TO ARTICULATE THE PROBLEM OF ATTRIBUTION IN CYBER WARFARE INCIDENTS

J. CHAKRADHAR M.Tech(CSE) Scholar,Aditya Engineering College(A),Surampalem,India
Dr. M. VAMSI KRISHNA Professor in IT ,Aditya Engineering College(A),Surampalem,India

ABSTRACT

As increase of hard work done by intelligence and law enforcement organizations to identify the enforcer which made heavy effort. The tools and techniques of attribution are used for harmful activities on the Internet and still present for pratical measurements, the source of harmful code, and non- theoretical assessments of attack and attacker features to link an attack activity to single or groups. It is commonly done through a tough manual process that relies on both technical analysis and ground intelligence. As a result, this high effort able process of attribution and it is mainly occupied for cyber attacks and those attacks are conducted against fully equipped organizations. Gradually,the ability of attribution have been improved, however, this improvement is a two-side coin: as improvement of attribution capabilities and privacy of internet is decreasing low. In this we discuss about attribution of two distinct types of attacks that are important to cyber clash today. Those are network intrusions and social bot. In this the state of the art with respect to attribution abilities across both types of attacks, give advice for improvement of attribution.

Keywords: Attribution, cyber attacks, malicious , bots, cyber space.

1] INTRODUCTION

Now a days to deal with such attacks many countries are attributing cyberattacks details to detect attackers and in propose paper author is suggesting to use MACHINE LEARNING or AI (Artificial Intelligence) algorithms can be used to mitigate such attacks with high success rate compare to existing techniques such as dynamic and static network packet analysis.

In this paper we describe concept of Attribution cyber-attack which is known for IDNTIFYING or blaming criminals for attack.In network attack malicious users can hack network banking data to steal password and make fake transactions and this attack can be targeted on any server.

In propose paper author is staging two different types of attacks from cyber world such as Network Intrusion (refers to attack on network to steal data or corrupt data or crashing servers or DOS attack by sending huge amount of requests) and Information Operations which is known to spread FAKE or False NEWS through social media by hiring BOTS (robot program to spread fake opinions on social media) application.

For any country or company this types of attacks can be dangerous for example if attackers can spread fake opinion or reviews on GENUINE political person which can influence normal users and divert their votes to corrupt politicians and sometime some corrupt companies can hire BOTS to spread good reviews on their BAD quality products and normal users can be influence from such reviews and purchase that BAD quality product.

Here we use Random Forest algorithm to detect cyber attacks and fake news, so as extension we have used combination of 3 different latest classifiers such as AdaBoost, XGBoost and Bagging classifier to make hybrid ensemble algorithm with the help of Voting Classifier.

2] LITERATURE SURVEY:

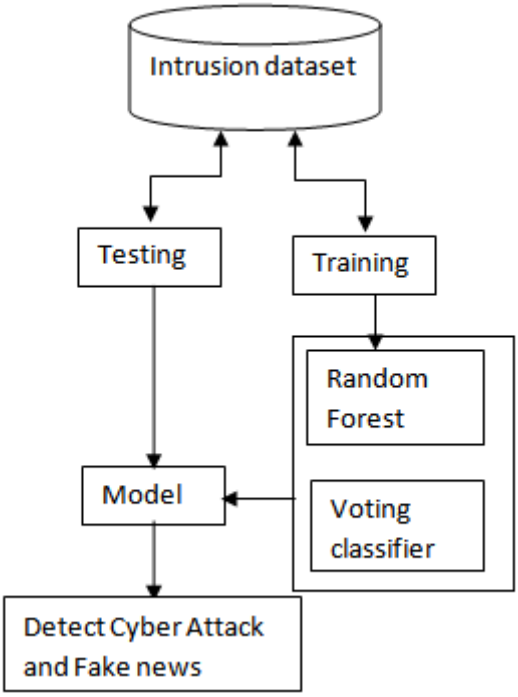
Factual rendering of an event and activities associated with development of early internet would be a valuable contribution. These contribution make the best that only factual material will include into this document [1] It provides geographies of global internet censorship firstly it maps severity of worldwide censorship then second it explores the severity of levels and then approximation of an ideal speech situation [2] The emerging trends, threats and concepts are discussed and employed to gain positional advantage [3] The attribution of cyber attacks found to be abundance. Here we differentiate between machine and manual responsible for activity [4]cyber confilict is a common problem which target a strategic choice based on its ability for attribution of attacks[5] Attribution is centre point of view for response to cyber intrusions here they examine the use of public attribution. It shows the importance of meaning making [6] The examination of cyber attack conditions can be pointed out here when there is no state is applied in cyber attack then victim state directly take self defence action against non-state actor [7] In this to perform attribution of computer attackers who exploit the data we use various techniques and source tracking is used instead of Attribution [8] Two different types of attacks have taken as main concept they are network intrusion and information operations[9] The anonymity is a main concept in internet and it used to identity of users [10]

3] PROPOSED SYSTEM:

A single approach was incorporated into our proposed system which include RANDOM FOREST algorithm

The attribution of cyber attacks is proposed model for detecting cyber attacks and fake news. The architecture of proposed model is shown below. The algorithms is collected and implemented in python software for the use of random forest and combination 3 classifiers(voting classifier).The system model for this study was developed using jupyter notebook.

