการจำกัดให้เฉพาะผู้ที่จำเป็นต้องใช้งาน เช่น ผู้ดูแลระบบเท่านั้นที่สามารถเข้าถึงระบบควบคุมหลักได้ และผู้ใช้ทั่วไปสามารถใช้พื้นที่ ไฟล์ข้อมูล ฐานข้อมูล และโปรแกรมระบบฯ เฉพาะที่ผู้ดูแลระบบอนุญาตเท่านั้น ความมั่นคงปลอดภัยคอมพิวเตอร์เป็นกระบวนการที่เกี่ยวข้องกับการป้องกันและตรวจสอบการเข้าใช้งานเครื่องคอมพิวเตอร์โดยไม่ได้รับอนุญาต ขั้นตอนการป้องกันจะช่วยให้ผู้ที่ใช้งานสกัดกั้นไม่ให้เครื่องคอมพิวเตอร์ถูกเข้าใช้งานโดยผู้ที่ไม่ได้รับสิทธิ์ ส่วนการตรวจสอบจะทำให้ทราบได้ว่ามีใครกำลังพยายามที่จะบุกรุกเข้ามาใช้ระบบหรือไม่ การบุกรุกสำเร็จหรือไม่ และผู้บุกรุกทำอะไรกับระบบบ้าง

2.5.1 แนวคิดเกี่ยวกับการรักษาความมั่นคงปลอดภัยไซเบอร์

การรักษาความมั่นคงปลอดภัยไซเบอร์ มีแนวคิดเกี่ยวกับการรักษาความมั่นคงปลอดภัยที่หน่วยงานควรคำนึงถึงดังนี้

1. ความเป็นส่วนบุคคล (Privacy) หมายถึงข้อมูลถูกใช้ตามวัตถุประสงค์ที่เจ้าของข้อมูลระบุในช่วงที่เก็บรวบรวมเท่านั้น

2. การระบุตัวตน (Identification) หมายถึงระบบสามารถระบุตัวตนของผู้ใช้แต่ละคนที่ใช้งานระบบได้เพื่อให้สามารถเข้าถึงข้อมูลที่มีการกำหนดระดับความลับข้อมูลตัวอย่าง เช่น การระบุชื่อผู้ใช้

3. การพิสูจน์ทราบตัวตน (Authentication) หมายถึงระบบสามารถพิสูจน์ได้ว่าผู้ใช้คือผู้ที่ได้รับอนุญาตจริง ตัวอย่างเช่นป้อนรหัสผ่านที่ตรงกับการระบุชื่อผู้ใช้ได้ถูกต้องและสามารถผ่านเข้าใช้ข้อมูลได้

4. การอนุญาตให้ใช้งาน (Authorization) หมายถึงการตรวจสอบสิทธิ์ของผู้ใช้ว่าสามารถใช้งานระบบได้ในระดับใดหลังจากระบุรายชื่อผู้ใช้และป้อนรหัสผ่านถูกต้องแล้ว เช่น การ

เข้าถึง การอ่าน การแก้ไข การลบข้อมูล การอนุญาตให้ใช้งานจะครอบคลุมถึงการรวมกลุ่มของผู้ใช้ในระบบ เช่น กลุ่มผู้ใช้ทั่วไป ผู้ดูแลระบบ

5. การตรวจสอบได้ (Accountability) หมายถึงความสามารถในการตรวจสอบการใช้งานระบบได้ว่าผู้ใช้แต่ละคนได้ทำกิจกรรมใดบ้างกับระบบ ตัวอย่างเช่นการจัดเก็บแฟ้มบันทึกข้อมูลการใช้เครือข่าย (Log File) เกี่ยวกับกิจกรรมต่าง ๆ ผู้ใช้แต่ละคนใช้งานระบบ

2.5.2 องค์ประกอบของการรักษาความมั่นคงปลอดภัยไซเบอร์

องค์ประกอบของการรักษาความมั่นคงปลอดภัยไซเบอร์เพื่อให้คุณสมบัติของข้อมูลที่มีความสำคัญและถือว่าเป็นทรัพย์สินอันมีค่าขององค์กรคงความสำคัญ 3 ด้านคือ

1. ความลับ (Confidentiality) หมายถึงข้อมูลที่มีความสำคัญและเป็นทรัพย์สินอันมีค่าไม่ได้ถูกเปิดเผยให้ผู้ไม่หวังดีรับรู้

2. ความคงสภาพ (Integrity) หมายถึงข้อมูลไม่ได้ถูกเปลี่ยนแปลงไปจากสภาพเดิมยังคงมีความถูกต้องของเนื้อหาและแหล่งที่มา

3. ความพร้อมใช้งาน (Availability) หมายถึงข้อมูลและอุปกรณ์ของระบบเครือข่ายคอมพิวเตอร์และอินเทอร์เน็ตสามารถใช้งานตามต้องการ

2.6 เทคนิคเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) คือกระบวนการค้นหาสารสนเทศหรือข้อความรู้ที่อยู่ในฐานข้อมูลขนาดใหญ่ที่ซับซ้อน เพื่อนำข้อความรู้ที่ได้ไปใช้ประโยชน์ในการตัดสินใจ สารสนเทศที่ได้อาจนำมาสร้างการพยากรณ์หรือพัฒนาตัวแบบสำหรับการจำแนกหน่วยหรือกลุ่มหรือแสดงความสัมพันธ์ระหว่างหน่วยต่างๆ หรือให้ข้อสรุปของสาระในฐานข้อมูล การทำเหมืองข้อมูลประกอบขึ้นด้วยการนำกระบวนการทางสถิติและการเรียนรู้ผ่านระบบคอมพิวเตอร์ เพื่อพัฒนาตัวแบบ กฎเกณฑ์ รูปแบบ การพยากรณ์และข้อความรู้ จากฐานข้อมูลขนาดใหญ่ โดยการทำเหมืองข้อมูลมีขั้นตอนการดำเนินงานหลายขั้นตอนซึ่งต้องอาศัยเทคนิคหรือวิธีการต่างๆ เช่น วิธีการรวมกลุ่ม (Cluster Analysis) การค้นหากฎความสัมพันธ์ (Association Rules) การพยากรณ์ (Prediction) เป็นต้น

2.6.1 ขั้นตอนหลักในการทำเหมืองข้อมูล

ทั้งนี้ขั้นตอนหลักในการทำเหมืองข้อมูลมี 5 ขั้นตอนดังนี้

1. การวิเคราะห์ความต้องการของธุรกิจ โดยการวิเคราะห์ความต้องการธุรกิจจะช่วยให้เข้าใจถึงประเด็นที่ผู้บริหารต้องตัดสินใจ ซึ่งเกี่ยวกับการสร้างความสำเร็จของธุรกิจ การวิเคราะห์ความต้องการของธุรกิจจะทำให้เกิดความเข้าใจถึงสถานะในปัจจุบันของธุรกิจ และทำให้ผู้บริหารสามารถกำหนดเรื่องที่ต้องตัดสินใจได้ดียิ่งขึ้น

2. การวิเคราะห์ความต้องการข้อมูล เนื่องจากคลังข้อมูลขนาดใหญ่ที่มีอยู่นั้น มีข้อมูลที่หลากหลาย ซึ่งมีทั้งข้อมูลที่จำเป็นต้องใช้ในการทำเหมืองข้อมูลกับข้อมูลอื่นซึ่งไม่เป็นที่ต้องการในขณะนี้ จึงต้องมีขั้นตอนการกำหนดรายการและประเภทของข้อมูลที่จะนำมาใช้ โดยมีการตรวจสอบในด้านของคุณภาพของข้อมูล จำนวน ปริมาณเนื้อหาและการเข้าถึงข้อมูล เพื่อกำหนดเป็นข้อมูลที่ต้องการทำเหมืองในกรณีที่มีข้อมูลจำนวนมาก อาจใช้การเลือกตัวอย่างข้อมูลมาทำเหมืองก่อนได้เพื่อลดค่าใช้จ่าย

3. การแปลงข้อมูล เมื่อกำหนดข้อมูลที่จะใช้ในการทำเหมืองได้แล้ว อาจต้องทำการแปลงหรือปรับเปลี่ยนข้อมูล ให้อยู่ในรูปแบบที่เหมาะสมกับวิธีการที่จะใช้ในการทำเหมืองข้อมูลซึ่งโดยปกติแล้วการแปลงข้อมูลจะถูกกำหนดโดยเงื่อนไขของการปฏิบัติงานและวิธีการทำเหมืองข้อมูล

4. การดำเนินการทำเหมืองข้อมูล เป็นขั้นตอนที่นำเอาวิธีการหรือเทคนิคการทำเหมืองข้อมูลตั้งแต่หนึ่งวิธีขึ้นไป มาทำการสกัดสาระสำคัญออกจากฐานข้อมูลที่มี เช่นการตอบคำถามว่าลูกค้าจะซื้อสินค้าต่อไปหรือไม่ อาจต้องทำการวิเคราะห์ตั้งแต่การรวมกลุ่มของลูกค้าและการจำแนกหน่วยหรือลูกค้าแต่ละคนว่าจะซื้อหรือไม่ซื้อสินค้าต่อไป ทั้งนี้ ในขณะที่ทำเหมืองข้อมูลอาจมีความจำเป็นต้องเข้าถึงข้อมูลอื่นในคลังข้อมูล รวมทั้งต้องแปลงข้อมูลรายการอื่นด้วยก็ได้

5. การแปลผล หรือการประยุกต์ใช้กับธุรกิจ เป็นขั้นตอนที่นำเอาสารสนเทศที่ทำเหมืองได้มาวิเคราะห์เพื่อตอบคำถามที่ผู้ตัดสินใจมีอยู่ ซึ่งการวิเคราะห์ในส่วนนี้จะครอบคลุมการกรองสารสนเทศที่เหมาะสมกับการส่งให้ผู้ใช้ และการแปลผล เช่น ถ้าวัตถุประสงค์ของการทำเหมืองข้อมูลคือการพัฒนาตัวแบบการจำแนกหน่วย ในขั้นตอนการแปลผล ก็จะต้องพิจารณาความเชื่อถือได้ของตัวแบบที่ได้ด้วยวิธีเช่น Cross-Validation เป็นต้น ถ้าผลที่ได้ไม่เป็นที่พอใจ ก็จะต้องทำขั้นตอนนี้ซ้ำอีก รวมทั้งขั้นตอนก่อนหน้าด้วย การแปลผลนี้อาจมองได้เป็นการประยุกต์วิทยาศาสตร์และเทคโนโลยีที่มีในการทำเหมืองข้อมูล ให้เป็นผลทางธุรกิจ ทำให้สามารถทำการประเมินผลที่ได้จากการทำเหมืองข้อมูล

2.6.2 การปฏิบัติการในการทำเหมืองข้อมูล (Data Mining Operations)

กิจกรรมหลักของการทำเหมืองข้อมูลคือ การสกัดข้อความรู้หรือรูปแบบที่มีอยู่ในฐานข้อมูลเพื่อสร้างสารสนเทศที่เป็นประโยชน์ต่อการตัดสินใจ การทำเหมืองข้อมูลโดยทั่วไปจะมีการปฏิบัติการ 4 ด้านด้วยกัน ได้แก่

1. การพัฒนาตัวแบบการพยากรณ์และการจำแนก เป็นการปฏิบัติการที่ใช้สาระที่มีอยู่ในฐานข้อมูลซึ่งเป็นข้อมูลที่เกิดขึ้นแล้ว มาพัฒนาตัวแบบเพื่อทำนายสิ่งที่จะเกิดในอนาคต วิธีการพัฒนาตัวแบบจะใช้วิธีการวิเคราะห์เชิงสถิติเป็นพื้นฐานและเสริมด้วยวิธีปฏิบัติของการทำเหมืองข้อมูลเพื่อให้ตัวแบบที่หาได้ง่ายต่อการทำความเข้าใจมากขึ้น ด้วยการใช้วิธีการเสนอตัวแบบในลักษณะการให้ทางเลือก การพัฒนาตัวแบบการพยากรณ์และการจำแนกนี้ อาจใช้วิธีการจัดหน่วยเข้ากลุ่ม (Classification) ซึ่งเป็นวิธีการจำแนกหน่วยว่าเป็นสมาชิกของกลุ่มใดโดยมีการกำหนดกลุ่มไว้ล่วงหน้าแล้ว หรือการวิเคราะห์ความถดถอย (Regression) ซึ่งเป็นวิธีการหารูปแบบความสัมพันธ์ระหว่างตัวแปรต่างๆ เพื่อใช้ในการพยากรณ์หรืออธิบายตัวแปรที่สนใจ หรือการพยากรณ์ (Prediction) ซึ่งเป็นการกำหนดค่าที่คาดว่าจะเกิดขึ้นของตัวแปรที่สนใจโดยอาศัยรูปแบบที่ได้จากข้อมูลที่มีอยู่เดิม

2. การวิเคราะห์ความเชื่อมโยงหรือความสัมพันธ์ เป็นการปฏิบัติการเพื่อสร้างความสัมพันธ์ระหว่างหน่วยต่าง ๆ ในฐานข้อมูล เช่น ผู้จัดการฝ่ายผลิตภัณฑ์ต้องการทราบว่าควรจัดสินค้าใดคู่กับสินค้าใดเพื่อการสั่งซื้อสินค้าเข้าร้าน เป็นต้น

3. การแบ่งส่วนฐานข้อมูล เมื่อฐานข้อมูลมีขนาดใหญ่ขึ้นๆ และมีความหลากหลายมากขึ้น จะมีความจำเป็นต้องแบ่งฐานข้อมูลออกเป็นส่วนๆ โดยข้อมูลในแต่ละส่วนมีความเกี่ยวพันกันอย่างเหมาะสม ก่อนที่จะใช้วิธีปฏิบัติการอื่นต่อไป ทั้งนี้ เพื่อความสะดวกในการทำเหมืองข้อมูล รวมทั้งทำให้การสรุปสาระในแต่ละส่วนมีความหมายและใช้ประโยชน์ได้ตรงกับความต้องการ การแบ่งส่วนนี้อาจใช้การรวมกลุ่ม (Clustering) ซึ่งเป็นวิธีการกำหนดกลุ่มหรือประเภทที่มีความแตกต่างกันจำนวนหนึ่ง

4. การค้นพบหน่วยผิดปกติ เป็นการปฏิบัติการเพื่อค้นพบว่ามีหน่วยที่แตกต่างจากหน่วยอื่นๆ ในฐานข้อมูลเป็นอย่างมาก หรือมีหน่วยผิดปกติหรือไม่ และสาเหตุของการเกิดหน่วยเช่นนั้นมาจากความผิดพลาดในการบันทึกข้อมูล หรือความผิดพลาดอื่นหรือเป็นลักษณะผิดปกติจริง การค้นพบหน่วยผิดปกติและจัดการกับหน่วยดังกล่าวอย่างเหมาะสมในการทำเหมืองข้อมูลจะทำให้ข้อความรู้ที่ได้มีคุณค่ามากขึ้น และการใช้ประโยชน์เป็นไปอย่างมีประสิทธิผลยิ่งขึ้น

2.6.3 เครื่องมือที่ใช้ในการทำเหมืองข้อมูล

การทำเหมืองข้อมูลโดยทั่วไปจะประกอบด้วยองค์ประกอบหลัก 3 ส่วน ได้แก่ การพัฒนาตัวแบบ การประเมินความเหมาะสมของตัวแบบ และวิธีการค้นหาสาระในฐานข้อมูล ตัวแบบต้องแสดงถึงข้อสมมติฐานได้อย่างชัดเจนเพื่อให้สามารถค้นพบรูปแบบที่มีอยู่ได้ ตัวแบบควรต้องมีความเหมาะสมในการพยากรณ์และตรวจสอบความเหมาะสมได้ และกระบวนการค้นหาควรทำให้เกณฑ์การประเมินความเหมาะสมของตัวแบบมีความถูกต้องมากที่สุด เครื่องมือในการทำเหมืองข้อมูล มักจะเป็นโปรแกรมอัจฉริยะที่เป็นระบบอัตโนมัติ โดยรวมปัญญาประดิษฐ์ในฐานข้อมูลเชิงสัมพันธ์ โปรแกรมเหล่านี้จะค้นพบรูปแบบที่ถูกกำหนดไว้ล่วงหน้าและเตือนให้ผู้ใช้ทราบถึงการเปลี่ยนแปลง

ทั้งนี้ เครื่องมือที่นำมาใช้ในการทำเหมืองข้อมูลมีหลายประเภท เทคนิคที่ใช้มากที่สุดในการทำเหมืองข้อมูลได้แก่

1. โครงข่ายประสาทประดิษฐ์ (Artificial Neural Networks) เป็นตัวแบบในการทำนายแบบไม่ใช่เชิงเส้นตรง ซึ่งหาได้จากการพัฒนาตัวแบบและลอกเลียนเครือข่ายประสาทเชิงชีวภาพ

2. การใช้รูปแบบการตัดสินใจแบบต้นไม้ (Decision Trees) ซึ่งจะแสดงทางเลือกในการตัดสินใจ ในรูปแบบของต้นไม้ ซึ่งทางเลือกเหล่านี้จะกำหนดกฎเกณฑ์ในการจำแนกข้อมูล ตัวอย่างของวิธีการคือ Classification and Regression Trees (CART) และ Chi-Square Automatic Interaction Detection (CHAID)

3. กระบวนการเชิงพันธุกรรม (Genetic Algorithms) เป็นวิธีที่ใช้กระบวนการเช่น การรวมการกลายพันธุ์ และการคัดเลือกโดยธรรมชาติ ในขั้นตอนการออกแบบที่ตั้งอยู่บนพื้นฐานของวิวัฒนาการ

4. วิธีรวมกลุ่มสมาชิกใกล้กัน (Nearest Neighbor) เป็นเทคนิคที่จำแนกข้อมูลแต่ละหน่วยในชุดข้อมูล โดยการรวมหน่วยที่คล้ายกันมากที่สุดเข้าเป็นกลุ่มเดียวกัน ซึ่งบางครั้งเรียกเทคนิคนี้ว่า K-Nearest Neighbor

