

Development and Implementation of fast Network Intrusion Detection Systems	العنوان:
Mahmoud, Mohamed Awad Mohamed	المؤلف الرئيسي:
Mohamed, Mohamed Abd Alazim, Abd Alfattah, Abd Alfattah Ibraheem, Tolba, Ahmed Said(Super., Super)	مؤلفين آخرين:
2012	التاريخ الميلادي:
المنصورة	موقع:
1 - 192	الصفحات:
536381	رقم MD:
رسائل جامعية	نوع المحتوى:
English	اللغة:
رسالة دكتوراه	الدرجة العلمية:
جامعة المنصورة	الجامعة:
كلية الهندسة	الكلية:
مصر	الدولة:
Dissertations	قواعد المعلومات:
لكشف الدخلاء، جرائم المعلومات، شبكات الحاسب	مواضيع:
https://search.mandumah.com/Record/536381	رابط:

تكنولوجيا نظم كشف أو مراقبة البيانات المتسلسلة من/إلى شبكات الحاسب الألى واحده من أكثر التكنولوجيات أهمية في حماية الشبكات و الذى يعمل بامتزاج مع الجدار الناري, Firewall . بحكم أن في وقتنا الحاضر قد ازداد الاحتكاك بالإنترنت حيث أصبح الركن الأساسي في الاتصال بالعالم الخارجي و ايضا حاجة الشركات -بغض النظر عن حجمها- بالسماح للموظفين أو للعملاء أيضا بالدخول إلى الشبكة الداخلية للشركة أو حتى اتصال الشركة نفسها بفروعها في أماكن مختلفة , فقد أصبح لزاما تحديد المناطق أو الأفراد الموثوق بهم و مع ذلك فإن الموضوع تعقد إلى أن المشكلة أصبحت ليست فقط تحديد مواطن الثقة بل أيضا كيفية التأكد من هذه الجهة الموثوق بها أو بمعنى آخر كيف أحمي نفسي أيضا ممن أثق بهم!! , صاحب تقديم هذه الخدمة هو نظام كشف التسلسل IDS. إكتشاف الأختراق تلقى اهتماما كبيرا من جانب كل من الصناعة والأوساط البحثية و سبب هذا الاهتمام هو الخسائر الهائلة التى تنتج عن أختراق وتسريب او تدمير المعلومات وأيضا براعة ومهارة المخترقين حيث انهم يقومون بتحديث وسائل وطرق الاختراق. أيضا التطور السريع فى سرعة تداول المعلومات على شبكات الحاسب الألى أدى الى ضرورة تطوير سرعة أنظمة إكتشاف الأختراق لتواكب سرعات التدفق العالية على تلك الشبكات.

لأختبار أى تقنية مبتكرة كان حتما واجب أختبارها على شبكة حقيقية ولاكن بات هذا صعبا لعدة أسباب من أهمها احتمال توقف مؤقت للشبكة قد يؤثر على سير العمل الذى قد يكون خطيرا أو

حرجا وأيضا اذا لم يحدث ذلك فقد تكون هناك معلومات سرية مثل كلمة السر وأسماء المستخدمين لا يستوجب الاطلاع عليها من قبل أفراد فريق الأختبار. كل هذا اجبر المجتمع البحثي على توفير قواعد بيانات تسهل على الباحثين تنفيذ أبحاثهم وكان من أشهرها DARPA Dataset و KDD'99 Dataset.

الخوazميات المطبقة في هذا البحث تم تنفيذها باستخدام مصفوفة البوابات المبرمجة حقليا (FPGA) للإستفادة من المميزات التي توفرها كسرعة الأداء لتتوافق مع السرعات العالية لتتدفق المعلومات على شبكات الحاسب الألى وقابليتها للبرمجة أكثر من مرة لتتوافق مع التحديثات المستمرة في أنواع وأساليب الأختراقات. FPGA تعني مصفوفات البوابات المنطقية القابلة للبرمجة، وهي عبارة عن دوائر متكاملة تتكون من صفوف من البوابات المنطقية كل بوابة يمكن التحكم في نوعها وتحويلها من نوع الى آخر مثلا من AND الى NAND ومن NOR الى NOT الى أخره ويمكن أيضا التحكم في توصيل كل بوابة بالبوابات الأخرى فيمكن عمل التوصيل الذي نريده ويمكن بذلك تحويل الدائرة المتكاملة من نوع الى آخر عن طريق البرنامج الخاص ببرمجتها. وهي تستخدم في التطبيقات التي تحتاج الى تغيير دوائرها باستمرار وبينما كان ذلك صعبا في الماضي فإنه الان يمكن أن يحدث بمجرد انزال برنامج جديد يغير من تركيب الدائرة.

ويهدف البحث إلي: (أ) مناقشة تقنيات إكتشاف الأختراق، (ب) محاولة التغلب على التحديات التي تواجه هذه التقنيات (ت) تطوير أساليب جديدة لأكتشاف الأختراق، (ث)

تطبيق اكتشاف الأختراق باستعمال VHDL و (ج) تنفيذ أفضل تقنيات اكتشاف الأختراق باستخدام مصفوفة البوابات المنطقية المبرمجة حقلياً (FPGA).

و تنقسم الرسالة إلى سبعة أبواب و ملحقين كما يلي:

الفصل الأول: تحت عنوان "Introduction" و هو مقدمة تاريخية عن الإختراق المعلوماتي وطرق اكتشافه على مدار العقود الماضية وأهداف وأسهمات الرسائله وأخيرا التقسيمه الاساسيه فى الرسائله.

الفصل الثاني: تحت عنوان "Network Security: An Overview" وفيه تم عرض مقدمه عن أنظمة تأمين الشبكات وخدمات وأليات الأمن المعلوماتي على شبكات الحاسب الألى وأخيرا عرض لأدوات تأمين شبكات الحاسب الألى والتي من ضمنها مكتشف الإختراق IDS.

