Sudan University of Science and Technology
College of Graduate Studies

Intrusion Detection System Using Computational Intelligence

نظام كشف التسلل باستخدام الذكاء الحسابي

Submitted for the degree of Doctor of Philosophy
in Computer Science

By
Asma Abbas Hassan Elnour

Supervisor
Mohammed Alhafiz, Dr

Co-Supervisor

Talat Wahbi, Dr

March 2018

Dedication

This thesis is dedicated to:

My great mother, who never stops giving of herself in countless ways,

My dearest husband, who leads me through the valley of darkness with light of hope and support,

My beloved brothers and sisters, my beloved kids, who encourage and support me

Acknowledgements

I owe a deep gratitude to our university for giving us an opportunity to complete this work.

I am grateful to some people, who worked hard with me from the beginning till the completion of the present research particularly my supervisor Dr. Mohammed Alhafiz, for his support and guidance, and I highly appreciate the efforts expended by my Co-supervisor Dr. Talaat Wahbi, who has always been generous during all phases of the research

I would like to take this opportunity to say warm thanks to all my beloved friends, who have been so supportive along the way of doing my thesis.

I am very appreciative to my colleagues at Sudan University, who participated in this study. And Without their support, this study would not have been possible.

Last but not least, deepest thanks go to all people who took part in making this thesis real.

Abstract

Traditional intrusion prevention techniques, such as firewalls, access control or encryption, failed to fully protect networks and systems from increasing attacks. Therefore, an intrusion detection system (IDS) has become an important component of security infrastructure and a key part of system defense to detect these attacks before they make a disaster in the system**,** intrusion detection based upon computational intelligence (CI) is currently attracting considerable interest from the research community. Characteristics of CI systems, such as adaptation, fault tolerance, high computational speed and error resilience in the face of noisy information fit the requirements of building a good intrusion detection model. The scope of this thesis will be on core methods of CI. One of the important research challenges for constructing high performance Network Intrusion Detection Models (NIDS) is dealing with data containing large number of features. Which make it harder to detect and classify the intrusion, causing slow training and testing process, higher resource consumption as well as poor detection rate. Therefore, feature selection methods were used to reduce the dimensionality of the dataset by removing redundant features. Feature selection improves the NIDS classification performance by searching for the subset of features, which decrease the size of the structure without significantly decreasing prediction accuracy of the classifier built using only the selected features. Therefore, applying feature selection as a preprocessing step when building NIDS is very important if real-time detection is desired. This thesis proposed different NIDS models; PCA-SVM model, GA-NB model and GA-IEM-C4.5 model; those models involves data preprocessing, data reduction and intrusion classification. All this implemented in Weka machine learning tools, with KDD CUP 99 and NSL-KDD intrusion detection data sets. Experimental results show the advantages of enhancing the detection accuracy and testing speed by reducing the feature dimension space.

المستخلص

تفشل التقنيات التقليدية المستخدمة في الحاسوب مثل برامج الحماية وضوابط الوصول او التشفير في توفير الحماية الكاملة للشبكات والأنظمة الحاسوبية ضد الهجمات، لذا فان أنظمة الكشف عن التطفل صارت عنصر هام في تأمين البنيات التحتية وتشكل الجزء الأساسي في أنظمة الحماية من الهجمات التي قد تشكل كارثة عليها، وتعتمد أنظمة الكشف عن التطفل والتي صارت حاليا جاذبة للباحثين في مجال الذكاء الاصطناعي، وتتناسب خصائص ومميزات أنظمة الكشف عن التطفل مثل الموائمات ونسبة الخطأ والسرعة الحاسوبية العالية ومقاومة الخطأ وثباته في مواجهة المعلومات المزعجة مع متطلبات أنظمة كشف التطفل الجيدة. تهدف هذه الاطروحة لتناول جوهر أنظمة كشف التطفل. ويعتبر التعامل مع البيانات التي تحتوي على ارقام كبيرة واحد من اهم التحديات في بناء أنظمة شبكات كاشفة للتطفل بأداء عالي، والتي يصعب عليها تحديد وتصنيف التطفل مما يتسبب بإبطاء أداء العملية المنفذة والاستهلاك العالي للموارد وضعف المعدل الكشفي. لذا فان مميزات أنظمة الاختيار المستخدمة لتقليل ابعاد العملية الاختبارية عبر إزالة السمات الفراغية. يحسن اختيار السمات من تصنيف أداء عمل أنظمة شبكات الكشف التطفلية من خلال البحث عن المميزات او السمات التي تقلل من الحجم البنيوي من دون انخفاض ملحوظ في دقة التوقع لبناء المصنف المستخدم فقط في تحديد السمات، ولذا فالتطبيق على اختيار السمة كخطوة إجرائية عند بناء أنظمة شبكات كاشفة التطفل فانه من المهم جدا معرفة عما إذا كان كاشف الوقت الحقيقي مرغوب به.

تقترح هذه الرسالة أنظمة مختلفة من شبكات كاشفة التطفل مثل المكونات الأساسية ونظام جهاز النقل الداعم وأنظمة اختزال الرسائل والأنظمة المعروفة ب GA-IEM-C4.5 حيث تشترك هذه الأنظمة في معالجة البيانات وتصنيف التطفل وتقليل تعرض البيانات للتطفل. وتم تطبيق كل أدوات التعلم ل Weka مع بيانات NSL-KDD,KDD CUP 99 الكاشفة للتطفل.

تظهر الاختبارات المعملية الأفضلية لتقوية دقة الكشف وزيادة السرعة عبر تقليل الفراغ بين السمات.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Terminology
IDS	Intrusion Detection System
HIDS	Host Based Intrusion Detection Systems
NIDS	Network Based Intrusion Detection Systems
IS	Intelligent Systems
CI	Computational Intelligence
EC	Evolutionary Computation
ML	Machine Learning
FL	Fuzzy Logic
SVM	Support Vector Machine
GA	Genetic Algorithm
ANN	Artificial Neural Network
GP	Genetic Programming
MLP	Multi Layer Perceptron
PCA	Principal Component Analysis
NB	Naive Bayes
NSM	Network Security Monitor
IEM	Information Entropy Maximization
DoS	Denial of service
R2L	Remote-to-Local
U2R	User-to-Root
RBF	radial basis function
CR	classification rate
CPE	Cost per example
FPR	False Positive Rate
DR	Detection Rate
PR	Precision Rate
AUC	Area under curve
RMSE	root mean square error
TN	True Negatives
TP	True Positives
FP	False Positives
FN	False Negatives

List of Publications

1. Intrusion Detection System Using Weka Data Mining Tool. International Journal of Science and Research (IJSR), Volume 6 Issue 9, September 2017, 337-342

2. Intrusion Detection Using Neural Network. International Journal of Science and Research (IJSR), Volume 6 Issue 9, September 2017, 343 - 347