Dedication

This work is dedicated to my family

To my husband, my daughters and sons

To my supervisor Dr. Mohamed Awad

Acknowledgements

To Dr. Noureldien Abdelrahman, who introduced me to this field of .research, for his aid and encouragement

To my supervisor Dr. Mohammed Awad Elshaikh for his valuable .observations, useful suggestions and beneficial criticism

To my uncle engineer Abdelgader Mohammed Ahmed for his deep .remarkable continuous scientific help

To my colleague Dr. Eihab Bashier for his practical cooperation and .effective support

A special thanks to all the members of my family who shouldered me and, for their unconditional love, support, and encouragement through .this battle

Abstract

Intrusion detection is an exemplary method designed to monitor the actions happening in a network. Then analyze them for suspected patterns that may identify a network or system violation from someone trying to penetrate and endanger the system. So an Intrusion Detection System (IDS) is software which is applied automatically as a procedure to stop the penetration and attacks of the intruders. It is applied as either Signature recognition or Anomaly detection methodologies. Most of existing IDS required reduction technique in order to minimize the features of data which is irrelevant or redundant. This is needed in case of high dimensionality in network traffic. It is also known that the reduction technique helps the classification algorithms to be very effective. As for the Classification, it achieves and executes the intrusion detection job practically. We realize that Kernel Principal Component Analysis (KPCA) is recognized as a robustification reduction method for standard Principal Component Analysis (PCA) [34]. This research adopts an optimal anomaly detection method to detect multivariate attacks. This method is going to be achieved by measuring the performance of different functions of KPCA as a reduction method applied to different classification algorithms to find out which function of KPCA is the best with any algorithm. Consequently we show that KPCA's methods will not always outperform standard PCA. The final detection's performance, in fact depends on the used classification algorithm. The experiments with NSL-KDD data set demonstrate that the adopted method achieves **98.048%** in detection rate and **98.261%** in precision with **1.484%** false positive rate, consequently outperforms all the other methods. Moreover the results prove that [PCA & K-Nearest Neighbor] outperform [KPCA (Gaussian) & K-Nearest Neighbor] and [KPCA (Quadratic) & K-Nearest Neighbor]. In addition [PCA and Discriminant Analysis] .[outperform [KPCA (Quadratic) & Discriminant Analysis

المستخلص

ان اكتشاف المتطفلين هى طريقة مثالية لمراقبة الاحداث التي تتم في الشبكة واختبارها بحثاً عن نماذج مشكوك فيها, و قد تعتبر هذه النماذج مهددات للشبكة اوالنظام من شخص ما يحاول اختراقها وبالتالي يهدد النظام. ان نظام اكتشاف المتطفلين برنامج يطبق اتوماتيكياً كإجراء لايقاف الاختراقات وهجوم المتطفلين. وهو يقع ضمن احدى المنهجيتين: منهجية التعرف على التوقيعات أو منهجية اكتشاف الشواذ. نجد ان معظم نظم اكتشاف المتطفلين الحالية تتطلب وجود طريقة تقنية لتقليل ميزات البيانات الزائدة عن المطلوب و ليست ذات علاقة بمجموعة البيانات. وهذه التقنية نحتاجها في حالة الابعاد الفائضة في حركة الشبكة. ان اسلوب التقليل الفني يساعد خوارزميات التصنيف لكي تكون شديدة الفعالية. اما فيما يتعلق بالتصنيف فإنه يحدد وظيفة اكتشاف المتطفلين عملياً. نلاحظ ان نظام KPCA يعتبر تقوية لنظام PCA. هذا البحث يتبنى طريقة مثالية لإكتشاف الشواذ في حركة الشبكة لتحديد الهجمات متعددة المتغيرات. هذه الطريقة يتم انجازها بقياس اداء وظائف مختلفة لـ KPCA كطريقة للتقليل مطبقة على خوارزميات تصنيف مختلفة لمعرفة اي وظائف KPCA هي الافضل ومع اي خوارزمية طبقت. وايضا نوضح ان KPCA لا يتفوق دائما على PCA. و اداء الاكتشاف النهائي يعتمد على خوارزمية التصنيف المستخدمة.

كل التجارب اجريت باستخدام برنامج ماتلاب ومجموعة البيانات NSL KDD وهي متوفرة عالمياً. هذه التجارب وضحت أن الطريقة المتبناة تحصلت على 98.048 % في نسبة اكتشاف المتطفلين , 98.261 % في الدقة و 1.484% في نسبة الخطأ الايجابي بذلك تفوق على كل الطرق الأخرى. بالاضافة الى ذلك النتائج تثبت ان نظام [PCA & KNN] يتفوق على نظام [KPCA (Gaussian) & KNN] و [KPCA (Quadratic) & KNN] وايضاً نظام [PCA and DA] يتفوق على نظام [KPCA (Quadratic) & DA].

Table of contents

List of figures

List of table