IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 5, May 2022

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.118

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|

||Volume 11, Issue 5, May 2022||

| DOI:10.15680/IJIRSET.2022.1105092 |

Malware Detection using Machine Learning

Akshaykumar Vaidya[1], Tanaji Kamble [2], Prof. Tushar Phadtare [3]

B.E Student, Department of Computer Engineering, JSPM'S Bhivarabai Sawant Institute of Technology and Research, Pune, Maharashtra, India[1,2]

Guide, Department of Computer Engineering, JSPM'S Bhivarabai Sawant Institute of Technology and Research, Pune, Maharashtra, India[3]

ABSTRACT: When users enter their credentials in a public place, they risk adversaries gaining access to their credentials. 2 An attacker can get a password by observing or recording a person's authentication session. Shoulder-surfing is a well-known danger, and it's especially risky when In public venues, you must authenticate yourself. The user's main line of defence against shoulder-surfing was until recently was his own alertness. The user is protected from shoulder-surfing by 2a password authentication system that is resistant to intrusion detection. Because the user is never obliged to click directly on password symbols, it allows users to authenticate in insecure locations by typing their password graphically. The usability testing of this technique indicated that novice users were able to type their graphical password accurately and remember it 3over a period of time 1However, the protection against shoulder-surfing comes at the cost of a longer identification process.

I. INTRODUCTION

A malware is a malicious code. Malware can be considered as an entity in that easily added new feature and that enhance its side effects in the form of various attacks. These malware can be dangerous with all side effects like break the system, corrupt data, etc. IOT Applications have led to the development of modern concept of the infor-mation society. However, security concerns pose a major challenge in realising the benefits of industrial revolution as cyber-criminals attack indi vidual PC's and net-works for stealing confidential data for financial gains and causing DOS attacks to systems. Such attackers make use of malicious software or malware to cause serious threats and vulnerabilities of system. Malware is a computer software that is designed to harm the operating system. Deep Learning is an artificial intelligence function that mimics the human brain's functions in data processing and pattern creation for decision-making. Deep Learning) is a form of Machine Learning in Ai Technology (AI) that uses neural networks to learn unsupervised from unorganized and unlabeled data. Deep Neural Learning or Convolutional Neural Network are other terms for the same thing.

II. LITERATURE SURVEY

1.Paper Name: Risk prediction of malware victimization based on user behavior
Author: Fanny Lalonde Levesque
Abstract : Understanding what types of users and usage are more prone to malware infestations is critical if we are to develop sufficient strategies for dealing with and minimising the effects of computer crime in all of its manifestations. Real-time usage data is thus critical for making better evidence-based judgments that will increase user security. To that purpose, we performed a 4-month field study with 50 individuals, collecting real-time data by detecting potential illnesses and obtaining information on user behaviour. We present a first attempt at forecasting the likelihood of malware victimisation based on user behaviour in this research. Using neural networks, we created a predictive model with an accuracy of up to 80 percent.

2.Paper Name: :- Multilevel Permission Extraction in Android Applications for Malware Detection
Author: Zhen Wan
Abstract : With the widespread use of Android applications in security-critical circumstances, an increasing amount of Android malware has been detected. Existing malware detection research fails to automatically learn effective feature interactions, which are important to the operation of many prediction models. In this work, we offer Multilevel Permission Extraction, an approach to automatically identifying permission interactions that are effective in discriminating between dangerous and benign programmes, in able to locate malware rapidly and reliably. The gathered information is can then use by machine learning-based classification algorithms to classify dangerous and benign programmes. We test our method on a huge data set that includes 4,868 benign and 4,868 malicious applications. The

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|

||Volume 11, Issue 5, May 2022||

| DOI:10.15680/IJIRSET.2022.1105092 |

experimental study suggests shown our malware detection approach may obtain a detection rate of much more than 95.8 achieve a better malware detection rate of 97.88.

3.Paper Name: Behavioral malware analysis algorithm comparison
Author name: Matu´s Uchn ˇ ar´
abstract : Malware analysis and analysis made on this is a significant part of computer security. Despite the enormous effort put forth by companies developing anti-malware solutions, it is usually not possible to respond to new malware in a timely manner, and some computers become infected. This disadvantage could be overcome in part by employing behavioural malware analysis. This work compares machine learning methods for such aim of behavioural malware analysis.

4.Paper Name:Securing Android App Markets via Modelling and Predicting Mal-ware Spread between Markets
Author::- Guozhu Meng†‡, Matthew Patrick§ , Yinxing Xue¶ , Yang Liu‡ , Jie Zhang
abstract : Recently, the Android ecosystem has dominated mobile devices. Android software platforms, both official Google Play and third-party shops, are becoming breeding grounds for malware. Mobile virus has been found to proliferate both between and within markets. If the distribution of Malware attacks between markets can be foreseen, market administrators can take appropriate actions to prevent malware outbreaks and mitigate malware damage. We make the first efforts to defend the Android ecosystem in this paper by modelling and predicting the dispersion of Malware detection across markets. To just that end, they analyze that social behaviours that influence malware propagation, model these spread behaviours using different epidemic models, and forecast infec-tion time and order among markets for well-known malware families.
We model malware spread behaviours in the following way to produce an accurate prediction of malware spread: 1) For a single market, we model within-market malware growth by taking into account both malware development and removal. 2) For many markets, we assess market relevance by computing mutual information among them, and 3) based on the previous two phases, we stochastically simulate a Susceptible Infected (SI) model for market spread. ANDRADAR, a publicly available well-labeled dataset, is used for model inference. We acquired a significant quantity (334,782) of malware samples from 25 Android markets around the world to conduct lengthy experiments to evaluate our approach. The experimental results suggest that our method can display and replicate the spread of Android malware on a wide scale, as well as anticipate infection time and order among markets with 0.89 and 0.66 precision, respectively.

