iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** VI **Month of publication:** June 2019

DOI: http://doi.org/10.22214/ijraset.2019.6110

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: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

Detection and Prevention for Ransomware using Machine Learning

Yamika Solanki[1], Mr. Mahesh Panchal[2]
[1, 2]*GTU- Graduate School Of Engineering And Technology*

Abstract: ***Ransomware Contains Ransome+ware. Ransomware attacks are increasing Everyday. Hackers are using various techniques for encrypt Your data/information, Hacker Exploits the Vulnerabilities of operating system, especially those who have Windows system. From May 2017 millions of computers around the world experienced this virus "Wannacry Ransomware". For this reason, the necessity of creating different mechanisms which act proactively, This proposed research will create a Model for ransomware detection .for training model purpose we took dataset of Wannacry, Notpetya and Locky Ransomware and etc and Benign dataset as an input and Apply Machine Learning Algorithm and then Find accuracy of result.***

Keywords :Malware, virus, worm ,ransomwarestatic analysis, malware analysis, machine learning

I. INTRODUCTION

In computer network, if vulnerability of system/data exploits occurs then security of data/system has been compromised. Malware is a short for malicious software[1] which is specially designned to disrupt, damage, or gain authorized access to a computer system or their resources. malware is any piece of software that was written with the intent of doing harm to data and devices .

Ransomware is a one Type of Malware ,which is created for gaining access of user's data and deny access to a system Nowadays, Ransomware increasing everyday, across the world ,Ransomware named Wannacry, Notpetya, ryuk and bad rabbit hitted Everywhere.

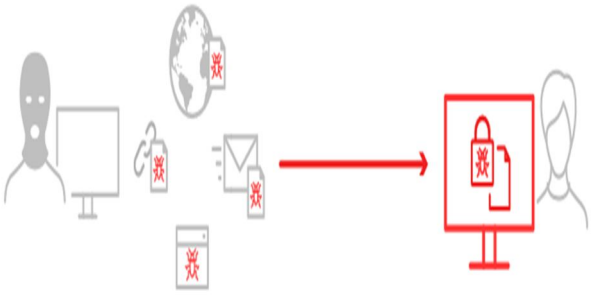


Fig 1 : how Ransomware spread[2]

II. RELATED WORK

In this section we represent the related work done by different researcher.

In [3] the author has done functional piece of software to stop this malware demonstrate inner working of wannacry Ransomware Finding hash value of wannacry file using Zoo Tool

Unpacked file using PEID and Dependency walker for overview of file and PEheaders.In [4] the author has done Give Overview of ransomware like **:** Glassberg (2016), and give some preventive steps .

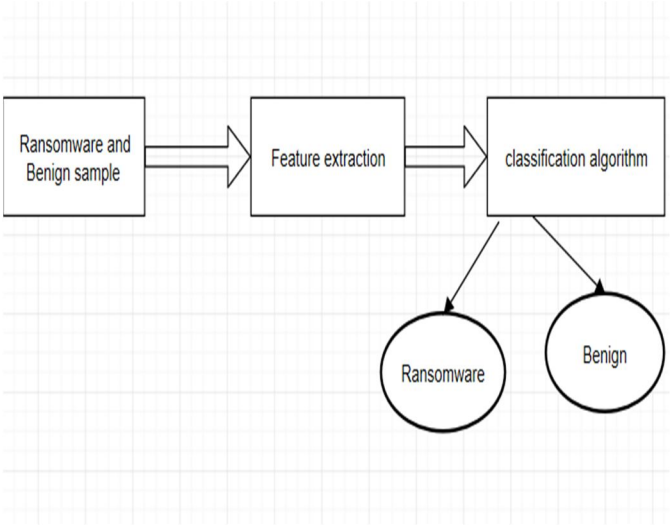
In[5] author uses, Perform static analysis and extracted features like- api calls, keywords, extension etc in[6]uses conduct experiments on two datasets Ubuntu 14.04 and Windows 7. In Linux dataset collect 100 malware from Virustotal and 400 benign application concluded that DeepMalware solved the problems of low accuracy and low-quality dataset in conventional machine learning based malware detection. In[7] Create two model :

1)malware run time model 2) malware prediction model.and concluded The Neural Network model outperforms the other models in terms of accuracy achieving over 92%.. in[8] CERT-MU published white paper of Wannacry Ransomware , author gives detail study of how Wannacry Ransomware works and prevention measures

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

III. METHODOLOGY

We begin with an overview of the proposed system, followed by a discussion of the generation of our data sets, features we extracted, and finally the set of machine learning classifiers we use to evaluate our methodology.
The experiments were performed within a virtual VMWare workstation environment running on windows XP with 4 GB RAM and The machine learning tool used was WEKA 3.8.3[10]



A. Sample

Taken Ransomware samples from Virusshare.com[9] (Wannacry,Locky,Petya,Cryptolocker)and benign sample (cam .exe, snippingtool.exe,visualstudio.dll,wordpad.exe etc).
Total : 2312 Samples
1000 – Benign sample and 1312- Ransomware sample.

