Int. J. Electron. Commun. (AEÜ) 74 (2017) 94–106

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

Regular paper

Implementation of Intrusion Detection System using Adaptive Neuro-Fuzzy Inference System for 5G wireless communication network

Reeta Devi [a], Rakesh Kumar Jha [a,*], Akhil Gupta [a], Sanjeev Jain [b], Preetam Kumar [c]

[a] Department of Electronics and Communication Engineering, Shri Mata Vaishno Devi University, Katra, J&K, India
[b] Department of Computer Science and Engineering, Shri Mata Vaishno Devi University, Katra, J&K, India
[c] Department of Electrical Engineering, Indian Institute of Technology, Patna, India

ARTICLE INFO

Article history:
Received 23 August 2016
Accepted 20 January 2017

Keywords:
5G
Adaptive Neuro-Fuzzy Inference System
Intrusion Detection System
Relay

ABSTRACT

In the present scenario, the demand for extended coverage, low latency and increased data rate is the need of the hour. Many emerging 5G technologies like massive Multiple Input Multiple Output and device to device communication are introduced in order to meet these increasing demands. For extending the coverage region, new wireless components are introduced in the network like relays, Small Cell Access Point and hotspots. But these components are highly vulnerable to security breaches and thus provides an entry point for the intruder to enter into the network. In this paper, a general 5G wireless communication network with an incorporated relay is proposed. This paper focuses on the implementation of Intrusion Detection System using Adaptive Neuro-Fuzzy Inference System using KDD cup 99 data set for detecting an attack on the relay. The effect of varying the membership function and learning algorithms have also been analyzed.

1. Introduction

Mobile wireless communication has seen a remarkable growth in the data traffic over the past few years because of the extensive use of intelligent devices and applications. The mobile data traffic is increasing on a larger rate. In the past 10 years, it has grown 4000-fold and in the coming years, it is expected that the monthly global mobile data traffic will reach up to 30.6 exabytes. It is expected that by 2020, the mobile-connected tablets will generate nearly eight times more traffic as compared to the traffic generated in 2015 [1]. In an era of 5G, researchers all over the world are looking for robust and efficient wireless transmission technologies. To achieve this, they are trying to incorporate new components and new technologies in the present network architectures [1].

In the coming years, user demand is going to be very high in terms of large coverage region, high speed and greater efficiency. Along with this, we are aspiring of a completely connected and mobile scenario, in which every person is connected with every other person by having an unbound access anywhere and anytime. This may result in many serious challenges including coverage region, spectrum utilization, security, latency, data rate etc. We have started the journey from 1G and now we have reached up to 4G, but still, the user is not satisfied, which has propelled us to migrate from 4G to 5G very soon. There are certain emerging technologies that need to be considered and will be an integral part of 5G like massive Multiple Input Multiple Output (MIMO), device to device communication (D2D), millimeter waves, moving networks etc.

5G is a heterogeneous wireless technology that includes relays, wi-fi hotspots, small cells, micro cells and macro cells. These divisions not only helps in increasing the coverage region but also in combating the near far problem. This will also reduce the load at the Base Station (BS) up to a considerable extent. But from the security point of view, these components are posing serious problems by providing an active site for the attacker to attack. Thus these sites are highly vulnerable to attacks because any unauthorized user can easily penetrate into the network through these sites.

In the wireless industry, we are evolving from closed hierarchical networks to flat networks. The flat networks are more vulnerable to security attacks. Instead of using expensive Radio Access Network (RAN) equipment's, the use of femto cells, small cells and Wi-Fi hot spots are cost effective but will increase the possibility of an attack. In the evolving wireless generation, we have been evolved from voice based networks to powerful data-centric

* Corresponding author.
E-mail addresses: reetarajput87@gmail.com (R. Devi), jharakesh.45@gmail.com (R.K. Jha), akhilgupta112001@gmail.com (A. Gupta), dr_sanjeevjain@yahoo.com (S. Jain), pkumar@iitp.ac.in (P. Kumar).

http://dx.doi.org/10.1016/j.aeue.2017.01.025
1434-8411/

devices which are visible from internet, and thus creating more entry points for attacks [2].

With the evolving wireless communication industry, the attackers are also evolving and they are searching for efficient ways to attack a network. These attacks can be categorized on the basis of access control, availability, authentication, confidentiality, and integrity [3]. It has now become very essential to curb these attack vulnerabilities for the reliable operation of the network. This can be achieved by using encryption techniques, imposing firewalls, secure coding, antivirus software, applying Intrusion Detection System (IDS) and many more. In the recent scenarios, the main problem has arrived when an external intruder is trying to intrude in the network. For mitigating this, an IDS is considered as the most suitable method.

In the different parts of the world, many researchers are using IDS for detecting an intruder. Hence, for comparing the feasibility of different works that has already been done in the field of IDS, the Lincon laboratory at MIT has developed and distributed the first standard dataset for the analysis of IDS, under the sponsorship of Defense Advanced Research Project Agency (DARPA) and Air Force Research Laboratory (AFRL). The fifth association for computing machinery's special interest group on knowledge discovery and data mining (ACM SIGKDD) has collected and generated TCP dump data that has been provided by the DARPA. It is in the form of training and testing sets of features which are defined for the connection records. This set of data is now named as KDD cup 99 data sets, which we are using in our simulations [4,16].

