A Project Report

INTRUSION DETECTION SYSTEM USING MACHINE LEARNING TECHNIQUES

Submitted in partial fulfillment of the requirements for the award of degree

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

by

Student Name	Roll Number	Mentor Name
Charan Teja	160118733027	Smt.G. Shanmukhi Rama
SaiKumar Kaleru	160118733045	Dr. Sangeeta Gupta
Yashwanth Javvaji	160118733060	Dr. Sangeeta Gupta

CHAITANYA BHARATHI

INSTITUTE OF TECHNOLOGY

స్వయం తేజస్విన్ భవ

1979

INDEX

SOFTWARE REQUIREMENT SPECIFICATION

1 Introduction

1.1 Purpose

The Project concept is to achieve the new method for extracting the information from the KDD Cup Dataset that will help to automatically detect the attack like Dos. How such attacks are Detected by ANN module and provide the security from any anomaly data which will slow your system.and provide security to the server.

1.2 Scope

One of the most important issues about our proposed architecture is the interaction between system-user and intrusion detection system, in order to verify predictions of the system. As means to reduce the number of interactions, system updates in presence of the user could be done in a periodic manner or at specified times that the number of wrong predictions reaches a predefined threshold.

1.3 Definitions, acronyms, and abbreviations

KDD - Knowledge Discovery and Data Mining
IDS - Intrusion Detection System
An Intrusion Detection System (IDS) is a network security technology originally built for detecting vulnerability exploits against a target application or computer.

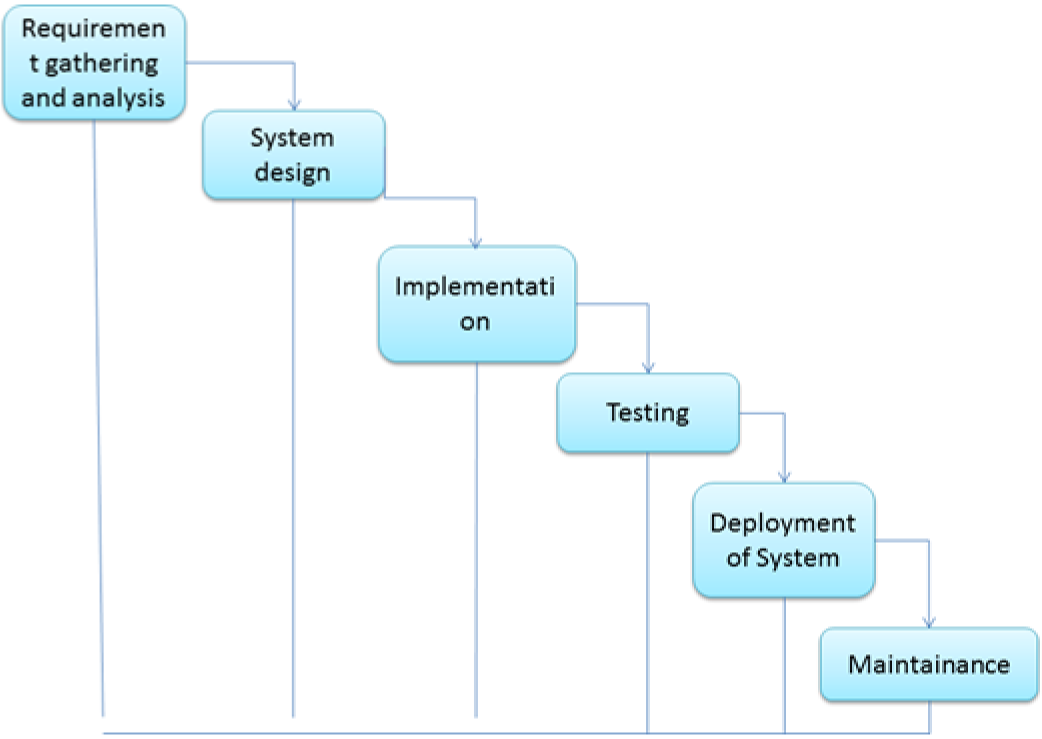
1.4 References

1. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-018-0145-4
2. https://ieeexplore.ieee.org/abstract/document/9215333
3. https://www.sciencedirect.com/science/article/pii/S1877050920311121
4. https://www.sciencedirect.com/science/article/abs/pii/S1353485820300568
5. https://medium.com/cuelogic-technologies/evaluation-of-machine-learning-algorithms-for-intrusion-detection-system-6854645f9211

1.5 Applying Software Engineering Approach

Software Development Model Used: Waterfall Model

There are various software development approaches defined and designed which are employed during development process of software, these approaches are also referred as software Development Process Models? Each process model follows particular life cycle in order to ensure success in the process of software development. One such approach used in software development is the waterfall model? It was the first process model to be introduced and followed widely in software engineering to ensure success of the project. In the waterfall approach, the whole process of software development is divided into separate process phases. The phases in the waterfall model are: Requirement specification phase, Software design, Implementation and maintenance. All these phases are cascaded to each other so that second phase is started as and when defined set of goals are achieved for first phase. General overview of the waterfall model is as follows.



