iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52419

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Mobile Botnet Detection: A Machine Learning Approach using SVM

Milind Bramhane[1], Madhuri Bhusare[2], Dhananjay Nimase[3], Prof Bhondave S. D.[4]
[1, 2, 3]*Student, ATC, SPPU, PUNE*
[4]*Assistant Professor, ATC, SPPU, PUNE, Maharashtra, INDIA*

Abstract: ***Android is now the most widespread mobile operating system worldwide. Over the years the volume of malware targeting Android has continued to grow. This is because it is easier and more profitable for malware authors to target an operating sys- tem that is open-source, more prevalent, and does not restrict the installation of apps from any possible source. Hence, in this paper we present a deep learning approach that leverages Support Vector Machine (SVM) for Android botnet detection. The SVM model employs 342 static features to classify new or previously unseen apps as either 'botnet' or 'normal.***
Keywords: ***Support Vector Machine, SQLite Database, botnet, dataset.***

I. INTRODUCTION

A. Overview

A botnet consists of a number of Internet-connected devices under the control of a malicious user or group of users known as bot-masters. It also consists of a Command and Control (CC) infrastructure that enables the bots to receive commands, get updates and send status information to the malicious actors. Since smartphones and other mobile devices are typically used to connect to online services and are rarely switched off, they provide a rich source of candidates for operating botnets. Thus, the term 'mobile botnet' refers to a group of compromised smartphones and other mobile devices that are remotely controlled by bot-masters using CC channels. Nowadays, malicious botnet apps have become a serious threat.

B. Project Scope

Mobile botnets typically contain code that can be used for sending messages, attacking other networks or conducting DDoS attacks. Researchers have also discovered that mobile botnets will only infect devices running on particular versions of the An-droid operating system

C. Motivation

As a result of using machine learning techniques (in basic mode) and adjusting its parameters manually, it may be able to generate a good answer for a single botnet dataset, but if the same technique is implemented on different datasets with the same parameters

D. Objective

Users are often exposed to malware infections unknowingly by hackers exploiting security issues in software or websites. Malware is commonly distributed through emails or other online communications.

E. Problem Statement

Our proposed botnet detection system is implemented as a SVM-based model that is trained on app features to distinguish between botnet applications and normal applications.

II. SYSTEM REQUIREMENTS

A. Database Requirements

SQLite is one of the most popular and easy-to-use relational database systems. It possesses many features over other relational databases. SQLite is an embedded, server-less relational database management system. It is an in-memory open-source library with zero configuration and does not require any installation. Also, it's less than 17.6mb in size, which is significantly lesser than other database management systems.

©IJRASET: All Rights are Reserved

B. Software Requirements

Anaconda Navigator: Anaconda is an open-source distribution of the Python and R programming languages for data science, aimed at simplifying package deployment and management. Package versions in Anaconda are managed by the Conda package management system, which scans the current environment before running an installation to avoid interfering with other frameworks and packages. The Anaconda distribution comes with more than 250 automatically installed packages. More than 7,500 additional open source packages can be installed from PyPI, as well as the Conda package and the Virtual Environment Manager. It also includes a GUI (graphical user interface), Anaconda Navigator, as a graphical alternative to the command line interface. Anaconda Navigator is included with the Anaconda distribution and allows users to launch applications and manage Anaconda packages, environments and channels without using command line commands. Navigator can search for packages, install them in an environment, run and update packages.

C. Hardware Requirement

RAM: 8 GB or higher
Hard Disk: 500 GB or above
Required Processor : Intel i5 Processor or above
IDE : Spyder, Software Anaconda Navigator : Tkinter

III. ANALYSIS MODELS: SDLC MODEL TO BE APPLIED

SDLC Models stands for Software Development Life Cycle Models. In this article, we explore the most widely used SDLC methodologies such as Agile. Each software development life cycle model starts with the analysis, in which the Also, here are defined the technologies used in the project, team load. One of the basic notions of the software development process is SDLC models which stands for Soft-ware Development Life Cycle models. SDLC – is a continuous process, which startsfrom the moment, when it's made a decision to launch the project, and it ends at the moment of its full remove from the exploitation. There is no one single SDLC model.

