Anomaly-Based Detection of Attack on SSL Protocol using Pyod

Rasmiya Najeem, Shahnaz K Nassar, Nima S Nair

***Abstract:** At the current era Security and Protection is of vital significance. While there is no system that is immune to attack, a steady and effective system-security-framework is fundamental to ensuring the protection of data. On a fundamental-level, we are living in a more technologically advanced world than we were as of a decade ago. This rise in the broad utilization of technology carried with it and ascent in network attack. Networked PCs have taken over practically all parts of our lives, they store and deal with a ton of data that whenever traded off could bring about critical results. Secure Sockets Layer and Transport Layer Security(SSL/TLS) are presently broadly embraced innovation to give security. SSL/TLS convention is made to give privacy to delicate data trade over the Web. They can be used to secure secrecy and protection however can in like manner be used to shroud vindictive exercises. Regardless, assurance right now in all the information being scrambled independent of whether the information is malignant or not. The SSL-Attackers don't utilize a particular system for the assault, the attacker may attempt to stick the system by making superfluous-traffic. Using Anomaly-Detection-technique we find the outliers by analyzing the data-captured using Wireshark and identify any possible attack on the network. For detecting outlier in the traffic we have used ABOD technique contained in the PyOD-library, which is an open-source toolbox provided in Python for identification of anomaly on multi-variate information. Anomaly-Based-Detection is a practical and realistic option for identification of attack against security-convention.*

***Keywords:** ABOD, Anomaly Detection, Network Traffic Analysis, Outliers, Python, SSL, TLS, Wire-Shark.*
Abbreviations: SSL, Secure Socket Layer; TLS, Transport Layer Security; SS, System Security; MSK, Master Secret Key;

I. INTRODUCTION

System Security-(SS) incorporates the methodologies and strategies got a handle on to hinder and screen unapproved to locate a decent pace, or repudiating of a PC system and structure available assets SS recalls a system that is required by the framework head is underwriting the exposure towards information. Clients choose or are assigned the ID or puzzle articulation or are then given another confirmatory information that allows them easy exposure to information and as well as activities in their key role.

SS incorporates an open and private blend of PC systems which can be employed in an equitable way: coordinating businesses and correspondences-between affiliations, various governmental work places and people. Frameworks, for example, within an affiliation and other, which may be open to the system, may be private / guaranteed. SS has affiliations, attempts and different kinds of foundations. It does what its title says: it asserts the structure, even as activities are assured and are coordinated.SLL and TLS is a cryptographic/encryptiion convention made to verify correspondence over the Internet/Web by giving protection and unwavering quality. There are various forms of SSL-and-TLS being made and conveyed over the previous years.

A. SSL Handshake Protocol

SSL/TLS capture is a questionable strategy to decode SSL communication. At the point when it is actualized in a professional workplace, if there isn't an enlightened arrangement, in some cases it might direct the organization to protection and security concerns. The main advantage of this method is the permeability given to the security specialist.

SSL/TLS decryption licenses analysis of communication payload. The principle downside of this method is the alteration of the end to end communication which could be recognized by the attacker. Modern malware can likewise distinguish that a server certificate isn't given by its own confided/restricted in Certificate Authority.

In order to exchange messages between an SSL-empowered server and SSL-empowered customers when they start to establish an SSL association, the Handshake protocol employs the SSL Record protocol.

This trade of messages is intended to empower the accompanying activities:

Server authentication to clients. To select cryptography algorithms and/or ciphers both serving the client and server.

Customer to the server-authenticate.

- To create common secret keys by using public key encryption.
- Set up an encrypted SSL-connection and open it.

A session/sessions-begins when a client connects to an SSL-server.

- An SSL-handshake starts when an SSL-session begins.
- Protocols to be utilized to communicate are established.
- Cryptography algorithms are chosen.
- Both the customer and server are authenticated.
- MSK is generated using public-key encryption.
- A MSK is created from a premaster-secret-key sent

Revised Manuscript Received on April 30, 2020.
***** Correspondence Author

Rasmiya Najeem*, PG Student, Dept of CS and IT, Amrita School of Arts and Sciences, Kochi, INDIA, Email: rasmiyanajeem76@gmail.com
Shahnaz K Nassar, PG Student, Dept of CS and IT, Amrita School of Arts and Sciences,Kochi, INDIA, Email: shahnazknassar@gmail.com
Nima S Nair, Asst Professor, Dept of CS and IT, Amrita School of Arts and Sciences, Kochi, INDIA, Email: nimasnair@gmail.com

from the client. It is utilized to generate four-session keys.

