บทที่ 3

วิธีดำเนินการวิจัย

การศึกษาและวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อศึกษาและวิเคราะห์หาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์ ศึกษาและวิเคราะห์การแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ พัฒนาตัวแบบในการสืบค้นผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ พร้อมทั้งพัฒนาตัวแบบในการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า รวมทั้งการพัฒนาตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ เพื่อที่จะค้นหากลยุทธ์การติดตามภัยจากการใช้เครือข่ายสังคมออนไลน์ ซึ่งการทราบถึงเหตุการณ์ปัจจุบันและประเมินหาการเชื่อมโยงในเครือข่ายสังคมออนไลน์เป็นประโยชน์มากในการวิเคราะห์ และความเข้าใจในเครือข่ายสังคมออนไลน์ดังกล่าว สามารถนำไปสู่การดำเนินงานที่มีประสิทธิภาพของเครื่องมือในการสืบค้นผู้กระทำผิดปกติหรือมีพฤติกรรมการโจมตี การค้นหาเป้าหมายที่อาจจะได้รับผลกระทบจากการโจมตีนั้นๆ รวมทั้งการระบุกลุ่มที่ซ่อนอยู่หรือการหาสมาชิกที่หายไปของกลุ่ม ฯลฯ ในเครือข่ายสังคมออนไลน์ ซึ่งเป็นปัญหาที่พบบ่อยที่สุดในการรักษาความมั่นคงปลอดภัยและการสอบสวนทางอาญา โดยมีขั้นตอนการดำเนินการวิจัยดังนี้

3.1 กำหนดขั้นตอนการวิจัย

3.2 กำหนดเครื่องมือและอุปกรณ์ที่ใช้ในการวิจัย

3.3 กำหนดประชากรและกลุ่มตัวอย่าง

3.4 กำหนดการวิเคราะห์ข้อมูล

3.5 สรุป

3.1 ขั้นตอนการวิจัย

3.1.1 กำหนดขอบเขตการศึกษาและกลุ่มข้อมูล

การศึกษาครั้งนี้ ได้ใช้ฐานข้อมูลความผิดปกติ (Anomaly) และการถูกโจมตี (Attack) ของการใช้เครือข่ายสังคมออนไลน์ขององค์กรที่ต้องการความมั่นคงปลอดภัยในระดับสูง เนื่องจากต้องป้องกันความเสี่ยงต่อข้อมูลและระบบสารสนเทศขององค์กรที่รองรับระบบงานทางด้านสุขภาพในระดับประเทศ (Public Health Sector in Thailand) จำนวน 2 ประเทศ ดังนี้

1. ข้อมูลของประเทศไทย ซึ่งประกอบไปด้วยสำนักงานใหญ่ที่กรุงเทพมหานคร 1 แห่ง สำหรับรองรับงานทางด้านสุขภาพของทั้ง 76 จังหวัดทั่วประเทศ

2. ข้อมูลของประเทศที่ร่วมทดสอบ ซึ่งประกอบด้วยสำนักงานใหญ่ที่เมืองหลวงของประเทศ 1 แห่งที่มีโครงการวิจัยร่วมกับประเทศไทยในประเด็นการปลูกถ่ายเซลล์ต้นกำเนิดเม็ดโลหิตสำหรับรองรับงานทางด้านสุขภาพของทั้งประเทศ

โดยผู้วิจัยวางแผนการเลือกตัวอย่างฐานข้อมูลความผิดปกติ (Anomaly) และการถูกโจมตี (Attack) ของการใช้เครือข่ายสังคมออนไลน์ซึ่งเป็น Traffic Log ด้วยหลักการของการเลือกตัวอย่างตามความน่าจะเป็น (Probability Sampling) ดังนี้

1. การเลือกตัวอย่างแบบกลุ่ม (Cluster Sampling)

การเลือกตัวอย่างแบบกลุ่ม (Cluster Sampling) เป็นการเลือกตัวอย่างแบบกลุ่มจะให้ภายในแต่ละกลุ่มมีลักษณะที่สนใจทุกลักษณะ หรือมีลักษณะที่แตกต่างคละกัน ดังนั้นการเลือกตัวอย่างแบบกลุ่มจึงเป็นการเลือกตัวอย่างเพียงบางกลุ่มมาทำการศึกษาเท่านั้น โดยผู้วิจัยได้กำหนดกลุ่มย่อย (Cluster Factor) ตามเดือน

