2016

Darknet and DoD Networks: Obfuscation, Spoof Detection, and Elimination

Gallup, Shelley P.; Anderson, Tom; Garza, Victor (Bob); Irvine, Nelson; Wood, Brian (Woodie)

Monterey, California. Naval Postgraduate School

Title: Darknet and DoD Networks: Obfuscation, Spoof Detection, and Elimination
Report Date: 10 March 2017 Project Number (IREF ID): WF801-WF900
Naval Postgraduate School / School: GSOIS
Report Type: Final Report

Period of Performance: 10/01/2016-03/30/2017

Project PI: Associate Research Professor Shelley P. Gallup, Ph.D. (GSOIS)

Additional Author/Authors: Tom Anderson, Ph.D., TRAC Monterey; Victor (Bob) Garza, Lecturer (GSOIS, CS); Associate Research Professor Nelson Irvine, Ph.D. (GSOIS); Brian (Woodie) Wood, Faculty Research Associate (GSOIS)

Student Participation: LT Kevin Dougherty, USN (Cyber Systems Operations)

Prepared for:

Topic Sponsor: N2/N6I

Research Sponsor Organization (if different): Fleet Cyber Command

Research POC Name: Capt Roy Petty USN

Research POC Contact Information: 443 634 4608

EXECUTIVE SUMMARY

Project Summary

There is no process or system capable of detecting obfuscated network traffic on DOD networks, and the quantity of obfuscated traffic on DOD networks is unknown. The presence of obfuscated traffic on a DOD network creates significant risk from both insider-threat and network-defense perspectives. This study used quantitative correlation and simple network-traffic analysis to identify common characteristics, relationships, and sources of obfuscated traffic. A set of concepts were identified and proposed as a set of testable Key Cyber Concepts (KCCs) for obfuscation behavior. Each characteristic was evaluated individually for its ability to detect obfuscated traffic and in combination in a set of Naive Bayes multi-attribute prediction models. The best performing evaluations used multi-attribute analysis and proved capable of detecting approximately 80 percent of obfuscated traffic in a mixed dataset. By applying the methods and observations of this study, the threat to DOD networks from obfuscation technologies can be greatly reduced (Abstract from LT Kevin Dougherty NPS 2017 thesis "Identification of low latency obfuscated traffic using multi-attribute analysis".)

Background

OPNAV N2 is seeking ways to mitigate network vulnerabilities and capabilities to meet the evolving cyber threat. A considerable vulnerability to cyber security is the unfettered network access to users with obfuscated identities. Currently obfuscation technologies allow users to act anonymously, without attribution, thus creating a hostile cyber culture where both inside threats and outside may snoop, act outside of protocol, or sabotage systems. Developing a capability to discern obfuscated traffic from non-obfuscated traffic would allow policies to be put in place to block obfuscated users where it is deemed appropriate. This effort produced LT Dougherty's master's thesis that provides details for the research discussed here, and was honored as an Outstanding Thesis.

We sought to identify obfuscation indicators that can be used to evaluate whether low-latency Transmission Control Protocol/Internet Protocol (TCP/IP), specifically HyperText Transfer Protocol Secure (HTTPS) traffic, is employing obfuscation techniques. The research questions pertaining to this explored were: Can low-latency obfuscated network traffic be identified in real time? What IP traffic indications can be used to identify obfuscated low-latency network traffic? Can multiple indications be incorporated into a multi-attribute analysis model to accurately identify obfuscated traffic? Can a multi-attribute analysis model be used in a tool to provide a real-time

processing capability to analyze obfuscated traffic data for automated response? Can the key cyber concepts be easily adapted to an architecture framework for ad hoc implementation of cyber solutions into an operational system?

Findings and Conclusions (to include Process):

We used a quantitative correlation approach to examine analysis techniques for identifying low-latency obfuscated network traffic. Real-world network data was analyzed using traditional network traffic analysis. We examined the characteristics, and the relationships between characteristics, associated with obfuscated traffic. We conducted independent statistical analysis of network traffic attributes first to determine each attribute's viability to function as a single discriminator. We then used a Naive Bayes classifier model for multi-attribute analysis with the assumption that all variables are independent. The Naïve Bayes classifier utilized the individual characteristics together, some with higher false-positive rates (FPRs), to detect obfuscated traffic.

A virtual lab was configured with an Internet-facing SharePoint page, that enabled testing and evaluation of obfuscated and non-obfuscated network traffic for a variety of popular operating systems. To generate non-obfuscated network traffic, a standard Firefox browser was used. To standardize the datasets, Selenium IDE, a Web browsing automation tool, was used to script 'normal' browsing activity on the webserver.[1] Baseline data consisting of several hours of network traffic from regionally distributed actors was used to evaluate and determine the performance of each indicator prior to constructing the multi-attribute detection model. A combination of physical and virtual machines was used to gather 24 hours of Tor and non-Tor traffic, respectively. All data was written to database, and subsequent statistical analysis was conducted on the data.