  - An encryption key to send/transfer data from client to server.
  - An encryption key to send-transfer data from server to client.

- An authentication key to send/transfer data from client to server.
- An authentication key to send data from server to client.

PyOD is an open source Python toolbox for doing adaptable anomaly identification on multivariate information. Strikingly, it offers access to wide/huge extent of exception recognition algorithms ,including built up oddity gatherings and ongoing neural framework based approaches, under a lone, very well reported/papered API proposed for use by both the professionals and analysts. Considering force and adaptability, best practices, for instance, unit testing, continuous coordination, code inclusion, viability checks, interactive models, and parallelization are emphasized as focus parts in the tool compartment's improvement.

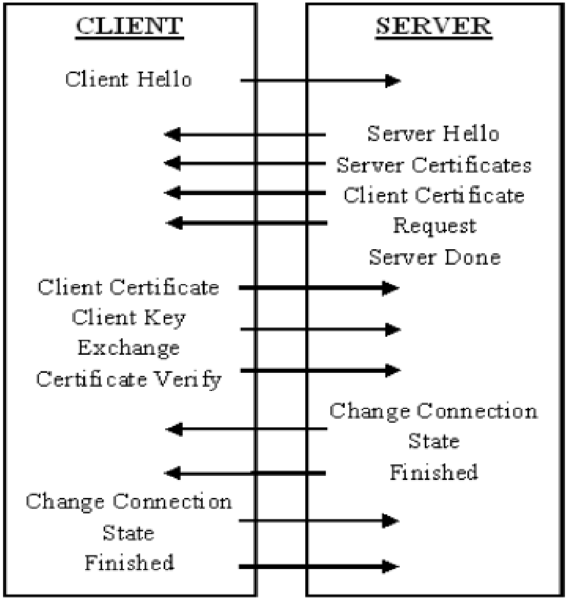

Fig 1.SSL handshakes

Wireshark is a well network analytics tool for progressively capturing and displaying packets in human readable format. Wireshark incorporates filters, colors and a wide variety of features which allow you to thoroughly investigate single packets and organize traffic.

II. BACKGROUND STUDY

This paper proposes how the malware hides its activities by moving to encrypted communication with protocols like SSL/TLS. The paper shows that the defenders still have passive methods to collect SSL/TLS artifacts and discover malicious activity. It shows how to perform the analysis in a controlled environment by building a cloud-based infrastructure concept as a contribution to the security community[3].

In [2] it explains about designing and implementing the PyOD library to fill the gap of lack of an outlier detection library in Python. It incorporates in more than 20 classical and emerging detection algorithms and is being utilized in both scholastic and business ventures. The library was implemented keeping in mind the need for scalability and robustness.

In [5],the author talks about the importance of security and privacy in today's world and how SSL/TLS is widely used for the same. The paper proposes that data being encrypted in SSL causes problems for IDS that relies on sniffers to analyze packets in the network. Therefore, the authors propose a host-based IDS that analyzes packets after decryption to detect such attacks. It also determines that anomaly-based detection is a viable alternative for detecting attacks against security protocols.

The authors in [1] talks about the role of SSL in cyber--security and how it can ensure regularly that a large amount of community information is passed down over the internet during web exchanges or classified data.The paper clarifies how SSL protects a transaction with the utilization of a mix of public-key and symmetric-key encryption to make sure about the security of possible connection across two computers which could be a Web or mail server or a client computer, having Internet or internal network communications.-This paper presents an exhaustive survey of well-known distance, density and other outlier detection techniques and compares them and then discusses its detection in conjunction with network anomaly detection based on supervised and unsupervised learning.in the context of network anomaly detection.

Relevance of outlier in the detection of anomalies [6].Data mining techniques are also quite popular in detecting attacks in secure traffic and in related research [4] the authors use data mining techniques and they suggested that attacks using SSL / TLS protocol be detected by Denial of Service(DoS). Based on the analysis of statistics extracted from packet headers, a model of normal user behavior is made by calculating Chi-square values and clustering flows with DBSCAN and is used to detect DDoS attacks.

III. PROBLEM STATEMENT

SSL is a double-edged sword that could be utilized to forestall and identify strange transactions. Information exfiltration, noxious correspondence with Command and Control and vindictive download use SSL/TLS scrambled traffic. Safeguards have inactive strategies for gathering SSL/TLS ancient rarities and find noxious activity[1].

SSL/TLS interception is a disputable technique to decode SSL communication. At the point when it is implemented in a professional workplace, if there isn't an edified arrangement, sometimes it might direct the organization to protection and security concerns. The primary favorable position of this technique is the visibility given to the security analyst.

SSL/TLS decryption grants analysis of the correspondence payload. The primary disadvantage of this strategy is the change of the start to finish communication which could be distinguished by the attacker.

Advanced malware can likewise distinguish that server authentication isn't given by its own trusted Certificate Authority[3].The SSL attackers don't utilize a particular system for the assault.

The attacker may attempt to stick the system by making superfluous traffic. This can be avoided by recognizing the outlier and blacklisting and later analyzing it[5]. Anomaly-based identification is a reasonable option for distinguishing assaults against security conventions and based on their discoveries confirms that the bogus positive rates are very low[5]. Anomalies in organized traffic are significant side effects of PC security issues and system blockage. System chairmen and analysts have been attempting to find a strategy that can precisely and speedily see inconsistencies in arranging traffic. For the most part, we can order anomaly detection strategies into two significant gatherings: signature-based-techniques and statistical-based-techniques.

The signature-based-techniques screen and contrast arrange bundles or associations and foreordained examples known as marks. This method is a basic and efficient handling of the review information. Despite the fact that the bogus positive pace of these strategies can likewise be low, contrasting parcels or associations and enormous arrangements of marks is a tedious undertaking and has restricted prescient abilities. The signature-based techniques can't identify novel irregularities that would not be defined in the marks, and along with these lines managers, much of the time need to refresh the framework marks. The statistical-based strategies, be that as it may, can learn the conduct of system traffic and the chance of distinguishing novel inconsistencies. The AI approach is one of the statistical-based techniques that have high abilities to consequently figure out how to perceive complex examples and settle on astute choices dependent on data[8].

IV. IMPLEMENTATION AND RESULT

As it has already been established, Anomaly-based identification is a sensible choice for recognizing attacks against security conventions.

PyOD is a Python scalable multiple data abnormality recognition toolkit. It enables access under one well-defined API to approx. 20 anomaly detection algorithms.

Here we choose the Angle Based Outlier Detection(ABOD) specified in the PyOD Library in Python.

In the comparison of Angular Based Outlier Detection(ABOD) with other variants in [11], ABOD achieved the maximum best performance possible.

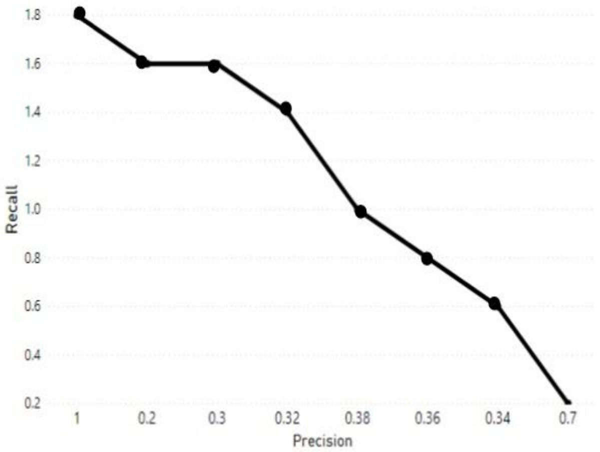

Fig .2.Precision-Recall Graph on 1000data points

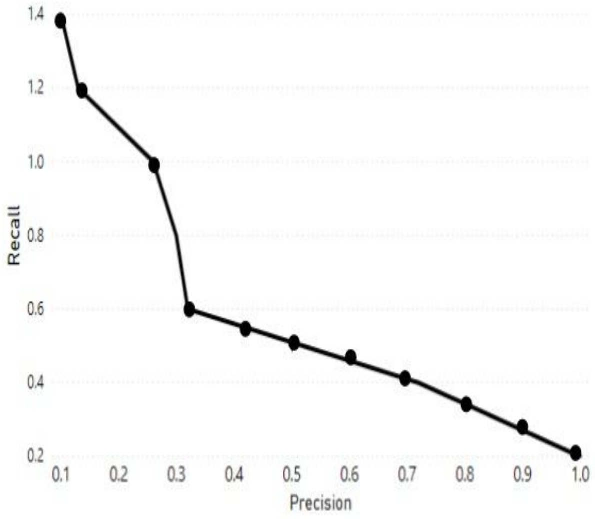

Fig .3.Precisiion-Recall Graph on 5000datapoints

ABOD fulfils the highest possible performance. An algorithm's outlier obtaining capability can be calculated with recall and precision graphs. Since the recall is the percentage of all outliers that had already been retrieval from the data set, for each additional outlier we can observe a new recall level. We are now measuring the precision that means how many items in the current result set are actually outliers for each of these recall rates. In Figure 2 and figure 3 the observed precision-recall graph shows 1000 and 5000 data point for two separate database sizes.

Using Wireshark SSL encrypted packets are captured. While analyzing the captured packets it was noted that an IP is trying to jam the network by constantly sending more than 1000 packets in a single timestamp, unlike the normal packet transfer which ranges from 20 to 30 in a single timestamp. Hence for the assurance, we have to find the outlier from the traffic.

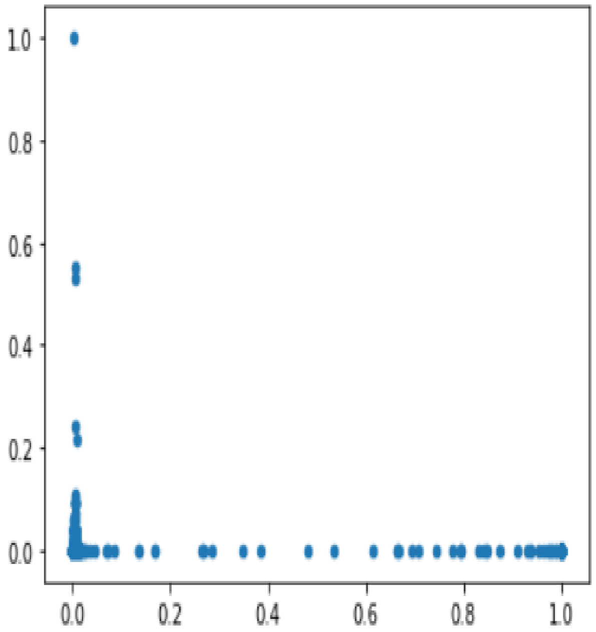

Fig 3.Scatter Plot for dataset

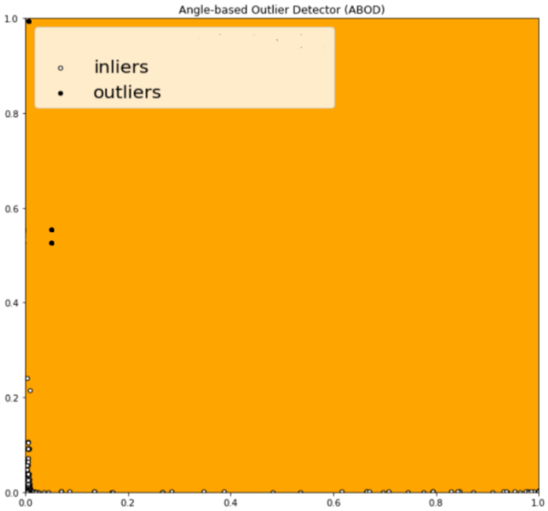

Fig 4.ABOD graph

The relationship between each point and their neighbours, considered by ABOD. The relations between these neighbors are not known. The disparity between their weighted cosine values and all neighbors can be known as the outlier scoring.On multi-dimensional data ABOD performs well. PyOD provides a two distinct types of ABOD[2]. First, Quick ABOD that uses k-closest neighbors(KNN) for estimating. Furthermore, an initial ABOD that takes into consideration all high-time dynamic training points.The outlier detected is then marked for later analysis. The later analysis tells whether the outlier found from ABOD should be blacklisted or not.

V. CONCLUSION

As discussed in the introduction, attackers hide their activities by moving to encrypted communication with protocols like SSL/TLS. Due to the increased use of SSL/TLS in organizations, attackers have started to modify their strategies to slow down network forensics analysis.

- Here we detected those attacks by analyzing the network traffic data captured using Wire-shark tool, a network analysis tool to capture packets. Anomaly based detection is a suitable alternative for detection of attacks against security protocols. We used the Angle Based Outlier Detection in PyOD for finding the outliers as it achieved the best performance in comparison to others.
- Using this algorithm the outliers were detected and helps to identify any possible attack that could occur based on the spurious transactions in the network.
- We do not have confirmed proof to quantify the false positives (which will form part of the future works) and whether it will identify all attacks, but based on our findings we assume that the false positives rate will be very low.

VI. FUTURE WORK

An enhanced version of the ABOD algorithm can be used to improve the accuracy of the results. Furthermore, the anomalous traffic can be blacklisted, and an algorithm can be devised which can be implemented to block transactions from suspicious sources. As future work it would be intriguing to see the actual result of executing the algorithm for identifying the attacks and blacklisting the source. This could be implemented as an independent tool which runs on the host.

REFERENCES

1. Er. Kaur P, Er. Kaur G, May 2017"Review of Role of SSL in Cyber Security," International Journal of Advanced Research in Computer Science(IJARCS) Volume 8, No. 4s
2. Zhao Y, Nasrullah Z, Li Z, May 2019 "PyOD: A Python Toolbox for Scalable Outlier Detection," Journal of Machine Learning Research (JMLR) Vol. 20,
3. Ngoc Huy Nguyen, March 2019" SSL/TLS Interception Challenge from the Shadow to the Light ," The SANS Institute 2019
4. Zolotukhin M, Hämäläinen T, Kokkonen T, Niemelä A,Siltanen J, August 2015 "Data Mining Approach for Detection of DDoS Attacks Utilizing SSL/TLS Protocol," Springer International Publishing Switzerland 2015
5. Kazi S, May 2010 "Anomaly based Detection of Attacks on Security Protocols ,"
6. Prasanta Gogoi, D K Bhattacharyya, B Borah and Jugal K Kalita, February 2011 "A Survey of Outlier Detection Methods in Network Anomaly Identification,"
7. Tran A T , September 2017 "Network Anomaly Detection," Network Architectures and Services
8. Limthong K and Tawsook T, 2012 "Network Traffic Anomaly Detection using Machine Learning Approaches," 2 IEEE Network Operations and Management Symposium (NOMS)
9. Dokas P, Ertoz L, Kumar V, Lazarevic A R, Srivastava J, Tan P N , 2002 "Data Mining for Network Intrusion Detection,"
10. Shubh T, Sharma S, June 2016 "Man-In-The-Middle-Attack Prevention Using HTTPS and SSL," International Journal of Computer Science and Mobile Computing, Vol.5 Issue.6
11. Kriegel H P, Schubert M, Zimek A, August 2008 "Angle-Based Outlier Detection in High-dimensional Data," International conference on Knowledge discovery and data mining
12. Patcha, A. and Park, J.M, August 2007 "An overview of anomaly detection techniques: Existing solutions and latest technological trends," Comput. Networks 51, 12, 3448– 3470
13. Wagner D, Schneier B B, November 1996 "Analysis of the SSL 3.0 protocol," Proceedings of the Second USENIX Workshop on Electronic Commerce Oakland
14. Shou Z, Tian H, Li S and Zou F, October 2018 "Outlier detection with enhanced Angle Based Outlier Factor in high dimensional data," International Journal of Innovative Computing, Volume 14, Number 5
15. Dominguesa R, Filipponea M, Michiardi P, Zouaoui J, September 2017 "A comparative evaluation of outlier detection algorithms,"
16. Kurniabudi, Purnama B, Sharipuddin, Darmawijoyo, Stiawan D, Samsuryadi, Heryanto A, Budiarto R, March 2019 "Network anomaly detection research: a survey," Indonesian Journal of Electrical Engineering and Informatics (IJEEI) Vol. 7, No. 1
17. Kumar N and Kumar U (2018) "Anomaly-Based Network Intrusion Detection: An Outlier Detection Techniques,"
18. Xiaojie Li, Lv J C, Cheng D(January 2015) "Angle-Based Outlier Detection Algorithm with More Stable Relationships," Springer International Publishing Switzerland.

AUTHORS PROFILE

Rasmiya Najeem is a PG student in Amrita School of Arts and Sciences , Kochi, INDIA currently pursuing Masters in Computer Application.

Shahnaz K Nassar is a PG student in Amrita School of Arts and Sciences , Kochi, INDIA currently pursuing Masters in Computer Application.

Nima S Nair is an Assistant Professor in Dept of Computer Science and IT, Amrita School of Arts and Science, Kochi, INDIA.