الفصل الثالث: تحت عنوان "Intrusion Detection Systems Overview" وفيه تم عمل أستعراض كامل لنظام إكتشاف الإختراق من حيث تعريفه ومكان تواجده فى الشبكة المطلوب حمايتها ومكانه بالنسبه للجدار النارى وأيضا الأجزاء المكونه له وأخيرا قدم الفصل طرق تنفيذ أنظمة إكتشاف الأختراق المختلفة.

الفصل الرابع: تحت عنوان "Conventional IDS Systems" في هذا البحث تم تنفيذ ثمانية أنواع من خوارزميات الشبكات العصبية في محاولة لتحديد ما إذا كان التدفق تدفق طبيعي أو تدفق فيه إختراق وإذا كان به إختراق فماتصنيف هذا الإختراق وما أسمه حيث أن هذا يسهل لمدير النظام اتخاذ اجراء مناسب حيال هذا الإختراق. أيضا تم تنفيذ اثنين من خوارزميات التققيب المعلوماتي وهما Decision Tree و Naïve Bayes لتحقيق نفس الهدف. نتائج ما سبق أظهرت أن الخوارزميات السابقة لها قدرة كبيرة على اكتشاف ما إذا كان هناك إختراق ام لا ولكن تحديد تصنيفه وأسمه فما زالت مهمة صعبة.

الفصل الخامس: تحت عنوان "The Proposed IDS Systems" هذا ماتم تقديمه في هذا البحث حيث قدم فكرة مستواحة من تطبيق مكتشف متعدد المراحل/مختلط من الأنواع التقليدية السابق تنفيذها في الباب السابق. هذا الأسلوب استطاع بنجاح التعرف على جميع أصناف الأختراق بالإضافة الى التعرف على جميع أنواع أسماء الإختراقات. البيانات المستخدمة تعتبر ضخمة جدا ولتسهيل عملية تنفيذ الخوارزميات السابقة كان لابد من تنفيذ عمليات ضغط وإختصار لتلك المعلومات من الناحية الكمية وذلك بإزالة البيانات المتشابهة

أو من الناحية الكيفية بإختصار عناصر البيانات إما بإختيار أهم تلك العناصر أو دمجها مع بعضها وإستتباط عناصر جديدة وتم الأستقرار على طريقة تحليل المكونات الرئيسية Principal Components Analysis PCA التي قلصت الـ 41 عنصر المكون للتدفق الواحد فى قاعدة البيانات الى 22 عنصر.

الفصل السادس: تحت عنوان " Hardware Implementation of IDS " يحتوي هذا الفصل على تنفيذ بعض الأساليب المقترحة الجديدة فى الفصل السابق على شريحة مصفوفة البوابات المبرمجة حقليا FPGA للإستفادة من المميزات الهائلة التي توفرها وخصوصا في سرعة الأداء. أيضا تم تنفيذ منظومات متعددة المراحل منهم مستهدفه التعرف على صنف الهجوم. عملية التنفيذ على الشريحة واجهت صعوبات عديدة منها ضيق المساحة المتاحة على الشريحة وطرق تداول البيانات على الشريحة وأيضا تنفيذ الدالة الغير خطية المستخدمة فى الخلية العصبية المكونه للشبكة العصبية كل هذه الصعوبات تم التغلب عليها بنجاح حيث تم ضغط الـ 22 عنصر السابق التحدث عنهم فى الفصل الثامن الى عدد عناصر أقل. الضغط كان مبنى على إستبعاد قيم الـ PCs ذات الأقل مرتبة مع الاحتفاظ بنسبة تعرف

عاليه. قدم هذا الفصل دالة جديدة لتمثيل الدالة الأصلية الغير خطيه السابق الحديث عنها ولكنها كانت صعبة التنفيذ لوجود عملية قسمة فيها والتي تُغلب عليها بأستخدام دالة القسمة المتوفرة في برنامج الشركة المنتجه للشريحة كملكية فكرية لها.

الفصل السابع: تحت عنوان " Conclusion " يحتوي هذا الفصل على ملخص الرسالة.

Abstract

Intrusion Detection Systems (IDSs) have recently received a blooming attention and interest from the scientific community as well as from the public. The interest from the public is mostly due to the recent events of terror around the world, which has increased the demand for useful security systems. There has been significant progress in improving the performance of computer-based intrusion detection systems algorithms over the last decade. Although algorithms have been tested and compared extensively with each other, there has been remarkably little work comparing the accuracy of computer-based IDSs with humans.

Our long term goal is to design and build an intelligent IDS that is accurate (low false negative and false positive rates), flexible, not easily fooled by small variations in intrusion patterns, adaptive in new environments, modular with both misuse and anomaly detection components, robust, able to detect known and unknown intrusions and real-time.

To achieve our goal, two phases of implementation were introduced; software and hardware phases. The software phase is divided into two sub-phases; the first introduces conventional IDS implementation which is based on neural networks and Data mining techniques. The results obtained were unsatisfactory. The second sub-phase introduces a proposed IDS which is based on hybridizing and/or multistaging more than one conventional IDS. The proposed models were successfully able to detect all signatures and all classes of intrusions, in addition to the type of each stream (normal or attack) in the tested database. The first stage of

the multistage classifier is a perfect 2-types anomaly classifier, which ensures securing the network against attack (known and unknown attacks), whereas an alarm could be raised to the system administrator to take the proper action. Feature and topology reduction were one of most important challenges in this thesis due to massiveness of tested dataset and the huge structure of the implemented models. The proposed models were repeated with different neural networks topologies with simpler structure which are based on trial and error. Feature reduction process proceeds in two ways; record reduction by eliminating duplications and feature reduction. By applying Principal Components Analysis (PCA), the original 41-features were transformed to 22 ranked principal components (PCs) and the generated PCs were minimized by eliminating less ranked PCs until reaching a compromise situation between the smallest number of PCs and best detection efficiency.