We examined four separate indicators (KCCs) to determine whether network traffic was obfuscated. Key Cyber Concept One: Low TTL Count, this examined the time-to-live (TTL) field of the IP header to determine whether incoming traffic originated from Tor or routine network traffic. Key Cyber Concept Two: Common Tor Packet Sizes, this is a measure to discern common Tor packet sizes. Key Cyber Concept Three: High TCP Offset, we analyzed the average TCP offset of Tor and non-Tor Web traffic to test whether it was possible to categorize each new instance of traffic into the webserver as either obfuscated or non-obfuscated. Key Cyber Concept Four: Known Tor Exit Node, a

[1] Selenium IDE is a Firefox extension that is also compatible with the Tor Browser. A full description of Selenium IDE capabilities is available at http://www.seleniumhq.org/projects/ide/.

blacklist of published Tor nodes was used to verify whether any network traffic originated from a known Tor exit node.

Analysis of 702,376 unique packets showed distinct attributes for the key cyber concepts (KCCs) in this study. Each KCC was tested both with and without filtering to determine its ability to discern Tor traffic. The testing was conducted in two ways. First, using the list of attributes as a pre-filter, the datasets were filtered to include only those rows that exhibited the observed attribute, and second, the datasets were evaluated in their unfiltered states. After individual testing, each KCC's attributes were applied in a multi-attribute analysis model to increase the probability of classifying Tor traffic.

KCC1: Analysis showed 56 unique IP time-to-live (TTL) values present in Tor traffic and 77 unique TTL values present in non-Tor traffic. Interestingly, most Tor TTL values were concentrated between 30 and 60 with very few observations above 100. Non-Tor TTL values were also observed between 30 and 60, but many exhibited a value above 60, albeit in lower quantities.

The observed Tor TTL is consistent with the default operating system TTL count of 64, which is normally attributed to Linux-based systems. Notable concentrations were observed in Tor traffic at 44, 45, 46, 47, 48, 49, 50, 52, 53, and 54 while non-Tor IP TTL values were concentrated at 51, 57, 116 and 128. Confirming these TTL values enables their use as discriminants in both single and multi-attribute analyses.

KCC2: Analysis of packet sizes showed Tor packets were observed only at sizes 52 and 1500 and accounted for 54.5 percent of all Tor packet sizes. However, additional analysis observed a low number of Tor and non-Tor packets with a size of 1500 and a very high number of both exhibiting a packet size of 52. Thus, it is determined that solely using packet sizes of 52 and 1500 will not serve as a good discriminator in either single attribute or multi-attribute analysis.

KCC3: Analysis showed that 92 percent of Tor packets and 92.6 percent of non-Tor packets exhibited a TCP offset value of either 5 or 8. Based on this observation, only TCP offset values greater than 9 could be viable discriminators for traffic type. There was very little difference between the mean and standard deviation of Tor and non-Tor packets. This suggests the additional TCP header data required by Tor has a negligible effect on the overall TCP offset.

KCC4: We used R-script to identify unique IP addresses in both Tor and non-Tor datasets separately. The unique rows were then compared to a listing of known Tor exit nodes active during the testing period. Results showed 1,218 unique Tor IP addresses, 2,091

unique non-Tor IP addresses, and 1,106 unique Tor exit node IP addresses. Each set of unique addresses was compared to identify intersections, or collisions, with the other datasets. There were only 178 intersections between the known Tor IP address dataset and the known Tor exit node dataset. These results confirm that a majority of known Tor traffic did not originate from a published exit node. Further comparisons showed zero intersections between the known Tor exit node dataset and known non-Tor dataset. Based on these two data points, when packets originate from a known Tor exit node, they are obfuscated.

The results from this research provide the basis for DoD to adapt to the DarkNet cyber threat using available cyber sensors and information (log files). We demonstrate data analytics of Network behavior to establish knowledge, and then did the preliminary development of an operational model based application. The data analysis of network traffic confirm expectations that a multi-attribute approach to classification improves certainty of classification. The exploitation and analysis of the TTL and packet length demonstrates that basic rules for classifying network entities and behavior may be developed through empirical analysis of available information, as in this case Snort log files.

The study of each KCC did result in either the confirmation of previously observed indicators or in new identifiable characteristics. In the second best performing test, error rates were roughly 20 percent for both Tor and non-Tor traffic. Although the observed False Negative Rates and False Positive Rates were high, use of this type of analysis on a network in real-time could result in a reduction of Tor traffic by approximately 80 percent—significantly reducing the overall threat level.

We implemented the obfuscation rules learned in the KCC analysis into the USAF's Behavior Based Network Management model for operational testing. Due to emphasis on data analytics, a preliminary implementation of the DarkNet KCCs into the GINA framework and AF BBNM model was accomplished, however, evaluation was left for future work.

Recommendations for Further Research

We recommend research and analysis on additional indicators of Tor traffic. 1. High RTT. Based on Tor's geographically dispersed architecture and routing schema, the linear correlation between geographic distance and round-trip time (RTT) may be present in Tor traffic; 2. HTTP Flow Analysis: Determine if there is a sufficient separation in the sizes for flow packet three; 3. Varied Source IP: Explore the possibility to identify Tor

traffic by monitoring for a IP address change from an authenticated user during a session.

Further research into Vector Relational Data Modeling in the GINA Framework and Joint Cyber model BBNM in the Joint Network Model Development and Testing: To facilitate real-time detection. Recommended future research is the operational testing of the axioms in the BBNM model that was extended with KCCs from this work. This work would be a logical extension of this effort, and may be either MS level or PI executed research.