5. วิธีการจำแนกหน่วยในฐานข้อมูล (Rule Induction) เป็นกระบวนการที่ทำการพัฒนาตัวแบบจำแนกหน่วยในฐานข้อมูลออกเป็นกลุ่มๆ โดยตัวแบบที่ได้จะมีรูปแบบที่มีประโยชน์

ในการแยกกลุ่มซึ่งอาจใช้วิธีการ Neural Induction หรือ Symbolic เช่น Decision Trees หรือใช้กฎของ If-Then ในการแบ่งส่วนฐานข้อมูลออกเป็นกลุ่มโดยใช้เกณฑ์ที่มีนัยสำคัญเชิงสถิติ

6. การสร้างภาพข้อมูล (Data Visualization) เป็นการสรุปผลเกี่ยวกับลักษณะสำคัญของข้อมูลออกมาเป็นภาพ ซึ่งในหลายกรณีจะแสดงสาระสำคัญในฐานข้อมูลได้ดีกว่าการพิจารณาข้อมูลเป็นตารางตัวเลข โดยเฉพาะเมื่อพิจารณาลักษณะที่ปรากฏอยู่ในข้อมูลเพียงส่วนเล็กๆ ส่วนหนึ่งเท่านั้น ซึ่งหากใช้วิธีการวิเคราะห์ข้อมูลเชิงสถิติกับข้อมูลทั้งหมดจะไม่สามารถทำให้สังเกตรูปแบบของข้อมูลในส่วนนี้ได้ เนื่องจากปริมาณของข้อมูลนั่นเอง ข้อดีอีกประการหนึ่งของการสร้างภาพในการทำเหมืองข้อมูลก็คือ ผู้วิเคราะห์ไม่จำเป็นต้องทราบลักษณะหรือประเภทของเหตุการณ์ที่ต้องการค้นหาเพื่อที่จะพบสิ่งที่แปลกแยกออกไป

2.6.4 เทคนิคในการทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) คือ รูปแบบการค้นหาเทคนิคในการค้นหารูปแบบที่ผิดปกติ (Regularities) ในชุดข้อมูลขนาดใหญ่ มีเทคนิคต่าง ๆ ในการทำเหมืองข้อมูล เช่น การสรุปความถี่ข้อมูลด้วยสถิติ การแสดงภาพประกอบในการนำเสนอข้อสรุปแบบกราฟิกของข้อมูล การค้นพบกฎความสัมพันธ์เพื่อกำหนดกิจกรรมปกติ และการค้นพบของความผิดปกติ (Anomalies) และการทำนายประเภท ดังนั้นการทำเหมืองข้อมูลจึงเป็นการช่วยในการแยกข้อมูลที่มีความหมายและความรู้จากข้อมูลที่ไม่มีโครงสร้าง

2.6.4.1 เทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) เป็นเทคนิคหนึ่งของการเรียนรู้ของเครื่อง (Machine Learning) โดยการพัฒนาตัวแบบที่เหมาะสมกับข้อมูล การเรียนรู้แบบนี้แตกต่างจากการเรียนรู้แบบมีผู้สอน (Supervised Learning) คือ จะไม่มีการระบุผลที่ต้องการหรือประเภทไว้ก่อน การเรียนรู้แบบนี้จะพิจารณาวัตถุเป็นเซตของตัวแปรสุ่ม แล้วจึงพัฒนาตัวแบบความหนาแน่นร่วมของชุดข้อมูลเช่น เทคนิคการรวมกลุ่ม (Cluster Analysis)

การเรียนรู้แบบไม่มีผู้สอนสามารถนำไปใช้ร่วมกับการอนุมานแบบเบย์ เพื่อหาความน่าจะเป็นแบบมีเงื่อนไขของตัวแปรสุ่มโดยกำหนดตัวแปรที่เกี่ยวข้องให้ นอกจากนี้ยังสามารถนำไปใช้ในการบีบอัดข้อมูล ซึ่งโดยพื้นฐานแล้ว ขั้นตอนวิธีการบีบอัดข้อมูลจะขึ้นอยู่กับการแจกแจงความน่าจะเป็นของข้อมูลไม่อย่างชัดแจ้ง

1. เทคนิคการรวมกลุ่ม (Clustering Analysis)

เทคนิคการรวมกลุ่ม (Clustering Analysis) เป็นเทคนิคทั่วไปสำหรับการวิเคราะห์ข้อมูลทางสถิติ รวมถึงการเรียนรู้ของเครื่องจักร (Machine Learning) การทำเหมืองข้อมูล (Data Mining) เป็นเทคนิคที่มีประโยชน์สำหรับตรวจจับการบุกรุกของเครือข่ายสังคมออนไลน์ ซึ่งถือว่าเป็นกิจกรรมที่เป็นอันตราย เงื่อนไขของการจัดประเภทของวัตถุที่คล้ายคลึงกันเป็นกลุ่มต่าง ๆ โดยพิจารณาระหว่างวัดความเหมือน (Similarity Measure) และวัดระยะทาง (Distance Measure)

เทคนิคการรวมกลุ่ม (Clustering Analysis) ที่นิยมมากมีอยู่ 2 ชนิดคือ การรวมกลุ่มข้อมูลแบบลำดับชั้น (Hierarchical Cluster Analysis) และการรวมกลุ่มข้อมูลแบบไม่ใช่ลำดับชั้น (Non Hierarchical Cluster Analysis) โดยวิธีการรวมกลุ่มข้อมูลแบบลำดับชั้น (Hierarchical Cluster Analysis) แบ่งย่อยเป็น Agglomerative Method และ Division Method และการรวมกลุ่มข้อมูลแบบไม่ใช่ลำดับชั้น (Non Hierarchical Cluster Analysis) แบ่งย่อยเป็น K-Means Clustering Method

2. การวัดระยะทาง (Distance Measure)

เพื่อที่จะอธิบายถึงความคล้ายกันของสองคุณลักษณะเวกเตอร์ การวัดระยะทาง (Distance Measure) ถูกใช้เป็นฟังก์ชันว่าชุดของจุดอยู่ใกล้พอที่จะถือว่าคลัสเตอร์เดียวกัน เราใช้วัดระยะทาง D (x, y) ที่บอกระยะห่างแค่จุด x และ y ใดๆ โดยระยะห่างระหว่างจุดสองจุดใดๆ ดังสมการที่ 2.1

กำหนดให้ x= $[x_1, x_2, ..., x_k]$ และ y = $[y_1, y_2, ..., y_k]$

จะได้ว่า $\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$ …(2.1)

3. K-Means Algorithm

K-means Clustering คือ วิธีการวิเคราะห์การรวมกลุ่มซึ่งมุ่งหวังที่จะสังเกต n พาร์ติชันเป็นคลัสเตอร์ k โดยค่าสังเกตการณ์ของแต่ละเป็นสมาชิกของคลัสเตอร์จะมีค่าเฉลี่ยที่ใกล้ที่สุด และเป็นวิธีหนึ่งหนึ่งของอัลกอริทึมการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) การเรียนรู้ที่ง่ายที่สุดที่แก้ปัญหาออกเป็นจินตภาพรู้จักกันดีคือ K-Means Algorithm

2.6.4.2 เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning)

การเรียนรู้แบบมีผู้สอน (Supervised Learning) เป็นเทคนิคหนึ่งของการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งสร้างฟังก์ชันจากข้อมูลสอน (Training Data) โดยข้อมูลสอนประกอบด้วยวัตถุเข้า (Vector) และผลที่ต้องการ ผลจากการเรียนรู้จะเป็นฟังก์ชันที่อาจจะให้ค่าต่อเนื่อง (จะเรียกวิธีการว่า การถดถอย หรือ Regression) หรือ ใช้ทำนายประเภทของวัตถุ (เรียกว่าการแบ่งประเภท หรือ Classification) ภารกิจของเครื่องเรียนรู้แบบมีผู้สอนคือการทำนายค่าของฟังก์ชันจากวัตถุเข้าที่ถูกต้องโดยใช้ตัวอย่างสอนจำนวนน้อย (Training Examples หรือคู่ของข้อมูลเข้าและผลที่เป็นเป้าหมาย) โดยเครื่องเรียนรู้จะต้องวางนัยทั่วไป (Generalize) จากข้อมูลที่มีอยู่ไปยังกรณีที่ไม่เคยพบอย่างมีเหตุผล

เทคนิคในการเรียนรู้แบบมีผู้สอน (Supervised Learning) จำแนกได้ดังนี้

1. เทคนิคการจำแนกประเภท (Classification Analysis)

เป็นเทคนิควิธีการทางสถิติในการตรวจสอบกลุ่มย่อยของตัวแปรอิสระแต่ละตัวว่ามีผลอย่างไรต่อตัวแปรตาม และตรวจสอบลักษณะผสม (Combinations) ของตัวแปรอิสระหลายๆ ตัว ว่ามีผลอย่างไรต่อตัวแปรตาม โดยที่ตัวแปรตามเป็นข้อมูลที่อยู่ในมาตราช่วงหรือมาตราอัตราส่วน และตัวแปรอิสระเป็นข้อมูลในมาตรานามบัญญัติหรือมาตราอันดับ

2. เทคนิคการสร้างกฎความสัมพันธ์ (Association Rules)

กฎความสัมพันธ์เป็นที่นิยม และเหมาะสมสำหรับการค้นหาระหว่างตัวแปรในฐานข้อมูลขนาดใหญ่ โดยกฎความสัมพันธ์จะต้องเป็นไปตามการสนับสนุนต่ำสุด (User-Specified Minimum Support) ที่ระบุผู้ใช้และการระบุความมั่นใจของผู้ใช้ขั้นต่ำ (User-Specified Minimum Confidence) ในเวลาเดียวกัน ในการสร้างกฎความสัมพันธ์นี้ จะประกอบด้วยกระบวนการที่มีสองขั้นตอน ขั้นแรก นำค่าสนับสนุนต่ำสุด (Minimum Support) ที่กำหนดไว้นำไปใช้ในการค้นหา Item Sets ทั้งหมดในฐานข้อมูล ในขั้นตอนที่สอง Item Sets ที่ค้นหาได้เหล่านี้จะต้องค้นหาความถี่ตามความเชื่อมั่นที่ต่ำสุด (Minimum Confidence) ที่ใช้กับการสร้างกฎความสัมพันธ์ (Association Rules)

มีอัลกอริทึมอยู่หลายอัลกอริทึมสำหรับการสร้างกฎความสัมพันธ์ บางอัลกอริทึมที่รู้จักกันดีคือ Apriori, Eclat และ FP-Growth

กำหนดให้ รายการข้อมูลตั้งค่า X = {X 1, X 2, …, X_m } ดังนั้นกฎความสัมพันธ์ (Association Rules) ที่สามารถแสดงเป็นรูปแบบของ “X ⟶Y” ถ้ากฎ X ⟶Y เป็นจริง แล้วก็มีการสนับสนุนและค่าความเชื่อมั่น (Support and Confidence Values) สามารถอธิบายได้เป็นดังสมการที่ 2.2 และ 2.3

$$\text{Support}(X \longrightarrow Y) = P(X \cup Y) \quad \ldots (2.2)$$

$$\text{Confidence}(X \longrightarrow Y) = P(Y/X) \quad \ldots (2.3)$$

3. อัลกอริทึม Apriori (Apriori Algorithm)

ทั้งนี้ Apriori เป็นอัลกอริทึมที่รู้จักกันเป็นอย่างดีในเหมืองของกฎความสัมพันธ์ เนื่องจากสามารถนำมาใช้เป็นกลยุทธ์ต้นๆ ที่ใช้ในการค้นหากฎความสัมพันธ์จากการ Mining ความถี่ของ Item Sets โดย Apriori จะใช้การทำซ้ำ มีความใกล้เคียงกับ Level-wise Search ซึ่ง K-Item จะใช้สร้าง (k+1) Item ครั้งแรกจะเป็นเซตของความถี่ของ 1 Item Set ที่พบ

เซตจะแสดงโดย L_1 และ L_1 จะใช้หา L_2 เป็นเซตของ 2 Item Sets และใช้หา L_3 ไปจนกระทั่งไม่มีความถี่ K-Item ที่สามารถหาพบได้ ในการหา L_k แต่ละครั้งจะต้องสแกนทั้งฐานข้อมูล Apriori Property คือจะลด Search Space ลงโดยใช้คุณสมบัติที่เรียกว่า Anti-Monotone (ลดการซ้ำกันในความรู้สึก) นั่นคือถ้าเซตที่ไม่ผ่านการทดสอบ ซุปเปอร์เซตทั้งหมดก็จะไม่ผ่านด้วยเช่นกัน

ดังนั้น จะได้กระบวนการที่ L_{k-1} ใช้หา L_k เป็น 2 ขั้นตอน คือ การ Join และการ Prune การสร้างกฎความสัมพันธ์จากความถี่ Item Set คือ การผ่านทั้ง Minimum Support และ Minimum Confidence ดังสมการที่ 2.4

$$\text{Confidence}(A \Rightarrow B) = P(B|A) = \text{Support Count}(A \cup B) / \text{Support Count}(A) \quad \ldots (2.4)$$

การแก้ปัญหาการเรียนรู้แบบมีผู้สอน (เช่น การเรียนรู้เพื่อรู้จำลายมือ) มีขั้นตอนที่ต้องพิจารณา ได้แก่

1. กำหนดชนิดของตัวอย่างสอน ก่อนจะเริ่มทำอย่างอื่น จะต้องตัดสินว่าข้อมูลชนิดใดที่จะใช้เป็นตัวอย่าง เช่นในกรณีการรู้จักลายมือ ตัวอย่างอาจจะเป็นตัวอักษรตัวเดียว คำหรือบรรทัด

2. เก็บตัวอย่าง ชุดตัวอย่างสอนจะต้องมีลักษณะเป็นตามที่ใช้จริง ดังนั้น ชุดข้อมูลตัวอย่างและผลที่สอดคล้องจะต้องถูกจัดเก็บจากผู้เชี่ยวชาญหรือจากการวัด

3. กำหนดวิธีการแทนลักษณะ (Feature) ของข้อมูลเข้า ความถูกต้องของฟังก์ชันจะขึ้นอยู่กับการแทนข้อมูลอย่างมาก โดยทั่วไปวัตถุเข้าจะถูกแปลงเป็นเวกเตอร์ของลักษณะใช้อธิบายวัตถุที่ต้องการแบ่งประเภท จำนวนลักษณะจะต้องไม่มากจนเกินไป เพราะจะทำให้เกิดเนื่องจากมิติที่กว้างเกินไปจนทำให้มีพื้นที่ว่างมากจนเคร่องเรียนรู้ไม่สามารถวางนัยทั่วไปได้ แต่จำนวนลักษณะก็จะต้องมากพอที่จะทำให้สามารถทำนายผลได้แม่นยำ

4. กำหนดโครงสร้างของฟังก์ชันที่ต้องการ และขั้นตอนวิธีการเรียนรู้ที่สอดคล้อง เช่น อาจจะต้องเลือกว่าจะใช้เครือข่ายงานประสาทเทียม หรือ ต้นไม้ตัดสินใจ

5. ทำการออกแบบให้สมบูรณ์ แล้วใช้ขั้นตอนวิธีการเรียนรู้กับตัวอย่างที่เก็บมา อาจจะปรับพารามิเตอร์ต่างๆ ของขั้นตอนวิธีให้เหมาะที่สุดโดยใช้ชุดย่อยของชุดตัวอย่างเรียกว่า ชุดตรวจสอบ (Validation Set) หรือ ใช้การตรวจสอบไขว้ (Cross-Validation) หลังจากปรับค่าต่างๆ แล้ว อาจจะวัดประสิทธิภาพของขั้นตอนวิธีโดยใช้ชุดทดสอบ (Test Set) ซึ่งแยกต่างหากจากชุดสอน

2.7 โครงสร้างกราฟ (Graph-based Structure)

เทคนิคการแสดงโครงสร้างกราฟ (Graph-based Structure) ที่ใช้ในการวิเคราะห์เครือข่ายทางสังคม (Social Network Analysis: SNA) จะขึ้นอยู่กับหลักการของทฤษฎีกราฟ (Graph Theorem) ซึ่งประกอบด้วยชุดของสูตรทางคณิตศาสตร์และแนวความคิดในการศึกษารูปแบบของเส้น บุคคลจะเป็นโหนดหรือจุดในกราฟ (Node) และความสัมพันธ์ระหว่างบุคคลหรือจุดจะเป็นเส้นกราฟ (Edge) ระหว่างโหนดหรือจุดนั้นๆ ของเครือข่ายสังคมออนไลน์ที่เรียกว่า Sociograms

ทั้งนี้ กราฟเป็นโครงสร้างที่นำมาใช้เพื่อแสดงความสัมพันธ์ระหว่างวัตถุ โดยแทนวัตถุด้วยเวอร์เท็กซ์ (Vertex) หรือโหนด (Node) และเชื่อมโยงความสัมพันธ์ด้วยเส้นเชื่อมหรือเอดจ์ (Edge) สามารถเขียนในรูปของสัญลักษณ์ได้เป็น $G = (V, E)$ โดยที่

$V(G)$ คือ เซตของเวอร์เทกซ์หรือโหนด ที่ไม่ใช่เซตว่าง และมีจำนวนจำกัด

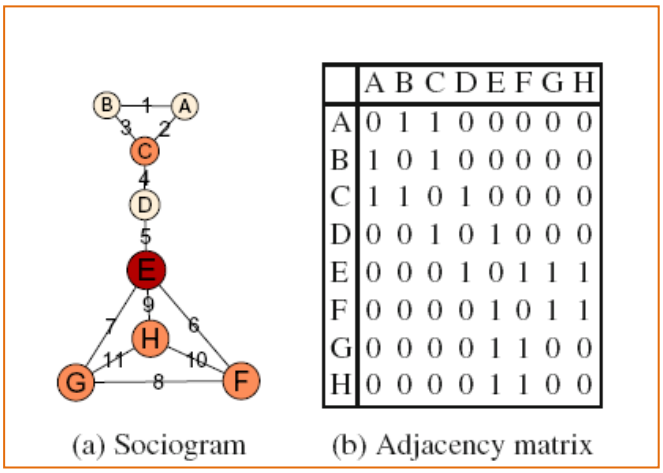
$E(G)$ คือ เซตของเส้นเชื่อม ซึ่งเขียนด้วยคู่ของเวอร์เท็กซ์หรือโหนด