As the demand on more network speed increases and new network protocols emerge, Network Intrusion Detection Systems (NIDSs) are increasing in importance and are being integrated in network processors. Currently, most IDSs are software running on a general purpose processor. Unfortunately, it is becoming increasingly difficult for software based IDSs to keep up with the increasing network speeds (10Gbps at backbone networks). This has underscored the need for the specialized hardware-based solutions which are portable and operate at wire speeds. Field Programmable Gate Arrays (FPGAs) are generic pieces of hardware

that can be reconfigured to perform any task. FPGA-based platforms can exploit the fact that the NIDS rules change relatively infrequently, and use reconfiguration to reduce implementation cost. In addition, FPGA-based systems can exploit parallelism in order to achieve satisfactory processing robustness.

To achieve robustness of IDS; FPGA-based IDSs were introduced, the hardware phase, based on neural networks and data mining intrusion detection techniques. Through MultiLayer perceptron (MLP) design, we faced common challenge in MLP hardware implementation which is tansig activation function which is difficult to implement due to its nonlinearity. Three techniques for tansig representation were presented, one based on Kwan approximation [172] and the other based on Piecewise Linear Approximation of a Nonlinear function (PLAN) approximation [173]. A third approximation, which is based on symbolic regression and genetic algorithm, is proposed. New challenge arisen in implementing the proposed tansig approximation which is the division process in FPGA. This was solved by using (Intellectual Property) IP CORE available in Xilinx ISE. FPGA-based IDSs, MLP and Decision Tree DT-j48 based, were implemented using two different programming strategies and data formats. Both techniques gave perfect results to detect normal and attack stream. Both implemented single stage IDS was followed by second stage which was able to detected stream's class after detecting stream's type.

Development and Implementation of fast Network Intrusion Detection Systems	العنوان:
Mahmoud, Mohamed Awad Mohamed	المؤلف الرئيسي:
Mohamed, Mohamed Abd Alazim, Abd Alfattah, Abd Alfattah Ibraheem, Tolba, Ahmed Said(Super., Super)	مؤلفين آخرين:
2012	التاريخ الميلادي:
المنصورة	موقع:
1 - 192	الصفحات:
536381	رقم MD:
رسائل جامعية	نوع المحتوى:
English	اللغة:
رسالة دكتوراه	الدرجة العلمية:
جامعة المنصورة	الجامعة:
كلية الهندسة	الكلية:
مصر	الدولة:
Dissertations	قواعد المعلومات:
لكشف الدخلاء، جرائم المعلومات، شبكات الحاسب	مواضيع:
https://search.mandumah.com/Record/536381	رابط:

Contents

	Page
Chapter-1: Introduction To NIDS	
1.1 Network Intrusions And Network Security.....	1
1.2 Intrusion Detection History & Foundation Work.....	2
1.3 Aims Of The Work.....	6
1.4 Work Organization.....	7
Chapter-2: Network Security: An Overview	
2.1 Information Security	8
2.2 The OSI Security Architecture.....	9
2.3 Security Attacks.....	10
2.3.1 Passive Attacks.....	10
2.3.2 Active Attacks.....	12
2.4 Security Services.....	14
2.5 Security Mechanisms.....	16
2.5.1 Specific Security Mechanisms.....	16
2.5.2 Pervasive Security Mechanisms.....	17
2.6 A Model for Network Security.....	18
2.7 Security System tools.....	20
Chapter-3: Intrusion Detection Systems Overview	
3.1 Intrusion Detection Systems (IDS).....	21
3.1.1 IDS in General.....	21
3.1.2 Why IDS?.....	22
3.2 Intrusion Detection Frameworks.....	23
3.3 Other Network Security Blocks.....	24
3.4 Function of Intrusion Detection Systems.....	25
3.4.1 False Positive and False Negative.....	25
3.5 Efficiency of intrusion-detection systems.....	26
3.6 Taxonomy of IDS.....	26
3.7 Knowledge-based versus behavior-based intrusion detection.....	28
3.7.1 Knowledge-based Intrusion Detection (Signature/misuse based).....	28
3.7.2 Behavior-based Intrusion Detection (Anomaly based).....	30
3.8 Host-based versus network-based intrusion detection.....	33
3.8.1 Host-Based Information Sources.....	34
3.8.2 Network-based information sources.....	36
3.8.3 Application log files.....	37
3.9 Detection paradigm.....	38
3.9.1 State-based vs. Transition-based IDSs.....	38

3.9.2	Nonperturbing vs. pro-active analysis of state/transition.....	39
3.10	Additional properties.....	39
3.10.1	Continuous Monitoring versus Periodic Analysis.....	39
3.10.2	Protection of the IDS.....	40
3.11	Methodologies for IDSs.....	41
3.11.1	Colored Petri Nets	42
3.11.2	State Machines	42
3.11.3	Signature Analysis.....	42
3.11.4	Genetic Algorithms (GA).....	42
3.11.5	Immune Based	43
3.11.6	Neural Networks	43
3.11.7	Agents	44
3.11.8	Fuzzy Logic	44
3.11.9	Bayesian Inference.....	44
3.11.10	Expert Systems	45
3.11.11	Data Mining	45
3.12	Data Mining Algorithms.....	46
3.12.1	Statistical techniques.....	46
3.12.2	Machine Learning	46
3.12.3	Ensemble Approaches	48
3.12.4	Predictive Analysis	49
3.12.5	Existing Systems	50
Chapter-4: Conventional IDS Systems		
4.1	Data Collection.....	53
4.2	Data Conditioning.....	54
4.2.1	Confusion Matrix (CM).....	55
4.3	Neural Network Based IDS Classification Techniques.....	56
4.3.1	Multilayer Perceptron (MLP).....	57
4.3.2	Generalized Feed-Forward (GFF).....	59
4.3.3	Modular Neural Network (MNN).....	60
4.3.4	Jordan/Elman network (JAN).....	62
4.3.5	Principal Component Analysis (PCA) Network.....	63
4.3.6	Radial Basis Function (RBF) Networks.....	65
4.3.7	Self Organized Maps (SOM).....	67
4.3.8	Time Lagged Recurrent Networks (TLRNs).....	68
4.3.9	Recurrent Network (RN).....	70
4.4	Data Mining Based IDS Classification Techniques.....	72
5.4.1	DT-J48 IDS Based.....	72
5.4.2	Naïve Bayes IDS Based	73
4.5	Discussion.....	74

