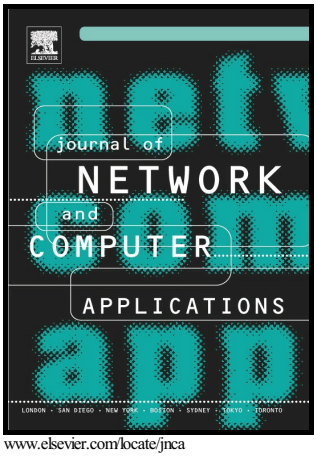
Frequent Traffic Flow Identification through Probabilistic Bloom Filter and its GPU-based Acceleration

Sisi Xiong, Yanjun Yao, Michael Berry, Hairong Qi, Qing Cao



www.elsevier.com/locate/jnca

PII: S1084-8045(17)30108-X
DOI: http://dx.doi.org/10.1016/j.jnca.2017.03.006
Reference: YJNCA1878

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

Received date: 13 September 2016
Revised date: 21 January 2017
Accepted date: 7 March 2017

Cite this article as: Sisi Xiong, Yanjun Yao, Michael Berry, Hairong Qi and Qing Cao, Frequent Traffic Flow Identification through Probabilistic Bloom Filter and its GPU-based Acceleration, *Journal of Network and Computer Applications* http://dx.doi.org/10.1016/j.jnca.2017.03.006

Frequent Traffic Flow Identification through Probabilistic Bloom Filter and its GPU-based Acceleration

Sisi Xiong[a,*], Yanjun Yao[a], Michael Berry[a], Hairong Qi[a], Qing Cao[a]

[a] *University of Tennessee, Knoxville, TN, USA*

Abstract

Traffic measurement and monitoring are critical for many network applications, for example, bandwidth management and detecting security threats such as DoS (Denial of Service) attacks. In these cases, traffic is usually modeled as a collection of flows. One central problem is to identify those frequent flows, which account for a large percentage of total traffic, or whose frequency exceeds the user-specified threshold. In this paper, we describe a novel data structure called the probabilistic Bloom filter (PBF), which extends the classical Bloom filter into probabilistic direction, so that it can effectively estimate flows' frequencies, and identify frequent flows. We analyze the performance, tradeoffs, and capacity of this data structure, and investigate how to calibrate this data structure's parameters. We further develop one extension of the PBF for dynamic data stream needs. We implement our PBF on GPUs to gain more time efficiency. By testing with real network traces collected from a backbone router, we demonstrate that our method can keep track of flows' frequencies with adjustable accuracy, so that frequent flows can be identified with constant computational time complexity and low memory overhead.

Keywords: Network Measurement, Traffic Analysis, Bloom Filter.

1. Introduction

Accurate traffic monitoring and measurements are crucial in order to manage today's complex internet backbones, both for short-term purposes such as security needs

*Corresponding author
Email address: sxiong@vols.utk.edu (Sisi Xiong)