Dynamic Score-Based Network Segmentation

David Maluf (dmaluf@cisco.com)
Nancy Cam-Winget (ncamwing@cisco.com)

Abstract

With insider threats still being a significant factor in the growing attack surface, technology and industry is adopting analytics and machine learning techniques to improve the automation and time to detect a vulnerability or potential threat. This paper proposes the additional use of a risk scoring mechanism based on differential analysis as a means to dynamically assess potential risk incurred by a system.

Introduction

Insider threat statistics from various security reports show that over 50% of the security incidents observed or incurred by organizations were due to employees. More importantly, over 53% of the attacks resulted in significant financial damages [1]. This highlights businesses strong concern about employees inappropriately sharing data or company assets or plain carelessness (e.g. from suspicious downloads, asset loss or being inadvertently phished).

The average network today, while using monitoring and security tools, is insufficiently equipped to automatically detect that they are being probed, let alone to understand the risks. Nonetheless, both classical and software defined (SDN) networks still require a human-in-the-loop to affect configuration, control and final assessment of a *threat* from controlled monitoring and observed security events. This human factor is the element of strength and yet a *point of failure*. That is, with the existence of security tools, the wealth of available data to analyze, and Security Information and Event Management (SIEM) tools, humans can still "miss" critical events. Additionally, improving the efficacy of detection often requires a faster and more dynamic closed loop function.

With the need for such a dynamic closed loop and improved automation, the cybersecurity industry is already embracing and leveraging the use of analytics and machine learning whether it be referred to as Intent Based Network or Secure Defined Networks [2]. While it is a growing trend, and their efficacy for detection is improving, there is still much room for improvement in both greatly reducing the false positives and the time to detection [3].

For such a closed loop function, we propose another element to the dynamic system: the use of differential analysis. While diverse assessments including the use of machine learning techniques are used, they lack the sensitivity analysis needed to measure how baseline network metrics change given the input. Thus, control can be managed by both, the expected behavior and the underlying sensitivity of the network. A good example is the rate of change of bandwidth in a network being more meaningful for the closed loop control than the measure of bandwidth over time.

This paper proposes to leverage quantitative measurements as a means to determine the network sensitivity as a means to dynamically improve the access controls in real-time. The controls are another means of affecting a dynamic network segmentation through a two-step process:

1) first, measure behavior quantitatively to derive a trust score (Trust Degradation Factor), and

2) second, enabling improved access control through the trust score influence.

Further details of the proposed Trust Degradation Factor and its use are described in this paper.

Measurements and Metrics

In a world of finite resources and time, assessment of a network will follow the aspects of any system from the point of view of the perceived *Risk* at the time of design or deployment. In other words, these guarantees are limited to the amount of accepted risk. As in most risk assessments, risk is decomposed into terms of likelihood and impact:

$$Risk = Likelihood \times Impact$$

Likelihood: For many cases, the likelihood (e.g. probability) is the method that correlates the present to the *known*-part from the past. This paper further augments the likelihood for the *unknown* as an exponential cost growth function to follow Solow's economic growth impact [7] (degradation over time). In other words, the likelihood will change purely with time regardless of the operational data.

Impact: The impact is associated with the existence of a cost. Impact is typically defined as the utility value of the system. As a consequence, when the impact of a threat is fully materialized, it is quantified to be limited by the utility value of the system it threatens. Simply put, the maximum threat to a $1-dollar investment is $1 whereas the maximum threat to $1B is $1B.

Risk: *Risk is the product of Likelihood and Impact*. Many standards govern the computation of risk. For example, in economic point of views, risk is viewed as an optimal answer to a cost function across the whole system. Comparatively, today's cyber-security strategies are a collection series of human defined scenarios such as SDN (not optimally minimizing the overall risk).

Use Case MW-1: virus attack based on network modulation

A "speculative" example of a new class of risk and therefore the mitigation strategy addressed in this document is as follows: Algebraic modulation (e.g. non-random) are made in the time delay between packets in TCP sessions. For example, a network sniffer can be used to decode the algebraic modulations somewhere between the end-points if not at the opposite peer. The modulation measurements serve as the means to establish a data point that can be used to determine if a network is being overloaded or under attack unbeknownst to either peer.

Network Trust Degradation Concepts

The choice of using the OSI model is intentional as it points to meet the stated MW-1 use case. MW-1 takes advantage from the main innovation of the Internet nature which is artificial time. Artificial time occurs with the reordering of data arrival frames on a new sequence such as demultiplexing. This novelty renders the communication dimensions asynchronous and as such, a manifestation of a frame (i.e., packet) can arrive any time and out of order.

A major outcome of the derived arbitrary sense of time is the need to analyze the newly created data-time and thus their frequency spectrums. With demultiplexing, it is often expected that the derived spectrums

will be different from the true input to accommodate for the allowed data reordering among the data sets. The newly created spectrums are artifacts of demultiplexing. Demultiplexing observations are linear in their spectrum mappings. This spectrum assumption also holds inversely on a multiplexing process.