They are divided into main groups, each with its features and weaknesses.

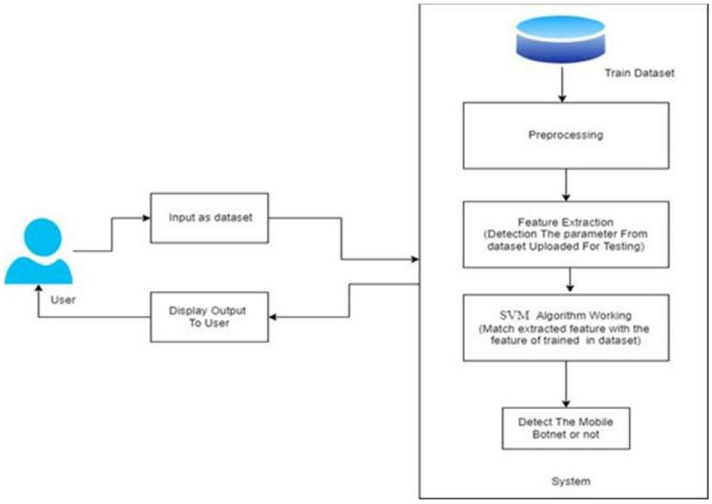
1) *Requirement Analysis* - Requirement Analysis is the most important and necessarystage in SDLC. The senior members of the team perform it with inputs from all thestakeholders and domain experts or SMEs in the industry. Planning for the quality assurance requirements and identifications of the risks associated with the projects isalso done at this stage. Business analyst and Project organizer set up a meeting withthe client to gather all the data like what the customer wants to build, who will be the end user, what is the objective of the product. Before creating a product, a core understanding or knowledge of the product is very necessary.
2) *System Design* - The next phase is about to bring down all the knowledge of requirements, analysis, and design of the software project. This phase is the product of the last two, like inputs from the customer and requirement gathering.
3) *Implementation* - In this phase of SDLC, the actual development begins, and the programming is built. The implementation of design begins concerning writing code. Developers have to follow the coding guidelines described by their management and programming tools like compilers, interpreters, debuggers, etc. are used todevelop and implement the code.
4) *Testing* - After the code is generated, it is tested against the requirements to makesure that the products are solving the needs addressed and gathered during the requirements stage. During this stage, unit testing, integration testing, system testing, acceptance testing are done.
5) *Deployment* - Once the software is certified, and no bugs or errors are stated, then it is deployed.
6) *Maintenance* - Once when the client starts using the developed systems, then thereal issues come up and requirements to be solved from time to time. This procedurewhere the care is taken for the developed product is known as maintenance.

B. Proposed Algorithm

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms. which is used for Classification as well as Regression problems. SVM chooses the extreme points/vectors that help in creating the hyperplane Theseextreme cases are called as support vectors, and hence algorithm is termed as Sup- port Vector Machine

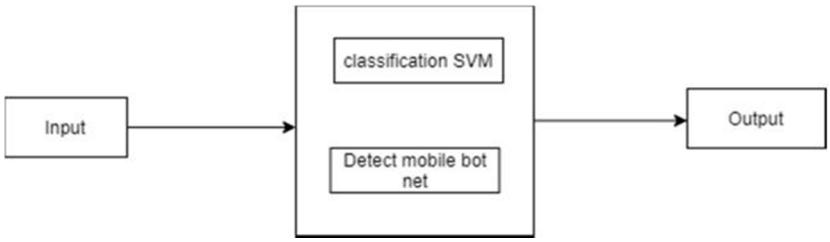
©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV SYSTEM ARCHITECTURE



A. Data Flow Diagram

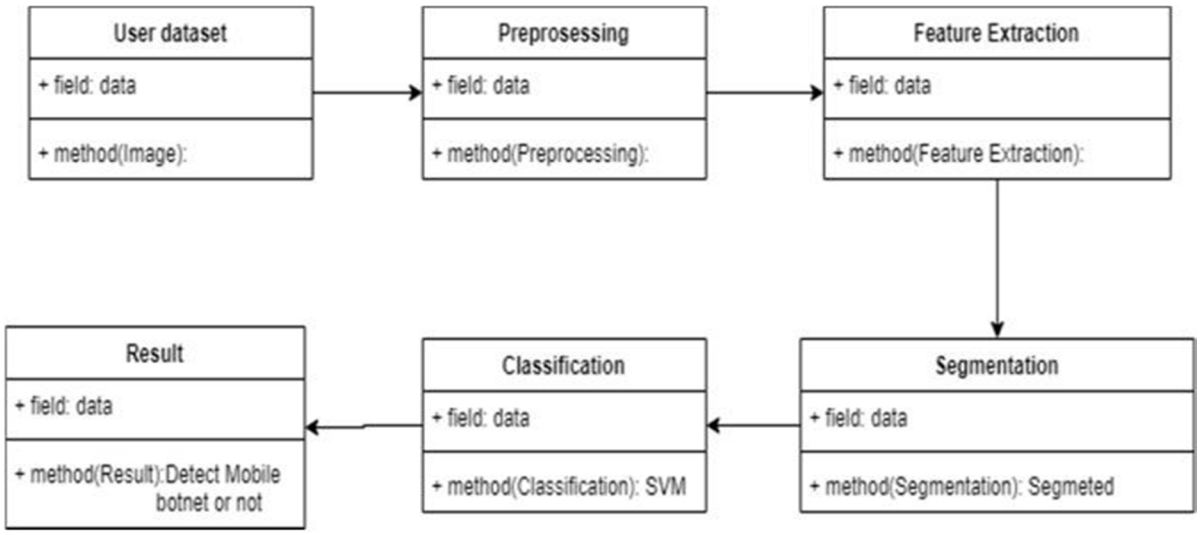
In Data Flow Diagram, we Show that flow of data in our system in DFD0 , we show that base DFD in which rectangle present input as well as output and circle show our system, In DFD1 we show actual input and actual output of system input of our system , is text or image and output is rumor detected likewise in DFD 2 we present operation of user as well as admin.



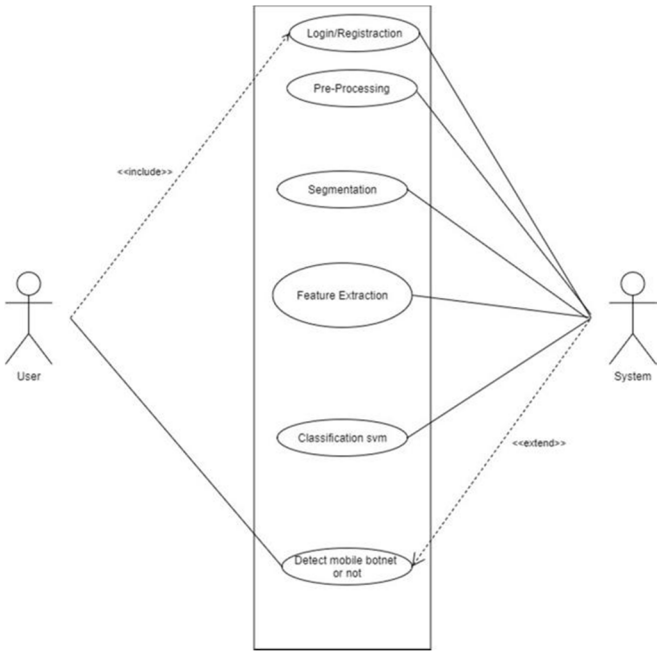
B. UML Diagrams

Unified Modelling Language is a standard language for writing software blueprints. The UML may be used to visualize, specify, construct and document the artefacts of a software intensive system. UML is process independent although optimally it should be used in process that is use case driven, architecture centric, iterative, and incremental. The Number of UML Diagram is available.

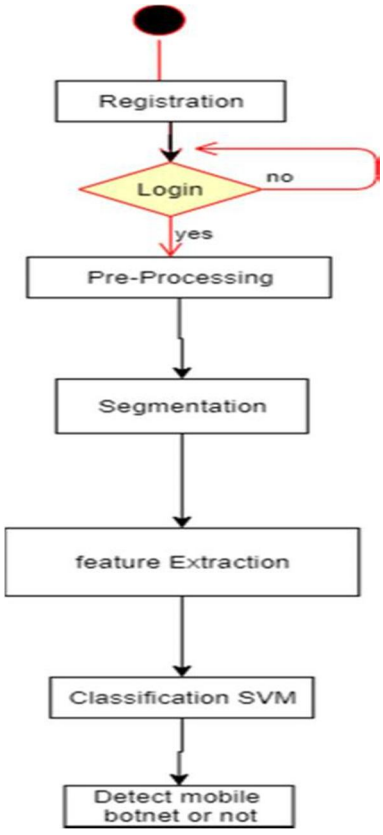
1) Class Diagram



2) Use Case Diagram

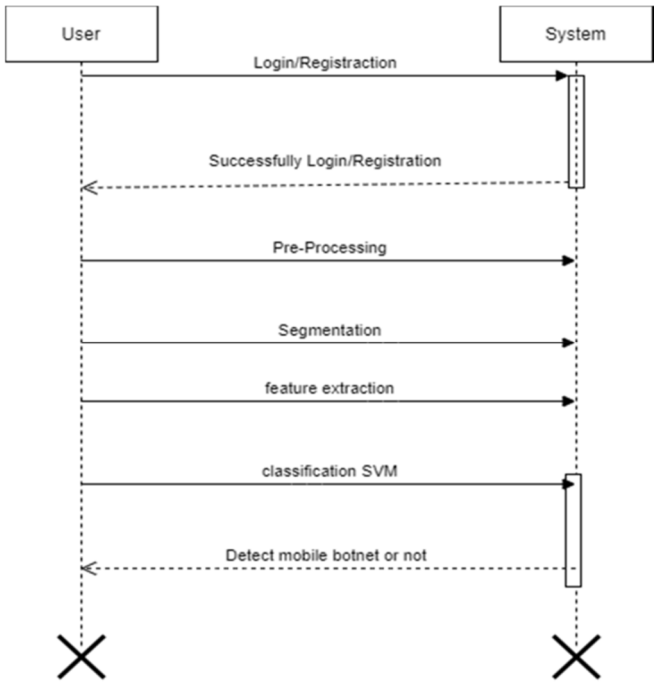


3) Activity Diagram



©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

4) Sequence Diagram

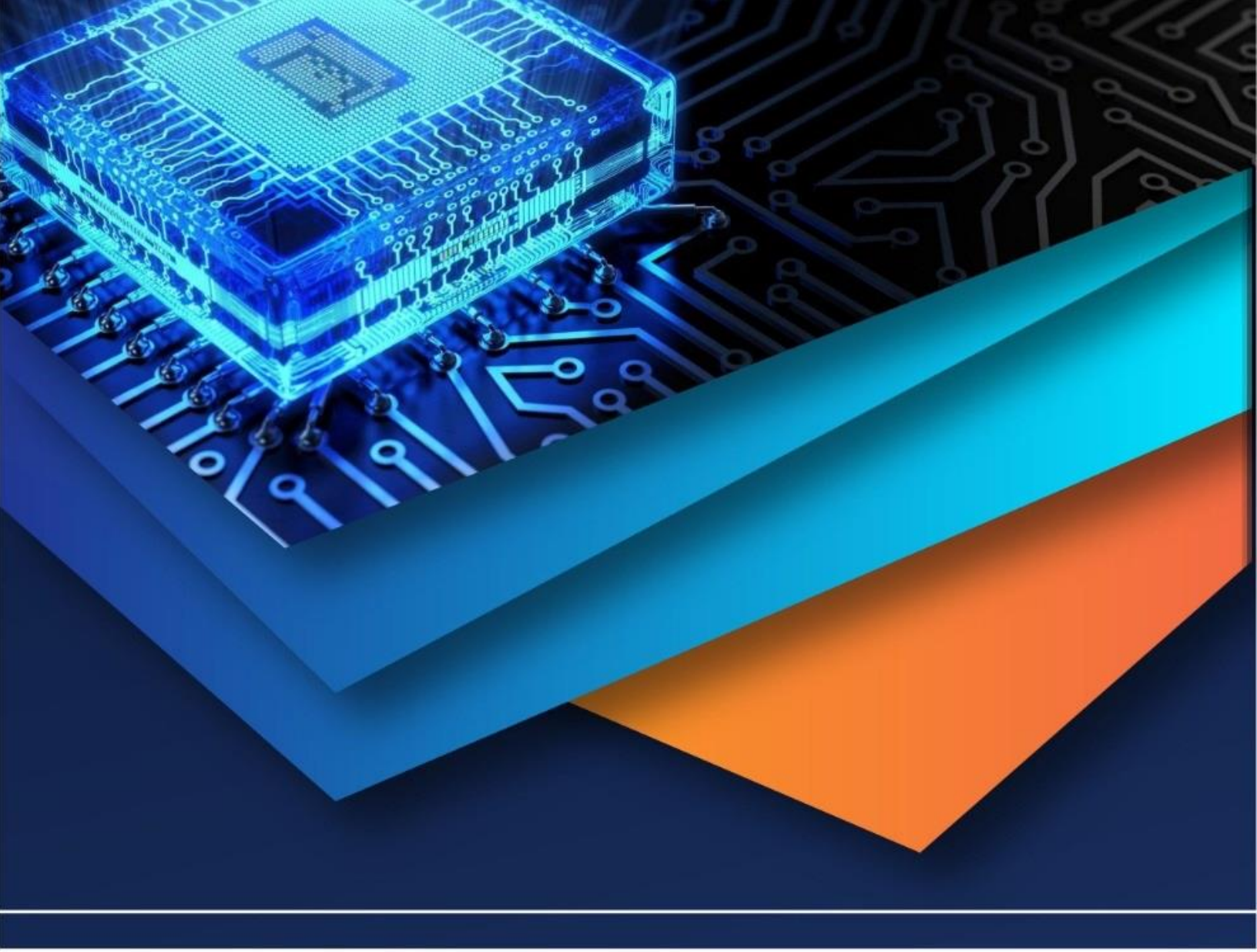


V. CONCLUSION

Botnets are a Dangerous evolution in the malware world. They are being used to damage systems, steal information and Comprise Systems. They are very hard to detect and eliminate. So Our System Is very Useful To detect Mobile Botnet.

REFERENCES

[1] S. Y. Yerima and S. Khan “Longitudinal Perfomance Anlaysis of Machine Learning based Android Malware Detectors” 2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security), IEEE.
[2] H. Pieterse and M. S. Olivier, ”Android botnets on the rise: Trends and charac- teristics,” 2012 Information Security for South Africa, Johannesburg, Gauteng, 2012, pp. 1-5.
[3] Letteri, I., Del Rosso, M., Caianiello, P., Cassioli, D., 2018. Performance of botnet detection by neural networks in software-dened networks, in: CEUR WORKSHOP PROCEEDINGS, CEUR-WS.