โดยความเชื่อมโยงของเครือข่ายสังคมออนไลน์จะแสดงด้วยเมทริกซ์ประชิด (Adjacency Matrix) ซึ่งเป็นเมทริกซ์สมมาตรขนาด n x n โดยที่ n คือ จำนวนของโหนด (Node/Person) ในเครือข่ายสังคมออนไลน์ ทั้งนี้เมทริกซ์ประชิดจะประกอบด้วยสมาชิก A_{ij} ในแต่ละตำแหน่งจะแทนค่าด้วย ‘0’ และ ‘1’ กล่าวคือถ้าโหนด (Node/Person) ใดมีความเชื่อมโยงกัน จะกำหนดค่าสมาชิกเป็น ‘1’ แต่ถ้าโหนด (Node/Person) ใดที่ไม่มีความเชื่อมโยงกัน จะกำหนดค่าสมาชิกเป็น ‘0’ โดยแนวทางการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) มีรายละเอียด ดังนี้

2.7.1 ความเชื่อมโยงของเครือข่ายสังคมออนไลน์

ความเชื่อมโยงของเครือข่ายสังคมออนไลน์จะแสดงด้วยเมทริกซ์ประชิด (Adjacency Matrix) ซึ่งเป็นเมทริกซ์สมมาตรขนาด n x n โดยที่ n คือ จำนวนของโหนด (Node/Person) ในเครือข่ายสังคมออนไลน์ และสมาชิกเมทริกซ์ประชิด A มีค่าเป็นดังแสดงในสมการที่ 2.5 และรูปที่ 2.5

$$A_{ij} = \begin{cases} 1 & \text{ถ้าโหนดที่ } i \text{ และ } j \text{ ต่อเชื่อมกัน} \\ 0 & \text{อื่นๆ} \end{cases} \quad \ldots (2.5)$$



รูปที่ 2.5 ภาพแสดงผังสังคมมิติ (Sociogram) และเมทริกซ์ประชิด (Adjacency Matrix)

2.7.2 ความเป็นศูนย์กลาง (Degree Centrality: *D*)

ซึ่งคือจำนวนของ Node/Person ที่มีการติดต่อกับ Node/Person ใดๆ โดยที่ Degree D_i ของโหนด i แสดงได้ดังแสดงในสมการที่ 2.6

$$D_i = \sum_{j=1}^{n} A_{ij} \quad \ldots (2.6)$$

2.7.3 ความเป็นการเชื่อมโยง (Betweenness Centrality: *B*) แสดง Node อยู่ใน Shortest Path ระหว่าง Node คู่อื่นๆ ทุก Nodes ทั้งนี้บุคคลที่มี Betweenness Centrality สูงสุดคือบุคคลที่บุคคลอื่นจะผ่านเข้าถึงเครือข่ายมากที่สุด

กำหนดให้ B_a คือจำนวนของ Geodesics (Shortest Paths ระหว่างสองโหนดใดๆ) และ $g_{ij}(a)$ คือค่า Shortest Path ระหว่าง 2 โหนด *i* และ *j* ใดๆ ที่ผ่านโหนด a จะได้ดังแสดงในสมการที่ 2.7

$$B_a = \sum_{j}^{n} \sum_{i}^{n} g_{ij}\,(a) \qquad \ldots (2.7)$$

2.7.4 ความเป็นศูนย์รวมความใกล้ชิด (Closeness Centrality: *C*) คือจำนวนรวมของความเชื่อมโยง (Links) ที่บุคคลต้องผ่านเพื่อเข้าถึงบุคคลอื่นๆ ในเครือข่าย บุคคลที่มี Closeness Centrality สูงสุดคือบุคคลที่จะผ่านเข้าถึงเครือข่ายด้วยการเชื่อมโยงที่น้อยที่สุด

กำหนดให้ C_a คือผลรวมของ Length ของ Geodesics ระหว่างโหนดที่กำหนด (Particular Node: *a)* และ โหนดใดๆ ที่อยู่ในเครือข่าย จะได้ $l(i, a)$ คือ Length ของ Shortest Path ที่เชื่อมต่อระหว่าง nodes *i* และ *a*: เป็นดังแสดงในสมการที่ 2.8

$$C_a = \sum_{i=1}^{n} l\,(i, a) \qquad \ldots (2.8)$$

2.7.5 โหนดศูนย์กลางการรับข้อมูลข่าวสาร (In-Degree Nodes: deg_{in}) แสดงจำนวนเส้นเชื่อมที่เข้ามายังโหนดนั้นๆ ในเครือข่ายสังคมออนไลน์ โหนดศูนย์กลางการรับข้อมูลข่าวสาร (In-Degree Nodes) หมายถึงบุคคลในเครือข่ายสังคมออนไลน์ที่มีการติดต่อกับบุคคลอื่นในเครือข่ายสังคมออนไลน์ ซึ่งเป็นบุคคลที่เป็นศูนย์กลางของการรับข้อมูลข่าวสารหรือผลการกระทำผิดปกติและการโจมตีในเครือข่ายสังคม ดังแสดงในสมการที่ 2.9

$$deg_{in}(v_i) = \sum_{j=1}^{|V|} V_{ji} \qquad \ldots (2.9)$$

2.7.6 โหนดศูนย์กลางการกระจายข้อมูลข่าวสาร (Out-Degree Nodes: $OutDg$) แสดงจำนวนเส้นเชื่อมที่ออกจากโหนดนั้นไป โดยในเครือข่ายสังคมออนไลน์ โหนดศูนย์กลางการกระจายข้อมูลข่าวสาร (Out-Degree Nodes) หมายถึงบุคคลในเครือข่ายสังคมออนไลน์ที่มีการติดต่อ

กับบุคคลอื่นในเครือข่ายสังคมออนไลน์ ซึ่งเป็นบุคคลที่เป็นศูนย์กลางของการกระจายข้อมูลข่าวสาร หรือผลการกระทำผิดปกติและการโจมตีในเครือข่ายสังคมออนไลน์ ดังแสดงในสมการที่ 2.10

$$deg_{out}(v_i) = \sum_{j=1}^{|V|} V_{ij} \qquad \ldots(2.10)$$

2.7.8 การแพร่กระจายข้อมูลข่าวสาร (Information Diffusion)

ตัวแบบการกระจ่ายข้อมูลข่าวสาร (Models of Information Diffusion in Social Networks) มี 3 ชนิด คือ

1. **The Contagion Model** โดยใช้แนวความคิดจากการแพร่กระจายของโรคระบาด **(**Contagious Diseases Spread) ในกลุ่มของประชาชน

กำหนดให้ $f(t)$ คือ Fraction ของโหนดในเครือข่าย

λ_{int} คือ Internal Rate ของการรับข้อมูลข่าวสาร

λ_{ext} คือ External Rate ของการแพร่กระจายข้อมูลข่าวสาร

ดังนั้นการแพร่กระจายข้อมูลข่าวสารของตัวแบบนี้ ณ เวลา t จะเป็นดังสมการอนุพันธ์ที่ 2.11 และ 2.12

$$\frac{df(t)}{dt} = (\lambda_{int} f(t) + \lambda_{ext})\big(1 - f(t)\big) \qquad \ldots (2.11)$$

$$f(t) = \Big(1 - c\lambda_{ext} e^{-(\lambda_{int}+\lambda_{ext})t}\Big)\Big(1 + c\lambda_{int} e^{-(\lambda_{int}+\lambda_{ext})t}\Big) \qquad \ldots(2.12)$$

2. **The Social Influence Model** โดยใช้แนวคิดว่าข้อมูลข่าวสารแพร่ไปในเฉพาะกลุ่มผู้อยู่ใกล้เคียงในเครือข่าย (Neighbors) จึงแตกต่างจาก The Contagion Model โดยตัวแบบนี้มีสมมติฐานที่ว่าโหนดที่มีอิทธิพลจะไม่รับข้อมูลข่าวสารใดๆ เว้นแต่จำนวนผู้อยู่ใกล้เคียงในเครือข่ายของโหนดที่มีอิทธิพลเข้ามามีส่วนเกี่ยวข้องกับข้อมูลข่าวสารนั้นมากเกินค่าที่กำหนดไว้ (Certain Threshold)

3. **The Social Learning Model** ตัวแบบนี้มีความแตกต่างจาก 2 ตัวแบบดังกล่าวข้างต้น โดยมีแนวคิดว่าผู้อยู่ใกล้เคียงในเครือข่ายของโหนดที่มีอิทธิพลจะใช้การตัดสินใจในการร่วมเข้ามาเกี่ยวข้องในการรับหรือกระจายข้อมูลข่าวสาร **(**Rational Decision Makers**)**

กำหนดให้ $p(0)$ คือ ค่าเริ่มต้นของเวกเตอร์ความเชื่อ (Belief Vector)

n คือ จำนวนโหนดในเครือข่าย

T เป็น เมทริกซ์ n x n โดยที่ $T_{i,j}$ คือ ค่าความเชื่อระหว่าง nodes i และ nodes j

$T_{i,j}$ มีค่าเท่ากับ $1/d_i$ เมื่อ d_i คือจำนวนของเส้นเชื่อมที่ page i ไปยัง pages อื่นใด

ดังนั้นโหนดที่อยู่ใกล้เคียงในเครือข่ายของโหนดที่มีอิทธิพลจะเข้ามามีส่วนเกี่ยวข้องกับข้อมูลข่าวสารนั้น $p(n)$ ณ เวลาปัจจุบัน เป็นดังสมการที่ 2.13

$$p(n) = Tp(n-1) = T^n p(0) \quad \ldots (2.13)$$

2.7.9 การค้นหาแหล่งต้นทางการแพร่กระจายข้อมูลข่าวสาร (Locating the Source of Diffusion)

1. แหล่งต้นทางการแพร่กระจายข้อมูลข่าวสาร 1 แหล่ง

กำหนดให้ค่าประมาณ Maximum Likelihood Estimator (MLE) เป็นตัวกำหนดแหล่งต้นทางการแพร่กระจายข้อมูลข่าวสารของโหนดที่ได้รับอิทธิพล n โหนด

กำหนดให้ G_n คือกราฟที่มีการถูกโจมตี

$P(G_n|s)$ คือ ความน่าจะเป็นของกราฟที่มีการถูกโจมตี

S คือ เซตของโหนดที่มีการกระจายข้อมูลข่าวสาร

s คือ โหนดที่มีการกระจายข้อมูลข่าวสาร และ $s \in G_n$

จะได้ว่า แหล่งต้นทางการแพร่กระจายข้อมูลข่าวสาร (S^*) ดังสมการ 2.14

$$s^* = \arg\max_{s \in G_n} P(G_n|s) \quad \ldots (2.14)$$

กำหนดให้ $R(s, G_n)$ คือ จำนวนของ Paths ที่สามารถผ่านไปยังศูนย์กลางการกระจายข้อมูลข่าวสาร

$\tau_{bfs}(s)$ คือ ต้นไม้ BFS ที่เกี่ยวข้องกับต้นไม้ที่เป็นแหล่งกระจายข้อมูลข่าวสาร s

ในขณะที่การท่องไปในกราฟเป็นแบบ BFS (Breadth-First Search in Graph) จะได้ว่าเส้นทางสั้นที่สุดจากแหล่งต้นทางเดียวของการแพร่กระจายข้อมูลข่าวสาร (Single Source Shortest Path) เป็นดังสมการที่ 2.15

$$s^* = \arg\max_{s\epsilon\{S\cap G_n\}} \mathrm{R}\left(s, \tau_{bfs}(s)\right) \quad \ldots(2.15)$$

2. แหล่งต้นทางการแพร่กระจายข้อมูลข่าวสารหลายแห่ง

กำหนดให้ค่าประมาณ Maximum Likelihood (ML) เป็นตัวกำหนดแหล่งต้นทางการแพร่กระจายข้อมูลข่าวสารของโหนดที่ได้รับอิทธิพล n โหนด

กำหนดให้ G_n คือกราฟที่มีการถูกโจมตี

S คือ เซตของโหนดที่มีการกระจายข้อมูลข่าวสาร

s คือ โหนดที่มีการกระจายข้อมูลข่าวสาร

A คือ เซตของโหนดที่มีการกระจายข้อมูลข่าวสารร่วมกัน

$P(G_n|s)$ คือ ความน่าจะเป็นของกราฟที่มีการถูกโจมตี

A_n คือ ส่วนของกราฟที่มีการกระจายข้อมูลข่าวสารร่วมกัน

จะได้ว่า แหล่งต้นทางการแพร่กระจายข้อมูลข่าวสารแบบหลายแหล่ง (S^*) ดังสมการ 2.16

$$S^* = \arg\max_{S} P(G_n|S) \quad \ldots (2.16)$$

ในขณะที่การท่องไปในกราฟเป็นแบบ BFS (Breadth-First Search in Graph) จะได้ว่าการหาเส้นทางสั้นที่สุดจากแหล่งต้นหลายแหล่ง (All-Pairs Shortest Path) ของการแพร่กระจายข้อมูลข่าวสาร เป็นดังสมการที่ 2.17

$$A_n^*(S^*) = \arg\max_{A} P(A|S^*, G_n) \quad \ldots (2.17)$$

2.7.10 ทุนทางสังคม (Social Capital)

ทุนทางสังคม หมายถึงผลลัพธ์ทั้งหมดของศักยภาพ/ทรัพยากรที่ทำให้ Node/Person เพิ่มขึ้นจากผลการสร้างความสัมพันธ์ระหว่างบุคคล (Interpersonal Relationship) ซึ่งหมายถึงโหนด (Node/Person) ที่ได้รับการชื่นชอบ ยอมรับและได้รับการตอบสนองต่อการร้องขอใดๆ ได้อย่างรวดเร็ว และสามารถเข้าถึงโหนดที่มีอิทธิพล (Influence Node) ได้ ทั้งนี้ทุนทางสังคม (Social Capital) อาจเกิดขึ้นจากการที่มีเครือข่ายสังคมออนไลน์ในลักษณะการได้รับความสนับสนุนที่หนาแน่น หรือที่มีบทบาทพิเศษต่อบุคคลอื่น (Structure Holes)

2.7.11 หลุมโครงสร้าง (Structure Holes)

หลุมโครงสร้างจะเกิดขึ้นเมื่อโหนด (Node/Person) มีลักษณะดังต่อไปนี้

1. มีความสัมพันธ์กับบุคคลอื่นที่ติดต่อกับกลุ่มเครือข่ายย่อย (Cluster Network) อื่น

2. ตนเองต้องไม่มีการเชื่อมต่อทั้งโดยตรงและโดยอ้อมกับบุคคลอื่นในกลุ่มเครือข่ายย่อย (Cluster Network) นั้น ซึ่งหมายถึงบุคคลที่จะได้รับความหลากหลายของข้อมูล (Information) และโอกาส (Opportunities) มากกว่าบุคคลอื่นในเครือข่ายเดียวกัน ดังนั้นบุคคลที่มี Structure Holes จำนวนมากจะมีแนวโน้มในการได้รับการสนับสนุน (Promote)ให้ด้รับอำนาจและอิทธิพล โดยจะทำหน้าที่เป็น Brokerage ระหว่างกลุ่มอื่นๆ ต่อไป

2.7.12 นายหน้า (Brokerage)

นายหน้า คือโหนด (Node/Person) ที่เชื่อมโยงกับโหนด (Node/Person) ใดๆ ที่ไม่ได้ติดต่อกัน มีบทบาทในการจัดการการแพร่กระจายของข้อมูล (Information Flow) ทั้งนี้ Brokerage จะแตกต่างจาก Structure Holes ในประเด็นเหล่านี้

1. ไม่มีความเกี่ยวข้องในกลุ่ม
2. เป็นผู้ที่มีบทบาทต่อการถ่ายโอนหรือแพร่กระจายของข้อมูล

2.7.13 ความเป็นศูนย์กลาง (Centrality)

ความเป็นศูนย์กลาง หมายถึงโหนด (Node/Person) ที่เป็นเหมือนจุดศูนย์กลางของเครือข่าย (Network) ซึ่งจะเป็นโหนดที่มีอิทธิพลต่อเครือข่าย ทั้งนี้สิ่งที่ใช้วัดความเป็นศูนย์กลาง (Centrality) ส่วนใหญ่ คือ

1. ระดับของความเป็นศูนย์กลาง (Degree Centrality) คือจำนวนของโหนดที่มีการติดต่อกับโหนดใดๆ

2. การคั่นกลางของคนที่เป็นศูนย์กลาง (Betweenness Centrality) คือการแสดงขอบเขตที่โหนดอยู่ใน Shortest Path ระหว่างโหนดคู่อื่นๆ ทุกโหนด ทั้งนี้บุคคลที่มีการคั่นกลางของคนที่เป็นศูนย์กลาง สูงสุดคือบุคคลที่บุคคลอื่นจะผ่านเข้าถึงเครือข่ายมากที่สุด

3. ความใกล้ชิดของการเป็นศูนย์กลาง (Closeness Centrality) คือจำนวนของลิงค์ (Links) ที่บุคคลต้องผ่านเพื่อเข้าถึงบุคคลอื่นๆ ในเครือข่าย โดยบุคคลที่ความใกล้ชิดของการเป็นศูนย์กลาง สูงสุดคือบุคคลที่จะผ่านเข้าถึงเครือข่ายด้วยการเชื่อมโยงที่น้อยที่สุด

2.7.14 ความหนาแน่น (Density)

ความหนาแน่นจะบอกในเรื่องของความใกล้ชิดกันของกลุ่มย่อย และเป็นการคำนวณถึงสัดส่วนของจำนวนของความสัมพันธ์ที่แท้จริงในกลุ่มกับจำนวนความสัมพันธ์ที่เป็นไปได้ทั้งหมดในกลุ่ม โดยที่ความหนาแน่นสามารถคำนวณได้ทั้งภายในกลุ่ม และระหว่างกลุ่ม จึงสามารถบอกความใกล้ชิดของแต่ละโหนด (Node) ได้ด้วย

