ชื่อเรื่องวิทยานิพนธ์ การจัดกลุ่มข้อมูลจราจรบอตเน็ตโดยใช้ขั้นตอนวิธีฟัซซีซีมีนและการเลือกฟีเจอร์ที่เหมาะสม

ผู้เขียน นาย ขจรเกียรติ อักษรเสือ

ปริญญา วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมคอมพิวเตอร์)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ รศ.ดร.ตรัสพงศ์ ไทยอุปถัมภ์

บทคัดย่อ

ในปัจจุบัน มีผู้นิยมใช้อินเทอร์เน็ตกันอย่างแพร่หลาย ผู้ใช้งานอินเทอร์เน็ตอาจประสบปัญหาจากภัยคุกคามทางอินเทอร์เน็ต อาทิเช่น สแปม ฟิชชิ่ง เวิร์ม สปายแวร์ โทรจัน มัลแวร์ และบอตเน็ต ปัญหาการทำงานที่ผิดปกติทางอินเทอร์เน็ต เหล่านี้ นำมาสู่การศึกษา ค้นคว้า การตรวจจับการทำงานที่ผิดปกติทางอินเทอร์เน็ต โดยการศึกษาครั้งนี้ ได้ศึกษาการทำงานของบอตเน็ตเป็นกรณีศึกษา จากการทำงานของบอตเน็ต มีรูปแบบการทำงาน การส่งข้อมูลที่มีความซับซ้อนในการส่งข้อมูลคล้ายกับการทำงานของผู้ใช้ปกติ โดยพิจารณาจากขั้นตอนการทำงานของบอตเน็ต ที่มีการสั่งงานจากบอตมาสเตอร์ ไปยังบอตเน็ต แต่ละตัวที่อยู่ในความควบคุมของบอตมาสเตอร์

ในงานวิจัยนี้นำเสนอวิธีการ ตรวจจับการทำงานที่ผิดปกติของบอตเน็ต โดยการจัดกลุ่มให้กับบอตเน็ต เริ่มจากการนับจำนวนความถี่การใช้งาน การส่งข้อมูลอักขระ ASCII ซึ่งมีจำนวน 256 ตัว นำความถี่การใช้งานอักขระ ASCII จำนวน 256 ตัว ที่ได้นี้ มาเป็น Feature เพื่อทำการแยกประเภทหรือจัดกลุ่ม โดยงานวิจัยนี้นำเสนอวิธีการ Supervised Fuzzy C-Means ที่ทำการเลือก Feature ที่เหมาะสม เปรียบเทียบการทำงานด้วยวิธีการของ K-Means และ Fuzzy C-Means ที่มีการใช้งาน Feature จำนวน 256 Feature โดยใช้ข้อมูลการทำงานบอตเน็ต ที่ปรับการทำงานจากเดิม ทำงานบน IRC ให้มาอยู่บน TCP และอีกกลุ่มการทำงานคือ กลุ่มข้อมูลของผู้ใช้งานปกติ ในสภาพแวดล้อมการทำงานที่แตกต่างกัน

Thesis Title Clustering of Botnet Traffic Using Fuzzy C-Means with Features Selection

Author Mr. Khajornkiat Agsornsue

Degree Master of Engineering (Computer Engineering)

Thesis Advisor Assoc.Prof Dr. Trasapong Thaiupathump

Abstract

Nowadays, Internet users face lots of threats such as spam, phishing, worm, spyware, Trojan, and botnet. Many studies have been conducted in order to find suitable ways to detect these threats. This article focuses on detecting the botnet traffic by analyzing botnet behavior where its control command traffic between botmaster and botnet are in specific patterns.

The study investigates botnet detection method using Supervised Fuzzy C-Means with proper feature selection. The features in this study are frequency of 256 ASCII in a data block. In addition, the comparison with K-means and Fuzzy C-Means using 256 ASCII frequency features are also investigated. Traffic data used in the experiments are from simulation models of 2 sources, 1) modified botnet traffic clusters where IRC transmission channel is replaced with TCP channel, and 2) general internet user traffic information in various working environments.