[4] Kadir, A.F.A., Stakhanova, N., Ghorbani, A.A., 2015. Android botnets: What urls are telling us, in: International Conference on Network and System Secu-rity, Springer. pp. 78–91.
[5] ISCX Android botnet dataset. Available from https://www.unb.ca/cic/datasets/android- botnet.html. [Accessed 03/03/2020]
[6] M. Eslahi, M. V. Naseri, H. Hashim, N. M. Tahir, and E. H. M. Saad, ”BYOD: Current State and Security Challenges,” presented at the IEEE Symposium on Computer Applications Industrial Electronics, Peneng, Malaysia, 2014
[7] S. S. C. Silva, R. M. P. Silva, R. C. G. Pinto, and R. M. Salles, ”Botnets: A survey,” Computer Networks, vol. 57, pp. 378-403, 2013.
[8] G. Gu, J. Zhang, and W. Lee, ”BotSniffer: Detecting botnet command and control channels in network traffic,” in Proceedings of the 15th Annual Net- work and Distributed System Security Symposium (NDSS’08), 2008
[9] C. Byungha, C. Sung-Kyo, and C. Kyungsan, ”Detection of Mobile Botnet Using VPN,” in Proceedings of the Seventh International Conference on Inno-vative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2013, pp. 142-148



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)