Chapter-5: The Proposed IDS Systems		
5.1	23-signatures' Multi-Stage MLP.....	78
5.2	5-Classes' multi-stage MLP.....	81
5.3	MLP-SVM hybrid approach.....	83
5.3.1	Applying Sequential minimal optimization (SMO) for SVM...	84
5.3.2	23-Signatures Hybrid MLP-Naïve Bayes Classifier.....	86
5.3.3	5-Classes Hybrid MLP-Naïve Bayes Classifier.....	87
5.3.4	5-Classes multistage J48 Classifier.....	87
5.3.5	23-Signatures Hybrid J48- GFF Multistage/hybrid Classifier...	89
5.3.6	5-Classes Hybrid J48- GFF Classifier.....	90
5.4	Neural Networks Architecture reduction.....	92
5.5	Minimize neurons in hidden layer MLP $41 \times 5 \times 2$	93
5.6	Data Feature Reduction (attribute selection).....	95
5.6.1	Feature (Subset) Selection.....	96
5.6.2	Dimensionality Reduction.....	98
5.7	Attribute selection schemes.....	99
5.7.1	Subset Evaluators.....	100
5.7.2	Attribute evaluators.....	101
5.7.3	Attribute transformers.....	103
5.7.4	Search methods.....	104
5.8	KDD Features Reduction.....	107
5.9	MLP-based IDS before & after features reduction.....	109
5.10	DT j48-based IDS before & after features reduction.....	110
5.11	Use PCA to do feature reduction.....	111
5.12	Conclusion.....	112
Chapter-6: Hardware implementation of IDS		
6.1	Synthesis of PCA	116
6.2	ANN-synthesis	119
6.2.1	Design issues.....	119
6.2.2	Taxonomy of FPGA implementations of ANNs.....	120
6.2.3	Design of Single neuron.....	122
6.2.4	Realization of activation function.....	125
6.2.5	Comparison between three Activation Functions.....	133
6.3	MLP synthesis.....	134
6.4	Synthesis of decision tree.....	137
6.5	Synthesis of MLP in the second stage	140
6.6	Synthesis of second stage DT.....	142
6.7	Mixed design.....	144
6.8	Summery.....	145
Chapter-7-: Conclusion		

7.1	Summary	150
7.2	Contributions	151
Appendix		
	Appendix A.....	155
	Appendix B.....	170
References		179
List of Publications		192

Development and Implementation of fast Network Intrusion Detection Systems	العنوان:
Mahmoud, Mohamed Awad Mohamed	المؤلف الرئيسي:
Mohamed, Mohamed Abd Alazim, Abd Alfattah, Abd Alfattah Ibraheem, Tolba, Ahmed Said(Super., Super)	مؤلفين آخرين:
2012	التاريخ الميلادي:
المنصورة	موقع:
1 - 192	الصفحات:
536381	رقم MD:
رسائل جامعية	نوع المحتوى:
English	اللغة:
رسالة دكتوراه	الدرجة العلمية:
جامعة المنصورة	الجامعة:
كلية الهندسة	الكلية:
مصر	الدولة:
Dissertations	قواعد المعلومات:
لكشف الدخلاء، جرائم المعلومات، شبكات الحاسب	مواضيع:
https://search.mandumah.com/Record/536381	رابط:



Mansoura University
Faculty of Engineering
Electronic & Comm. Eng. Dept.

Development and Implementation of Fast Network Intrusion Detection Systems

A Thesis Submitted In Partial Fulfillment for the
Requirements of

Ph.D. Degree
In
Electrical Communications Engineering

By

Eng. Mohamed Awad Mohamed Mahmoud
B.Sc. Electronic and Communications Engineering

Supervised by

Prof. Abdel-Fattah Ibrahim Abdel-Fattah
Professor at the Electronics and Communications Department-Faculty of Engineering-
Mansoura University

Prof. Ahmed Said Tolba
Professor at Faculty of Computers and Information Sciences
Mansoura University

Dr. Mohamed Abdel-Azim Mohamed
Professor at the Electronics and Communications Department-Faculty of Engineering-
Mansoura University

2012

Thesis Title:

Development and Implementation of Fast Network Intrusion Detection Systems

Researcher Name:

Mohamed Awad Mohamed Mahmoud

SUPERVISORS:

No.	Name	Position	Signature
1	Prof. Abdel-Fattah Ibrahim Abdel-Fattah	Department of Electronics and Communication Engineering Faculty of Engineering Mansoura University	
2	Prof. Ahmed Said Tolba	Professor at Faculty of Computers and Information Sciences- Mansoura University	
3	Prof. Mohammed Abdel-Azim Mohamed	Department of Electronics and Communication Engineering Faculty of Engineering Mansoura University	

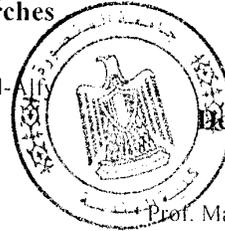
Head of Dept.



Prof. Ahmed S. Samra

**Vice Dean for Post Graduate
Studies and Researches**

Prof. Kasim Salah El-A



Dean of the faculty

Prof. Mahmoud Mohamed Elmelegi



Thesis Title:

Development and Implementation of Fast Network Intrusion Detection Systems

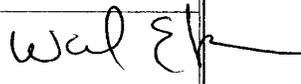
Researcher Name:

Mohamed Awad Mohamed Mahmoud

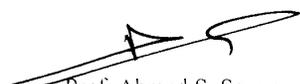
Supervisors:

No.	Name	Position	Signature
1	Prof. Abdel-Fattah Ibrahim Abdel-Fattah	Department of Electronics and Communication Engineering Faculty of Engineering Mansoura University	
2	Prof. Ahmed Said Tolba	Professor at Faculty of Computers and Information Sciences- Mansoura University	
3	Prof. Mohammed Abdel-Azim Mohamed	Department of Electronics and Communication Engineering Faculty of Engineering Mansoura University	