2.7.15 การติดต่อกัน (Cohesion)

การติดต่อกันมีอยู่หลายวิธีการ (Measures) แต่ที่นิยมคือ การวัดค่าเฉลี่ยของความสัมพันธ์ที่เข้าถึงคนอื่น หรือโหนด (Node) อื่นในกลุ่มเดียวกัน

2.7.16 การกำหนดกลุ่มย่อย (Subgroup Identification)

สามารถกำหนดจำนวนของความใกล้ชิดของเครือข่ายย่อย (Subgroup) จะทำให้เห็นโหนด (Node) ที่ซ้อนทับกันอยู่ (Overlapping Members) และถ้าเครือข่ายมีเครือข่ายย่อย (Subgroup) แยกกันอยู่สูง จะมีผลต่อการเครือข่ายโดยรวม (Integrated Network)

2.7.17 Minimum Spanning Tree ของ Prim's Algorithm

เป็นอัลกอริทึมในทฤษฎีบทเกียวกับกราฟ โดยจะทำการหา Minimum Spanning Tree จากน้ำหนักของเส้นเชื่อมต่อกราฟ หรือการหาเส้นเชื่อมระหว่างจุดที่มีค่าระยะทางที่น้อยที่สุด และไม่ทำให้เกิดลูปขึ้นในกราฟ อัลกอริทึมของพริมเป็นหนึ่งใน Greedy Algorithm

สมมติว่า $G = (V, E)$ คือกราฟแบบต่อเนื่อง ซึ่งแต่ละเส้นเชื่อมมีค่า Cost ประจำอยู่ Spanning Tree ของ G คือทรีอิสระที่เชื่อมต่อทุกๆ Vertices ใน V คือ Spanning Tree คือ ผลรวมของ Cost ของเส้นเชื่อมในทรีนั้น และในตอนนี้เราจะหาค่า Spanning Tree ที่มีค่าต่ำสุด

2.7.18 All-Pair Shortest Path ของ Floyd-Warshall Algorithm

เป็นขั้นตอนวิธีการวิเคราะห์กราฟเพื่อที่จะหาระยะทางของเส่นทางสั้นสุดในกราฟที่มีน้ำหนักของเส้นเชื่อมเป็นบวก หรือ น้ำหนักของเส้นเชื่อมเป็นลบ ก็ได้แต่ไม่สามารถหาได้

ถ้ามีวงจรลบ โดยการทำงานหนึ่งครั้งของขั้นตอนวิธีนี้จะได้คำตอบของระยะทางของเส้นทางสั้นสุดของทุกๆคู่ปมบนกราฟ อย่างไรก็ตามจะไม่สามารถคืนค่ารายละเอียดของเส้นทางสั้นสุดในแต่ละคู่ปมได้ ยกเว้นมีการเพิ่มเติมเข้าไป ขั้นตอนวิธีนี้เป็นตัวอย่างของกำหนดการพลวัตแบบด้านล่างขึ้นด้านบน โดยขั้นตอนวิธีนี้ถูกคิดขึ้นโดย โรเบิร์ต ฟลอยด์ ในปี 1962 อย่างไรก็ตามขั้นตอนวิธีนี้มีส่วนสำคัญเหมือนกับอัลกอริทึมของเบอร์นาร์ด รอยด์ในปี 1959 และของสตีเฟน วอร์แชล ในปี 1962 ในการค้นหา ความสัมพันธ์แบบถ่ายทอดของกราฟ

2.8 สถิติวิเคราะห์ (Statistic Analysis)

สถิติที่ใช้กันอยู่ในทางวิจัย แบ่งออกได้เป็น 2 ประเภทใหญ่ ๆ คือ

2.8.1 สถิติเชิงบรรยายหรือสถิติเชิงพรรณนา (Descriptive Statistics)

เป็นสถิติที่บรรยายคุณลักษณะของสิ่งที่ต้องการศึกษา จากกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ ซึ่งอาจจะเป็นกลุ่มเล็กหรือกลุ่มใหญ่ก็ได้ ผลที่ได้จากการศึกษาไม่สามารถนำไปอ้างอิงถึงกลุ่มประชากร (Population) ได้ สถิติที่ใช้ในการบรรยายคุณลักษณะของข้อมูล ได้แก่ ความถี่ (Frequency) ร้อยละ (Percentage) ค่าเฉลี่ย (Mean) มัธยฐาน (Median) พิสัย (Range) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

2.8.2 สถิติเชิงอ้างอิงหรือสถิติอนุมาน (Inferential Statistics)

เป็นสถิติที่ศึกษาข้อมูลจากกลุ่มตัวอย่าง (Sample) แล้วนำผลสรุปที่ได้จากกลุ่มตัวอย่าง สรุปอ้างอิงไปยังลักษณะประชากรหรือค่าสถิติ (Statistics) ที่ได้จากกลุ่มตัวอย่างสรุปไปยังค่าพารามิเตอร์ (Parameters) ของประชากร การได้มาซึ่งกลุ่มตัวอย่างมีความสำคัญยิ่งที่ใช้เป็นตัวแทนของประชากร โดยสถิติที่อ้างอิงจะเกี่ยวกับการประมาณค่า (Estimation) และการทดสอบสมมุติฐาน (Hypothesis Testing)

2.9 การแสดงภาพวิชวลไลเซชัน (Visualization)

การแสดงภาพวิชวลไลเซชัน (Visualization) คือ การนำเสนอสารสนเทศออกมาเป็นรูปภาพ โดยมีเป้าหมายหลักคือทำให้ผู้ชมเข้าใจเนื้อหาของสารสนเทศนั้นได้อย่างมีประสิทธิภาพ

ลักษณะของการแสดงภาพวิชวลไลเซชัน มีดังนี้

1. มีประสิทธิผล (Effective) ผู้ใช้สามารถเข้าใจสิ่งที่นำเสนอได้โดยง่าย

2. มีความแม่นยำ (Accurate) ข้อมูลที่นำเสนอไม่ควรมีความคลาดเคลื่อน ความแม่นยำเป็นสิ่งจำเป็นอย่างมากต่อผู้ใช้ในการประเมินค่าเชิงปริมาณของข้อมูล

3. มีประสิทธิภาพ (Efficient) ไม่นำเสนอให้ดูยุ่งจนเกินไป ลดจุดและเส้นที่ไม่จำเป็น

4. มีความสวยงาม (Aesthetics) ทำให้แสดงผลในรูปแบบที่สวยงามและทำความเข้าใจได้ง่าย

5. มีความยืดหยุ่น (Adaptable) ตอบสนองความต้องการของผู้ใช้ได้หลายรูปแบบ

2.10 งานวิจัยที่เกี่ยวข้อง

การศึกษาและวิจัยในครั้งนี้เป็นการวิเคราะห์หาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์ ศึกษารูปแบบการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ พัฒนาตัวแบบการสืบค้นผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ พร้อมทั้งพัฒนาตัวแบบการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ เพื่อที่จะค้นหากลยุทธ์การติดตามภัยจากการใช้เครือข่ายสังคมออนไลน์จากตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ ดังนั้นงานวิจัยชิ้นนี้จึงได้รวบรวมแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ต่อไปนี้

2.10.1 งานวิจัยของผู้วิจัยอื่นๆ

มีงานวิจัยอยู่จำนวนมากทำการศึกษาถึงลักษณะการใช้เครือข่าย (Network Behavior Analysis) [21, 32, 52, 90, 111] โดยผู้วิจัยได้ทำการพัฒนาตัวแบบ (Model) พฤติกรรมการใช้งานเครือข่าย ตัวแบบของ Network Forensics และผลกระทบต่อองค์ประกอบของเครือข่าย และสถาบันมาตรฐานและเทคโนโลยีแห่งชาติ (National Institute of Standards and Technology : NIST) [53] ได้พัฒนาคู่มือที่สอดคล้องกับข้อกฎหมายด้านการจัดการความมั่นคงปลอดภัยข้อมูลของรัฐบาลกลาง (Federal Information Security Management Act: FISMA) เพื่อช่วยให้องค์กรต่างๆ ได้เข้าใจถึงการบริหารระบบความปลอดภัยของเครือข่าย การติดตั้ง การดูแลระบบอย่างมีประสิทธิภาพ กระบวนการพัฒนาและประมวลผลเหตุการณ์เครือข่าย (Log Management) โดยผู้บริหารเครือข่ายศึกษาเหตุการณ์เครือข่าย (Logs) โดยใช้เครื่องมือในการวิเคราะห์หารูปแบบที่ผิดปกติ เช่นใช้ค่าความสัมพันธ์ (Correlation) ดูความสัมพันธ์ของเหตุการณ์เครือข่ายแต่ละประเภท

ทั้งนี้ ในประเทศไทยได้มีการกำหนดพระราชบัญญัติว่าด้วยการกระทำผิดเกี่ยวกับคอมพิวเตอร์ พ.ศ.2550 [2] ขึ้นเพื่อสร้างความเข้าใจในตัวบทกฎหมายแก่ประชาชนและผู้เกี่ยวข้อง ใช้เป็นคู่มือแนวทางในการปฏิบัติตามกฎหมายปกป้องสิทธิ โดยกระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร ซึ่งมีการกล่าวถึงลักษณะความผิดเกี่ยวกับคอมพิวเตอร์ประเภทต่างๆ ที่อาจนำภัยคุกคามอย่างอื่นแฝงมาด้วย เช่นไวรัส สปายแวร์ เป็นต้น จนส่งผลถึงขั้นทำให้ระบบคอมพิวเตอร์ไม่สามารถทำงานได้อีก โดยสามารถใช้เป็นเครื่องมือที่สำคัญในการกระทำความผิดทางคอมพิวเตอร์ต่อไป ดังนั้นเพื่อให้องค์กรสามารถใช้ระบบเทคโนโลยีสารสนเทศและการสื่อสารเป็นไปอย่างเหมาะสม มีประสิทธิภาพ มีความมั่นคงปลอดภัยและสามารถดำเนินงานได้อย่างต่อเนื่อง รวมทั้งป้องกันปัญหาที่อาจจะเกิดขึ้นจากการใช้งานในลักษณะที่ไม่เหมาะสม หรือถูกคุกคามจากภัยต่างๆ ซึ่งจะช่วยลดความเสียหายต่อการดำเนินงาน ทรัพย์สิน และบุคลากร จึงสมควรในการหาแนวทางเพื่อกำหนดนโยบายการรักษาความมั่นคงปลอดภัยของระบบเทคโนโลยีสารสนเทศขององค์กรให้เหมาะสมที่สุด

งานวิจัยจำนวนหนึ่งที่ทำการศึกษาเกี่ยวกับกรอบการวิจัยและวิธีดำเนินการวิจัยโดยการใช้เทคนิคเหมืองข้อมูล (Data Mining) เพื่อปรับปรุงประสิทธิภาพของระบบตรวจจับการบุกรุก (Intrusion Detection System) [28, 35, 45, 50, 82, 89] หรือเพื่อตรวจจับความผิดปกติของการบุกรุก [40, 63, 87] ส่งผลให้มีการใช้เทคนิคเหมืองข้อมูล (Data Mining) ในการวิจัยเป็นที่หลากหลายมาก

อีกทั้ง มีผู้วิจัยส่วนใหญ่ได้พยายามคิดค้นการผสมผสานวิธีการจากหลายแนวทางเพื่อใช้วิเคราะห์และค้นหาพฤติกรรมที่ผิดปกติจากการใช้งานเครือข่าย โดยการใช้เทคนิคทางสถิติเข้ามาผนวกเพิ่มเติมจากวิธีการวิเคราะห์ดั้งเดิมที่มีการนำเสนอไว้แล้ว ทั้งเพื่อมุมมองในการหาเทคนิควิเคราะห์ใหม่ๆ และการนำเสนอภาพของการรักษาความมั่นคงความปลอดภัยของระบบเครือข่าย [17, 27, 30, 33, 47, 68, 69, 70, 86, 103]

นอกจากที่กล่าวไว้ข้างต้น จะเห็นว่ามีงานวิจัยส่วนที่นำเสนอวิธีการหรือหลักการในการศึกษาเกี่ยวกับการแพร่กระจายของข้อมูลข่าวสาร (Information Diffusion) [22, 23, 25, 102] โดยงานวิจัยส่วนหนึ่งได้ทำการศึกษาเกี่ยวกับการแพร่กระจายของข้อมูลข่าวสารในเครือข่ายสังคมออนไลน์ [9, 15, 22] และงานวิจัยส่วนหนึ่งได้ทำการศึกษาเทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) กับการประยุกต์ใช้งานในลักษณะต่างๆ [9, 25] อีกทั้งมีนักวิจัยได้ทำการวิจัยเกี่ยวกับตัวแบบการวิจัยหรือเทคนิคที่ศึกษาเกี่ยวกับการศึกษาการพยากรณ์ความสัมพันธ์ (Link Prediction) ในเครือข่ายสังคมออนไลน์ [83, 102, 117] ไว้เช่นกัน

ทั้งนี้ ในงานวิจัยของ S.H. Sharif และคณะผู้วิจัย [95] ได้ศึกษาในงานวิจัยเรื่อง A Review on Search and Discovery Mechanisms in Social Networks ซึ่งเป็นการทบทวนวรรณกรรมงานวิจัย 35 ชิ้นที่เป็นวิธีเทคนิคที่ผู้วิจัยต่างๆได้นำมาใช้ในการค้นหาโหนดในลักษณะต่างๆในเครือข่ายสังคมออนไลน์ โดยสามารถจำแนกวิธีเทคนิคได้ทั้งสิ้น 3 กลุ่ม ดังนี้

1. People Search โดยมีวิธีเทคนิคที่นำมาใช้ในมิติ People Search พร้อมวิเคราะห์ข้อดี-ข้อจำกัด ดังตารางที่ 2.1 ดังนี้

ตารางที่ 2.1 วิธีเทคนิคที่นำมาใช้ในมิติ People Search

Mechanism	Main Idea	Advantages	Disadvantages
Context-base Search	Returns a ranking list of people by using query labels.	Has a good performance on the time efficiency.	Uses a lot of space in memory.
SN-based personalized social search	Investigated personalized social search based on the user's social relations.	It can be easily applied in very large scales	Suffers from lower quality.

2. Keyword Search โดยมีวิธีเทคนิคที่นำมาใช้ในมิติ Keyword Search พร้อมวิเคราะห์ข้อดี-ข้อจำกัด ดังตารางที่ 2.2 ดังนี้

ตารางที่ 2.2 วิธีเทคนิคที่นำมาใช้ในมิติ Keyword Search

Mechanism	Main Idea	Advantages	Disadvantages
Heap Union	Provides accurate cost models of the query operators.	Improves query processing efficiency.	Queries update wastes a lot of time.
T-DHT Algorithm	Solves hot spots and keyword search problems.	Makes the system fully scalable.	Implementation is only based on computation.

3. Web Service Discovery โดยมีวิธีเทคนิคที่นำมาใช้ในมิติ Web Service Discovery พร้อมวิเคราะห์ข้อดี-ข้อจำกัด ดังตารางที่ 2.3 ดังนี้

ตารางที่ 2.3 วิธีเทคนิคที่นำมาใช้ในมิติ Web Service Discovery

Mechanism	Main Idea	Advantages	Disadvantages
LinkedWS	Merges, substitute different web services into composition.	Expands existing compositions based on the recommendations; identifies the most appropriate substitutes in case the web services fail.	Implementation have not been used in practical fields.
Social Networks for Web Services Discovery	Uses social networks for web service discovery.	Improves scalability and implementations are based on experimental work.	Scalability does not affect navigation and discovery mechanism.
Composite Web Service Builder Algorithm	Selects an effective set of Web services that can collaborate in order to attain to the composite Web services.	Secure all the functions required by the composite Web service minimizing the communication overhead.	Implementation is a work on progress not in real test bed.