KDD cup 99 data sets were issued for the use in the KDDCUP 99' classifier-learning competition and is consisted of preprocessed version of 1998 DARPA evaluation Data [5]. In this data set, for each connection, there are a total of 41 features and are briefly described in [5,17]. These features are grouped into four categories:

1) *Basic features*: These features can be derived from the packet header without going through the payload. It includes features like 'duration', 'source bytes', 'destination bytes', 'flag status', 'protocol type' and 'type of service'.
2) *Content features*: These features are using domain knowledge for accessing the payload of the original TCP packets. It includes features like 'number of failed login attempts'.
3) *Time based traffic features*: These features are introduced for extracting the properties that matures over 2 s in a temporal window. It included features like 'number of connections to the same host over the 2 s interval'.
4) *Host-based traffic features*: These features will assess attacks that lasts for more than 2 s because they are not working based on time, instead they are using historical window based on number of connections. It includes features like 'destination host count' [6].

The KDD cup 99 dataset contains 24 types of attack and these attacks are classified under the below mentioned categories:

1) *Denial of service (DoS)*: In these types of attacks, the attacker is making resources unavailable for the legitimate users by excessively overloading the target system like SYN flood. The attacks that comes under this category are 'Land', 'Neptune', 'Smurf', 'Ping of death (PoD)', 'Tear drop' and 'Back'.
2) *Remote to user (R2L)*: In these types of attacks, the intruder aims to gain an access on the network or computer to which the user had a remote access earlier. Since the intruder does not possess any account on the targeted remote machine, so in order to gain an access of the machine, it will exploit some vulnerabilities by continuously transmitting data packets to the targeted machine. The attacks that comes under this category are 'Warezclient', 'Guess password', 'Warezmaster', 'Imap', 'Ftp write', 'Multihop', 'Phf', and 'Spy'.
3) *User to root (U2R)*: In these types of attacks, the intruder aims at achieving the root access of the target system. The attacker initiates the attack by having normal user account on the system and then by exploiting the vulnerabilities it penetrates into the system and gains the root access. The attacks that comes under this category are 'Buffer overflow', 'Rootkit', 'Loadmodule' and 'Pearl'.
4) *Probing attack*: In these types of attacks, the main goal of the attacker is to gather all the information about the configuration of a computer system or network. Then the attacker scans the complete network and tries to find the vulnerabilities. The attacks that comes under this category are 'Satan', 'Ipsweep', 'Portsweep' and 'Nmap' [7].

The category of an attack is determined by the ultimate goal of an attack and all the attacks in a particular category resembles each other. The DoS attack aims at disrupting the network by making resources unavailable to the legitimate users. Some of the DoS attacks like 'smurf', excessively inject the fake packets into the target network for making the resources unavailable for the legitimate users, while other DoS attacks like 'PoD' and 'Teardrop' creates malformed packets, which are wrongly handled by the target system. Other attacks of this category like 'back' is still taking advantage of software bugs. Probe attacks are induced by the programs which can automatically scan the whole network and looks for the vulnerabilities in the system. These attacks generally leads to more dangerous attack situations because they are providing mapping to machines and services, and can pinpoint the vulnerable connections. Some of the scanning tools that will be used for the probe attacks are 'saint', 'mscan', etc. [5]. In R2L attacks, the intruder has no account on the targeted computer and it tries to gain an access by sending packets to that computer. An R2L attack like 'Imap', exploits buffer overflow in network server software for gaining an access in to the targeted computer. In U2R attacks, attacker tries to get access to the services which are only reserved for the super-user. Some U2R attacks exploits the poorly developed system programs which operate at the root level and are vulnerable to buffer overflow, while others captures the bugs in the software [5,18].

In KDD cup 99 dataset, the attributes selected for making a connection consists of basic features of an individual TCP connection like 'number of bytes transferred', 'duration', 'protocol type', 'flag' etc. The intrinsic features will help in providing the information that will be helpful for general network traffic analysis. Attacks in the category of 'DoS' and 'probe' have frequent sequential patterns because they are sending a lot of packets to the same host in the same time. These patterns are usually different from the normal traffic patterns. For these patterns, a 'same host' feature will scan all the other connections that has been built in the previous 2 s, which has similar destination as the existing connection. Similarly, 'same service' feature will scan all the connections with similar service as the existing connection in the previous 2 s. These features are generally referred as time based traffic features. But for the case of 'probe' attacks, there are some attacks which may take more than 2 s for complete scanning of ports or hosts. For these types of attacks, a set of 'host base' traffic features has been introduced centered on the 100 connections [5].

Contribution: The present generation is facing many problems like load sharing, increased capacity demand, inference management, high data rate etc. Researchers all over the world are trying to overcome these hurdles by introducing 5G. The introduction of relays, small cells, D2D communication have gone a long way to