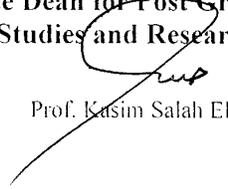
Exam Committee:

No.	Name	Position	Signature
1	Prof. Abdel-Fattah Ibrahim Abdel-Fattah	Department of Electronics and Communication Engineering Faculty of Engineering Mansoura University	
2	Prof. Hassan H. Soliman	Department of Electronics and Communication Engineering Faculty of Engineering Mansoura University	
3	Prof. Wail Shawki Elkilani	Department of Computer Systems Faculty of Computers and Information -Ain Shams University	

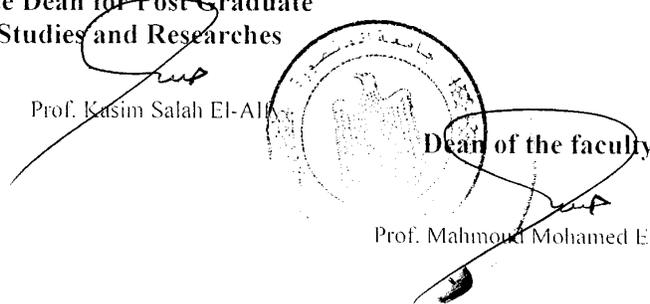
Head of Dept.

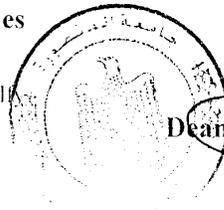

Prof. Ahmed S. Samra

Vice Dean for Post Graduate Studies and Researches


Prof. Kasim Salah El-All

Dean of the faculty


Prof. Mahmoud Mohamed Elmelegi



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

«سُبْحَانَكَ لَا عِلْمَ لَنَا إِلَّا مَا عَلَّمْتَنَا إِنَّكَ أَنْتَ الْعَلِيمُ الْحَكِيمُ»

صدق الله العظيم

(سورة البقرة- الآية 32)

Acknowledgement

First and foremost, I want to thank the source of every success to *ALLAH*;
Next I am deeply thankful to my advisor Prof. Dr. Prof. *Abdel-Fattah Ibrahim Abdel-Fattah* who guided me through this entire journey.

I want to express my gratitude to Dr. *Mohamed Abdel-Azim Mohamed*, whose not only served as my supervisor but also encouraged, valuable guidance, indispensable help and challenged me throughout my academic program. His words of advice, his trust, and his patience and understanding helped me to reach this work. He has been a role model to me. He has always provided support whenever I need.

I owe my deepest gratitude and heartily thankful to my wife and my daughters *Mahy and Lara* whose accompanied me in every stage of this work with their patience, friendship, openness and their caring.

Special thanks to my family, my father, my mother, my sister, and my brothers whose support and unconditional help were invaluable in completing this work. Finally, I would like to seize the opportunity to thank all of you who make up an essential part of my life. Although not directly involved in this work, you have contributed to its completion in many ways.

Mohamed Awad

Abstract

Intrusion Detection Systems (IDSs) have recently received a blooming attention and interest from the scientific community as well as from the public. The interest from the public is mostly due to the recent events of terror around the world, which has increased the demand for useful security systems. There has been significant progress in improving the performance of computer-based intrusion detection systems algorithms over the last decade. Although algorithms have been tested and compared extensively with each other, there has been remarkably little work comparing the accuracy of computer-based IDSs with humans.

Our long term goal is to design and build an intelligent IDS that is accurate (low false negative and false positive rates), flexible, not easily fooled by small variations in intrusion patterns, adaptive in new environments, modular with both misuse and anomaly detection components, robust, able to detect known and unknown intrusions and real-time.

To achieve our goal, two phases of implementation were introduced; software and hardware phases. The software phase is divided into two sub-phases; the first introduces conventional IDS implementation which is based on neural networks and Data mining techniques. The results obtained were unsatisfactory. The second sub-phase introduces a proposed IDS which is based on hybridizing and/or multistaging more than one conventional IDS. The proposed models were successfully able to detect all signatures and all classes of intrusions, in addition to the type of each stream (normal or attack) in the tested database. The first stage of

the multistage classifier is a perfect 2-types anomaly classifier, which ensures securing the network against attack (known and unknown attacks), whereas an alarm could be raised to the system administrator to take the proper action. Feature and topology reduction were one of most important challenges in this thesis due to massiveness of tested dataset and the huge structure of the implemented models. The proposed models were repeated with different neural networks topologies with simpler structure which are based on trial and error. Feature reduction process proceeds in two ways; record reduction by eliminating duplications and feature reduction. By applying Principal Components Analysis (PCA), the original 41-features were transformed to 22 ranked principal components (PCs) and the generated PCs were minimized by eliminating less ranked PCs until reaching a compromise situation between the smallest number of PCs and best detection efficiency.

As the demand on more network speed increases and new network protocols emerge, Network Intrusion Detection Systems (NIDSs) are increasing in importance and are being integrated in network processors. Currently, most IDSs are software running on a general purpose processor. Unfortunately, it is becoming increasingly difficult for software based IDSs to keep up with the increasing network speeds (10Gbps at backbone networks). This has underscored the need for the specialized hardware-based solutions which are portable and operate at wire speeds. Field Programmable Gate Arrays (FPGAs) are generic pieces of hardware

that can be reconfigured to perform any task. FPGA-based platforms can exploit the fact that the NIDS rules change relatively infrequently, and use reconfiguration to reduce implementation cost. In addition, FPGA-based systems can exploit parallelism in order to achieve satisfactory processing robustness.