จากแนวคิดในการทบทวนวรรณกรรมงานวิจัยของ S.H. Sharif และคณะผู้วิจัย [95] ซึ่งเป็นการสกัดแนวคิดในการคัดเลือกวิธีเทคนิคที่นำมาใช้ในการค้นหาโหนดในลักษณะต่างๆ ในเครือข่ายสังคมออนไลน์ โดยอธิบายถึงข้อดี ข้อจำกัดของวิธีเทคนิคต่างๆ ดังนั้น ผู้วิจัยจึงได้ทำการจัดหมวดหมู่งานวิจัยที่ได้ทำการศึกษา เพื่อจำแนกงานวิจัยให้เหมาะสมกับการวิจัย โดยวิเคราะห์ข้อดีและข้อจำกัดของงานวิจัยที่ได้ทำการศึกษา เพื่อกำหนดเป็นคุณลักษณะของงานวิจัยชิ้นนี้ ดังแสดงในตารางที่ 2.4-2.10

1. ตารางที่ 2.4 แสดงการสรุปรายการงานวิจัยที่เกี่ยวข้องจำแนกตามหมวดหมู่

2. ตารางที่ 2.5 แสดงการสรุปงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์พฤติกรรมการใช้งานในเครือข่ายคอมพิวเตอร์

3. ตารางที่ 2.6 แสดงการสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์

4. ตารางที่ 2.7 แสดงการสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการติดตามการแพร่กระจายข้อมูลข่าวสาร

5. ตารางที่ 2.8 แสดงการสรุปงานวิจัยที่เกี่ยวข้องกับโครงสร้างกราฟ

6. ตารางที่ 2.9 แสดงการสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์

7. ตารางที่ 2.10 แสดงการสรุปข้อดีและข้อจำกัดของงานวิจัยที่ได้ทำการศึกษา

ตารางที่ 2.4 ตารางสรุปรายการงานวิจัยที่เกี่ยวข้องจำแนกตามหมวดหมู่

ตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ				
(1) การวิเคราะห์พฤติกรรมการใช้งานในเครือข่ายคอมพิวเตอร์	(2) เทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์	(3) เทคนิคการติดตามการแพร่กระจายข้อมูลข่าวสาร	(4) โครงสร้างกราฟ	(5) เทคนิคการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์
1. C. Guang และคณะ [21] 2. D. V.Klein [32] 3. J. TIMOFTE และคณะ [52] 4. S. Kang และคณะ [90] 5. W. Ren และคณะ [111]	1. สำนักงานปลัดกระทรวงเทคโนโลยีสารสนเทศและการสื่อสารและศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ [2] 2. A. Adhikari และคณะ [8] 3. B. Bringmann และคณะ [18] 4. D. Fu และคณะ [28] 5. H. Anh Nguyen และคณะ [35] 6. H. Yang และคณะ [40 7. J. Song และคณะ [50] 8. K. Kent และคณะ [53] 9. M. Awad และคณะ [62] 10. M. Bordie และคณะ [63] 11. R. Smith และคณะ [82] 12. S. Feruza Yusufovna [87] 13. S. Jian-hua และคณะ [89]	1. A. Apolloni และคณะ [9] 2. C. Haythornthwaite [22] 3. C. lo Storto [23] 4. H. Xia, และคณะ [39] 5. M. Kimura และคณะ [66] 6. T. Fushimi และคณะ [99]	1. A. Pablo และคณะ [15] 2. H. Xia และคณะ [25] 3. J. Lei และคณะ [46] 4. L. Lü และคณะ [57] 5. R.N. Lichtenwalter และคณะ [83] 6. T. Tylenda และคณะ [102] 7. W.K. Sharabati และคณะ [113] 8. Z. Huang [117]	1. A. D. Amico และคณะ [12] 2. A. Soule และคณะ [17] 3. D. Farid และคณะ [27] 4. E. Keogh และคณะ [33] 5. J. Lin และคณะ [47] 6. M. Mizutani และคณะ [68] 7. M. Ruiz และคณะ [69] 8. M. Weber และคณะ [70] 9. S. C. Chin และคณะ [86] 10. T.Oates และคณะ [103]

ตารางที่ 2.5 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์พฤติกรรมการใช้งานในเครือข่ายคอมพิวเตอร์

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
C. Guang และคณะ [21] ตัวแบบการสุ่มข้อมูลจราจรทางเครือข่ายสำหรับตรวจวัดโดยใช้การระบุแพ็กเก็ต (A Traffic Sampling Model for Measurement using Packet Identification)	แนะนำตัวแบบการสุ่มตัวอย่าง (PSAMP) ซึ่งสามารถรองรับการทำงานสำหรับวัตถุประสงค์ทั้งหมดของโปรแกรมประยุกต์ โดยทำการประเมินโปรแกรมประยุกต์สองโปรแกรมใหญ่ คือการวิเคราะห์ลักษณะการทำงานของข้อมูลจราจร และการวิเคราะห์ลักษณะการทำงานของเครือข่ายคอมพิวเตอร์	หลังจากทำการวิเคราะห์ข้อมูลจราจรขนาดใหญ่โดยการสุ่มตัวอย่างที่ผ่าน CERNET Backbone พบว่า bits ใน IP แพคเก็ตส่วนหัวไม่เปลี่ยนแปลงระหว่างการขนส่งทั้งหมด โดยสามารถพัฒนาตัวแบบเพื่อสุ่มตัวอย่างและพิสูจน์ประสิทธิภาพของตัวแบบในมุมของลักษณะการสุ่ม (Randomicity) และ ความเป็นอิสระของการวิเคราะห์การวัดตัวอย่าง (Independency) ได้
D. V.Klein [32] การวิเคราะห์แบบสอบกลับการโจมตีของสแปมตามรูปแบบการกระจายสองขั้นตอน (A Forensic Analysis of the Distributed Two-Stage Web-Based Spam Attack)	ใช้คำสั่งการตรวจสอบ (Audit Scripts) เพื่อเฝ้าระวังระบบเครือข่ายคอมพิวเตอร์	การศึกษาถึงผลกระทบของการถูกโจมตีในรูปแบบของอีเมล และเสนอแนวทางแก้ไขปัญหาพร้อมทั้งแนวทางการเพิ่มระดับความมั่นคงปลอดภัยของเครือข่ายโดยการใช้คำสั่งการตรวจสอบ (Audit Scripts) เพื่อเฝ้าระวังระบบเครือข่ายคอมพิวเตอร์
J. TIMOFTE และคณะ [52] การรักษาความมั่นคงปลอดภัยขององค์กรด้วยการวิเคราะห์พฤติกรรมการใช้เครือข่ายคอมพิวเตอร์ (Securing the organization with Network	ใช้เทคนิคการวิเคราะห์พฤติกรรมการใช้เครือข่าย (Network Behavior Analysis: NBA) เพื่อทำการจำแนกพฤติกรรมการใช้เครือข่ายคอมพิวเตอร์	นำเสนอแนวคิดในการสร้างฮาร์ดแวร์ หรือซอฟแวร์เพื่อรองรับการวิเคราะห์พฤติกรรมการใช้เครือข่ายที่สามารถจำแนกพฤติกรรมการใช้เครือข่ายคอมพิวเตอร์ที่ผิดปกติ เพื่อให้ได้

ตารางที่ 2.5 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์พฤติกรรมการใช้งานในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
Behavior Analysis)	ที่ผิดปกติ เช่น DoS, โปรแกรมม้าโทรจัน และวอร์ม เป็นต้น	ระบบรักษาความมั่นคงปลอดภัยที่ระบุถึงรูปแบบข้อมูลจราจร (Traffics) ที่ผิดปกติได้อย่างมีประสิทธิภาพ
S. Kang และคณะ [90] การวิเคราะห์แบบสอบกลับเครือข่ายคอมพิวเตอร์โดยใช้การทำให้เห็นภาพ (Network Forensic Analysis Using Visualization Effect)	งานวิจัยชิ้นนี้ได้เน้นทางด้านผลกระทบต่อองค์ประกอบของเครือข่ายคอมพิวเตอร์ โดยการสร้างระบบการสอบกลับ (Forensic System) ที่มีประสิทธิภาพ โดยทำการเปรียบเทียบระหว่างวิธีที่นิยมใช้กัน (Traditional Methods) และวิธีที่นำเสนอ (Proposed Method) กับรูปแบบการถูกจู่โจมระบบเครือข่ายคอมพิวเตอร์ในลักษณะต่างๆ ผู้วิจัยได้นำเทคนิคการทำให้เห็นภาพ (Visualization Technique) เป็นเครื่องมือทดสอบทางด้านเครือข่ายเพื่อทำการทดสอบข้อมูลจราจรในระบบเครือข่ายคอมพิวเตอร์	ในงานวิจัยชิ้นนี้ได้นำเสนอระบบการสอบกลับการใช้งานเครือข่ายคอมพิวเตอร์ และได้ทำการทดสอบ โดยการแสดงภาพประกอบการวิเคราะห์เครือข่ายคอมพิวเตอร์ โดยพบว่าวิธีที่นำเสนอมีประสิทธิภาพสูงสุด
W. Ren และคณะ [111] ตัวแบบการทำงานของการสอบกลับพฤติกรรมการใช้งาน	1. กำหนดกระบวนการวิเคราะห์ดังนี้ 1.1 การจับภาพ: Capture	ได้ศึกษาถึงการสอบกลับลักษณะการทำงานของเครือข่ายคอมพิวเตอร์ ทั้งในมุมมองด้านเทคนิคและมุมมองด้านกฎหมาย

ตารางที่ 2.5 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์พฤติกรรมการใช้งานในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
เครือข่ายคอมพิวเตอร์ (Modeling the Network Forensics Behaviors)	1.2 การทำสำเนา: Copy 1.3 การโอนย้าย: Transfer 1.4 การวิเคราะห์: Analysis 1.5 การตรวจสอบ: Investigate 1.6 การนำเสนอ: Presentation 2. กำหนดเครื่องมือในการใช้งาน (Deployment) 2.1 Computer forensics tools 2.2 Network Forensics Machine	เพื่อสร้างมาตรฐานของการสอบกลับลักษณะการทำงานของเครือข่ายคอมพิวเตอร์ขึ้น โดยผู้วิจัยได้ดำเนินการตามกระบวนการที่กำหนด

ตารางที่ 2.6 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
สำนักงานปลัดกระทรวงเทคโนโลยีสารสนเทศและการสื่อสารและศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ [2] ความรู้เกี่ยวกับพระราชบัญญัติว่าด้วยการกระทำความผิดเกี่ยวกับคอมพิวเตอร์ 2550	สร้างความเข้าใจในตัวบทกฎหมายแก่ประชาชนและผู้เกี่ยวข้องใช้เป็นคู่มือแนวทางในการปฏิบัติตามกฎหมายปกป้องสิทธิ	กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร ซึ่งมีการกล่าวถึงลักษณะความผิดเกี่ยวกับคอมพิวเตอร์ประเภทต่างๆ ที่อาจนำภัยคุกคามอย่างอื่นแฝงมาด้วย เช่นไวรัส สปายแวร์ เป็นต้น จนส่งผลถึงขั้นทำให้ระบบคอมพิวเตอร์ไม่สามารถทำงานได้อีก และเป็นเครื่องมือที่สำคัญในการกระทำความผิดทางคอมพิวเตอร์ต่อไป

ตารางที่ 2.6 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
A. Adhikari และคณะ [8] การค้นหาข้อมูลจากหลายแหล่งข้อมูลขนาดใหญ่ (Mining Multiple Large Data Sources)	จากข้อจำกัดของ Local Pattern Analysis ที่ไม่สามารถค้นหาข้อมูลที่ต้องการจากหลายแหล่งข้อมูลขนาดใหญ่ได้ ผู้วิจัยจึงได้เสนอแนวทางไว้ดังนี้ 1. Partition Algorithm 2. IdentifyExPattern Algorithm 3. Rule Synthesizing Algorithm 4. Association-Rule-Synthesis (ARS) algorithm 5. Mining Multiple Databases for the Purpose of Studying a Set of Items 6. Mining Multiple Databases Using Pipelined Feedback Technique	ผู้วิจัยได้นำเสนอ Multi-Database Mining Technique (MDMT) ซึ่งเป็นการใช้เทคนิคร่วมระหว่าง Pipelined Feedback Technique (PFT) และ Simple Pattern Synthesizing (SPS) Algorithm [MDMT: PFT+SPS] โดยผลการวิจัยสามารถยืนยันได้ว่า Database Mining Technique (MDMT) เป็นวิธีที่มีประสิทธิภาพ

ตารางที่ 2.6 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
B. Bringmann และคณะ [18] การเรียนรู้และการทำนายวิวัฒนาการของเครือข่ายสังคม (Learning and Predicting the Evolution of Social Networks)	ผู้วิจัยนำเสนอแนวคิดของการใช้คุณสมบัติของกราฟในการวิเคราะห์การทำนายในเครือข่ายสังคมโดยศึกษาความเชื่อมโยงของความสัมพันธ์ระหว่างโหนด	ผลการวิจัยพบว่าการศึกษาความเชื่อมโยงความสัมพันธ์ของโหนดโดยใช้ทฤษฎีกราฟ สามารถคาดกาณ์การเกิดความสัมพันธ์ระหว่างโหนดที่มีอยู่เดิม หรือระหว่างโหนดเดิมและโหนดใหม่ที่อาจจะเกิดขึ้นเมื่อเครือข่ายสังคมมีวิวัฒนาการตามเวลา โดยไม่สามารถทำนายความเชื่อมโยงตวามสัมพันธ์ระหว่างโหนดใหม่ที่อาจจะเกิดขึ้นในวิวัฒนาการของเครือข่ายสังคมได้
D. Fu และคณะ [28] การออกแบบและการดำเนินการของระบบการตรวจจับการบุกรุกเครือข่ายคอมพิวเตอร์แบบกระจายโดยใช้เทคนิคเหมืองข้อมูล (The Design and Implementation of a Distributed Network Intrusion Detection System Based on Data Mining)	การทำเหมืองข้อมูลถูกนำไปใช้กับระบบการตรวจสอบการบุกรุก โดยการสร้างตัวแบบตามวิธีการของอัลกอริทึม Improved FP-Growth และเทคนิค FCM เพื่อสร้างการตรวจจับการบุกรุกเครือข่ายคอมพิวเตอร์ตามหลักการทางสถิติ	ผลการวิจัย มีดังนี้ 1. การใช้เทคนิคเหมืองข้อมูล ในวิธีการของอัลกอริทึม FP-Growth และเทคนิค FCM พบว่าสามารถจำแนกรูปแบบความผิดปกติได้อย่างแม่นยำ 2. ประสิทธิภาพการทำงานร่วมกันระหว่าง FP-Growth และเทคนิค FCM มีประสิทธิภาพการตรวจจับการบุกรุกดีกว่าการใช้ FP-Growth เพียงอย่างเดียว

ตารางที่ 2.6 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
H. Anh Nguyen และคณะ [35] การประยุกต์ใช้เทคนิคเหมืองข้อมูลเพื่อตรวจจับการบุกรุกในเครือข่ายคอมพิวเตอร์ด้วยตัวแบบการจำแนกกลุ่ม (Application of Data Mining to Network Intrusion Detection: Classifier Selection Model)	ผู้วิจัยได้ประยุกต์ใช้เทคนิคเหมืองข้อมูลเพื่อตรวจจับการบุกรุกเครือข่ายคอมพิวเตอร์ด้วยตัวแบบการจำแนกกลุ่ม (Classifier Selection Model)	การประยุกต์ใช้เทคนิคเหมืองข้อมูลเพื่อตรวจจับการบุกรุกในเครือข่ายคอมพิวเตอร์ด้วยตัวแบบการจำแนกกลุ่ม (Classifier Selection Model) สามารถทำการจำแนกกลุ่มของรูปแบบการบุกรุกในเครือข่ายคอมพิวเตอร์ได้อย่างมีประสิทธิภาพ
H. Yang และคณะ [40] การรวมกลุ่มและการจำแนกกลุ่มเพื่อการตรวจจับความผิดปกติ (Clustering and Classification Based Anomaly Detection)	นำเสนอวิธีการตรวจหาความผิดปกติตามเทคนิคการรวมกลุ่ม (Clustering) และเทคนิคการจำแนกประเภท (Classification) สำหรับการบุกรุก (ID)	ผลการวิจัย มีดังนี้ 1. ตัวแบบที่ผู้วิจัยนำเสนอมีประสิทธิภาพของการตรวจจับความผิดปกติได้เป็นอย่างดี โดยเฉพาะ PROBE และ DoS 2. ไม่สามารถจำแนก U2R และ R2L จากความผิดปกติ
J. Song และคณะ [50] แนวทางการตรวจจับการถูกโจมตีที่ไม่รู้จักผ่านระบบการแจ้งเตือนการบุกรุก (A Comprehensive Approach to Detect Unknown Attacks Via Intrusion Detection Alerts)	ผู้วิจัยนำเสนอวิธีการ (Approach) เพื่อกำหนดลักษณะ (Features) ในการตรวจจับการถูกโจมตี จากนั้นทำการจำแนกการถูกโจมตีด้วยเทคนิคการรวมกลุ่ม (Cluster Analysis) เพื่อสร้างระบบการแจ้งเตือนการบุกรุก	ผลการวิจัย พบว่า วิธีการที่ผู้วิจัยนำเสนอสามารถตรวจจับการถูกโจมตีที่ไม่เคยรู้จักผ่านระบบผ่านระบบการแจ้งเตือนที่ได้พัฒนาขึ้น โดยพบ False Posive Alerts ต่ำ

ตารางที่ 2.6 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
K. Kent และคณะ [53] คู่มือการรักษาความมั่นคงปลอดภัยโดยการบริหารจัดการข้อมูลจราจรทางคอมพิวเตอร์ (Guide to Computer Security Log Management)	สถาบันมาตรฐานและเทคโนโลยีแห่งชาติ (National Institute of Standards and Technology : NIST) ได้พัฒนาคู่มือที่สอดคล้องกับข้อกฎหมายด้านการจัดการความมั่นคงปลอดภัยของรัฐบาลกลาง (Federal Information Security Management Act: FISMA) เพื่อช่วยให้องค์กรต่างๆ ได้เข้าใจถึงการบริหารระบบความมั่นคงปลอดภัยของเครือข่าย การติดตั้ง การดูแลระบบอย่างมีประสิทธิ-ภาพ กระบวนการพัฒนาและประมวลผลเหตุการณ์เครือข่าย (Log Management)	โดยผู้บริหารเครือข่ายศึกษาเหตุการณ์เครือข่าย (Logs) โดยใช้เครื่องมือในการวิเคราะห์หารูปแบบที่ผิดปกติ เช่นใช้ค่าความสัมพันธ์ (Correlation) ดูความสัมพันธ์ของเหตุการณ์เครือข่าย (Logs) แต่ละประเภท
M. Awad และคณะ [62] การทำนายของอนุกรมเวลาด้วย RBF และ โครงข่ายประสาทเทียม: แนวทางใหม่ของการรวมกลุ่ม (Prediction of Time Series	ผู้วิจัยนำเสนอวิธีการใหม่สำหรับการพยากรณ์ของอนุกรมเวลาชุดข้อมูลโดยใช้ฟังก์ชันรัศมีพื้นฐาน (Radial Basis Functions) ร่วมกับเครือข่ายงานประสาทเทียม	ผลการวิจัยเป็นดังนี้ 1. การพยากรณ์ในระยะสั้น (Short-Term Prediction) ผลการทำนายในระยะสั้นของวิธี RBFNNs มีความแม่นยำมากกว่าวิธีอื่นๆ ที่นำมาเปรียบเทียบ

