Analysis of Software-Defined Network Traffic Using Entropy

Anastasiia Yatsenko, Valeriy Dubrovin, Maksym Pecherskyi and Maksym Chornobuk

December 18, 2022

Analysis of software-defined network traffic using entropy

Yatsenko Anastasiia, Dubrovin Valeriy, Pecherskyi Maksym, Chornobuk Maksym
Department of Software
Zaporizhzhia Polytechnic National University
Zaporizhzhia, Ukraine
nastya.yatsenko.zp@gmail.com, vdubrovin@gmail.com, m.pech.pr@gmail.com, 05643mak@gmail.com

***Abstract*— Entropy-based anomaly detection is being studied to improve the traditional approaches to network flow analysis based on volume and rules. In this work, the main threats to software-defined networks are considered, and the possibilities of using entropy for traffic analysis are also studied.**

***Keywords*— anomaly detection, entropy, netflow, network traffic measurement**

I. Introduction

A software-defined network (SDN) is an approach to creating a network, which uses software controllers or application programming interfaces (API) to communicate with basic hardware infrastructure and traffic direction in the network instead of physical routers and switches.

There are three parts of the typical SDN architecture:

- Programs that provide resource requests or information about the network as a whole.
- Controllers that use information from programs to solve how to route the data package.
- Network devices that receive information from the controller about where data were moved.

Software-defined networks use a centralized controller, hence providing the reliability of its work has a very important significance for network functioning. [1]

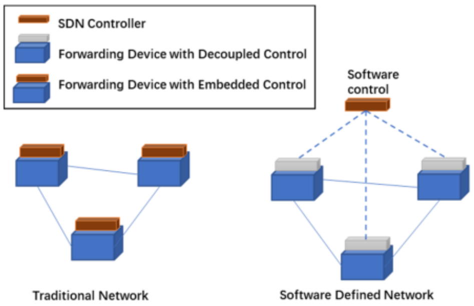
SDN Controller
Forwarding Device with Decoupled Control
Forwarding Device with Embedded Control
Software control
Traditional Network
Software Defined Network

Figure 1 - Decoupled control in SDN vs. traditional network [2]

II. Security issues in software-defined networks

The issue of security becomes especially acute when the number of network users increases. One of the largest and most common threats for software-defined networks is the Distributed denial of Service (DDoS) attack.

DDoS attack– a hacker attack on a computer system with the aim of bringing it to a state of failure. It's creating such conditions under which honest users will not be able to access system resources (servers), or this access will be complicated [3]. It can destroy available user network services, seriously threatening the network. When attackers send harmful data packages to the network, the usual traffic cannot be processed due to network resource consumption. As a result, networks and servers are blocked and conventional services are interrupted. DDoS-attackers are often targeted on SDN mainly due to the centralization of the controller.

III. The concept of entropy

To detect network anomalies based on statistical methods, the following approaches are generally used: clustering algorithms (of which the K-mean method is the most popular), Markov models, wavelet analysis, neural networks, artificial immune systems.

To detect network attacks, you can use its statistical characteristics as parameters of network traffic, such as sample mean, sample variance, Pearson's chi-squared test, and information-theoretic measure – entropy. Quantitatively, entropy is characterized using the entropy of probability distribution by C. Shannon.

The goal is to enhance the efficiency of IDS (intrusion detection systems), ADS (anomaly detection system) and information security management systems, perform theoretical and experimental research on the possibility of using real-time computed values of information entropy as a basic attack indicator on network services. [4].

Information entropy is a measure of uncertainty associated with random magnitude. Entropy characterizes the probability P with which one or another state is established, it is a measure of chaos or irreversibility. [5]. The greater the chaos of the system, the higher the value of entropy, and vice versa.

$$H_S(X)=\sum_{i=1}^{n} p(x_i)\log_a \frac{1}{p(x_i)} \qquad (1)$$

where X is a sign that can take the value of $\{x_1, ..., x_n\}$,

$p(x_i)$ is the probability function of the xi result,

n is the number of possible states. For the purpose of detecting anomalies, sample probabilities are commonly used, assessed by the number of x_i cases in the time window *t*. The value of entropy depends on the randomness

(it reaches the maximum when the probability *p(x_i)* is the same for all x_i), but also on the value *n*. In order to measure only randomness, normalized forms should be applied.

Alfréd Rényi [6] introduced an entropy metric of the order α as a mathematical generalization of the Shannon entropy to quantify the randomness of a system. Considering the discrete probability distribution $P = \{p_1, p_2, p_3 ... p_n \}$, , $p_i \geq 0$. Then the Renyi entropy of order α is determined as follows (2):

$$H_\alpha(x) = \frac{1}{1-\alpha} log_2 (\sum_{i=1}^{n} p_i^{\alpha}), \qquad (2)$$

where $\alpha \geq 0$, $\alpha \neq 1$, $p_i \geq 0$.

If the values of p_i are same, the maximum value of entropy is reached, known as Hartley entropy [7] (3):

$$H_0(x) = log_2(\mathrm{n}) \qquad (3)$$

When $\alpha \rightarrow 1$, H_α converges to the Shannon entropy (4):

$$H_1(x) = -\sum_{i=1}^{n} (p_i * log_2 p_i) \qquad (4)$$

Quadratic entropy, sometimes called collision entropy, is the Rényi entropy with parameter $\alpha = 2$ (5):

$$H_2(x) = -log_2 \sum_{i=1}^{n} p_i^2 \qquad (5)$$

Finally, when $\alpha \rightarrow \infty$, reaches the minimum value of information entropy. So we say that generalization of information entropy is a non-increasing function of order α, i.e. $H_{\alpha 1}(x) \geq H_{\alpha 2}(x)$, $\alpha_1 < \alpha_2$, $\alpha_1 > \alpha_2$ [8].

IV. Entropy to analyze traffic

Software complexes based on entropy analyze network data recorded by NetFlow sensors. Typical sensors, such as routers or special sensors, such as Softlowd [9], are connected to TAP or SPAN ports on switches. Streams are analyzed during fixed time intervals (every 5 minutes). Collected threads are registered in the database and then analyzed. The anomaly filters are predicted by direction, protocol and subnetwork for restricting the search area. Further, the value of the entropy of positive and negative values of α is calculated for the distribution of traffic characteristics.

In the detection phase, the observed entropy is compared to the minimum and maximum values stored in the profile and the anomality threshold is determined. Threshold values less than 0 or greater than 1 indicate abnormal concentration or dispersion, respectively. This abnormal dispersion or concentration for different trait distributions is characteristic of anomalies. [10]

Cong Fan et al. [2] proved that one of the solutions for detecting such attacks is using a fusion entropy. This method allows for detecting DDoS attacks in time close to real, and entropy values for normal and harmful traffic vary by 90%. This makes a fusion entropy effective in struggling with network threats.

V. Conclusion

This paper considers the possibilities of using entropy for traffic analysis of a software-defined network. Entropy analysis is a powerful tool in the fight against network threats such as DDoS attacks.

Software complexes detect harmful software by computing the values of entropy of traffic characteristics and comparison of the resulting values with a normal traffic profile.

Among the varieties of entropy analysis, analyzing based on fusion entropy is promising, since it is very sensitive to harmful traffic and allows for detecting threats in a mode that is close to real-time.

References

[1] What is Software-Defined Networking (SDN)? [Electronic resource]. – Access mode: https://www.vmware.com/topics/glossary/content/software-defined-networking.html

[2] C. Fan, N.M. Kaliyamurthy, S. Che, H. Jiang, Y. Zho and C. Campbell "Detection of DDoS Attacks in Software Defined Networking Using Entropy" 2022, 12, 370.

[3] DoS attack [Electronic resource]. – Access mode: uk.wikipedia.org/wiki/DoS attack (in Ukrainian)

[4] Burlakov M. E. The algorithm for detecting invasions in information networks based on artificial immune system: diss. ... Ufa, 2017 (in Russian).

[5] Shannon, C. A Mathematical Theory of Communication. Bell Syst. Tech. J. 1948, 27, 379–423.

[6] Renyi, A. Probability Theory; Enlarged version of Wahrscheinlichkeitsrechnung, Valoszinusegszamitas and Calcul des probabilites. English translation by Laszlo Vekerdi; North-Holland: Amsterdam, The Netherlands, 1970, p. 573.

[7] Hartley, R.V.L., Transmission of Information, Bell System Technical Journal, 7: 3. July 1928 pp 535-563.

[8] M. Bhuyan, D. Bhattacharyya, J. Kalita, "Information metrics for low-rate DDoS attack detection: A comparative evaluation" // Seventh International Conference on Contemporary Computing – 2014, P. 80-84.

[9] Damien Miller, Hitoshi Irino — Softflowd — Traffic flow monitoring. Available online: https://www.freebsd.org/cgi/man.cgi?query=softflowd& sektion=8&manpath=freebsd-release-ports

[10] Yatsenko A.K., Dubrovin V.I., Tverdochlieb Yu.V. Analysis of network traffic using entropy // Combinatorial configurations and their applications: Materials of the XXIII International Scientific and Practical Seminary named after A.Y. Petriuk, dedicated to the 70th anniversary of the National Aviation University's Flight Academy (Zaporizhia–Kropyvnytsky, 13–15 May 2021) / Kropyvnytsky: PPP "Exclusive-Sit", 2021. (in Ukrainian)