B. Features Extraction

Extract features like - Debug Size, Debug RVA, Major Image Version, Major OS Version, Export Size, IAT RVA, Major Linker Version, Minor Linker Version, Number Of Sections, Size Of Stack Reserve, Dll Characteristics, after finding features give input file(.csv) to WEKA tool, The WEKA machine learning tool is used for finding accuracy rate of detection of Ransomware.

C. Classsification Algorithm

Classification algorithm like Random forest, J48, Naïve bayes, Logistic regression, SVM and KNN is used.

IV. CALCULATION AND RESULT

Evaluated the performance using the Two standard WEKA metrics: true positive rate (TPR), false positive rate (FPR).

Metric	Calculation	Value
True positive rate (TPR)	TP/(TP+FN)	Correct classification of predicted malware
False positive rate (FPR)	FP/(FP+TN)	Benign file incorrectly predicted as malware

TP= True positive, FN=false negative, TN= true negative, FP=false positive
Comprative analysis:

A. Ransomware Detection Evaluation Result

Classifier	TPR (%)	FPR (%)
Naive bayes	85.80	17.70
SVM	87.10	15.72
Logistic regression	89.30	12.53
Random forest	91.70	9.33
Bayes Network	94.50	6.66
LMT	92.20	8.10
J48	97.70	2.40
KNN	98.10	1.80

Sample output:

```
IB1 instance-based classifier
using 1 nearest neighbour(s) for classification


Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        2268               98.1393 %
Incorrectly Classified Instances        43                1.8607 %
Kappa statistic                          0.9621
Mean absolute error                      0.019
Root mean squared error                  0.1363
Relative absolute error                  3.8614 %
Root relative squared error             27.5226 %
Total Number of Instances             2311

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.982    0.019    0.975      0.982    0.979      0.962    0.985     0.969     Benign
                 0.981    0.018    0.986      0.981    0.984      0.962    0.985     0.985     Ransomware
Weighted Avg.    0.981    0.018    0.981      0.981    0.981      0.962    0.985     0.978

=== Confusion Matrix ===

    a    b   <-- classified as
  981   18 |    a = Benign
   25 1287 |    b = Ransomware
```