ตารางที่ 2.6 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
Using RBF Neural Networks: A New Approach of Clustering)	(Neural Networks) จึงเรียกว่า Radial Basis Function Neural Networks (RBFNNs) โดยทำการศึกษาประสิทธิภาพเปรียบเทียบกับวิธีต่างๆ ดังนี้ 1. Linear Prediction Model 2. Auto Regressive Model 3. Casecade Correlation 4. 6^{th}-Order Polynomial 5. Back-Prop NN เป็นต้น	2. การพยากรณ์ในระยะยาว (Large-Term Prediction) ผลการทำนายในระยะยาวของวิธี RBFNNs มีความแม่นยำมากกว่าวิธีอื่นๆ ที่นำมาเปรียบเทียบ
M. Bordie และคณะ [63] การตรวจจับการบุกรุกโดยใช้เทคนิคเหมืองข้อมูล (Using Data Mining for Intrusion Detection)	ผู้วิจัยนำเสนอวิธีการเพื่อตรวจจับการบุกรุกในเครือข่ายคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูล	ผลการวิจัย พบว่าวิธีการที่ผู้วิจัยนำเสนอสามารถตรวจจับการบุกรุกในเครือข่ายคอมพิวเตอร์ได้อย่างแม่นยำ
R. Smith และคณะ [82] การใช้เทคนิคการเรียนรู้แบบไม่มีผู้สอนสำหรับการแจ้งเตือนความสัมพันธ์ของเครือข่าย (Using Unsupervised Learning for Network Alert Correlation)	ผู้วิจัยนำเสนอระบบวิเคราะห์ความสัมพันธ์การแจ้งเตือนโดยใช้เทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Machine Learning Algorithms)	ระบบจะดำเนินการโดยการวิเคราะห์สหสัมพันธ์ใน 2 ขั้นตอน ดังนี้ 1. สหสัมพันธ์ ในขั้นตอนที่หนึ่งคือการรวมกลุ่มข้อความแจ้งเตือนเข้าด้วยกันซึ่งแต่ละกลุ่มจะมีรูปแบบข้อความเตือนของการถูกโจมตี

ตารางที่ 2.6 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
		2. สหสัมพันธ์ ในขั้นตอนสองเป็นการรวมกลุ่มที่สร้างขึ้นจากขั้นตอนที่หนึ่ง ซึ่งการรวมกลุ่มแต่ละชุดจะประกอบด้วยข้อความแจ้งเตือนของการถูกโจมตีที่มีลักษณะใกล้เคียงกันอีก 3. ผลการทดสอบการใช้งานพบว่าระบบสามารถดำเนินการตามวัตถุประสงค์ได้อย่างมีประสิทธิภาพ โดยพบการแจ้งเตือนผิดลดลงอย่างมาก
S. Feruza Yusufovna [87] การบูรณาการระบบการตรวจจับการบุกรุกด้วยเทคนิคเหมืองข้อมูล (Integrating Intrusion Detection System and Data Mining)	ผู้วิจัยได้รวบรวมวิธีการของเทคนิคเหมืองข้อมูลเพื่อทำการตรวจจับการบุกรุก โดยได้รวบรวมกระบวนการต่างๆ ไว้เพื่อเป็นแนวทางกับผู้วิจัยท่านอื่นเพื่อเป็นแนวทางในการเลือกวิธีการในเทคนิคเหมืองข้อมูล ซึ่งแต่ละวิธีการจะต้องดำเนินการตามขั้นตอนเทคนิคเหมืองข้อมูลดังนี้ 1. การคัดเลือกข้อมูล (Selection)	วิธีการของเทคนิคเหมืองข้อมูลที่ผู้วิจัยได้รวบรวมไว้จากผลงานวิจัยชิ้นอื่นๆ ที่มีผู้นำวิธีการมาใช้เพื่อการตรวจนับการบุกรุกเป็นดังนี้ 1. Feature Selection 2. Clustering Techniques 3. Statistical Techniques 4. Predictive Analysis 5. Other Techniques ได้แก่ NIDS with Random Forests, Exploiting the Infrastructure เป็นต้น

ตารางที่ 2.6 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
	2. การประมวลผลเริ่มต้น (Preprocessing) 3. การแปลงข้อมูล (Transformation) 4. การทำเหมืองข้อมูลรูปแบบ(Patterns Data Mining) 5. การแปลผลและการประเมิน (Interpretation/Evaluation)	
S. Jian-hua และคณะ [89] MA-IDS: ระบบการตรวจจับการบุกรุกแบบกระจายด้วยเทคนิคเหมืองข้อมูล (MA-IDS: A Distributed Intrusion Detection System Based on Data Mining)	ผู้วิจัยได้นำเสนอตัวแบบเพื่อพัฒนาระบบการตรวจจับการบุกรุกแบบกระจายด้วยเทคนิคเหมืองข้อมูล	ตัวแบบเพื่อพัฒนาระบบการตรวจจับการบุกรุกแบบกระจายด้วยเทคนิคเหมืองข้อมูลที่ผู้วิจัยนำเสนอมีประสิทธิภาพที่มีความแม่นยำ

ตารางที่ 2.7 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการติดตามการแพร่กระจายข้อมูลข่าวสาร

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
A. Apolloni และคณะ [9] การศึกษาการแพร่กระจายข้อมูลผ่านตัวแบบเครือข่ายสังคม (A Study of Information Diffusion over a Realistic Social Network Model)	ใช้ Sociological Models ในการวิเคราะห์และศึกษาการแพร่กระจายข้อมูลผ่านตัวแบบเครือข่ายสังคมออนไลน์	ผลการวิจัย สามารถสรุปได้ดังนี้ 1. ถ้าเป็นการส่งผ่านข้อมูลระหว่างบุคคลที่มีความคุ้นเคยกัน พบว่าข้อมูลจะสามารถส่งในระยะเวลาสั้น เช่น ระยะเวลา 10-20 นาที โดยที่แหล่งซื้อสินค้าเป็นข้อมูลแพร่กระจายมากที่สุด 2. ถ้าเป็นการส่งผ่านข้อมูลระหว่างบุคคลที่มีตำแหน่งที่สูงกว่า พบว่าข้อมูลจะสามารถส่งได้ในระยะเวลาที่มากและนาน
C. Haythornthwaite [22] การวิเคราะห์เครือข่ายสังคม: วิธีการและเทคนิคเพื่อการศึกษาการแลกเปลี่ยนข้อมูล (Social Network Analysis: An Approach and Technique for the Study of Information Exchange)	ผู้วิจัยได้นำเทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) มาทำการศึกษารูปแบบการแลกเปลี่ยนข้อมูล	ผลการวิจัยยืนยันว่าแนวคิดที่นำเสนอซึ่งเป็นการทำงานของเทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) สามารถใช้ศึกษารูปแบบการแลกเปลี่ยนข้อมูลในเครือข่ายสังคมออนไลน์ได้
C. lo Storto [23] การตรวจสอบการไหลข้อมูลผ่านขั้นตอนการพัฒนาผลิตภัณฑ์ที่ซับซ้อนด้วยการวิเคราะห์เครือข่ายสังคม (Investigating Information	ผู้วิจัยนำเสนอแนวคิด (Approach) ในการวิเคราะห์ประสิทธิภาพของกระบวนการพัฒนาสินค้าโดยใช้ Dependency Structure Matrix (DSM)	ผลการวิจัยยืนยันว่าแนวคิดที่นำเสนอซึ่งเป็นการทำงานร่วม-กันระหว่าง DSM และ SNA สามารถใช้เป็นเครื่องมือในการการปรับปรุงพฤติกรรมการพัฒ-นาสินค้าได้อย่างมีประสิทธิภาพ

ตารางที่ 2.7 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการแพร่กระจายข้อมูลข่าวสาร (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
Flows Across Complex Product Development Stages by Using Social Network Analysis)	ร่วมกับเทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) โดยใช้ DSM ในขั้นตอนพัฒนาตัวแบบในการแพร่กระจายข้อมูลข่าวสาร และ SNA ในขั้นตอนการตรวจสอบโครงสร้างของการแพร่กระจายของข้อมูลข่าวสาร	
H. Xia, และคณะ [39] การวิเคราะห์เครือข่ายสังคมของการแพร่กระจายความรู้ในกลุ่มนักศึกษามหาวิทยาลัย (Social Network Analysis of the Knowledge Diffusion among University Students)	การวิเคราะห์เครือข่ายสังคม (SNA) คือ วิธีที่ใช้วิเคราะห์ความสัมพันธ์ของโหนด โดยบทความนี้จะใช้วิธีวิเคราะห์เครือข่ายทางสังคมที่มีอยู่แล้วเพื่อศึกษาการแพร่กระจายความรู้ในกลุ่มนักศึกษามหาวิทยาลัยใน 3 มุมมอง คือ มุมมองความหนาแน่น (Density) มุมมองของความเป็นศูนย์กลาง (Centrality) และมุมมองของความเป็นกลุ่มย่อย (Cohesive Subgroup)	ผลการวิจัย พบว่าในเครือข่ายสังคมของกลุ่มนักศึกษามหาวิทยาลัยนอกเหนือจากจะมีการแพร่กระจายความรู้แล้ว ยังมีความสัมพันธ์ในรูปแบบทางอารมณ์ ได้แก่ การปรึกษา เป็นต้น

ตารางที่ 2.7 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการแพร่กระจายข้อมูลข่าวสาร (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
M. Kimura และคณะ [66] การวิเคราะห์ชุมชนของโหนดที่มีอิทธิพลสำหรับการแพร่กระจายข้อมูลในเครือข่ายสังคม (Community Analysis of Influential Nodes for Information Diffusion on a Social Network)	ปัญหาของการค้นหาโหนดที่มีอิทธิพลสำหรับการแพร่กระจายข้อมูลข่าวสารในเครือข่ายสังคมออนไลน์ ซึ่งเป็นที่รู้จักกันว่า อัลกอริทึมเชิงละโมบ (Greedy Algorithm) สามารถใช้ในการแก้ไขปัญหาได้นั้น มีความเหมาะสมกับทุกลักษณะโครงสร้างเครือข่ายสังคมออนไลน์หรือไม่	ผลการทดลองสามารถพิสูจน์ได้ว่า SR-Community Structure มีความสัมพันธ์กันอย่างสูงกับอัลกอริทึมเชิงละโมบ (Greedy Algorithm) โดยแตกต่างจาก Community Structure อย่างมีนัยสำคัญ ดังนั้นในการวิเคราะห์โหนดที่มีอิทธิพลกับการแพร่กระจายข้อมูลในเครือข่ายสังคมออนไลน์ จึงควรใช้ SR-Community Structure วิเคราะห์ร่วมกับอัลกอริทึมเชิงละโมบ (Greedy Algorithm)
T. Fushimi และคณะ [99] ตัวแบบการแพร่กระจายข้อมูลบอกอะไรกับโครงสร้างเครือข่ายสังคม (What does an Information Diffusion Model Tell about Social Network Structure)	ผู้วิจัยต้องการนำเสนอแนวคิดในเรื่องตัวแบบการแพร่กระจายข้อมูลในเครือข่ายสังคมออนไลน์มีความสัมพันธ์กับโครงสร้างทางเครือข่ายสังคมออนไลน์ที่แตกต่างกันอย่างไร	ผลการวิจัยพบว่า โครงสร้างทางเครือข่ายสังคมที่แตกต่างกันจะให้ตัวแบบของการแพร่กระจายข้อมูลในเครือข่ายสังคมที่แตกต่างกัน กล่าวคือโครงสร้างทางเครือข่ายสังคมออนไลน์แบบ Community Structure จะมีตัวแบบในการแพร่กระจายข้อมูลในเครือข่ายสังคมออนไลน์แบบ Independent Casedade (IC Model) แทนที่จะเป็น Linear Threshold (LT Model)

ตารางที่ 2.8 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับโครงสร้างกราฟ

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
A. Pablo และคณะ [15] การเลือกโหนดที่มีอิทธิพลมากที่สุดในเครือข่ายสังคม (Selecting the Most Influential Nodes on Social Networks)	ผู้วิจัยนำเสนอเทคนิคสำหรับการเลือกโหนดที่มีอิทธิพลมากที่สุดในเครือข่ายสังคมออนไลน์ ด้วยเทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA)	ปัจจัยสำหรับการจำแนกโหนดที่มีอิทธิพลมากที่สุดในเครือข่ายสังคมออนไลน์ ด้วยเทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) ได้แก่ ความหนาแน่น (Density) ความเป็นศูนย์กลาง (Centrality) และความเป็นผู้เชื่อมโยงเครือข่ายย่อย (Cohesive Subgroup)
H. Xia และคณะ [25] ความซับซ้อนของเครือข่ายสังคม: ผลเชิงทฤษฎีและผลเชิงประจักษ์ในเครือข่ายสังคม (The Complexity of Social Networks: Theoretical and Empirical findings in Social Networks)	ผู้วิจัยทำการศึกษาผลเชิงทฤษฎีและผลเชิงประจักษ์ในการศึกษาความซับซ้อนของเครือข่ายสังคมออนไลน์	ผู้วิจัยอธิบายผลการทำการศึกษาผลเชิงทฤษฎีและผลเชิงประจักษ์ในการศึกษาความซับซ้อนของเครือข่ายสังคมออนไลน์ เพื่อเป็นความรู้ประกอบสำหรับการวิเคราะห์เครือข่ายสังคมออน-ไลน์ต่อไป
J. Lei และคณะ [46] การทำนายการถูกโจมตีในอนาคตโดยใช้กราฟการถูกโจมตีเครือข่าย (Using Network Attack Graph to Predict the Future Attacks)	ผู้วิจัยนำเทคนิคเหมืองข้อมูลร่วมกับกราฟการถูกโจมตีเครือข่ายคอมพิวเตอร์มาทำการวิเคราะห์การทำนายการถูกโจมตีในอนาคต	ผลการวิจัยยืนยันว่าการนำเทคนิคเหมืองข้อมูลร่วมกับกราฟการถูกโจมตีเครือข่ายคอมพิวเตอร์มาทำการวิเคราะห์การทำนายการถูกโจมตีเครือข่ายคอมพิวเตอร์ในอนาคตได้อย่างแม่นยำ

ตารางที่ 2.8 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับโครงสร้างกราฟ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
L. Lü และคณะ [57] ผลการสำรวจการทำนายความเชื่อมโยงในเครือข่ายที่มีความซับซ้อน (Link prediction in Complex Networks: A Survey)	ผู้วิจัยนำเสนอผลการสำรวจเทคนิคที่นำมาใช้ในการทำนายความเชื่อมโยงในเครือข่ายที่มีความซับซ้อน	มีผลการสำรวจเทคนิคที่นำมาใช้ในการทำนายความเชื่อมโยงในเครือข่ายนำเสนอ เพื่อใช้เป็นเนื้อหาประกอบการวิจัยอื่นๆ
R.N. Lichtenwalter และคณะ [83] มุมมองใหม่และวิธีในการทำนายความเชื่อมโยง (New Perspectives and Methods in Link Prediction)	งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อตรวจสอบปัจจัยสำคัญสำหรับการพยากรณ์การเชื่อมโยงในเครือข่าย และแสดงแนวความคิดของงานที่มีประสิทธิภาพสูงสำหรับงานพยากรณ์ โดยใช้ Flow-based Predicting Algorithm ร่วมกับ Classification Algorithm	ผลการประเมินประสิทธิภาพการทำงานร่วมกันระหว่าง Flow-based Predicting Algorithm และ Classification Algorithm วิธี Naïve Bayes และ C4.5 พบว่าการทำงานของ Flow-based Predicting Algorithm และ Naïve Bayes ให้ผลแม่นยำสูงถึง 99.99%
T. Tylenda และคณะ [102] ความตระหนักถึงกรอบเวลาของการทำนายความเชื่อมโยงในวิวัฒนาการของเครือข่ายสังคม (Towards Time-aware Link Prediction in Evolving Social Networks)	ผู้วิจัยเน้นการตรวจสอบความเชื่อมโยงความสัมพันธ์ของเครือข่ายสังคมออนไลน์ในปัจจุบัน เพื่อทำนายความเชื่อมโยงในเครือข่ายสังคมออนไลน์ที่จะเกิดขึ้น โดยจะทำการวิจัยถึงกรอบเวลาที่เหมาะสมของเครือข่ายสังคมออนไลน์ที่จะใช้ในการทำนายความเชื่อมโยงความสัมพันธ์	ผลการวิจัยแสดงให้เห็นว่าการทำนายความเชื่อมโยงความสัมพันธ์ด้วยวิธีการ Graph-based Link Prediction Techniques ร่วมกับ Probabilistic Model สามารถทำนายความเชื่อมโยงในเครือข่ายสังคมออนไลน์ได้อย่างมีนัยสำคัญ

ตารางที่ 2.8 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับโครงสร้างกราฟ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
	ได้อย่างแม่นยำ ด้วยวิธีการ Graph-based Link Prediction Techniques ร่วมกับ Probabilistic Model	
W.K. Sharabati และคณะ [113] การทำนายโหนดและความเชื่อมโยงในเครือข่าย (Predicting Edges And Vertices In A Network)	งานวิจัยนี้มุ่งเน้นที่จะหาวิธีการหาโหนดหรือความสัมพันธ์ที่หายไปจากเครือข่ายสังคม โดยใช้ความน่าจะเป็น (Probability) ภายใต้หลักการของเรขาคณิต (Geometry) และตรรกศาสตร์คลุมเครือ (Fuzzy Logic)	ผู้วิจัยได้ใช้ความสามารถในการแลกเปลี่ยน (Iinterchange Ability) ระหว่างโหนดและความสัมพันธ์ในกราฟเพื่อประเมินค่าความน่าเป็นของโหนดในเครือข่ายที่หายไป ทั้งนี้ผู้วิจัยได้ใช้วิธีประเมินค่าความแปรปรวนร่วม (Covariates) ของเวกเตอร์ เพื่อประเมินความน่าเป็นของความสัมพันธ์ภายใต้หลักการของเรขาคณิต (Geometry) และตรรกศาสตร์คลุมเครือ (Fuzzy Logic) พบผลการวิจัยมีความแม่นยำ
Z. Huang [117] การทำนายความเชื่อมโยงโดยใช้ทฤษฎีกราฟ: การทำนายจากค่าสัมประสิทธิ์การรวมกลุ่ม (Link Prediction Based on Graph Topology: The Predictive Value of the	เอกสารนี้แสดงถึงความพยายามเบื้องต้นเพื่อสำรวจความเกี่ยวเนื่องระหว่างการพยากรณ์การเชื่อมโยงความสัมพันธ์และรูปแบบของกราฟ โดยเน้นเพื่อการพัฒนาตัวแบบเพื่อประเมินค่า	การประเมินประสิทธิภาพการพยากรณ์การเชื่อมโยงนั้น ผู้วิจัยได้สร้าง Receiver Operating Characteristics (ROC) เป็นเปอร์เซ็นต์ของผลรวมที่ได้ความเชื่อมโยงที่อาจจะเกิดขึ้นและเปอร์เซ็นต์ของการเชื่อมโยงใหม่