2. การเลือกตัวอย่างสุ่มแบบง่าย (Simple Random Sampling)

เป็นการเลือกตัวอย่างที่ให้แต่ละหน่วยในประชากรมีโอกาสถูกเลือกเท่าๆกันในแต่ละครั้งของการเลือก โดยผู้วิจัยได้ทำการเลือกตัวอย่างแบบง่ายจากการเลือกตัวอย่างแบบกลุ่มได้เดือนเมษายน และธันวาคม

เนื่องจากงานวิจัยนี้ได้ทำการศึกษาตั้งแต่พฤศจิกายน 2554-2556 จึงทำการเลือกตัวอย่างในขั้นตอนนี้กำหนดขอบเขตการสุ่มตัวอย่างฐานข้อมูลความผิดปกติ (Anomaly) และการ

ถูกโจมตี (Attack) ของการใช้เครือข่ายสังคมออนไลน์ของ 2 ประเทศ คือ เดือนธันวาคม 2554 เดือนเมษายน 2555 เดือนธันวาคม 2555 และเดือนเมษายน 2556

ทั้งนี้ ข้อมูลที่ใช้ในการวิจัยนี้คือ ข้อมูลบันทึกการใช้งานเครือข่ายสังคมออนไลน์ขององค์กรภาคสาธารณสุขที่ให้ความสำคัญต่อความเป็นส่วนตัว (Privacy) และความปลอดภัย (Security) ของระบบข้อมูลอย่างสูง และต้องการหลีกเลี่ยงจากอันตรายของโปรแกรมประยุกต์ของเครือข่าย ซึ่งเป็นข้อมูลบันทึกเหตุการณ์ (Log File) ของระบบตรวจจับผู้บุกรุก (IDS) ขององค์กรจำนวน 2 ประเทศ คือประเทศไทย และประเทศที่ร่วมทดสอบที่ได้รับความผิดปกติ(Anomaly) และการถูกโจมตี (Attack) เมื่อใช้งานเครือข่ายสังคมออนไลน์ และฐานข้อมูลความสัมพันธ์ของสมาชิกที่ได้รับความผิดปกติ (Anomaly) และการถูกโจมตี (Attack) เมื่อใช้งานเครือข่ายสังคมออนไลน์นั้นโดยเป็นข้อมูลบันทึกการใช้งานเครือข่าย (Log File) ในงานวิจัยนี้ใช้ข้อมูล 4 ชุดคือชุดข้อมูลในเดือนธันวาคม 2554 เดือนเมษายน 2555 เดือนธันวาคม 2555 และเดือนเมษายน 2556 และตัวอย่างข้อมูลของบันทึกการใช้งานเครือข่ายดังแสดงในตาราง 3.1

ตารางที่ 3.1 ตัวอย่างข้อมูลของบันทึกการใช้งานเครือข่าย

date	2011-12-13
time	13:03:39
attack_id	10873
severity	High
src	172.16.10.96
dst	172.16.10.137
src_port	2156
dst_port	80
msg	The remote attackers can gain control of vulnerable systems

3.1.2 ปรับลักษณะข้อมูลให้เหมาะสมกับการวิเคราะห์

จากตัวอย่างข้อมูลของบันทึกการใช้งานเครือข่ายในตารางที่ 3.1 นอกเหนือจากวันและเวลา ยังมีคุณลักษณะของแฟ้มข้อมูลบันทึกการใช้งานเครือข่าย (Log File) ดังนี้

1. รหัสการโจมตี (Attack_ID)

2. ลักษณะของการโจมตีของผู้บุกรุก (Signature)

3. ความรุนแรง (Severity) แสดงระดับของความเสี่ยง

4. แอดเดรสต้นทาง (SRC)

5. แอดเดรสปลายทาง (DST)

6. พอร์ตต้นทาง (SRC_Port)

7. พอร์ตปลายทาง (DST_Port)

