A Propose Neuro-Fuzzy-Genetic Intrusion Detection System

Ibrahim Goni
Federal Polytechnic Kaura Namoda
Department of Computer Science

Ahmed Lawal
Federal Polytechnic Kaura Namoda
Department of Computer Science

ABSTRACT
The exponential growth and development of the internet has created many problems on network security. Current intrusion detection system has failed to fully protect system against sophisticated attacks. This research work explores some dedicated methodologies such as Artificial Neural Network (ANN), Fuzzy Logic, and Genetic Algorithms applied to Intrusion Detection Systems but attacks against networks and information systems are still successful. We proposed Neuro-fuzzy Genetic Intrusion Detection System which is a fusion of the three Artificial Intelligence techniques. We foresee they would stand a fighting chance against any sophisticated attack, improve accuracy, precision rate and reduce the false positive rate and would protect data integrity, confidentiality and availability. We also discuss the dataset for evaluating the system. In this work we have identified a new research direction in the related field.

Keywords
Neuro-fuzzy, Genetic algorithm, Artificial Neural Network, Fuzzy logic, intrusion detection system and Dataset.

1. INTRODUCTION
The rapid expansion of computer networks and mostly the internet has created many stability and security problems[4]. With the tremendous growth of network-based services and sensitive information on attacks, network security is becoming more and more importance than ever before [16]. Information system security is the integrity and safety of its resources and activities [30]. In the cyber world, it can be almost impossible to trace sophisticated attacks to their true source. The anonymity enjoyed by todays cyber attackers poses a grave threat to the global information society, the progress of an information based international economy and the advancement of global collaboration and cooperation in all areas of human endeavor [30]. Network security is of primary concerned nowadays for large organizations the Intrusion Detection System are becoming indispensable for effective protection against attacks that are constantly changing in magnitude and complexity [2]. Computer systems are turning out to be more and more susceptible to attack due to its extended network connectivity [11]. Traditional intrusion prevention techniques such as firewalls, access control and encryption have failed to fully protect networks and systems from increasingly sophisticated attacks and malware [12]. At the same time various researchers have performed studies using soft based computing for network intrusion detection including fuzzy logic, artificial neural network, probabilistic reasoning, and genetic algorithms. [22]. Despite the plethora of intrusion prevention techniques available attacks against computer system are still successful [16].

In the current scenario, information system security is an important issue for all the companies and institutions, government, finance and telecommunication [1]. A number of recent events we finally take up the challenge of intrusion detection system problems. Intrusion detection for computer system is a key problem of today's internet connected society [8]. An effective network security strategy require identifying threats and then choosing the most effective set of tools to combat them [5]. Computer security can be very complex and may be very confusing to many people [7]. It can be even a controversial subject [7]. To circumvent such limitations although some dedicated methodologies have already been proposed using Artificial intelligence techniques. Recent research and application employing non-analytical methods of computing such as Fuzzy logic, evolutionary computation and neural networks have demonstrated the utility and potential of these paradigms for intelligent control of complex system [20].Zadeh proposed the denomination soft computing to address the field of Neural Networks Genetic algorithms, Fuzzy logic and combination of those [23].soft computing is the fusion of methodologies that were designed to model and enable solutions to real world problems which are not modeled or too difficult to model mathematically [28]. Soft computing is a general term for describing a set of optimization and processing techniques that exploit the tolerance for imprecision, uncertainty, partial truth and approximation to achieve robustness and low solution cost [28]. Various data mining machine learning and soft computing are applied to Intrusion Detection System to detect attacks on the network. The adoption of probabilistic representation and statically learning methods has led to a large degree of integration and cross-fertilization between artificial intelligence machine learning, statistics control theory, Neuroscience and other fields [25]. An IDS has received a lot of attention in the past decade. The aim of this research work is to propose Neuro-fuzzy Genetic Intrusion detection System and also to explore the research contributions on the application of Artificial Intelligence techniques in Intrusion Detection system.

Intrusion Detection systems are becoming indispensable for effective protection against attacks that constantly changing and complexity [2]. Intrusion detection system adopts mainly two strategies for detection of threat, the signature based detection technique and anomaly based detection technique [10]. An intrusion detection system can be defined as the tools, methods, and resources to help identify asses and report unauthorized or unapproved network activity [28]. To be specialized tool that knows to parse and interpret network traffic and/or host activities [28]. Information Technology Security it concern with protection of computer systems, the networks interconnecting such systems, and the information stored, processed and transmitted within the systems and networks against intentional attacks[15]. IDS often store a database of known signatures can compare patterns of activity traffic, or behavior it sees. In the data it's monitoring agent those signatures to recognized when a close match between a signature and current and recent behavior occurs [28] at that point the IDS can issue alarms or alert take various kind of automated actions ranging from shutting down internet link or

specific servers to lunching back trace and make other active attempts to identify attacker and collect evidence of their reform activities [28].Intrusion detection system (IDS) is split into categories; Misuse detection and anomaly detection is used to identify intrusion that match known attack scenario. However anomaly detection is an attempt to search for malicious behaviors that deviate from established normal pattern [3].

Anomaly detection is based on the normal behavior of a subject (e.g. a user or a system) any action that significantly deviates from the normal behavior is considered intrusive [6]. Misuse detection is based on the characteristics of known attacks or system vulnerabilities, which are also called signatures. Any action that matches the signature is considered intrusive [6]. Misuse based detection detect attacks based on signature (known attacks signatures) at which the traffic pattern compared with this signature, if a match is formed, then it reported as an attack, otherwise it is not [6] so misuse detection cannot detect novel attack [6]. The anomaly based intrusion detection approaches responds to deviation from normal behavior, which typically involves the creation of knowledge based that contains the profile of monitored activities [14]. Our objective is to provide a roadmap of the available approaches of IDS in the field of AI.

[7] divide intrusion into six (6) main types;

1. Attempts break-ins which are detected by typical behavior profiles or violation of security constraint.
2. Masquerade attacks, which are detected by typical behavior profile or violation of security constraint
3. Penetration of security control system, which are detected by monitoring for specific patterns of activities.
4. Leakage, which is detected by typical use of system resources
5. Denial of service, which is detected by typical use of system resources.
6. Malicious use, which is detected by a typical behavior profiles, violation of security constraints, or use of special privilege.

All the above mentioned attacks have different ways and tactics to penetrate into information system networks among others. Current Intrusion detection system may not have all the means to protect against different attacks it may protect system against masquerade for instance but cannot protect system against denial of service attack and so on as [13] dictate that it is not possible to protect system completely. And [19]dictate that it is important to note that for each type of attack, a different set of features dominated the distance calculation for the attack traffic.

[7]Identified the characteristics of intrusion detection system regardless of what mechanism and IDS is based, it must do the following;

1. Run continuously without human supervision
2. Be fault tolerant and survivable
3. Resist subversion
4. Impose minimal overhead
5. Observe deviation from normal behavior
6. Be easily tailored to specific network
7. Adopt to changes overtime and
8. Be difficult to fool.

1.1 Challenges for Intrusion Detection System

Intrusion Detection System (IDS) products have failed to keep up with the rapid advancement in switching and bandwidth growth and increased sophisticated attacks that need to be handle today. Current IDS products often operate in a monitoring-only mode for example “Sniffers” which can detect attacks but cannot effectively and reliably block malicious traffic before the damage is done [30].

[30] In her Book identified nine (9) challenges of IDS:

1. Inaccurate detection
2. Incomplete attack coverage
3. More detection, less prevention
4. Designed primarily for sub-100Mbps network
5. Performance challenged
6. High availability deployment not available
7. Scalability issues
8. Security policy enforcement related issues
9. Require significant information technology.

2. ARTIFICIAL INTELLIGENCE TECHNIQUES

In this research work three Artificial Intelligence techniques were used these are Artificial Neural Network, Fuzzy Logic and Genetic algorithms.

2.1 Artificial Neural Network

Artificial Neural Network (ANN) often just called a “Neural Network” (NN) are sets of mathematical models based on biological neural networks (nerve cells) assigning different weights to connections between elements within the neural network similarly to how electrical potentials for neurons are built up at synaptic junctions based on their frequency of firing [28]. It consists of interconnected group of artificial neurons and process information using connectionist approach to computation [28]. Artificial Neural Network is an adaptive based on internal information that flows through the network during the learning phase [28]. ANN consists of nodes connected together with links between them, where there is a directed flow of data. These connections have numeric weights to determine how much one node will affect the other [14]. The numeric weight can be manipulated until desired outputs are attained [14].

Artificial neural network can be most adequately characterized as computational models with particular properties such as ability to adopt or learn, to generalize, or to cluster or organized data and which operation is based on parallel processing [18]. Artificial neural network consist of a pool of simple processing units which communicate by sending to each other over a large number of weighted connections [18]. Advantage of Neural Networks, However, includes their high tolerance of noisy data as well as their ability to classify patterns on which they have not been trained. They are well suited for continuous valued inputs and outputs [31].

[24] Identified the tasks for which the connectionist approach is well suited include;

1. Classification, deciding the category or grouping to which an input value belongs
2. Pattern recognition, identifying structures in sometimes noisy data
3. Memory recall, including the problem of content addressable memory
4. Optimization, finding the "best" organization of constraints
5. Noise filtering, or separating signal from background, factoring out the irrelevant components of a signal. s

2.2 Fuzzy logic

Fuzzy Logic is regarded as one of the artificial Intelligence techniques from which "conventional" expert system, Neural Network and Genetic algorithms are well known[23]. Fuzzy logic is a superset of conventional logic that has been extended to handle the concept of partial truth values between the Boolean dichotomy of true or false [26]. Fuzzy logic system is unique in that, it is able to simultaneously handle numeric data and linguistic knowledge [17]. It is also a nonlinear mapping of an input data vector into a scalar output [17]. Fuzzy logic systems have been applied to successfully to a broad range of problems in different application domains [29]. Fuzzy logic is an intelligent method that has been successfully employed for many intrusions detection system [2]. The concept of fuzzy logic in solving the problem of intrusion detection because fuzzy logic is an effective tool for introducing the concept of membership degree that determines the "strength" in which an object belong to different classes [2]. Fuzzy intrusion Recognition Engine (FIRE) is a network intrusion detection system that uses fuzzy system to assess malicious activity against computer network [5]. The result of [5] shows that fuzzy system can easily identify port scanning and denial of service attacks. The system can be effective at detecting some types of backdoor and Trojan horse attacks. According to Zadeh the essential characteristics of Fuzzy logic are;

1. Exact reasoning is viewed as a limiting case of approximate reasoning
2. Everything is a matter of degree
3. Any logic system can be fuzzified
4. Knowledge is interpreted as collection of variables
5. Inference is viewed as a process of propagation of fuzzy constraints [21].

2.3 Genetic Algorithm

The process of a genetic algorithm usually begins with a randomly selected population of chromosomes. The chromosome are representations of the problem to be solved according to the attributes of the problem different position of each chromosomes are encoded as a bits, characters or number, this position are sometimes referred to as genes and are changed randomly within a range during evolution[28]. Genetic algorithm can be used to keep the number of iterations as well as possible. Genetic algorithm randomize values for each sets and mix and differentiate values from each of these this continued until the desired result is reached or the maximum number of generation has passed [27].

[10] Identified four (4) parameters used in genetic algorithms;

1. Fitness Function: The fitness function evaluates the quality of a particular solution.
2. Selection: Selection is the process of choosing solution with better fitness function than their counterparts.
3. Crossover: Crossover is the phase in which two solutions exchange one of their characteristics with the other in the pair at a randomly selected crossover point, where the crossover probability is between 0.6 and 0.9
4. Mutation: Mutation is a process by which some random bits in a solution are changed. This is done mainly to maintain the genetic diversity of the solutions..

3. KDD CUP 99 DATASET

In 1998 DARPA in concert with Lincoln laboratory at MIT launched the DARPA 1998 dataset for evaluating IDS [11]. It contain seven weeks for training and also two weeks for testing data in total there are 38 attacks in training data as well as testing data [11]. The third international Knowledge Discovery and Data Mining tools competition was held in colligation with KDD-99[11]. KDD-99 has been the most widely used data set for evaluations of anomaly detection methods [9]. In the first step the KDD data set will be taken as input, in the next step dataset be refined [9].

KDD-99 consist of approximately 494020 data instances, each of which is a vector of extracted feature values from a connection record obtained from the row network data gathered during the simulated intrusions of 42 various quantitative and qualitative feature [5]. KDD-99 attacks fall into four categoriesDenial of Service Attack (DOS), User to Remote Attack (U2R), Remote to Local Attack (R2L) and Probing Attack [9]. Examples of DOS are Mail bomb, ping of death and apache [9]. Examples of R2L are Dictionary, Gest, Imap and Named [9]. Examples of U2R are Eject, Perl and Loadmodule [9].

4. PROPOSED SYSTEM

Neuro-fuzzy Genetic intrusion detection system were proposed in this research work which consist of Artificial neural network, Fuzzy logic and Genetic algorithm, they are artificial intelligence techniques that already been applied to complex problems that would otherwise require the human expert as in Intrusion Detection Systems. One or combination of the two has been applied to Intrusion Detection Systems from different researchers as in [1], [2], [3], [4], [5], [6], [8], [9], [10], [11], [12], [13], [14], [15], [16], [19]. But they have certain limitations as in [13] shows that it is not possible to protect system completely. Perhaps the combination of the three techniques they would stand a fighting chance against intruders and protect data integrity, confidentiality, and availability. This approach would provide high sense of fault tolerance, high computational speed, observe deviation from normal behavior and also adapt to changes overtime. As in [12] we also see a trend to applying Soft Computing to intrusion detection problems. Tightly or loosely assembling different methods in a cooperative way definitely improves the performance of IDS. The most popular combinations are genetic-fuzzy and genetic-neuro systems. The interest in integrating fuzzy sets as a part of these solutions is noticed. [12] also state that Soft computing exploits tolerance for imprecision, uncertainty, low solution cost, robustness, and partial truth to achieve tractability and better correspondence

to reality. Their advantages, therefore, boost the performance of intrusion detection systems. Evolutionary computation and artificial neural networks automatically construct fuzzy rules from training data, and present knowledge about intrusion in a readable format; evolutionary computation designs optimal structures of artificial neural networks. These methods in soft computing collectively provide understandable and autonomous solutions to IDS problems. In [12] also pointed out that tightly coupled soft computing systems are also known as hybrid systems. In a hybrid system, approaches are mixed in an inseparable manner. Neuro-fuzzy systems, genetic-fuzzy systems, genetic-neuro systems and genetic-fuzzy-neuro systems are the most visible systems of this type. Comparatively, loosely coupled soft computing systems, or ensemble systems, assemble these approaches together.

5. CONCLUSION

Neuro-fuzzy Genetic Intrusion Detection System were proposed in this research work, the research contributions on the application of Artificial Neural Network, Fuzzy logic Genetic algorithm and combination of any of the two were explore and identified some of their limitations. Then we proposed Neuro-fuzzy Genetic Intrusion Detection System which would be a new research direction and are literature for the related field. The dataset that would be used to evaluate the system were also briefly discussed in this work.

6. ACKNOWLEDGEMENT

We are extremely thankful to our parents, colleagues, friends Mentors, and well-wishers, for their encouragement and support morally and otherwise.

In recognition of this, here are the individuals whom we thank sincerely; Ibrahim Yahaya, Aishatu Ibrahim, Alh. Mahmoud Jandanis, Jerome M. Gumpy, Dr. Adamu Wakili, Murtala Mohammed, Kabir Mohammed, A'smau Wasuru Ahmad Wasuru, Shehu Tukur, and Auwal Mohammed S.

7. REFERENCE

[1] Barejpal Signh and Ahlawat 2013. Intrusion Detection of Network Attacks using Artificial Neural Network & Fuzzy Logic. International Journal of engineering and Management Technology. Vol. I pp. 53-66.

[2] Ben B. Sujithu, R. Roja Ramania and Parameswari 2012. Intrusion Detection System using Fuzzy Genetic approach. International Journal of Advanced Research in Computer and Communication Engineering. Vol. I. ISSN: 2278-1021.

[3] Gang Wang, Jinxing Hao, Jian Ma and Lihua Huang 2010. A new approach to Intrusion Detection using Artificial Neural Networks and Fuzzy clustering Expert systems with Applications. Available @ www.elsevier.com/locate/eswa

[4] Iftikhar Ahmad, Azween B. Abdullah, and Abdullah S. Alghamdi, Artificial Neural Network approach to Intrusion Detection. A Review Proceeding of the WSEAS International Conference on Telecommunication and Informatics. ISSN: 1790-5117, ISBN:978-474-084-0. Pp. 101-111.

[5] John E. Dickerson, Juka Juslin, Ourania Koukousoula and Jlie A. Dickerson, Fuzzy Intrusion Detection. Electrical and Computer Engineering Department Lowa State University Ames USA.

[6] Khattab M. Ali Lheeti and Raed I. Hameed, 2012. Application of Fuzzy Neural Network combined with an Expert Petri Net System to Intrusion Detection System. 13th International Arab Conference on Information Technology ACIT ISSN 1812-0857.

[7] Kamlesh Lahre, Tarun Duwan, Surresh Kumar, K. and Pooja Agrawal, 2013. Analysis Different approaches for using KDD-99 Dataset. International Journal on Recent and Innovation Trend in Computing and Communication. Vol. I. ISSN: 2321-8169 available @ *http//*www.ijritcc.org.

[8] Major Denis J.I., H. Steven K. and Neil C. Rowe. Distributed Intrusion Detection for Computer Systems using Communicating agent.

[9] Prabhdeep Kaur and Sheveta Vanshisht, 2013. Mingle Intrusion Detection System using Fuzzy Logic. International Journal of Engineering and Advanced Technology. ISSN: 2249-8958, Vol. II.

[10] Rajdeep Borgohain. FUGeIDS: Fuzzy Genetic Paragms in Intrusion Detection System.

[11] R. Shanmugavadivu and N. Nagaran. Network Intrusion Detection System using Fuzzy Logic. Indian Journal of Computer Science and Engineering. Vol. II. No. I. ISSN: 0976-5166.

[12] Shelly Xiaonan Wu and Wolfgana, 2008. The Use of Computational Intelligence in Intrusion Detection System: A Review Technical Report# 2008-05. Department of Computer Science Memorial University of Newfoundand st. John's NL Canada.

[13] S. Revathi and A. Malathi, 2014. Network Intrusion Detection Based on Fuzzy Logic. International Journal of Computer Application. Vol. I. available online @ http// www.rspublication.com/ijca_index .htm

[14] Tinny Magabile, Ishmel S. Msiza and Erick Dube, 2012. Anomaly Based Intrusion Detection for a Biometric Identification System using Neural Networks. International Conference on Artificial Intelligence and Image Processing (ICAIIP 2012) Oct. 6-7, 2012 Dubai.

[15] Ulf Lindqvist, 1999. On the Fundamentals of Analysis and Detection of Computer Misuse. ISBN 91-7197-832-1 Ph.D. Thesis.

[16] Moham Banerjee and Roopali Soni, 2013. Design and Implementation of Network Intrusion Detection System by using K-means Clustering and Naïve Byes. International Journal of Science, Engineering and Technology Research (IJSETR) Vol. II. ISSN: 2278-7798. Pp. 756-760.

[17] Jerry M. Mendle, 1995. Fuzzy Logic System for Engineering. A Tutorial Proceeding of the IEEE 83(3); 345-377.

[18] Ben Krose and Patrick Vander Smagt, 1996. An Introduction to Neural Networks. Eight Edition Faculty of Mathematics and Computer Science University of Amsterdam.

[19] Chong Eik Loo, Mun Yong Ng, Christopher Leckie and Marimuthu Palaniswani, Intrusion Detection for Routing Attacks in Sensor Network, NICTA Victoria Laboratory Department of Computer Science and Software Engineering University of Melbourne Parkville Australia.

[20] Vamsi Mottan P. 2005. Fuzzy Logic Controller for an Automous Mobile Robot, M.Sc. thesis Electrical Engineering Cleveland State University.

[21] Mohd Junaize Mohd Noor, 2004. Application of Knowledge-Based Fuzzy Inference System on High Voltage Transmission line Maintenance Master of Engineering thesis in Electric and Electronic System Engineering Queensland University of Technology.

[22] Brian Eugene Lavender, 2010. Implementation of Genetic algorithms into a Network Intrusion Detection System (netGA) and Integration into nProbe. M.Sc. Dissertation California State University Sacramento.

[23] Rene Jager, 1995. Fuzzy Logic in Control. Ph.D. thesis Technische Universiteit Deft Amsterdam

[24] George F. Luger and William A. Stubblefield, Artificial Intelligence Structures and Strategies for complex problem solving. Addison Wesley Longman, Inc one Jacob way Reading ISBN 0-805-31196-3 Third Edition.

[25] Research Priorities for Robust and Beneficial Artificial Intelligence. Last Update 2015.

[26] Michele Pirovano, 2012. The Use of Fuzzy logic for Artificial Intelligence in Games. Department of Computer Science, University of Milano, Milano Italy.

[27] Matti Manninen. Using Artificial Intelligence in Intrusion Detection System. Helsinki University of Technology.

[28] V. Bapuji, R. Naveen Kumar, A. Govardham and S.S.V.N. Sama, 2012. Soft Computing and Artificial Intelligence Techniques for Intrusion Detection system. Network and Complex Sytem ISSN: 2225-0603.

[29] Majid Almarashi, 2012. Learning of Type-2 Fuzy Logic Systems using Simulated Annealing. Ph.D. thesis in Artificial Intelligence Department of Informatics DE MONTFORT University.

[30] Nina Godbole, 2009. Information Systems Security. Security Management, Metrics, Frameworks, and best practice. Wiley India P.V.T. L.td.

[31] Jiawei Han and Micheline Kamber, 2006. Data Mining Concepts and Techniques. Morgan Kamfman publishers. An imprint of Elsevier second edition.