ตารางที่ 2.8 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับโครงสร้างกราฟ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
Generalized Clustering Coefficient	การทำนายด้วยการวัดค่าสัมประสิทธิ์การรวมกลุ่ม (Clustering Coefficient Measure) และตัวแบบความน่าจะเป็น (Probability Model)	ที่แท้จริงที่เกิดขึ้น พบว่าการประเมินค่าการทำนายด้วยการวัดค่าสัมประสิทธิ์การรวมกลุ่ม (Clustering Coefficient Measure) และตัวแบบความน่าจะเป็น (Probability Model) สามารถพยากรณ์ความเชื่อมโยงที่จะเกิดขึ้นใหม่ในเครือข่ายสังคมออนไลน์ได้

ตารางที่ 2.9 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
A. D. Amico และคณะ [12] การแสดงภาพการค้นพบในการป้องกันเครือข่ายคอมพิวเตอร์ (Visual Discover in Computer Network Defense)	การทำให้เห็นภาพการค้นพบในการป้องกันเครือข่ายคอมพิวเตอร์เป็นจุดประสงค์ของผู้ทำการวิจัย โดยใช้เทคนิค Computer Network Defense (CND) ร่วมกับ Cognitive Task Analysis (CTA)	การใช้เทคนิค Computer Network Defense (CND) ร่วมกับ Cognitive Task Analysis (CTA) ทำให้ผู้วิจัยสามารถนำเสนอภาพการค้นพบในการป้องกันเครือข่ายคอมพิวเตอร์
A. Soule และคณะ [17] การคัดกรองและวิธี การทางสถิติสำหรับการตรวจ จับความผิดปกติ (Combining Filtering and Statistical Methods for Anomaly Detection)	ผู้วิจัยนำเสนอการใช้เทคนิคการคัดกรองและวิธีการทางสถิติสำหรับการตรวจจับความผิดปกติ	การใช้เทคนิคการคัดกรองและวิธีการทางสถิติสำหรับการตรวจจับความผิดปกตินั้นสามารถนำมาใช้เพื่อนำเสนอภาพการตรวจจับความผิดปกติได้อย่างมีประสิทธิภาพ

ตารางที่ 2.9 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
D. Farid และคณะ [27] การจำแนกการถูกโจมตีในการตรวจจับการบุกรุกแบบปรับ-ปรุงด้วยต้นไม้การตัดสินใจ (Attacks Classification in Adaptive Intrusion Detection using Decision Tree)	ผู้วิจัยนำเสนอการใช้วิธีการจำแนกการถูกโจมตีในการตรวจจับการบุกรุกแบบปรับปรุงด้วยต้นไม้การตัดสินใจ	การใช้วิธีการจำแนกการถูกโจมตีในการตรวจจับการบุกรุกแบบปรับปรุงด้วยต้นไม้การตัดสินใจมีความแม่นยำถึง 98% และทำให้การนำเสนอภาพมีความชัดเจน
E. Keogh และคณะ [33] การค้นหารูปแบบที่ไม่รู้จักในฐานข้อมูลอนุกรมเวลาแบบเวลาเส้นตรง (Finding Surprising Patterns in a Time Series Database in Linear Time and Space)	ผู้วิจัยนำเสนอเทคนิคการค้นหารูปแบบที่ไม่รู้จักในฐานข้อมูลอนุกรมเวลาแบบเวลาเส้นตรงโดยใช้ตัวแบบมาร์คอฟ (Markov Model) ร่วมกับเทคนิคการตัดสินใจด้วยโครงสร้างต้นไม้ (Decision Tree)	การใช้ตัวแบบมาร์คอฟ (Markov Model) ร่วมกับเทคนิคการตัดสินใจด้วยโครงสร้างต้นไม้ (Decision Tree) สามารถค้นหารูปแบบที่ไม่รู้จักในฐานข้อมูลอนุกรมเวลาแบบเวลาเส้นตรงเพื่อทำการนำเสนอภาพได้อย่างมีประสิทธิภาพ
J. Lin และคณะ [47] การแสดงภาพและการค้นพบรูปแบบที่ไม่รู้จักในฐานข้อมูลอนุกรมเวลาขนาดใหญ่ (Visulizing and Discovering Non-Trivial Patterns in Large Time Series Databases)	ผู้วิจัยได้นำเสนอแนวทางการทำให้เห็นภาพของการค้นพบรูปแบบที่ไม่รู้จักในฐานข้อมูลอนุกรมเวลาขนาดใหญ่ ด้วยการพัฒนา VizTree ซึ่งเป็นการพัฒนาต่อเนื่องจาก Augmentig Suffix Tree	ทำการทำให้เห็นภาพของการค้นพบรูปแบบที่ไม่รู้จักในฐานข้อมูลอนุกรมเวลาขนาดใหญ่ ด้วยการพัฒนา VizTree ซึ่งเป็นการพัฒนาต่อเนื่องจาก Augmentig Suffix Tree ทำให้เห็นความถี่ของกลุ่มข้อมูลที่มีความเหมือนหรือแตกต่างกัน

ตารางที่ 2.9 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
M. Mizutani และคณะ [68] กฎพฤติกรรมจากการตรวจจับการบุกรุก (Behavior Rule based Intrusion Detection)	ผู้วิจัยเสนอแนวทางการตรวจจับการบุกรุกด้วยการสร้างกฎสร้างความสัมพันธ์ของพฤติกรรมการเข้าใช้งานเครือข่ายคอมพิวเตอร์	การสร้างกฎสร้างความสัมพันธ์ของพฤติกรรมการเข้าใช้งานเครือข่ายสามารถใช้เป็นแนวทางการตรวจจับการบุกรุกได้อย่างมีประสิทธิภาพ ทำให้สามารถนำเสนอภาพพฤติกรรมจากการตรวจจับการบุกรุก
M. Ruiz และคณะ [69] การผนวกวิธีการควบคุมกระบวนการทางสถิติและนโยบายการจำแนกกลุ่มสำหรับกระบวนการตรวจสอบสถานะการเข้าถึงแบบกลุ่ม (Combination of Statistical Process Control (SPC) Methods and Classification Strategies for Situation Assessment of Batch Process)	ผู้วิจัยนำเสนอการใช้วิธีการควบคุมกระบวนการทางสถิติและนโยบายการจำแนกกลุ่มสำหรับกระบวนการตรวจสอบสถานะการเข้าถึงแบบกลุ่ม	การใช้วิธีการควบคุมกระบวนการทางสถิติและนโยบายการจำแนกกลุ่มสำหรับกระบวนการตรวจสอบสถานะการเข้าถึงแบบกลุ่ม สามารถตรวจจับพฤติกรรมที่ผิดปกติในเครือข่ายคอมพิวเตอร์และนำเสนอภาพสถานะการเข้าถึงแบบกลุ่มได้อย่างมีประสิทธิภาพ
M. Weber และคณะ [70] การแสดงภาพอนุกรมเวลาแบบก้นหอย (Visualizing Time Series on Sprials)	ผู้วิจัยได้นำเสนอแนวทางการแสดงภาพอนุกรมเวลาแบบก้นหอยในลักษณะสามมิติ	การนำเสนอแนวทางการแสดงภาพอนุกรมเวลาแบบก้นหอยในลักษณะสามมิติทำให้สามารถเห็นภาพการเคลื่อนไหวของข้อมูลอนุกรมเวลาได้อย่างมีประสิทธิภาพ

ตารางที่ 2.9 ตารางสรุปงานวิจัยที่เกี่ยวข้องกับเทคนิคการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
S. C. Chin และคณะ [86] การวิเคราะห์อนุกรมเวลาสำหรับการตรวจจับความผิดปกติ: การประเมินการเปรียบเทียบ (Symbolic Time Series Analysis for Anomaly Detection: A comparative Evaluation)	ผู้วิจัยนำเสนอการใช้วิธีการวิเคราะห์อนุกรมเวลาสำหรับการตรวจจับความผิดปกติโดยใช้ตัวแบบมาร์คอฟ (Markov) พร้อมทั้งประเมินการเปรียบเทียบประสิทธิภาพกับเทคนิคการรวมกลุ่มตัวแปร (Principle Component Analysis) และโครงข่ายประสาทเทียม (Neural Networks)	ผลการวิจัยพบว่าวิธีการวิเคราะห์อนุกรมเวลาเพื่อตรวจจับความผิดปกติโดยใช้ตัวแบบมาร์คอฟ (Markov) ร่วมกับโครงข่ายประสาทเทียม (Neural Networks) มีประสิทธิภาพที่ดีกว่าการใช้ตัวแบบมาร์คอฟ (Markov) ร่วมกับเทคนิคการรวมกลุ่มตัวแปร (Principle Component Analysis)
T.Oates และคณะ [103] การกำหนดกระบวนการที่แตกต่างในอนุกรมเวลาแบบหลายตัวแปรโดยเทคนิคการรวมกลุ่ม (Identifying Distinctive Subsequences in Multivariate Time Series by Clustering)	ผู้วิจัยได้นำเสนอแนวทางการรวมกลุ่มข้อมูลอนุกรมเวลาแบบหลายตัวแปรด้วยเทคนิคการรวมกลุ่ม (Clustering Analysis)	การรวมกลุ่มข้อมูลอนุกรมเวลาแบบหลายตัวแปรด้วยเทคนิคการรวมกลุ่ม สามารถทำการรวมกลุ่มข้อมูลอนุกรมเวลาตามรูปแบบที่เหมือนและแตกต่างได้ ส่งผลให้การนำเสนอภาพสามารถแยกข้อมูลอนุกรมเวลาตามรูปแบบที่เหมือนและแตกต่างได้อย่างชัดเจน

ตารางที่ 2.10 ตารางสรุปข้อดีและข้อจำกัดของงานวิจัยที่ได้ทำการศึกษา

ข้อดี-ข้อจำกัดที่ค้นพบ				
(1) การวิเคราะห์พฤติกรรมการใช้งานในเครือข่ายคอมพิวเตอร์	(2) เทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์	(3) เทคนิคการติดตามการแพร่กระจายข้อมูลข่าวสาร	(4) โครงสร้างกราฟ	(5) เทคนิคการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์
ข้อดี ใช้หลักสถิติทำการวิเคราะห์พฤติกรรมการใช้งานเครือข่ายคอมพิวเตอร์ในลักษณะการสอบกลับ (Forensics) ภายหลังการเกิดเหตุการณ์	**ข้อดี** ใช้เทคนิคเหมืองข้อมูลทั้งการเรียนรู้แบบมีผู้สอน และการเรียนรู้แบบไม่มีผู้สอน พร้อมวิธีการทางสถิติในการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์	**ข้อดี** ค้นพบว่าการศึกษาโครงสร้างกราฟสามารถใช้อธิบายรูปแบบการแพร่กระจายข้อมูลข่าวสารในเครือข่ายสังคมออนไลน์	**ข้อดี** ค้นพบว่าสามารถใช้โครงสร้างกราฟเพื่อศึกษาการแพร่กระจายข้อมูลข่าวสารและใช้ทฤษฎีกราฟในการค้นหาหรือพยากรณ์ในเครือข่ายสังคมออนไลน์	**ข้อดี** ใช้เทคนิคเหมืองข้อมูลทั้งการเรียนรู้แบบมีผู้สอน และการเรียนรู้แบบไม่มีผู้สอน ซึ่งใช้รองรับการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์
ข้อจำกัด ไม่ได้ใช้เพื่อการแจ้งเตือนล่วงหน้า แต่เป็นการศึกษาและวิเคราะห์ภายหลังเมื่อเกิดเหตุการณ์	**ข้อจำกัด** ยังไม่มีการประยุกต์ใช้เทคนิคนี้ในกรณีเครือข่ายสังคมออนไลน์อย่างกว้างขวาง	**ข้อจำกัด** ไม่ค่อยมีการประยุกต์ใช้โครงสร้างกราฟกับเทคนิคอื่นๆ	**ข้อจำกัด** ไม่ค่อยมีการประยุกต์ใช้โครงสร้างกราฟและทฤษฎีกราฟกับเทคนิคอื่นๆ	**ข้อจำกัด** ยังไม่มีการประยุกต์ใช้เทคนิคนี้ในกรณีเครือข่ายสังคมออนไลน์อย่างกว้างขวาง
คุณลักษณะของงานวิจัย				
1. ใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟในการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์ 2. ทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์ เพื่อรองรับการหาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์ และรูปแบบการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ 3. พัฒนาตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์ เพื่อรองรับการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริง และการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า				

2.10.2 งานวิจัยของผู้วิจัยและคณะ

จากการวิเคราะห์ข้อดีและข้อจำกัดของงานวิจัยในแต่ละหมวดหมู่ที่ได้ทำการศึกษา ดังแสดงในตารางที่ 2.10 จะได้คุณลักษณะของงานวิจัยที่ควรพัฒนา จากนั้นทางผู้วิจัยและคณะ [73-78] ได้ทำการพัฒนางานวิจัยเพื่อประเมินผลการทดลองเบื้องต้น ดังแสดงในตารางที่ 2.11 และเพื่อสร้างความเชื่อมโยงในกรอบแนวคิดการวิจัย ดังแสดงในตารางที่ 2.12 ดังนี้

ตารางที่ 2.11 ตารางสรุปงานวิจัยของผู้วิจัยและคณะ

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
การติดตามโหนดที่มีอิทธิพลต่อการเกิดความผิดปกติและการถูกโจมตีในเครือข่ายสังคม (Tracking the Influencing Nodes of Anomaly and Attack Patterns in Social Networks) [73]	งานวิจัยนี้มีขั้นตอนวิธีการรวมกลุ่ม (Cluster Analysis) และใช้ในการวิเคราะห์ข้อมูลเข้าสู่ระบบ IDS ที่จะค้นพบความผิดปกติและโอกาสการถูกโจมตีในเครือข่ายสังคมออนไลน์ โดยเทคนิคการรวมกลุ่ม (Cluster Analysis) เป็นเทคนิคการทำเหมืองข้อมูลที่ใช้ในการจัดหมวดหมู่ข้อมูลลงในองค์ประกอบของกลุ่มที่เกี่ยวข้อง ส่วนการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) จะใช้ในการวิจัยนี้เพื่อแสดงภาพการติดตามโหนดที่มีอิทธิพลต่อรูปแบบความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์	ผลการวิเคราะห์รูปแบบผิดปกติและการถูกโจมตีเมื่อใช้เครือข่ายสังคมออนไลน์ ดังนี้ 1. มีผู้โจมตีพยายามในการดำเนินการเปิดโปรแกรมบนระบบที่ติดไวรัส (The attackers attempt to execute an arbitrary program on infected systems) 2. พบความผิดปกติกับ HTTP จากการใช้เครือข่ายสังคมออนไลน์ (An anomaly HTTP) 3. ผู้โจมตีจากระยะไกลสามารถเข้าควบคุมระบบที่มีความเสี่ยงเมื่อใช้งานเครือข่ายสังคมออนไลน์ (The remote attacker from Facebook can gain control of vulnerable systems) 4. พบรูปแบบการปฏิเสธการให้บริการ (DoS) จากการใช้งานเครือข่ายสังคมออนไลน์ (Denial of Service)

ตารางที่ 2.11 ตารางสรุปงานวิจัยของผู้วิจัยและคณะ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
การติดตามโหนดที่มีอิทธิพลต่อการเกิดความผิดปกติและการถูกโจมตีในเครือข่ายสังคม (Tracking the Influencing Nodes of Anomaly and Attack Patterns in Social Networks) [73] (ต่อ)		ผลวิเคราะห์โหนดที่มีอิทธิพล (Influencing Nodes) ในเครือข่ายสังคมออนไลน์ มีผลดังนี้ 1. พบโหนดที่มีอิทธิพล (Influencing Nodes) จำนวน 42 โหนด 2. สมาชิกในเครือข่ายสังคมออนไลน์ที่มีค่าความเป็นศูนย์กลางสูงสุดแบบ In-Degree ห้าอันดับแรก (Top Five In-Degree Nodes) คือโหนดหมายเลข 60, 87, 3, googalz และ 224 3. สมาชิกในเครือข่ายสังคมออนไลน์ที่มีค่าความเป็นศูนย์กลางสูงสุดแบบ Out-Degree ห้าอันดับแรก (Top Five Out-Degree Nodes) คือโหนดหมาย เลข 234, J7, 217, 229 และ 216 4. โหนดที่มีความเป็นผู้เชื่อมโยงสูงสุดห้าอันดับแรก (Top Five Betweenness Nodes) คือโหนดหมายเลข 216, 119, 234, C26 และ J7

ตารางที่ 2.11 ตารางสรุปงานวิจัยของผู้วิจัยและคณะ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
การติดตามโหนดที่มีอิทธิพลต่อการเกิดความผิดปกติและการถูกโจมตีในเครือข่ายสังคม (Tracking the Influencing Nodes of Anomaly and Attack Patterns in Social Networks) [73] (ต่อ)		5. โหนดที่ค่าความใกล้ชิดสูงสุดห้าอันดับแรก (Top Five Closeness Nodes) คือ โหนดหมายเลข 119, 217, 216, 171 และ 234
การทำนายความเชื่อมโยงในเครือข่ายสังคมด้วยเทคนิคการวิเคราะห์เครือข่ายสังคมและเทคนิคการเรียนรู้แบบมีผู้สอน (Links Prediction in Social Networks by SNA and Supervised Learning) [74]	ในงานวิจัยนี้มุ่งเน้นไปที่การทำนายการเชื่อมโยงในเครือข่ายสังคมออนไลน์ โดยการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) และ การเรียนรู้แบบมีผู้สอน (Supervised Learning) ของตัวแบบของข้อมูลชนิดเรียงซ้อนอย่างอิสระ (Independent Cascade Model)	1. สามารถใช้การวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) ทำการแบ่งกลุ่มโหนดต่างๆ (Nodes) เป็น 2 กลุ่ม คือ กลุ่มโหนดที่มีอิทธิพล (Influencing Nodes: IN) และโหนดอื่นๆ (Another Nodes: AN) 2. สามารถใช้ เทคนิคการจัดประเภท (Classification) ทำการทำนายการเชื่อมโยงความสัมพันธ์ใหม่ที่อาจเกิดขึ้นระหว่างโหนดในกลุ่ม **IN** ซึ่งหมายถึงกลุ่มของโหนดที่เป็นโหนดอิทธิพล (Influencing Nodes) 3. การทดสอบความแม่นยำในการทำนายการเชื่อมโยงความสัมพันธ์ของโหนดอิทธิพลในการแพร่กระจายข่าวสาร