Jupyter Notebook:It is a group of cells in which we can run and write a code in sequence or in parallel. As it loads necessary libraries and execute machine learning tasks and It is a dependable source for ML tasks..We take all necessary procedures to ensure that our data pre-processing, creation of model and model evaluation are successful. Later, we save and use the most appropriate estimation in our web - based application.



Architecture diagram

4 IMPLEMENTATION

The study is carried out in this by using below steps

The steps taken in this are mentioned below

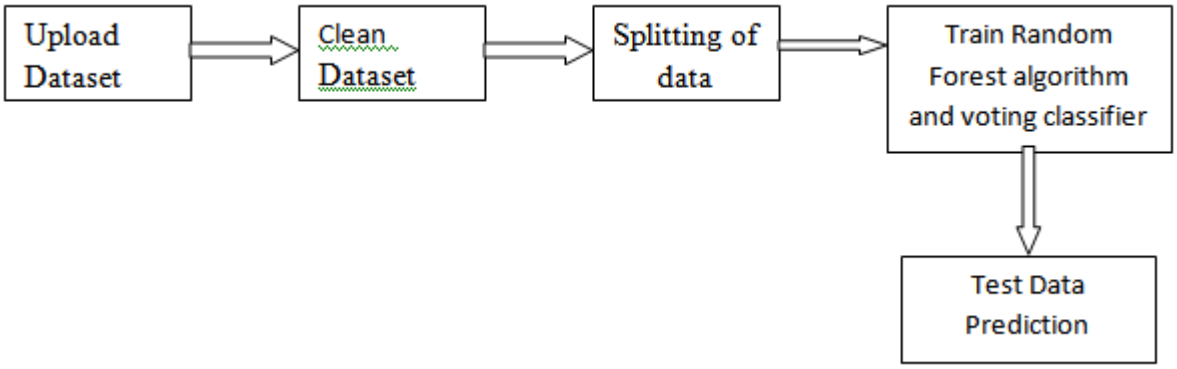
Step 1: First we load intrusion and fake news datasets then it display normal and attack packets.

Step 2: Now clean the datawhich contains missing and non-numeric data and machine learning algorithms accept only numeric dataset as input so we need to remove missing values and replace non-numeric data to numeric and then clean data.

Step3: After cleaning of data then split the dataset into 80% training and 20% testing then predicted values compare with original data to get algorithm prediction accuracy.

Step 4: Then after we train both Random Forest model and voting classifier using this module we will input train data to RF algorithm and voting classifier to trained model

Step 5: Lastly test data prediction we will apply on test data on trained model to predict whether test data is NORMAL or contains attack/fake news



Formulas to predict the below metrics used in this

Precision: The precision is the ratio of correctly predicted values to the total predicted values.

$$\text{Precision} = TP/TP + FP$$

Recall: It is the ratio which correctly predicts positive values among all values in real data.

$$\text{Recall} = TP/TP + FN$$

F1-score: It is the weighted average of both Precision and Recall. Therefore, this score takes both false positives and false negatives into count. This is very difficult to understand as accuracy.

$$\text{F1 Score} = 2(\text{Precision Recall}/\text{Precision} + \text{Recall})$$

RANDOM FOREST ALGORITHM:

It is an ensemble learning method for classification, regression and other tasks which is operated by constructing a number of decision trees at time of training.The class which is selected by most trees is the output of random forest for classification task.

It is a classifier that have number of decision trees on various subsets of the given dataset and to improve the prediction accuracy of that taken dataset it takes average

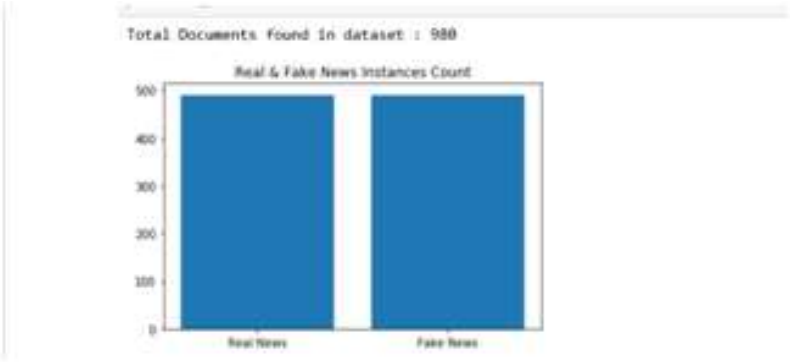
```
plt.xlabel('Predicted Class')
plt.show()

Random Forest Cyber Attack Detection Accuracy: 99.68247668188133

Random Forest Cyber Attack Detection Precision: 99.6887627867459

Random Forest Cyber Attack Detection Recall: 99.66544175754703

Random Forest Cyber Attack Detection FMeasure: 99.68132443562993
```

Metrics of RF algorithm

VOTING CLASSIFIER:

It takes multiple classifier algorithms as input and then train all those algorithms and then voted out or select algorithm with best accuracy. So by applying this algorithm on both Intrusion and Fake news giving better accuracy compare to Random Forest.

Metrics of Hybrid Ensemble Algorithm



Real and Fake news instances count

CONCLUSION

Here two different types of attacks from cyber space are used such as Network Intrusion (refers to attack on network to steal data or corrupt data or crashing servers or DOS attack by sending huge amount of requests) and Information Operations which is known to spread FAKE or False NEWS through social media by hiring BOTS (robot program to spread fake opinions on social media) application.

Thus we conclude that network attacks and malicious users can hack network banking data to steal password and make fake transactions and this attack can be targeted on any SERVER. So we use Random Forest algorithm to detect cyber-attack from Intrusion dataset and then detecting FAKE news by using FAKE dataset and this algorithm wrongly predict 15 and 1 records so total 16 records wrongly predicted. so as extension we have used combination of 3 different latest

classifiers such as AdaBoost, XGBoost and Bagging classifier to make hybrid ensemble algorithm with the help of Voting Classifier.

BIBILOGRAPHY

1. B. M. Leiner et al., "A brief history of the Internet," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 5, pp. 22–23, 1997.
2. H. L. Armstrong and P. J. Forde, "Forde Internet anonymity practices in computer crime," Inf. Manag. Comput. Security, vol. 115, no. 5, pp. 209–215, 2003.
3. B. Warf, "Geographies of global Internet censorship," GeoJournal, vol. 76, no. 1, pp. 1–23, 2011.
4. B. Krekel, P. Adams, and G. Bakos, "Occupying the information high ground: Chinese capabilities for computer network operations and cyber espionage," Int. J. Comput. Res., vol. 21, no. 4, p. 333, 2014.
5. N. J. Shallcross, "Social media and information operations in the 21st century," J. Inf. Warfare, vol. 16, no. 1, pp. 1–12, 2017.
6. Weedon, W. Nuland, and A. Stamos. (2017). Information Operations and Facebook. Facebook.Online Available:https://www.mm.dk/wpcontent/uploads/2017/05/facebook-and-information-operations-v1.pdf
7. S. W. Brenner, "At light speed: Attribution and response to cybercrime/terrorism/warfare," J. Crim. Law Criminol., vol. 97, p. 379, Mar. 2007.
8. H. S. Lin "Attribution of malicious cyber incidents: From soup to nuts," Legal Perspectives Inf. Syst. J., to be published.
9. E. M. Mudrinich, "Cyber 3.0: The department of defense strategy for operating in cyberspace and the attribution problem," AFL Rev., vol. 68, p. 167, Jul. 2012
10. B. Edwards, A. Furnas, S. Forrest, and R. Axelrod, "Strategic aspects of cyberattack, attribution, and blame," Proc. Nat. Acad. Sci. USA, vol. 114, no. 11, pp. 2825–2830, 2017.
11. H. Berghel, "On the problem of (cyber) attribution," IEEE Comput., vol. 50, no. 3, pp. 84–89, May 2017.
12. F. J. Egloff, "Public attribution of cyber intrusions," J. Cybersecurity, vol. 6, no. 1, 2020, Art. no. tyaa012.
13. F. J. Egloff and M. Smeets, "Publicly attributing cyber attacks: a framework," J. Strategic Stud., to be published.
14. M. Mueller, K. Grindal, B. Kuerbis, and F. Badiei, "Cyber attribution," Cyber Defense Rev., vol. 4, no. 1, pp. 107–122, 2019.
15. T. Rid and B. Buchanan, "Attributing cyber attacks," J. Strategic Stud., vol. 38, nos. 1–2, pp. 4–37, 2015.

16. J. Healey, Beyond Attribution: A Vocabulary for National Responsibility for Cyber Attacks. Vienna, VA, USA: Cyber Conflict Stud. Assoc., 2010.

17. J. Canfil, "Honing cyber attribution: A framework for assessing foreign state complicity," J. Int. Affairs, vol. 70, no. 1, pp. 217–226, 2016. [Online]. Available: https://www.jstor.org/stable/90012607

18. D. Wheeler and G. Larsen, Techniques for Cyber Attack Attribution, Inst. Defense Anal., Alexandria, VA, USA, 2003

19. N. Tsagourias, "Cyber attacks, self-defence and the problem of attribution," J. Conflict Security Law, vol. 17, no. 2, pp. 229–244, 2012.

20. K. E. Eichensehr, "The law and politics of cyberattack attribution," UCLA Law Rev., vol. 67, p. 520, Jan. 2020