To achieve robustness of IDS; FPGA-based IDSs were introduced, the hardware phase, based on neural networks and data mining intrusion detection techniques. Through MultiLayer perceptron (MLP) design, we faced common challenge in MLP hardware implementation which is tansig activation function which is difficult to implement due to its nonlinearity. Three techniques for tansig representation were presented, one based on Kwan approximation [172] and the other based on Piecewise Linear Approximation of a Nonlinear function (PLAN) approximation [173]. A third approximation, which is based on symbolic regression and genetic algorithm, is proposed. New challenge arisen in implementing the proposed tansig approximation which is the division process in FPGA. This was solved by using (Intellectual Property) IP CORE available in Xilinx ISE. FPGA-based IDSs, MLP and Decision Tree DT-j48 based, were implemented using two different programming strategies and data formats. Both techniques gave perfect results to detect normal and attack stream. Both implemented single stage IDS was followed by second stage which was able to detected stream's class after detecting stream's type.

List of Abbreviations

List of abbreviations

Symbol	Details
LMS	Least Mean Square rule
Login	Log in Unix command
ls	List Unix command
LVF	Las Vegas Filter
LVQ	Learning Vector Quantization
Madaline	Multiple ADaptive Linear Element
MADAM ID	Mining Audit Data for Automated Models for Intrusion Detection
MIB	Management Information Base
MIB	Management Information Base
MINDS	the Minnesota Intrusion Detection System
MIT	Massachusetts Institute of Technology
ML	Machine Learning
MLP	Multilayer Perceptron
MLS	Multi-Level Security
MNN	Modular Neural Network
NFR	Network Flight Recorder
NFS	Network File System
NID	Network Intrusion Detection
NNTP	Network News Transfer Protocol
NSM	Network Security Monitor
OSI	Open Systems Interconnection
PCA	Principal Component Analysis
PCs	principal components
PE	Processing unit (or Element)
PNN	Probabilistic Neural Network
POP	Post Office Protocol
ps	Active processes Unix command
R2L	Remote to Local
RBF	Radial Basis Function Networks
Relieff	Recursive Elimination of Features
RFC	Request For Comments
RN	Recurrent Network
SATAN	Security Administrator Tool for Analyzing Networks
Simnet	simulation network

List of abbreviations

Symbol	Details
SMO	Sequential Minimal Optimization
SMTP	Simple Mail Transfer Protocol
SNMP	The Simple Network Management Protocol
SNMP	Simple Network Management Protocol
SNMPv1	Simple Network Management Protocol Version 1.
SNMPv2	Simple Network Management Protocol Version 2.
SNMPv3	Simple Network Management Protocol Version 3.
SOM	Self-Organizing Maps
SQ	SQuare
Sudo	Super User Do
SVMs	Support Vector Machines
TCB	Trusted Computing Base
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
TCSEC	Trusted Computer System Evaluation Criteria
TDNN	Tapped Delay Neural Network
TDNN	Time Delay Neural Network
TELNET	TELEcommunications NETwork
TLRNs	Time Lagged Recurrent Networks
TTL	Time To Live
U2R	User to Root
UDP	User Datagram Protocol

List of abbreviations

Symbol	Details
a.k.a. system logs	Authentication and Key Agreement
ADAM	Audit Data Analysis and Mining
ADC	Approximate Distance Clustering
AI	Artificial Intelligence
AIX	Advanced Interactive eXecutive
API	Application Programming Interfaces
ARP	Address Resolution Protocol
ASIC	Application-Specific Integrated Circuit
ATM	Asynchronous Transfer Mode
ATP	Advanced Technology Program
BSM	Basic Security Model
BSM	Basic Security Module
C2	Class 2
CAM	Content Addressable Memories
CANFIS	CoActive Neuro-Fuzzy Inference System
CERT	Computer Emergency Response Team
CFS	Correlation-based Feature Subset
CGI	Common Gateway Interface
CIDF	Common Intrusion Detection Framework
CM	Confusion Matrix
COPS	Computer Oracle and Password System
CPU	Central Processing Unit
CSI	Crime Scene Investigation
C-SVC	regularized support vector classification
DARPA	Defense Advanced Research Projects Agency
DFA	Deterministic Finite Automata
DIDS	Distributed Intrusion Detection System
DIDS	Distributed Intrusion Detection System
DNS	Domain Name System
DoS	Denial of Service
DR	Dimensionality Reduction
DT	Decision Tree
EMERALD	Electronic Management Research Library Database
EPT	Effective Process Time

Symbol	Details
FBI	Federal Bureau of Investigation
FCBI	Fast Correlation-Based
FPGAs	Field Programmable Gate Arrays
FS	Feature subset Selection
ftp	File transfer Protocol
FTPD	File Transfer Protocol Daemon
GA	Genetic Algorithm
GFF	Generalized Feed-Forward
Gido	Generalized intrusion detection objects
GIDS	the Graph based Intrusion Detection System
GRNN	Generalized Regression Neural Network
GUI	Graphical User Interface
HMM	Hidden Markov Model
HTML	Hyper Text Markup Language
HTTP	HyperText Transfer Protocol
ICMP	Internet Control Message Protocol
ID3	Iterative Dichotomiser 3
IDDM	Intrusion Detection using Data Mining
IDES	Intrusion Detection Expert System
IDIOT	Intrusion Detection In Our Time
IDPSs	Intrusion Detection and Prevention System
IDSs	Intrusion Detection Systems
IP	Internet Protocol
IP	Intellectual Property
IPSs	Intrusion Prevention Systems
ISOA	Information Security Officer's Assistant
ISS's	Internet Security Systems
ITU-T	International Telecommunication Union Telecommunication standardization sector
JAM	Java Agents Metal-learning
JEN	Jordan/Elman Networks
KDDCUP' 99	Knowledge Discovery and Data mining tools competition
Klaxon	Klaxon audio signal
LAN	Local Area Network