5.Paper Name:A3CM: Automatic Capability Annotation for Android Malware
Author:JUNYANG QIU 1 , JUN ZHANG 2 , WEI LUO1 , LEI PAN 1 , SURYA NEPAL 3 , YU WANG 4 , AND YANG XIANG
Abstract: Android malware offers significant privacy risks to mobile users. Traditional virus detection and family classification technologies are becoming less successful as the malware field evolves rapidly, only with emergence of so-called zero-day malware families. To solve that problem, we provide Malware Capability Annotation, a fresh research challenge on automatically recognising the security/privacy related capabilities of any discovered malware (MCA). Motivated by the analysis that known and zero-day-family malware families share security/privacy-related capabilities, MCA presents a new alternative method for effectively analysing zero-day-family malware (computer viruses which does not belong to every original relatives) by examining re-lated knowledge and information from known malware families. We create a novel MCA hungry solution, Automatic Capability Annotation for Android Malware, to overcome the MCA problem (A3CM). The four steps of A3CM are as follows: 1) To characterise malware samples, A3CM automatically extracts a set of semantic features such as permissions, API calls, and network addresses from raw binary APKs; 2) A3CM uses a statistical embedding method to map the features into a joint feature space, allowing malware samples to be represented as numerical vectors; 3) A3CM uses a multi-label classification model to infer malicious capabilities; 4) The trained multi-label model is used to annotate the malicious capabilities. To make it easier for MCA to do new research.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 8.118|

||Volume 11, Issue 5, May 2022||

| DOI:10.15680/IJIRSET.2022.1105092 |

III. PROPOSED SYSTEM

a) SYSTEM ARCHITECTURE

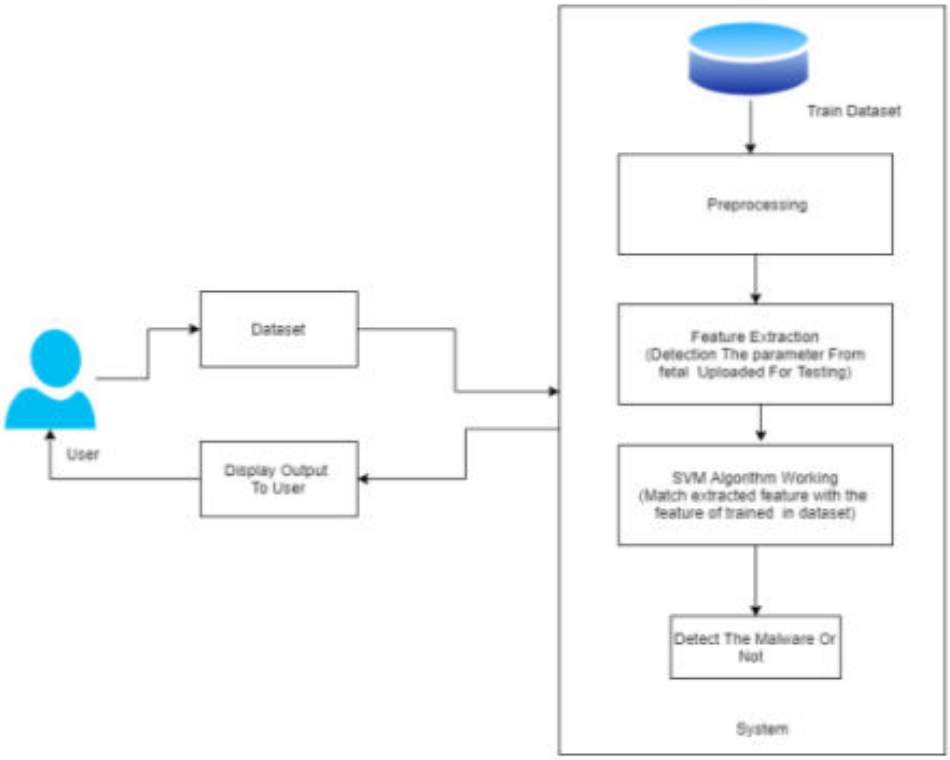

Fig.System Architecture

IV. ALGORITHM

Svm Classifier, or SVM, is a prominent Supervised Learning technique that is used for both classification and regression issues. However, it is mostly utilised in Machine Learning for Classification difficulties. The SVM algorithm's purpose is to find the optimum line or decision boundary for categorising n-dimensional space so that we may simply place fresh data points in the correct category as in later. A hyperplane is the optimal choice boundary. SVM selects the extreme points/vectors that aid in the creation of the hyperplane. These extreme examples are referred to as support vectors, and the technique is known as the Svm Classifier.

V. CONCLUSION

SVM is a superior classification algorithm that can be used to detect malware. Attention is required to build a better feature representation for better generalisation.

REFFERENCES

1 Gavrilut D., Cimpoesu M., Anton D., Ciortuz L., “Malware Prediction Using Machine Learning”, International Multiconference on Computer Science and Information Technology, 2009.
2 Rhode, M., Burnap, P., Jones, K., “Early-stage malware prediction using re- current neural networks”, computers security, 2018.
3 Baset, M, “Machine Learning For Malware Detection”, 2016.
4 Yeo, M., Koo, Y., Yoon, Y., Hwang, T., Ryu, J., Song, J., Park, C.,“Flow- based malware detection using convolutional neural network”,2018 Interna- tional Conference on Information Networking, 2018.
5 “Features”, LightGBM Documentation.[online] Available https://lightgbm.readthedocs.io/en/latest/Features.html.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

Scan to save the contact details