Stages of Waterfall Model:

1. **Requirements Gathering:** Requirements from customers are collected by communicating with customers.
2. **Planning and Analysis:** Analysis of gathered requirements is performed and planning and estimate of project cost and schedule is done.
3. **Modelling and Design:** Model and Design of system is created as per analysis of requirements.
4. **Implementation:** Actual system is implemented using 2 phases, coding and testing.
5. **Deployment and Feedback:** System is deployed on the user's machine and feedback is taken from the user.

2 General description

2.1 Product perspective

This feature will give the user a secure and simple login screen.This means rather than creating try catches for a handful of error types, it just has only a handful of available and possible inputs, to prevent any improper logging in, which might cause unexpected errors, and therefore limiting the systems capabilities and also client attack on the server by sending multiple selecting multiple attributes from KDD Cup Dataset.

2.2 Product functions

In this extraction framework, intermediate output of IDS is stored so that only the improved component has to be deployed to the entire database KDD Cup data set. Extraction is then performed on both the previously processed data from the unchanged components as well as the updated data generated by the improved component. Performing such kind of incremental extraction can result in a tremendous reduction of processing time. To realize this new information extraction framework, the project proposes to choose database management systems over file-based storage systems to address the dynamic extraction needs.

The proposed key phrase extraction method consists of four primary components: Document preprocessing, Candidate phrase identification, Information Extraction from Database Elements of the system with their functions as follow:

1. User management-username, password, add, update, login
2. Attack On Server by Providing query
3. Query suggestion-process query, map equivalent query
4. Checking source
5. Attack detection
6. Log generation-user records, result, add, update, search
7. Data management-attributes, user detail, add, search
8. Data extraction-query, search, extract
9. Request management-request accept, block, unblock

2.3 User characteristics

User classes will be Database(KDD Cup dataset), Administrator, User, Server.

2.4 General constraints

There are no constraints at this point in time.

2.5 Assumptions and dependencies

We assume that extra documentation beyond this SRS would not be necessary in order for the user to utilize this product.

3 Specific requirements

3.1 Functional requirements

3.1.1 Introduction

The Project concept is to achieve the new method for extracting the information from the KDDCUP Dataset that will help to automatically detect the attack like Dos.How such attack are Detected by ANN module and provide the security from the any anomaly data which is slow your system.and provide security to the server.

3.1.2 Input

KDDCUP dataset

3.1.3 Processing

Processing the dataset and getting the ML model for analysing new data

3.1.4 Output

Detecting intrusion ,validation ,logging.

3.2 External Interface Requirements

3.2.1 User interface

The first interface is the log-in screen of the Application. This is where the user and Admin have a specific User-name and Password so that they can gain access to the database. Next is the Search Hints interface. Using this interface users can get hints for searching databases for particular domains. Also client attacks on the server by providing anomaly queries. Another is admin login to view all the logs of detected attacks.

3.2.2 Hardware interface

Though not necessarily interfacing with the hardware, the system must make use with an internet connection.

3.2.3 Software interface

Along with the internet connection, the system makes indirect use of an internet browser. KDD Cup 99 data set is a new Database.
Operating System: Windows XP/7/8/10

3.2.4 Communication interfaces

The system uses an internet connection to connect to the database.

3.3 Performance Requirements

Considering our project is totally based on the client server architecture . so that the client and server should be client to serve the request as well as to send the request. Also as the number of clients are going to be larger than that indirectly or directly the server is overloaded .So that the server should serve all the requests coming from the clients. So the hardware or the software as the server must have the networking capability. The network architecture should be such that the request/response time is measured .So that the time between request and the response should be as minimum as possible. Also the network should be scalable so that the number of clients can be increased as needed.

3.4 Security Requirements

Access to the database should be restricted to people that are required to view information about users. Passwords and IDs should be regulated to be at least a certain length and must contain non-alphanumeric characters in both the password and ID. Access to the database should be restricted to people that are required to view information about users. Passwords and IDs should be regulated to be at least a certain length and must contain non-alphanumeric characters in both the password and ID. As we are giving the control of the whole system to the IDS and the server . So our overall data or the database could be totally accessed by the IDS or the system/server administrator. So any secret key and the other information about the server could not be told elsewhere. Any security system has limitations so that our IDS could not prevent them totally . Our software could not get the full control of the system. So try to avoid the system calls. We will also try to implement the jre7 to take kernel level privileges to try to differentiate between the http,tcp and other types of packets

3.5 Safety requirements

Other requirements should be the power supply should be uninterrupted. The networking devices should be properly connected . And faulty networking devices should be removed as early as possible such as router ,switch and the hub etc.