List of Figures

Figure No.	Figure Caption	Page No.
Fig.(1.1)	The evolution of attack sophistication and devolution of attack skills	2
Fig.(2.1)	Passive Attacks: (a) release of message content, (b) traffic analysis	11
Fig.(2.2)	Active Attacks: (a) Masquerade, (b) replay, (c) Modification of messages, and (d) DoS	13,14
Fig.(2.3)	Model for Network	18
Fig.(2.4)	Network Access Security Model	20
Fig.(3.1)	Components of Intrusion Detection framework	23
Fig.(3.2)	The Placement of the IDS in the Network	24
Fig.(3.3)	A taxonomy of IDS proposed by Debar et al.	27
Fig.(3.4)	Revised Taxonomy of IDS proposed by Debar et al.	28
Fig.(3.5)	State versus transition	38
Fig.(3.6)	Methodologies for IDSs	41
Fig.(4.1)	System Workflow	53
Fig.(5.1)	The intrusion identification steps	76
Fig.(5.2a)	2-cascaded MLPs 23-sig. IDS	79
Fig.(5.2b)	2 nd two-cascaded MLPs 23-sig. IDS	80
Fig.(5.2c)	3 rd two-cascaded MLPs 23-sig. IDS	80
Fig.(5.2d)	4 th two-cascaded MLPs 23-sig. IDS	81
Fig.(5.3a)	2-cascaded MLPs 5-classes IDS	81
Fig.(5.3b)	2 nd two-cascaded MLPs 5-classes IDS	81
Fig.(5.4a)	Directing R2L class to 41×50×50×8 MLP	82
Fig.(5.4b)	Directing DoS class to 41×50×50×6 MLP	82
Fig.(5.4c)	Directing PRB class to 41×50×50×4 MLP	82
Fig.(5.4d)	Directing U2R class to 41×50×50×4 MLP	82
Fig.(5.5)	Hybrid MLP-SMO-based SVM 23-sig. classifier	83
Fig.(5.6a)	Hybrid MLP-(SMO/SVM) 23-sig. classifier	84
Fig.(5.6b)	2 nd stage MLP-(SMO/SVM) 23-sig. classifier	85
Fig.(5.7)	Hybrid MLP-Naïve-Bayes 23-sig. classifier	86
Fig.(5.8)	Hybrid MLP-Naïve Bayes 5-classes classifier	87
Fig.(5.9)	Hybrid MLP-J48 5-classes classifier	87
Fig.(5.10)	Hybrid MLP-J48 23-sig. classifier	88

Figure No.	Figure Caption	Page No.
Fig.(5.11a)	Hybrid multistage GFF-J48 23-sig. classifier, 1 st stage	89
Fig.(5.11b)	Hybrid multistage GFF-J48 23-sig. classifier, 2 nd stage	90
Fig.(5.12)	Hybrid multistage GFF-J48 5-classes classifier	91
Fig.(5.13)	GFF per each class classifiers	91
Fig.(5.14a)	Hybrid GFF-J48 23-sig. classifier, 1 st stage	92
Fig.(5.14b)	Hybrid GFF-J48 23-sig. classifier, 2 nd stage	92
Fig.(5.14c)	Hybrid GFF-J48 23-sig. classifier, 3 rd stage	93
Fig.(5.14d)	Hybrid GFF-J48 23-sig. classifier, 3 rd stage	93
Fig.(5.14e)	Hybrid GFF-J48 23-sig. classifier, 4 th stage	93
Fig.(5.14f)	Hybrid GFF-J48 23-sig. classifier, 5 th stage	93
Fig.(5.15a)	2 cascaded MLPs, 1 st stage	94
Fig.(5.15b)	2 cascaded MLPs, 2 nd stage	94
Fig.(5.15c)	2 cascaded MLPs, 3 rd stage	94
Fig.(5.15d)	2 cascaded MLPs, 4 th stage	94
Fig.(5.15e)	2 cascaded MLPs, 5 th stage	94
Fig.(5.15f)	Multistage and hybrid concepts for detecting more intrusions in low complexity MLP	95
Fig.(6.1a)	7 PCs Model Simulation	118
Fig.(6.1b)	16 PCs Model Simulation	118
Fig.(6.2)	Mathematical model of artificial neuron.	122
Fig.(6.3)	VHDL structural diagram for neuron implementation.	123
Fig.(6.4)	The equivalent hardware architecture of the neuron	123
Fig.(6.5a)	Timing diagram Simulation result of single neuron without activation function	124
Fig.(6.5b)	Hardware circuit of single neuron without activation function tansig activation function	124
Fig.(6.6)	Comparison between original tansig and Kwan approximation.	126
Fig.(6.7)	Block diagram of the Kwan tansig activation function	126
Fig.(6.8)	Hardware circuit of implementing neuron with Kwan tansig activation function.	127
Fig.(6.9)	Simulation result of Kwan tansig activation function.	127
Fig.(6.10)	Comparison of Tangent-Sigmoid Function and PLAN Approximation	128

Figure No.	Figure Caption	Page No.
Fig.(6.11a)	Single neuron with The PLAN approximation activation function timing diagram.	128
Fig.(6.11b)	Single neuron with The PLAN approximation activation function RTL circuit diagram.	129
Fig.(6.12)	Comparison between original tansig, PLAN approximation and new proposed function	130
Fig.(6.13)	Hardware blocks for implemented tansig function	131
Fig.(6.14)	The simulation of the proposed function.	132
Fig.(6.15)	Area report comparison between three models	133
Fig.(6.16)	7×5×2 MLP blocks	135
Fig.(6.17)	7×5×2 MLP timing diagram for fixed-point data representation.	135
Fig.(6.18)	7×5×2 MLP timing diagram for integer data representation	136
Fig.(6.19a)	First stage Normal/Attack classifier based on DT	137
Fig.(6.19b)	Part of original DT.	138
Fig.(6.20)	Single Stage DT simulation result.	139
Fig.(6.21)	Two stages MLPs	141
Fig.(6.22)	Two Stages MLP simulation	141
Fig.(6.23a)	Second Stage DT	142
Fig.(6.23b)	Part of second Stage DT	143
Fig.(6.24)	Two stage DTs simulation	144
Fig.(6.25)	Mixed MLP-DT simulation	145

