Real-Time Detection of DNS Exfiltration and Tunneling from Enterprise Networks

J. Ahmed, H. H. Gharakheili, Q. Raza, C. Russell and V. Sivaraman, "Real-Time Detection of DNS Exfiltration and Tunneling from Enterprise Networks," 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), 2019, pp. 649-653.

—----------

Mohamed Amr Mohamed Nagib Alansary
uOttawa ID: 300273142

Summary

Problem

Exfiltration of data over DNS and maintaining tunneled command and control communications for malware is one of the critical attacks exploited by cyber-attackers against enterprise networks to fetch valuable and sensitive data from their networks since DNS traffic is allowed to pass through firewalls by default, attackers can encode valuable information in DNS queries without fear of being detected.

Solution

Introducing a real-time mechanism to detect exfiltration and tunneling of data over DNS through training a machine learning model that is capable of detecting anomalies in DNS queries and deploying the model on a huge traffic stream with malicious DNS queries to test the model accuracy in production environment.

Experiments

The authors collected their dataset from the DNS traffic of two enterprise networks, then identified the attributes for the FQDN (i.e. Fully Connected Domain Name) that characterizes benign and malicious queries. They used only stateless attributes that can be obtained from separate DNS query packets regardless of the time-series characteristics or

host DNS activity. That way it will not introduce any overhead in computation in real-time. The authors categorize the characteristics into three categories; namely, characters count, measure of randomness/non-readability in string (entropy) and the length of discrete labels in the query name.

The more characters the query has, the more probable the query is malicious and carries exfiltrated messages to Command and Control Server, they added count of characters in subdomain as another attribute because it is the part of the FQDN that carries these messages. Also the more uppercase characters and numerical characters the query has, the more probability the query has encoded/encrypted data and thus most probably will be malicious but without any guarantee; that's why the considered adding the count of uppercase characters and the count of numerical characters to their attributes. They also added the number of labels, maximum label length and average label length as attributes without specifying clear evidence of the reason they are adding these attributes.

The authors trained their model on benign data deriving the ground truth from top 10,000 primary ranked domains from Majestic Million which may contain some malicious domains as well due to the high number of infected clients querying them.

Results

The authors obtained very high accuracies for both the datasets involved in the experiment after tuning the hyperparameters.

Source: "Source: Real-Time Detection of DNS Exfiltration and Tunneling from Enterprise Networks," 2019

TABLE III
ANOMALY DETECTION FOR RESEARCH INSTITUTE.

Input	Output	Days 1-4	Days 5-7
Benign domains	normal	98.44%	98.30%
	anomalous	1.56%	1.70%
Others	normal	78.43%	77.55%
	anomalous	21.57%	22.45%

TABLE IV
ANOMALY DETECTION FOR UNIVERSITY CAMPUS.

Input	Output	Days 1-4	Days 5-7
Benign domains	normal	97.99%	97.99%
	anomalous	2.01%	2.01%
Others	normal	70.57%	70.59%
	anomalous	29.43%	29.41%

Conclusion

Cyber-attackers always seek valuable and sensitive information in organizations, a critical approach attackers may take is DNS exfiltration and tunneling. The authors developed, tuned, trained and validated a mechanism for real-time detection of malicious DNS queries. They evaluated the scheme on live 10 Gbps traffic streams by injecting malicious DNS

queries generated by an exfiltration tool. Finally they made their tools and datasets available to the public for more investigation.

Critical Review

Research Goal

What is the research goal?

Training a machine learning model that can be trained to detect the anomalies in DNS queries for detection of exfiltration and tunneling of data over DNS, model that can be deployed in real-time large traffic stream enterprise networks.

What question(s) is the author trying to answer? Explain.

The author is trying to detect if we can train a machine learning model that can detect malicious DNS queries in real-time with minimal computation needed and as fast as possible. The author trained such a model and deployed the model in a real-time environment and found that the accuracy was very high and the computational speed was very fast as well with lower computational power.

Clarity

How is the clarity of the paper?

The authors introduced the problem very well, they explained the terms in the research, introduced many examples of this attack in real life, and introduced the methodology of the attack in simple terms.

Is it written in such a way that an interested reader with a background in machine learning, but no special knowledge of the paper's subject, could understand and appreciate the paper's results?

Yes, of course, they introduced the problem in simple and straight forward terms, described the attack simply to the readers not in the cyber security field, they gave many examples on the attack and how critical it is for the organizations to have such a solution to detect this kind of attacks to protect their networks from being breached.

Related Work

Is the related work adequate?

The authors didn't provide clear evidence of the claim that signature-based classification is not sufficient for addressing the growing security issues, they also didn't provide any evidence of the claim that DNS tunnel detection using character frequency analysis that is based on threshold value can be easily tricked by attackers.

Is it complete and well written?

The authors provided claims that are not supported with a reference or a clear evidence, they should support their claims and therefore it is not complete but well written and included many references that describe the problem well.

Do the authors clearly acknowledge and identify the contributions of their predecessors?

They acknowledged the contributions of their predecessors; however, they didn't provide clear evidence of their claims that the other methods will not be sufficient and can be easily tricked by attackers.

Methods

What methods are being applied?

The authors used Isolation Forest (iForest) because of its efficiency in anomaly detection in high-dimensional datasets with minimal memory and time complexities where it is highly likely an instance will be anomalous if the forest of random trees collectively produces shorter path lengths for that instance. They tuned three parameters; namely, the number of trees (n_estimators), height limit of trees (max_samples) and contamination rate. Regarding the ground-truth malicious instances, the authors generated DNS exfiltration queries using DNS Exfiltration Toolkit to exfiltrate CSV files containing 1000 samples of random credit card details.

