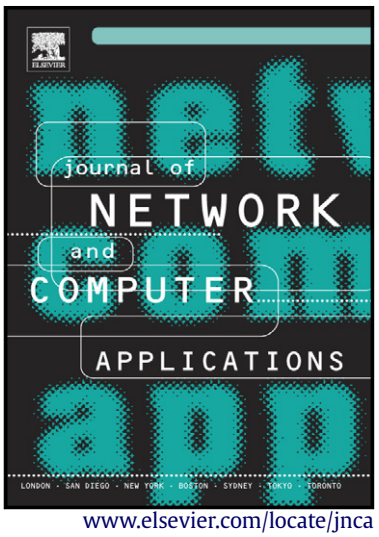
Author's Accepted Manuscript



www.elsevier.com/locate/jnca

BENFORD'S LAW BEHAVIOR OF INTERNET TRAFFIC

Laleh Arshadi, Amir Hossein Jahangir

PII: S1084-8045(13)00195-1
DOI: http://dx.doi.org/10.1016/j.jnca.2013.09.007
Reference: YJNCA1123

To appear in: *Journal of Network and Computer Applications*

Received date: 10 November 2012
Revised date: 20 June 2013
Accepted date: 18 September 2013

Cite this article as: Laleh Arshadi, Amir Hossein Jahangir, BENFORD'S LAW BEHAVIOR OF INTERNET TRAFFIC, *Journal of Network and Computer Applications,* http://dx.doi.org/10.1016/j.jnca.2013.09.007

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

BENFORD'S LAW BEHAVIOR OF INTERNET TRAFFIC

LALEH ARSHADI (corresponding author)
Computer Engineering Department,
Sharif University of Technology,
Azadi Avenue, Tehran, 1458889694, Iran.
Tel: +98-21-6616619
la_arshadi@ce.sharif.edu

AMIR HOSSEIN JAHANGIR
Computer Engineering Department,
Sharif University of Technology,
Azadi Avenue, Tehran, 1458889694, Iran.
Tel: +98-21-66166610
Fax:+98-21-66019246
jahangir@sharif.edu

Abstract — In this paper, we analyze the Internet traffic from a different point of view based on Benford's law, an empirical law describing the distribution of leading digits in a collection of numbers met in naturally occurring phenomena. We claim that Benford's law holds for the inter-arrival times of TCP flows in case of normal traffic. Consequently, any type of anomalies affecting TCP flows, including intentional intrusions or unintended faults and network failures in general, can be detected by investigating the first digit distributions of the inter-arrival times of TCP SYN packets. In this paper we apply our findings to the detection of intentional attacks, whereas other types of anomalies can be studied in future works. We support our claim with the related researches that indicate the TCP flow inter-arrival times can be modeled by Weibull distribution with shape parameter less than one, and show the relation between Weibull distributed data and Benford's law. Finally, we validate our findings on real traffic and achieve encouraging results.

Keywords: Computer network traffic analysis; Benford's law; Weibull distribution; Anomaly detection.

1. Introduction

For many years scientists believed that data network traffic can be modeled by Poisson processes, i.e. processes made up of packets arriving with exponential rate. Poisson processes are known to be one of the most important processes for modeling voice traffic and by extension, data network traffic [1]. However, Leland et al. showed that data communication traffic has self-similar properties which cannot be captured by Poisson models [2]. This led to the understanding that IP packet arrival process behaves in agreement with long range dependent and asymptotically self-similar processes.

On the other hand, the TCP protocol provides an abstraction of TCP flows (TCP connections or TCP sessions in other words), each consisting of a set of IP packets, starting with a SYN packet and ending with a FIN or RST packet[1]. These flows are generated by a large population of users, and consequently can be considered mutually independent. This leads to the classic modeling of Poisson inter-arrivals for TCP flows[2] (as presented in [3] and referred to in [4]). However, according to [5] TCP flow inter-arrival times are statistically better modeled by distributions with heavy tails, especially the Weibull distribution. In addition, based on some empirical studies, Cao et al conclude in [6] that TCP flow inter-arrival time distribution follows the Weibull distribution with a shape parameter smaller than 1 and that as the traffic intensity increases, the Weibull shape parameter gets close to 1, so the Weibull distribution degrades to the exponential distribution. Some other empirical studies have come to the same results ([7], [8], and [9]). In our previous work, we showed that despite the number of active nodes in a network, the inter-arrival times of TCP flows conform to the Weibull distribution in case of normal traffic, while irregularities in the traffic can cause deviations in the distribution of the inter-arrival times [10]. Therefore, anomalies, including intentional penetrations or unintended faults and network failures, can be detected by testing the

[1] Similarly, the concept of "flows" can also be defined for non-TCP traffic as a group of related IP packets that are common in specific fields and close in time. We do not consider this type of traffic in this paper.

[2] It is important to emphasize on the difference between packet inter-arrival times and TCP connection inter-arrival times.

Accepted manuscript

Download English Version:

https://daneshyari.com/en/article/6885131

Download Persian Version:

https://daneshyari.com/article/6885131

Daneshyari.com