The focus is on the constant (that is pointed out when applying Shannon's theorem) on the information content carried by the packet reordering. It can be deduced that any transformation scheme cannot increase the total information exchange but can only decrease or equate it. Measuring this constant and its variability over time is essential in determining the maximum information that occurs at the interface [9][10].

For digital data, the discrete Fourier Transforms (DFT) is the desired analytical function to occur over an arbitrary time index [6]. For completeness in the introspection of the data, the signal phases are also considered. Digitization fails to account for the known impact signal phases have and they are therefore considered the lossy aspect of the digitization, the DFT analytical functions will therefore compensate phase shifts with the corresponding frequency adjustments [7][8].

The data set represents the **layer 2 behavior** over an IP network. The quasi-physical quantity parameters modeled in the table below are the differential times between real time and their corresponding offsets, or the delays. Lengths are also valid dimensions. In general, models for the data sets for I/O for physical entities are from the frames over serialization transmissions or other transmission/reception. This will also extend to all observation models.

The table below summarizes common data sets calculated for operational and their reflection of physical world measurements:

Operational Data Vectors	Variables	Original Spectrum	Dimensions
Dimension(s)	e.g. Bandwidth	Size	Size
Absolute(s)	e.g. Time	Differential	
Relativity(ies)	e.g. Offsets		OSI model suggests the option to reshuffle the measurements ordering for layer 3.
Type(s)	E.g. dimensions and mux/demux		OSI model suggests new virtual dimensions based on unique identifications (IP, protocol, ports).

Trust Erosion (Degradation) Factor (TEF)

The concept behind building a Trust Function or Factor (TF) score, is to assess the risk accrued by the deviations. Trust degradation is a precursor index to failure [4][5]. The use cases of scoring the trust degradation in a system can apply to almost every aspect in networking, edge and cloud included. A well devised TEF will cover many use cases: for example (1) better and adaptive asset management (e.g., software updates); (2) better and adaptive digital asset certifications; (4) troubleshooting; and (5) real-time scalability and risk assessment for extremely large network, for example in federated cloud environment. The features of a digital trust scoring will start to reflect the risk or trust created on day 0.

The goal of dynamic segmentation is to assert the foundational risk factors for a system interaction and therefore their dynamic grouping (segmentation). Policies becomes an artifact of continuous change. Either way, TEF makes one clear assumption beyond the underlying data; that any (i.e. digital) system grows proportionally with complexity and time [11].

Access Control using TEF as a factor

Dynamic access control can now be affected through the use of a TEF. In the MW-1 use case, a network policy to "block malware", in general, can now be affected once the TEF reaches a certain threshold. In a finer or dynamic network segmentation routing scenario, routing paths can be affected by enabling network devices to only connect to devices whose TEF is within an acceptable range.

References

1. https://itsecuritycentral.teramind.co/2018/04/03/insider-threat-research-reports-and-surveys-the-top-facts/
2. Combatting Advanced Cybersecurity Threats with AI and Machine Learning. Andrew B. Gardner, RSA Conference 2017
3. On the Effectiveness of Machine and Deep Learning for Cyber Security, 2018 10th International Conference on Cyber Conflict (CyCon), 29 May-1 June 2018
4. M. C. Shewry and H. P. Wynn, "Maximum Entropy Sampling." *Journal of Applied Statistics*, 14:165–170, 1987.
5. R. R. Zhou, N. Serban, and N. Gebraeel, "Degradation Modeling Applied to Residual Lifetime Prediction Using Functional Data Analysis," *Annals of Applied Statistics* 2011, Vol. 5, No. 2B, 1586-1610.
6. R. Brincker, L. Zhang, and P. Andersen, *Modal Identification from Ambient Responses using Frequency Domain Decomposition*. Proc. of the 18th International Modal Analysis Conference (IMAC), San Antonio, Texas, 2000.
7. R. M. Solow, "A Contribution to the Theory of Economic Growth," *The Quarterly Journal of Economics*, Vol 70, No. 1, pp. 65-94, 1956.
8. T. M. Cover, J. A. Thomas, "Elements of Information Theory," *Wiley Series in Telecommunications,* Donald L. Schilling, Editor, 1991.
9. P. Sebastiani and H. P. Wynn, *Bayesian Experimental Design and Shannon Information*. Proceedings of the Section on Bayesian Statistical Science, 1997.
10. K. Chaloner and I. Verdinelli, "Bayesian Experimental Design: A Review," *Statist. Sci.*, Volume 10, No. 3, 273-304, 1995.
11. D. Maluf and R. Sudhaakar. *Trust Erosion: Dealing with Unknown-Unknowns in Cloud Security,* IEEE Cloud Computing 5(4): (2018).