8. คำอธิบายเพิ่มเติมของลักษณะของการโจมตีของผู้บุกรุก (Msg) แสดงความหมายของการแจ้งเตือน

จากตารางที่ 3.1 แสดงตัวอย่างข้อมูลของบันทึกการใช้งานเครือข่าย ซึ่งมีความหมายดังนี้ รหัสการถูกโจมตีเลขหมาย 10873 จากแอดเดรสต้นทางหมายเลข 172.16.10.96 จากหมายเลขพอร์ตต้นทาง 2156 ไปยังแอดเดรสปลายทางหมายเลข 172.16.10.137 ที่หมายเลขพอร์ตปลายทาง 80 โดยมีลักษณะการโจมตีแบบ “ผู้โจมตีจากระยะไกลสามารถเข้าควบคุมระบบที่มีความเสี่ยง”

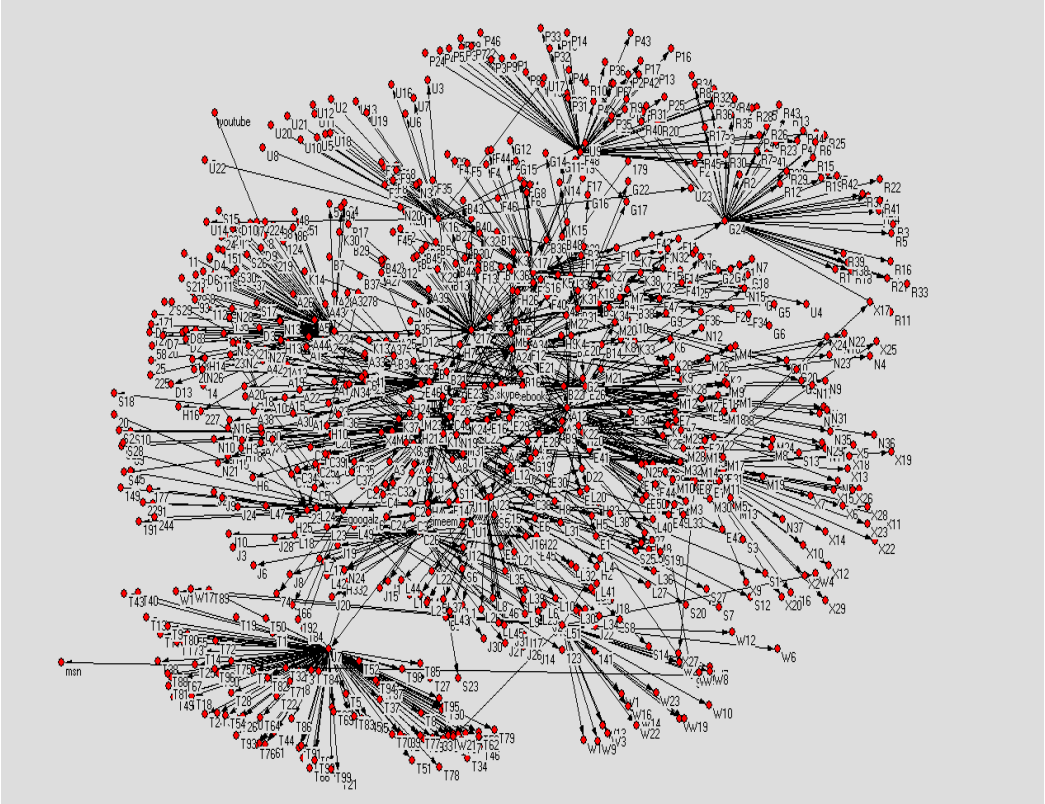
ตามที่กล่าวถึงข้างต้น การวิจัยนี้เน้นการค้นหาการตรวจจับการบุกรุกจากข้อมูลบันทึกการใช้เครือข่ายสังคมออนไลน์เพื่อระบุ และค้นพบรูปแบบไม่รู้จักแต่มีลักษณะที่อาจเป็นการบุกรุก โดยการเตรียมข้อมูลในการวิเคราะห์ ดังนี้