What methods are the authors applying to answer the question? Explain. Is the description provided adequate, detailed, and clear?

The authors described the method used; namely, Isolation Forest (iForest) but they did assume that the reader already knows about the technical details of the algorithm being used and so they only introduced the methodology of the algorithm without providing a proof that this method actually works; they added the reference of the algorithm though.

Results and Claims

What are the research results?

The authors obtained very high accuracies for both the datasets involved in the experiment after tuning the hyperparameters.

Source: "Source: Real-Time Detection of DNS Exfiltration and Tunneling from Enterprise Networks," 2019

TABLE III
ANOMALY DETECTION FOR RESEARCH INSTITUTE.

Input	Output	Days 1-4	Days 5-7
Benign domains	normal	98.44%	98.30%
	anomalous	1.56%	1.70%
Others	normal	78.43%	77.55%
	anomalous	21.57%	22.45%

TABLE IV
ANOMALY DETECTION FOR UNIVERSITY CAMPUS.

Input	Output	Days 1-4	Days 5-7
Benign domains	normal	97.99%	97.99%
	anomalous	2.01%	2.01%
Others	normal	70.57%	70.59%
	anomalous	29.43%	29.41%

A paper can contain many kinds of results. For instance, it may have applied results, theoretical results What claims are made in the paper?

The authors trained an Isolation Forest (iForest) model and after hyperparameter tuning they obtained very high accuracies on both datasets involved in the research.

What are the authors declaring to have accomplished?

The authors deployed a model that can detect malicious DNS queries with very high accuracy; the scheme can process about 1250 DNS queries per second and 800 μsec on average for feature extraction and prediction.

Support of Results and Claims

How are the claims supported?

The authors claimed that signature-based classification is not sufficient for addressing the growing security issues, especially having a ground truth on malicious queries to train the classifier; they didn't provide any clear evidence of this claim.

The authors claimed that DNS tunnel detection using character frequency analysis based on a threshold value can be easily tricked by attackers, again no evidence on such claim.

The authors added the number of labels as an attribute in their dataset and assumed that DNS exfiltration/tunneling traffic tend to have a pattern in labeling but they didn't provide any reference or clear evidence that support this assumption.

The authors added the maximum label length and the average label length to the list of attributes without providing neither the reason nor the evidence that this is even relevant to the classification problem in hand.

What experiments are conducted to support the claims? How are the experiments carried out?

The authors collected real data from two enterprise networks, identified the FQDN's attributes that will be beneficial for the model. They used stateless attributes that will not introduce an overhead in real-time computation. They categorized the features into characters count, entropy and length of discrete labels. They added the count of characters in subdomain, the count of uppercase characters, the count of numerical characters, the number of labels, the maximum label length and the average length. The authors then trained the Isolation Forest (iForest) algorithm on benign data deriving the ground truth from top 10,000 primary ranked domains from Majestic Million. They obtained very high accuracies for both the datasets involved in the experiment after tuning the hyperparameters.

Do the authors evaluate their work in an adequate way (theoretically and/or empirically)?

Yes, they used three ways to evaluate their work: (a) cross-validation and testing the accuracy of the trained model for benign instances, (b) malicious DNS queries detection

rate testing, (c) testing the performance in real-time on live 10 Gbps traffic streams from the two organizations.

If appropriate, have the authors implemented their work and demonstrated its utility on a significant problem?

AVG. TIME COMPLEXITY OF OUR SCHEME.

extracting attributes	54 μsec
detecting anomalies	746 μsec
Total time per each query name	800 μsec

Yes, the authors tested the performance of the model in real-time on live 10 Gbps traffic streams from the two organizations. The authors quantified the average time for feature extraction and prediction by testing more than 300 million DNS queries and it takes on average 800 μsec for final prediction. The scheme can process about 1250 DNS queries per second which is above the actual rate of DNS queries in both organizations.

Source: "Source: Real-Time Detection of DNS Exfiltration and Tunneling from Enterprise Networks," 2019

Missing Claims and Results

What reasonable claims and results are missing from the paper? What interesting experiments could be added to improve the paper?

The dataset of DNS queries collected from two research institutes; the machine learning model may fail to generalize for other domains since the distribution of research institutes DNS queries may differ from other organizations DNS queries.

The author could add other domains as well to the experiment, the author also may use different ground truth sources.

Discussion

Is the discussion adequate?

I think not. The discussion is missing many clarifications of features selection reasons and many proofs of assumptions about other approaches taken by other researchers. Also the authors didn't test their model in other organizations instead of research organizations to test the generalization of the model.

Is the discussion clear and well written?

Yes, the discussion is clear, easy to follow and well written but missing many clarifications and proofs.

Are strength, limitations and generality of the research adequately discussed?

The authors discussed the strength of their model and how it's better than other models but without providing any proof of their assumptions. The authors didn't discuss the limitations and generality of their model, maybe their model has a lack of generality and will not work very well with other organizations!

Future Work

What would be reasonable next steps for the research?

The authors should derive ground truth of benign domain from a more precise source cause they mentioned that top primary ranked domains from Majestic Million may include malicious domains as well due to the high number of infected clients querying them.

The authors should test their model in different organizations other than research institutes to test that the model is generalizing well.

The authors should consider collecting datasets from other domains as well.

Reference

J. Ahmed, H. H. Gharakheili, Q. Raza, C. Russell and V. Sivaraman, "Real-Time Detection of DNS Exfiltration and Tunneling from Enterprise Networks," 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), 2019, pp. 649-653.