ตารางที่ 2.11 ตารางสรุปงานวิจัยของผู้วิจัยและคณะ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
การทำนายความเชื่อมโยงในเครือข่ายสังคมด้วยเทคนิคการวิเคราะห์เครือข่ายสังคมและเทคนิคการเรียนรู้แบบมีผู้สอน (Links Prediction in Social Networks by SNA and Supervised Learning) [74] (ต่อ)		ข้อมูล การถูกโจมตีหรือการถูกภัยคุกคามจากการใช้เครือข่ายสังคมออนไลน์ กับเทคนิคการจัดประเภท (Classification) จำนวน 3 วิธีได้แก่ J48, oneR และ Naïve Bayes พบว่า OneR มีความแม่นยำในการทำนายการเชื่อมโยงของโหนดใดๆ โดยที่ oneR มีข้อมูลเวลาการเรียนรู้ (Learning Time) ที่มีความเสถียรมากที่สุด ดังนั้นการใช้ oneR ในการทำนายการเชื่อมโยงของโหนดใดๆ เป็นวิธีที่ดีที่สุด
การวิเคราะห์ความผิดปกติและการถูกโจมตีในเครือข่ายสังคม (Social Networks Anomaly and Attack Patterns Analysis) [76]	งานวิจัยได้พัฒนารูปแบบการวิเคราะห์ด้วยเทคนิคการรวมกลุ่ม (Cluster Algorithm) เพื่อทำการวิเคราะห์ความผิดปกติของการถูกโจมตีในเครือข่ายสังคมออนไลน์ โดยการคำนวณการวัดระยะทาง (Distance Measure) และ K-Means Algorithm	ผลงานวิจัยสามารถสรุปความผิดปกติและการถูกโจมตีได้ดังนี้ 1. ผู้โจมตีจากเครือข่ายสังคมออนไลน์ googalz พยายามในการดำเนินการเปิดโปรแกรมบนระบบที่ติดไวรัส 2. พบรูปแบบความผิดปกติกับ HTTP จากการใช้เครือข่ายสังคมออนไลน์ Facebook, imeem, YouTube, Hi5 3. มีผู้โจมตีจากระยะไกลสามารถเข้าควบคุมระบบที่มี

ตารางที่ 2.11 ตารางสรุปงานวิจัยของผู้วิจัยและคณะ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
การวิเคราะห์ความผิดปกติและการถูกโจมตีในเครือข่ายสังคม (Social Networks Anomaly and Attack Patterns Analysis) [76] (ต่อ)		ความเสี่ยงเมื่อใช้งานเครือข่ายสังคมออนไลน์ Facebook 4. พบการปฏิเสธของการให้บริการ (DoS) จากการใช้งานเครือข่ายสังคมออนไลน์ Skype และ MSN
การแสดงภาพวิชวลไลเซชันของการแพร่กระจายข่าวสารในเครือข่ายสังคม (Visualization of Information Diffusion in Social Networks) [75]	ผลงานวิจัยชิ้นนี้ใช้การวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) ที่มีประสิทธิภาพในการวิเคราะห์ปัจจัยสำคัญที่มีอิทธิพลต่อรูปแบบการแพร่กระจายข้อมูลเกี่ยวกับความหนาแน่น (Density) ความเป็นศูนย์กลาง (Centrality) และกลุ่มย่อยการเชื่อมโยงข้อมูล (Cohesive Subgroup) เผยให้เห็นว่าเทคนิคการรวมกลุ่มที่ใช้ในการจัดการเครือข่ายในโลกจริงกับเครือข่ายขนาดเล็กเรียกว่า Egocentric Networks	ในงานวิจัยนี้ มีพัฒนาระบบการทำเหมืองข้อมูล ด้วยเทคนิคการรวมกลุ่ม เพื่อใช้วิเคราะห์หารูปแบบความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ พบการถูกโจมตี ดังนี้ 1. มีผู้โจมตีพยายามในการดำเนินการเปิดโปรแกรมบนระบบที่ติดไวรัส (The attackers attempt to execute an arbitrary program on infected systems) 2. พบความผิดปกติกับ HTTP จากการใช้เครือข่ายสังคมออนไลน์ (An anomaly HTTP) 3. ผู้โจมตีจากระยะไกลสามารถเข้าควบคุมระบบที่มีความเสี่ยง

ตารางที่ 2.11 ตารางสรุปงานวิจัยของผู้วิจัยและคณะ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
การแสดงภาพวิชวลไลเซชันของการแพร่กระจายในเครือข่ายสังคม (Visualization of Information Diffusion in Social Networks) [75] (ต่อ)		เมื่อใช้งานเครือข่ายสังคมออนไลน์ (The remote attacker from Facebook can gain control of vulnerable systems) 4. พบรูปแบบการปฏิเสธการให้บริการ (DoS) จากการใช้งานเครือข่ายสังคมออนไลน์ (Denial of Service) ทั้งนี้เทคนิคการวิเคราะห์เครือข่ายทางสังคม (Social Network Analysis: SNA) ถูกใช้ในงานวิจัยนี้สำหรับการแสดงภาพการติดตามโหนดที่มีอิทธิพล (Influencing Node) ที่มีความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ โดย SNA ถูกใช้เพื่อวิเคราะห์ปัจจัยสำคัญที่มีอิทธิพลต่อรูปแบบการแพร่กระจายข้อมูลโดยใช้ค่าความหนาแน่น (Density) ความเป็นศูนย์รวม (Centrality) และความเป็นเครือข่ายย่อย (Cohesive Subgroup)

ตารางที่ 2.11 ตารางสรุปงานวิจัยของผู้วิจัยและคณะ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
การศึกษาการแพร่กระจายไวรัสคอมพิวเตอร์ในเครือข่ายสังคมกรณีศึกษาศูนย์บริการโลหิตแห่งชาติ สภากาชาดไทย (Study of Computer Virus Distribution in Social Network: A Case Study of National Blood Centre, Thai Red Cross Society) [77]	ใช้สถิติเชิงบรรยาย (Descriptive Statistics) และสถิติอนุมาน (Inferential Statistics) ในการวิเคราะห์พฤติกรรมการใช้งานเครือข่ายสังคมออนไลน์ กรณีศึกษาศูนย์บริการโลหิตแห่งชาติ สภากาชาดไทย	การศึกษาเรื่องการแพร่กระจายไวรัสในเครือข่ายสังคมออนไลน์ กรณีศึกษาศูนย์บริการโลหิตแห่งชาติ สภากาชาดไทยพบว่า 1. สัดส่วนการใช้งานเครือข่ายสังคมออนไลน์ของบุคลากรศูนย์บริการโลหิตแห่งชาติ สภากาชาดไทยนิยมใช้ MySpace มากที่สุด โดยมีสัดส่วน 40.58% รองลงมาคือ Hi5 สัดส่วน 35.12% และ Facebook 23.30% 2. จากการใช้งานเครือข่ายสังคมออนไลน์พบการแพร่กระจายของไวรัสชนิด JS/PackRedir.A!tr.dldr สูงถึง 90.36% ของการแพร่กระจายไวรัสทั้งหมด ซึ่งมาจากการใช้งานของ IP Address 172.16.10.96 โดยทำการแพร่กระจายไวรัส JS/PackRedir.A!tr.dldr สูงถึง 94% ของเครื่องคอมพิวเตอร์ที่มีการแพร่กระจายไวรัสชนิดดังกล่าวด้วยกัน

ตารางที่ 2.11 ตารางสรุปงานวิจัยของผู้วิจัยและคณะ (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	ผลการวิจัย
การวิเคราะห์ความผิดปกติและการถูกโจมตีในเครือข่ายสังคมด้วยการสร้างกฎความสัมพันธ์ (Social Networks Anomaly and Attack Patterns Analysis with Association Rules) [78]	งานวิจัยชิ้นนี้ได้ศึกษาถึงการวิเคราะห์รูปแบบในการบุกรุกเครือข่ายสังคมออนไลน์ (Intrusion Detection System: IDS) ด้วยการนำเสนอตัวแบบการทำเหมืองข้อมูล ด้วยขั้นตอนวิธี Apriori พร้อมทำการศึกษาความแม่นยำของวิธีการวิเคราะห์รูปแบบการบุกรุกในเครือข่ายสังคมออนไลน์ โดยการกำหนดพารามิเตอร์ของกฎความสัมพันธ์ (Parameters of Association Rules) ได้แก่ Associator คือ Apriori, Lower Bound for Minimum Support คือ 0.1 และ Minimum Confidence Metric คือ 0.9	งานวิจัยชิ้นนี้ได้ผลการวิจัยสอดคล้องกับงานวิจัย Social Network Anomaly and Attack Patterns Analysis และผลการวิเคราะห์ประสิทธิภาพของตัวแบบการ คือ 87.50% โดยผลงานวิจัยสามารถสรุปได้ว่า ความผิดปกติของการถูกโจมตีในเครือข่ายสังคมออนไลน์ เป็นดังนี้ 1. ผู้โจมตีจาก googalz พยายามในการดำเนินการเปิดโปรแกรมบนระบบที่ติดไวรัส 2. พบรูปแบบความผิดปกติกับ HTTP จาก Facebook, imeem, YouTube, Hi5 3. มีผู้โจมตีจากระยะไกลสามารถเข้าควบคุมระบบที่มีความเสี่ยงเมื่อใช้งาน Facebook 4. พบการปฏิเสธของการให้บริการ (DoS) จากการใช้งาน Skype และ MSN

ตารางที่ 2.12 ตารางสรุปความเชื่อมโยงในกรอบแนวคิดการวิจัย

งานวิจัย	ระเบียบวิธีวิจัย	วัตถุประสงค์การวิจัย
การศึกษาการแพร่กระจายไวรัสคอมพิวเตอร์ในเครือข่ายสังคมกรณีศึกษาศูนย์บริการโลหิตแห่งชาติ สภากาชาดไทย (Study of Computer Virus Distribution in Social Network: A Case Study of National Blood Centre, Thai Red Cross Society) [77]	สถิติเชิงบรรยาย (Descriptive Statistics) และสถิติอนุมาน (Inferential Statistics) ในการวิเคราะห์พฤติกรรมการใช้งานเครือข่ายสังคมออนไลน์	เพื่อศึกษาถึงความเสี่ยงขององค์กรที่อนุญาตให้บุคลากรใช้งานเครือข่ายสังคมออนไลน์ด้วยสถิติวิเคราะห์ (Statistic Analysis)
การวิเคราะห์ความผิดปกติและการถูกโจมตีในเครือข่ายสังคม (Social Networks Anomaly and Attack Patterns Analysis) [76]	เทคนิคเหมืองข้อมูล โดยใช้เทคนิควิเคราะห์การรวมกลุ่ม (Cluster Analysis)	เพื่อศึกษารูปแบบความผิดปกติและโอกาสการถูกโจมตีเมื่อมีการใช้งานเครือข่ายสังคมออนไลน์ด้วยเทคนิควิเคราะห์การรวมกลุ่ม (Cluster Analysis)
การวิเคราะห์ความผิดปกติและการถูกโจมตีในเครือข่ายสังคมด้วยการสร้างกฎความสัมพันธ์ (Social Networks Anomaly and Attack Patterns Analysis with Association Rules) [78]	เทคนิคเหมืองข้อมูล โดยใช้เทคนิคการสร้างกฎความสัมพันธ์ (Association Rules)	เพื่อศึกษารูปแบบความผิดปกติและโอกาสการถูกโจมตีเมื่อมีการใช้งานเครือข่ายสังคมออนไลน์ด้วยเทคนิคการสร้างกฎความสัมพันธ์ (Association Rules) โดยได้ผลการทดลองไม่แตกต่างจากการใช้เทคนิควิเคราะห์การรวมกลุ่ม (Cluster Analysis) ซึ่งมีเวลาในการประมวลผลข้อมูลที่เร็วกว่า

ตารางที่ 2.12 ตารางสรุปความเชื่อมโยงในกรอบแนวคิดการวิจัย (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	วัตถุประสงค์การวิจัย
การแสดงภาพวิชวลไลเซชันของการแพร่กระจายในเครือข่ายสังคม (Visualization of Information Diffusion in Social Networks) [75]	เทคนิควิเคราะห์การรวมกลุ่ม (Cluster Analysis) และ การวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA)	เพื่อศึกษารูปแบบการแพร่กระ-จายความผิดปกติและโอกาสการถูกโจมตีของการใช้งานเครือข่ายสังคมออนไลน์ ด้วยเทคนิควิเคราะห์การรวมกลุ่ม (Cluster Analysis) และ การวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) พบว่าการแสดงผลวิชวลไลเซชันสร้างความเข้าใจภาพการแพร่กระจายในเครือข่ายสังคมเป็นอย่างดี
การทำนายความเชื่อมโยงในเครือข่ายสังคมด้วยเทคนิคการวิเคราะห์เครือข่ายสังคมและเทคนิคการเรียนรู้แบบมีผู้สอน (Links Prediction in Social Networks by SNA and Supervised Learning) [74]	เทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) และเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning, Classification Analysis, J48, oneR, Naïve Bay)	เพื่อศึกษาถึงเทคนิคการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงและ เพื่อศึกษาถึงเทคนิคการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีเมื่อใช้เครือข่ายสังคมออนไลน์ พบว่าวิธี oneR มีประสิทธิภาพทางเวลาและการทำนายที่แม่นยำกว่าวิธีการ J48 และ Naïve Bay

ตารางที่ 2.12 ตารางสรุปความเชื่อมโยงในกรอบแนวคิดการวิจัย (ต่อ)

งานวิจัย	ระเบียบวิธีวิจัย	วัตถุประสงค์การวิจัย
การติดตามโหนดที่มีอิทธิพลต่อการเกิดความผิดปกติและการถูกโจมตีในเครือข่ายสังคม (Tracking the Influencing Nodes of Anomaly and Attack Patterns in Social Networks) [73]	เทคนิควิเคราะห์การรวมกลุ่ม (Cluster Analysis) และเทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA)	การศึกษารูปแบบความผิดปกติและโอกาสการถูกโจมตีเมื่อมีการใช้งานเครือข่ายสังคมออนไลน์ด้วยเทคนิควิเคราะห์การรวมกลุ่ม (Cluster Analysis) พร้อมทั้งศึกษาถึงเทคนิคการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ด้วยเทคนิคการวิเคราะห์เครือข่ายสังคม (Social Network Analysis: SNA) พบว่าการคัดกรองโหนดให้เหลือเฉพาะโหนดที่มีอิทธิพลต่อการแพร่กระจายข่าวสารในเครือข่ายสังคมออนไลน์ หรือโหนดที่เป็นผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ ด้วยเทคนิคการวิเคราะห์เครือข่ายสังคม ใช้สืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์

2.11 สรุป

ในบทที่ 2 ได้นำเสนอแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้องกับกรอบแนวคิดในการวิจัย เพื่อที่จะค้นหากลยุทธ์การติดตามภัยจากการใช้เครือข่ายสังคมออนไลน์จากตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ ซึ่งงานวิจัยนี้มีขั้นตอนการดำเนินงานวิจัยโดยเริ่มจากการศึกษาจากเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้องได้แก่ งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์พฤติกรรมการใช้งานในเครือข่ายคอมพิวเตอร์ งานวิจัยที่เกี่ยวข้องกับเทคนิคการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ งานวิจัยที่เกี่ยวข้องกับเทคนิคการติดตามการแพร่กระจายข้อมูลข่าวสาร งานวิจัยที่เกี่ยวข้องกับโครงสร้างกราฟ และงานวิจัยที่เกี่ยวข้องกับเทคนิคการทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายคอมพิวเตอร์ เพื่อพิจารณาข้อดีและข้อจำกัดของงานวิจัยเหล่านั้น สำหรับกำหนดคุณลักษณะของงานวิจัยชิ้นนี้ กล่าวคือ

1. ใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟในการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์
2. ทำให้เห็นภาพการวิเคราะห์ความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์ เพื่อรองรับการหาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์ และรูปแบบการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์
3. พัฒนาตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์ เพื่อรองรับการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริง และการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า

จากนั้นได้ทำการศึกษาถึงทฤษฎีเทคนิคเหมืองข้อมูลและโครงสร้างกราฟ เพื่อนำมาประยุกต์ใช้กับการพัฒนาตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ เพื่อศึกษาและวิเคราะห์หาความผิดปกติและ โอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์ ศึกษาและวิเคราะห์การแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ พัฒนาตัวแบบการสืบค้นผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ พร้อมทั้งพัฒนาตัวแบบการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า

สุดท้ายเป็นการพัฒนาตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ ซึ่งการทราบถึงเหตุการณ์ปัจจุบัน และประเมินหาความเชื่อมโยงในเครือข่ายสังคมออนไลน์เป็นประโยชน์มากในการวิเคราะห์ และความเข้าใจในเครือข่ายสังคมออนไลน์ดังกล่าว สามารถนำไปสู่การดำเนินงานที่มีประสิทธิภาพของเครื่องมือในการสืบค้นผู้กระทำผิดปกติหรือมีพฤติกรรมการโจมตี การค้นหาเป้าหมายที่อาจจะได้รับผลกระทบจากการโจมตีนั้นๆ รวมทั้งการระบุกลุ่มที่ซ่อนอยู่หรือการหาสมาชิกที่หายไปของกลุ่ม ฯลฯ ในเครือข่ายสังคมออนไลน์ ซึ่งเป็นปัญหาที่พบบ่อยที่สุดในการรักษาความมั่นคงปลอดภัยและการสอบสวนทางอาญา