1. ผู้วิจัยได้ทำการปรับเพิ่มเติมข้อมูลบันทึกการใช้งานเครือข่ายกับข้อมูลอื่น ๆ ที่เกี่ยวข้องที่ไม่พบในแฟ้มข้อมูลดิบของข้อมูลบันทึกการใช้งานเครือข่าย (Log File) เช่นชนิดของเว็บ (Web 1.0, Web 2.0) ชื่อเครือข่ายทางสังคมในแต่ละกลุ่มที่อยู่ปลายทาง IP จริง (Real IP) และอื่น ๆ

2. ข้อมูลบันทึกการถูกโจมตีในเดือนธันวาคม 2554 เดือนเมษายน 2555 เดือนธันวาคม 2555 และเดือนเมษายน 2556 จะถูกคัดกรองให้เหลือเฉพาะข้อมูลบันทึกของการใช้งานเครือข่ายสังคมออนไลน์ (Web 2.0) เพื่อวิเคราะห์และค้นหารูปแบบที่ไม่รู้จักโดยใช้เทคนิคเหมืองข้อมูลด้วย Cluster Analysis, Classification for Prediction OneR อัลกอริทึม และโครงสร้างกราฟเพื่อการวิเคราะห์เครือข่ายทางสังคม (Social Network Analysis: SNA)

3. งานวิจัยนี้ได้ผสานข้อมูลความสัมพันธ์หรือความเชื่อมโยงในเครือข่ายสังคมออนไลน์ของสมาชิกผู้ถูกโจมตี เมื่อใช้งานเครือข่ายสังคมออนไลน์ ความสัมพันธ์หรือความ

เชื่อมโยงในเครือข่ายสังคมออนไลน์ในโลกจริงมีขนาดใหญ่เกินไปที่จะแสดงให้เห็นอย่างเข้าใจถึงความสัมพันธ์ ดังแสดงในรูปที่ 3.1 ซึ่งเป็นตัวอย่างของความสัมพันธ์หรือความเชื่อมโยงในเครือข่ายสังคมออนไลน์กับโหนด (Nodes) เดือนธันวาคม 2554 จำนวน 7,035 และเส้นความสัมพันธ์ (Edges) จำนวน 25,778 ซึ่งจะเป็นเรื่องยากมากขึ้นอีกถ้าจะต้องทำการอ่านความสัมพันธ์หรือความเชื่อมโยงในเครือข่ายเมื่อสมาชิกเพิ่มขึ้น รวมทั้งทำการค้นหากลุ่มที่ซ่อนอยู่หรือหาสมาชิกที่หายไปของกลุ่ม ฯลฯ ซึ่งเป็นปัญหาที่พบบ่อยที่สุดในการรักษาความมั่นคงปลอดภัยและการสอบสวนทางอาญา



รูปที่ 3.1 ภาพความสัมพันธ์ของการใช้งานเครือข่ายสังคมโดยใช้การจำลองตามทฤษฎีกราฟ [73]

3.1.3 กำหนดวิเคราะห์ความมั่นคงปลอดภัยในการใช้เครือข่ายสังคมออนไลน์

การวิเคราะห์ความมั่นคงปลอดภัยในการใช้เครือข่ายสังคมออนไลน์ เพื่อนำมาเป็นฐานในการค้นหากลยุทธ์การติดตามภัยจากการใช้เครือข่ายสังคมออนไลน์ด้วยการพัฒนาตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคม

ออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ ในวัตถุประสงค์ดังต่อไปนี้ ศึกษาและวิเคราะห์หาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์ ศึกษาและวิเคราะห์การแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ พร้อมทั้งพัฒนาตัวแบบในการสืบค้นผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ และพัฒนาตัวแบบในการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า รวมทั้งพัฒนาตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ ซึ่งการทราบถึงเหตุการณ์ปัจจุบัน และประเมินหาการเชื่อมโยงในเครือข่ายสังคมออนไลน์เป็นประโยชน์มากในการวิเคราะห์ และความเข้าใจในเครือข่ายสังคมออนไลน์ดังกล่าว สามารถนำไปสู่การดำเนินงานที่มีประสิทธิภาพของเครื่องมือในการสืบค้นผู้กระทำผิดปกติหรือมีพฤติกรรมการโจมตี การค้นหาเป้าหมายที่อาจจะได้รับผลกระทบจากการโจมตีนั้นๆ รวมทั้งการระบุกลุ่มที่ซ่อนอยู่หรือการหาสมาชิกที่หายไปของกลุ่ม ฯลฯ ในเครือข่ายสังคมออนไลน์ ซึ่งเป็นปัญหาที่พบบ่อยที่สุดในการรักษาความมั่นคงปลอดภัยและการสอบสวนทางอาญา

โดยผู้วิจัยได้ศึกษาหลักการต่างๆ ได้แก่ สถิติวิเคราะห์, Social Network Analysis (SNA), Data Mining (Unsupervised Learning, Cluster Analysis, Distance Measure, K-Means Algorithm, Supervised Learning, Classification Analysis, oneR Algorithm)

3.1.4 ทำการวิเคราะห์ความมั่นคงปลอดภัยในการใช้เครือข่ายสังคมออนไลน์ โดยสามารถรองรับวัตถุประสงค์ดังนี้

1. เพื่อทำการวิเคราะห์หาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์
2. เพื่อศึกษารูปแบบการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์
3. เพื่อพัฒนาตัวแบบในการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์
4. เพื่อพัฒนาตัวแบบในการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า

3.1.5 การทดสอบประเมินความแม่นยำของผลการวิเคราะห์ฯ

การทดสอบประเมินความแม่นยำของผลการวิเคราะห์ตามที่ระบุในข้อ 3.1.4 คือ การแสดงรูปแบบการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ การพัฒนาตัวแบบในการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ การพัฒนาตัวแบบในการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในสังคมเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า โดยการทดสอบสมมติฐานที่มีความละเอียดถูกต้อง (Accuracy) อย่างมีนัยสำคัญภายใต้ระดับความเชื่อมั่น 95 %

3.1.6 สรุปผลการทดลอง

สรุปผลการทดลองและนำเสนอตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ

3.2 เครื่องมือและอุปกรณ์ที่ใช้ในการวิจัย

3.2.1 ฮาร์ดแวร์ที่ใช้ในการวิจัย ประกอบด้วย

1. ซีพียู Intel Core 2 Duo
2. ฮาร์ดดิสก์ความจุไม่ต่ำกว่า 320 GB
3. หน่วยความจำไม่ต่ำกว่า 4 GB

3.2.2 ซอฟต์แวร์ที่ใช้ในการวิจัย ประกอบด้วย

1. ระบบปฏิบัติการ Microsoft Windows 7
2. ระบบจัดการฐานข้อมูล MySQL
3. Pajek Software for Windows Version 1.07
4. Weka for Windows Version 3.4.11
5. SPSS for Windows Version 17.0
6. NBC Dataset เดือนธันวาคม 2554

3.2.3 การวิเคราะห์งานวิจัย

3.2.3.1 สถิติวิเคราะห์ (Statistic Analysis)

สถิติที่ใช้กันอยู่ในทางวิจัย แบ่งออกได้เป็น 2 ประเภทใหญ่ ๆ คือ

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

3.2.3.2 Social Network Analysis (SNA)

เทคนิคการวัดที่ใช้ในการวิเคราะห์เครือข่ายทางสังคมจะขึ้นอยู่กับหลักการของทฤษฎีกราฟ (Graph Theorem) ซึ่งประกอบด้วยชุดของสูตรทางคณิตศาสตร์และแนวความคิดในการศึกษารูปแบบของเส้น บุคคลจะเป็นโหนดหรือจุดในกราฟ (Node) และความสัมพันธ์ระหว่างบุคคลหรือจุดจะเป็นเส้นกราฟ (Edge) ระหว่างโหนดหรือจุดนั้นๆ ของเครือข่ายทางสังคมที่เรียกว่า Sociograms ทั้งนี้ในการประเมินเพื่อหาโหนดที่มีอิทธิพลต่อโหนดอื่นๆ (Influencing Nodes) โดยการค้นหาความเป็นศูนย์กลาง (Centrality)

3.2.3.3 Data Mining

การทำเหมืองข้อมูลหรืออาจจะเรียกได้ว่าการค้นหาความรู้ในฐานข้อมูล (Knowledge Discovery in Databases – KDD) เป็นเทคนิคเพื่อค้นหารูปแบบ (Pattern) จากข้อมูลจำนวนมหาศาลโดยอัตโนมัติ โดยใช้ขั้นตอนวิธีจากวิชาสถิติ (Statistics) การเรียนรู้ของเครื่อง (Machine Learning) และ การรู้จำแบบ (Pattern Recognition) หรือในอีกนิยามหนึ่ง การทำเหมือง

ข้อมูล คือ กระบวนการที่กระทำกับข้อมูล (โดยส่วนใหญ่จะมีจำนวนมาก) เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักสถิติ (Statistics) การรู้จำ (Pattern Recognition) การเรียนรู้ของเครื่อง (Machine Learning) และหลักคณิตศาสตร์ (Mathematics)

3.3 ประชากรและกลุ่มตัวอย่าง

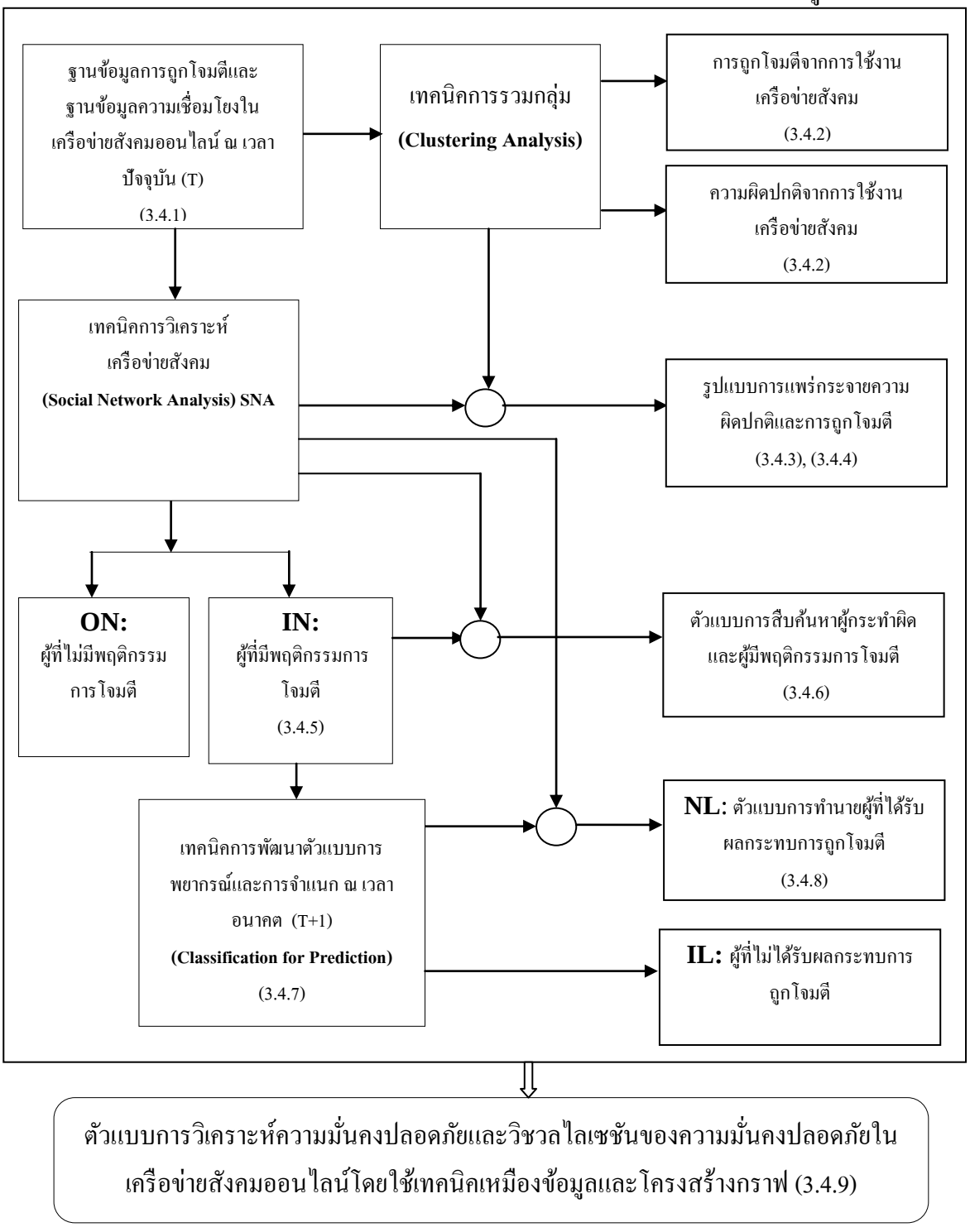
ข้อมูลที่ใช้ในการวิจัยนี้ ประกอบด้วย

3.3.1 ข้อมูลบันทึกการใช้งานเครือข่ายสังคมออนไลน์ขององค์กรภาคสาธารณสุขที่ให้ความสำคัญต่อความเป็นส่วนตัวและความปลอดภัยของระบบฐานข้อมูลด้านสุขภาพเป็นอย่างสูง และต้องการหลีกเลี่ยงจากอันตรายของโปรแกรมประยุกต์ของเครือข่าย เนื่องจากต้องป้องกันความเสี่ยงต่อข้อมูลและระบบสารสนเทศขององค์กรที่รองรับระบบงานทางด้านสุขภาพในระดับประเทศ (Public Health Sector in Thailand) จำนวน 2 ประเทศ คือประเทศไทย และประเทศที่ร่วมทดสอบ โดยข้อมูลบันทึกการใช้งานเครือข่ายในงานวิจัยนี้เป็นข้อมูลในเดือนธันวาคม 2554 เดือนเมษายน 2555 เดือนธันวาคม 2555 และเดือนเมษายน 2556 เพื่อใช้ในการวิเคราะห์ถึงความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์

3.3.2 งานวิจัยนี้ได้ผนวกข้อมูลความสัมพันธ์หรือความเชื่อมโยงการใช้งานเครือข่ายสังคมออนไลน์ของสมาชิกผู้ถูกโจมตี เมื่อเข้าใช้งานเครือข่ายสังคมออนไลน์ในเดือนธันวาคม 2554 เดือนเมษายน 2555 เดือนธันวาคม 2555 และเดือนเมษายน 2556 เพื่อใช้ในการศึกษาและวิเคราะห์รูปแบบการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ พัฒนาตัวแบบในการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ และพัฒนาตัวแบบในการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์

3.4 การวิเคราะห์ข้อมูล

จากกรอบแนวคิดในการวิจัยในบทที่ 1 จะได้วิธีดำเนินการวิจัยดังแสดงไว้ในรูปที่ 3.2



รูปที่ 3.2 วิธีดำเนินงานวิจัย

จากรูปที่ 3.2 เพื่อที่จะค้นหากลยุทธ์การติดตามภัยจากการใช้เครือข่ายสังคมออนไลน์จากตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟไว้ดังนี้ ผู้วิจัยได้กำหนดกระบวนการวิเคราะห์ข้อมูลดังนี้

3.4.1 นำข้อมูลจราจรการใช้งานเครือข่ายสังคมออนไลน์ขององค์กรและข้อมูลความสัมพันธ์ในการใช้งานเครือข่ายสังคมออนไลน์ของสมาชิกจำนวน 2 ประเทศในเดือนธันวาคม 2554 เดือนเมษายน 2555 เดือนธันวาคม 2555 และเดือนเมษายน 2556 เข้าทำการวิเคราะห์ตามวัตถุประสงค์ต่างๆ

3.4.2 ศึกษาและวิเคราะห์หาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์ ในหลักการของ Data Mining (Unsupervised Learning, Cluster Analysis, Distance Measure, K-means Algorithm)

3.4.3 ศึกษาและวิเคราะห์การแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ ในหลักการของ Data Mining (Unsupervised Learning, Cluster Analysis, Distance Measure, K-means Algorithm)

3.4.4 แสดงภาพการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ ในหลักการของ Social Network Analysis (SNA) ด้วย Minimum Spanning Tree ของ Prim’s Algorithm

3.4.5 วิเคราะห์หาวิธีการสืบค้นผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ ในหลักการของ Social Network Analysis (SNA) ด้วย Degree Centrality, Betweenness Centrality, Closeness Centrality เพื่อค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ (Influencing Nodes)

3.4.6 พัฒนาตัวแบบการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ ในหลักการของ Social Network Analysis (SNA) ด้วย Minimum Spanning Tree ของ Prim’s Algorithm และ All-Pair Shortest Path ของ Floyd-Warshall Algorithm เพื่อแสดงภาพผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์

3.4.7 วิเคราะห์หาวิธีการหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ ในหลักการของ Data Mining (Supervised Learning, Classification Analysis, oneR Algorithm)

3.4.8 พัฒนาตัวแบบการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ ในหลักการของ Social Network Analysis (SNA) ด้วย Minimum Spanning Tree ของ Prim's Algorithm และ All-Pair Shortest Path ของ Floyd-Warshall Algorithm เพื่อแสดงภาพเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์

3.4.9 การทดสอบประเมินความแม่นยำของตัวแบบที่ได้ คือ ตัวแบบการประเมินหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริงเมื่อใช้เครือข่ายสังคมออนไลน์ และตัวแบบการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์ ด้วยการทดสอบสมมติฐานระหว่างข้อมูลที่เกิดขึ้นจริงในอนาคตและข้อมูลที่ได้จากการประเมินหรือทำนายหาบนฐานข้อมูลในปัจจุบัน โดยมีความละเอียดถูกต้อง (Accuracy) อย่างมีนัยสำคัญภายใต้ระดับความเชื่อมั่น 95 %

3.4.10 สรุปผลการทดลองและนำเสนอตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ

3.5 สรุป

ในบทที่ 3 นี้เป็นการนำเสนอวิธีดำเนินการวิจัย ที่ประกอบด้วย ขั้นตอนการวิจัย เครื่องมือและอุปกรณ์ที่ใช้ในการวิจัย ขั้นตอนการสร้างเครื่องมือ ประชากรและกลุ่มตัวอย่าง การเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล เพื่อที่จะค้นหากลยุทธ์การติดตามภัยจากการใช้เครือข่ายสังคมออนไลน์จากตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ สามารถทำให้ทราบถึงเหตุการณ์ปัจจุบัน และประเมินหาการเชื่อมโยงในเครือข่ายสังคมออนไลน์ ซึ่งเป็นประโยชน์มากในการวิเคราะห์ และความเข้าใจในเครือข่ายสังคมออนไลน์ดังกล่าว สามารถนำไปสู่การดำเนินงานที่มีประสิทธิภาพของเครื่องมือในการสืบค้นผู้กระทำผิดปกติหรือมีพฤติกรรมการโจมตี การค้นหา

เป้าหมายที่อาจจะได้รับผลกระทบจากการโจมตีนั้นๆ รวมทั้งการระบุกลุ่มที่ซ่อนอยู่หรือการหาสมาชิกที่หายไปของกลุ่ม ฯลฯ ในเครือข่ายสังคมออนไลน์ ซึ่งเป็นปัญหาที่พบบ่อยที่สุดในการรักษาความมั่นคงปลอดภัยและการสอบสวนทางอาญา และเพื่อตอบคำถามตามวัตถุประสงค์และสมมติฐานที่กำหนดไว้ ได้แก่

1. เพื่อทำการวิเคราะห์หาความผิดปกติและโอกาสการถูกโจมตีจากการใช้เครือข่ายสังคมออนไลน์

2. เพื่อศึกษารูปแบบการแพร่กระจายความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์

3. เพื่อพัฒนาตัวแบบการวิเคราะห์ความมั่นคงปลอดภัยและวิชวลไลเซชันของความมั่นคงปลอดภัยในเครือข่ายสังคมออนไลน์โดยใช้เทคนิคเหมืองข้อมูลและโครงสร้างกราฟ รองรับการสืบค้นหาผู้กระทำผิดปกติและมีพฤติกรรมการโจมตีที่แท้จริง

4. พัฒนาตัวแบบการทำนายหาเป้าหมายในเครือข่ายสังคมออนไลน์ที่อาจจะได้รับผลกระทบจากความผิดปกติและการถูกโจมตีในเครือข่ายสังคมออนไลน์เพื่อทำการแจ้งเตือนล่วงหน้า