List of Tables

List of Tables

Table No.	Table Caption	Page No.
Table 4.1	23-signatures MLP	58
Table 4.2	5-classes MLP	58
Table 4.3	2-types MLP	59
Table 4.4	23-signatures GFF	59
Table 4.5	5-classes GFF	60
Table 4.6	2-types GFF	60
Table 4.7	23-signatures MNN	61
Table 4.8	5-classes MNN	61
Table 4.9	2-types MNN	62
Table 4.10	23-signatures JAN	62,63
Table 4.11	5-classes JAN	63
Table 4.12	2-types JAN	63
Table 4.13	23-signatures PCA	64
Table 4.14	5-classes PCA	65
Table 4.15	2-types PCA	65
Table 4.16	23-signatures RBF	65
Table 4.17	5-classes RBF	66
Table 4.18	2-types RBF	66
Table 4.19	23-signatures SOM	67,68
Table 4.20	5-classes SOM	68
Table 4.21	2-types SOM	68
Table 4.22	23-signatures TLRNs	69
Table 4.23	5-classes TLRNs	69
Table 4.24	2-types TLRNs	70
Table 4.25	23-signatures RN	70,71
Table 4.26	5-classes RN	71
Table 4.27	2-types RN	71

List of Tables

Table No.	Table Caption	Page No.
Table 4.28	23-signatures J-48 DT	72
Table 4.29	5-classes J-48 DT	73
Table 4.30	2-Types J-48 DT	73
Table 4.31	23-signatures Naïve-Bayes Classifier	73
Table 4.32	5-classes Naïve-Bayes Classifier	74
Table 4.33	2-types Naïve-Bayes classifier	74
Table 5.1	23-signatures MLP	79
Table 5.2	Attribute Selection Schemes	99
Table 5.3	Feature Reduction in Case of 23-sig. Classification	107
Table 5.4	23-sig. MLP	107,108
Table 5.5	23-sig. DT	108
Table 5.6	Feature Reduction in Case of 5-Classes Classification	109
Table 5.7	Feature Reduction in Case of 2-types Classification	110
Table 5.8	5-classes Classifiers Before and After Feature Reduction	110
Table 5.9	2-types MLP Classifiers Before and After Feature Reduction	110
Table 5.10	5-classes DT Classifiers Before and After Feature Reduction	111
Table 5.11	2-types DT Classifiers Before and After Feature Reduction	111
Table 5.12	22×50×23 MLP Based on New 22 PCA Features	112
Table 5.13	MLP 22×50×5 for 5-classes Based on 22 PCAs	112
Table 5.14	MLP 22×50×2 for 2-types Based on 22 PCAs	112
Table 6.1	Ranked PCAs	116
Table 6.2 (a)	Seven PCs Used in MLP Download Area Report	117
Table 6.2 (b)	16 PCs Used in DT Download Area Report	117
Table 6.3	Area Report of Single Neuron Without Activation Function	125
Table 6.4	Area Report for Approximated Kwan Tansig Function	127
Table 6.5	Single Neuron with PLAN Tansig Approximation Area Report	129
Table 6.6	Design Summary area Report of Proposed Function	132

List of Tables

Table No.	Table Caption	Page No.
Table 6.7	Mathematical Comparison Between Three Models	133
Table 6.8	7×5×2 MLP Area Report for Fixed-point Data Representation	136
Table 6.9	7×5×2 MLP Area Report for Integer Data Representation	137
Table 6.10	Single Stage DT Area Report	139
Table 6.11	PCA-Single Stage DT Area Report	140
Table 6.12	Two Stages MLPs Area Report	140
Table 6.13	Two Stages DTs Area Report	144
Table 6.14	Mixed MLP-DT Area Report	145

Development and Implementation of fast Network Intrusion Detection Systems	العنوان:
Mahmoud, Mohamed Awad Mohamed	المؤلف الرئيسي:
Mohamed, Mohamed Abd Alazim, Abd Alfattah, Abd Alfattah Ibraheem, Tolba, Ahmed Said(Super., Super)	مؤلفين آخرين:
2012	التاريخ الميلادي:
المنصورة	موقع:
1 - 192	الصفحات:
536381	رقم MD:
رسائل جامعية	نوع المحتوى:
English	اللغة:
رسالة دكتوراه	الدرجة العلمية:
جامعة المنصورة	الجامعة:
كلية الهندسة	الكلية:
مصر	الدولة:
Dissertations	قواعد المعلومات:
لكشف الدخلاء، جرائم المعلومات، شبكات الحاسب	مواضيع:
https://search.mandumah.com/Record/536381	رابط:



جامعة المنصورة
كلية الهندسة
قسم هندسة الإلكترونيات والاتصالات

تطوير وتنفيذ نظم سريعة لكشف الدخلاء علي شبكات الحاسب

توطئة للحصول على درجة دكتوراة الفلسفة
فى هندسة الإتصالات الكهربائية

رسالة مقدمة من

المهندس / محمد عوض محمد محمود

بكالوريوس هندسة الإلكترونيات والاتصالات
كلية الهندسة - جامعة المنصورة

إشراف

أ.د. عبدالفتاح إبراهيم عبدالفتاح
أستاذ متفرغ بقسم هندسة الإلكترونيات والاتصالات
كلية الهندسة - جامعة المنصورة

أ.د. أحمد سعيد طلبة
أستاذ بكلية الحاسبات ونظم المعلومات
جامعة المنصورة

د.م. محمد عبد العظيم محمد
مدرس بقسم هندسة الإلكترونيات والاتصالات
كلية الهندسة - جامعة المنصورة

2012



Mansoura University
Faculty of Engineering
Electronic & Comm. Eng. Dept.

Development and Implementation of Fast Network Intrusion Detection Systems

A Thesis Submitted In Partial Fulfillment for the
Requirements of

Ph.D. Degree
In
Electrical Communications Engineering

By

Eng. Mohamed Awad Mohamed Mahmoud
B.Sc. Electronic and Communications Engineering

Supervised by

Prof. Abdel-Fattah Ibrahim Abdel-Fattah
Professor at the Electronics and Communications Department-Faculty of Engineering-
Mansoura University

Prof. Ahmed Said Tolba
Professor at Faculty of Computers and Information Sciences
Mansoura University

Dr. Mohamed Abdel-Azim Mohamed
Professor at the Electronics and Communications Department-Faculty of Engineering-
Mansoura University

2012