Fig 2: 98.13% Accuracy achieved by KNN

B. Prevention measure For Ransomware

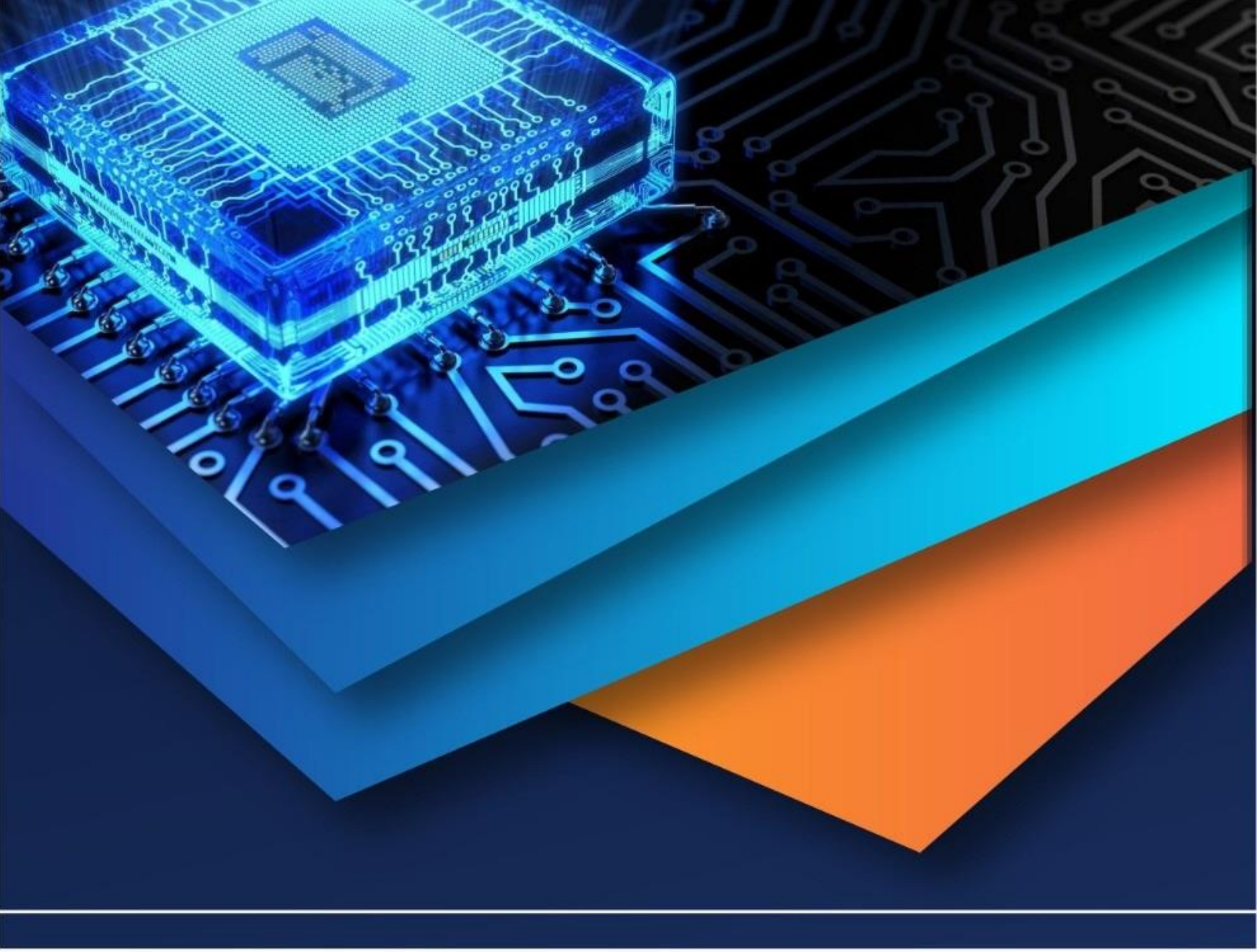
1) Apply all the patches Regularly.
2) Ensure anti-virus software is up-to-date.
3) Ensure anti-virus and anti-malware solutions are set to automatically conduct regular scans
4) Implement a data back-up and recovery plan to maintain copies of sensitive or proprietary data in a separate and secure location.
5) Backup copies of sensitive data should not be readily accessible from local networks.
6) Do not open attachments included in unsolicited emails.
7) Do not click on link in unsolicited emails
8) Only download software, especially free software - from sites you know and trust.

V. CONCLUSION AND FUTURE WORK

In Implemented system, highest detection rate accuracy achieved by KNN (K nearest neighbour) which is **98.10%(**TPR). we can do dynamic analysis and also increase dataset too. we can also more extract features too.

REFERENCES

[1] Techtarget "malware (malicious software)**"** accessed on 25 july 2018 https://searchsecurity.techtarget.com
[2] F- secure "how ransomware works"accessed on 30 july 2018 https://www.f-secure.com
[3] Justin Jones, Narasimha Shashidhar "Ransomware Analysis and Defense WannaCry and the Win32 environment" IJIS, 201
[4] Azad Ali "RANSOMWARE: A RESEARCH AND A PERSONAL CASE STUDY OF DEALING WITH THIS NASTY MALWARE" IISIT,2017
[5] May Medhat, Samir Gaber **"** A New Static-based Framework for Ransomware**.**", 2018
[6] Yuan, Xiaoyong. "PhD Forum: Deep Learning-Based Real-Time Malware Detection with Multi-Stage Analysis." *Smart Computing (SMARTCOMP), 2017 IEEE International Conference on*. IEEE, 2017.
[7] Kilgallon, Sean, Leonardo De La Rosa, and John Cavazos. "Improving the effectiveness and efficiency of dynamic malware analysis with machine learning." *Resilience Week (RWS), 2017*. IEEE, 2017
[8] Dick O'Brien,"the wannacry ransmware white paper" Advances Research in Computer Science-APA, CERT-MU,2017
[9] www. virusshare.com accessed on 12th march 2019.
[10] "weka", accessed on 6 may 2019, https://wekatutorial.com/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY