

HTTP/2 traffic characterization and classification

Motivation

- Understanding modern web traffic
- Building traffic models for benchmarking and simulations
- Network management and provisioning
- Track the adoption of HTTP/2

Challenges

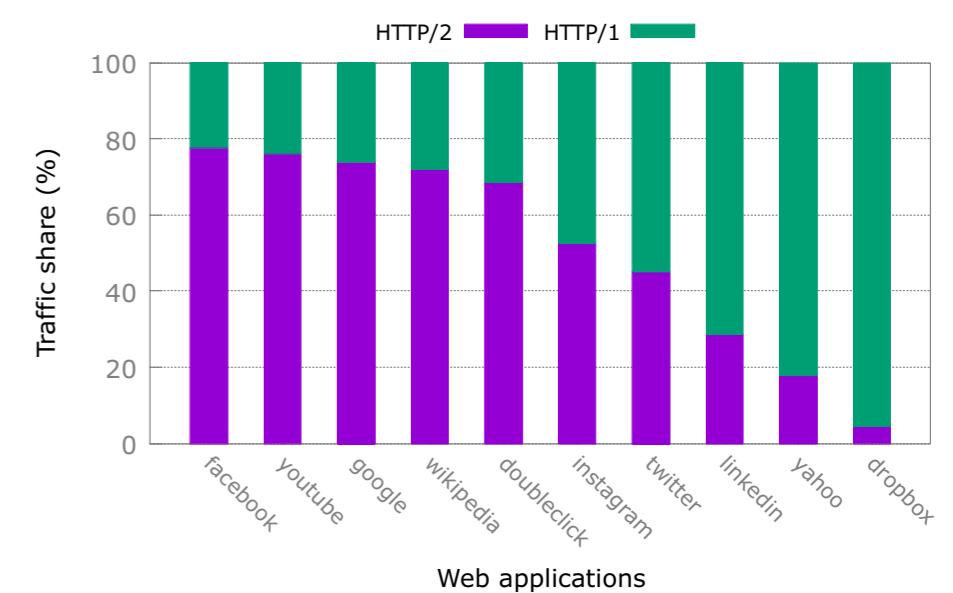
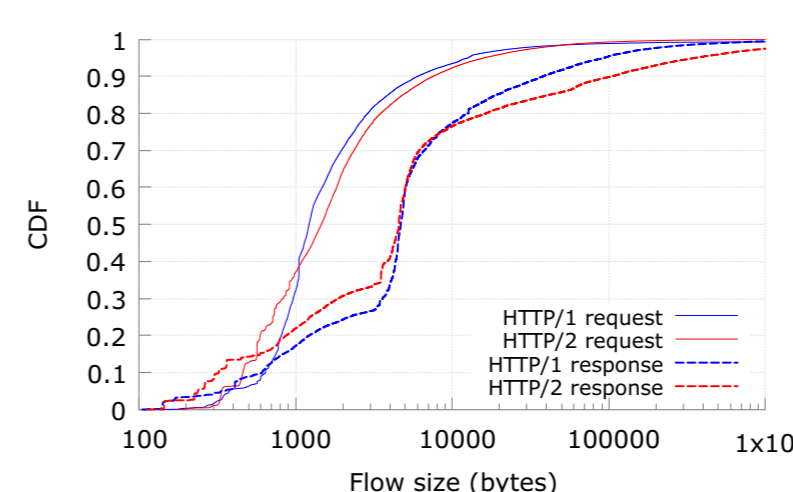
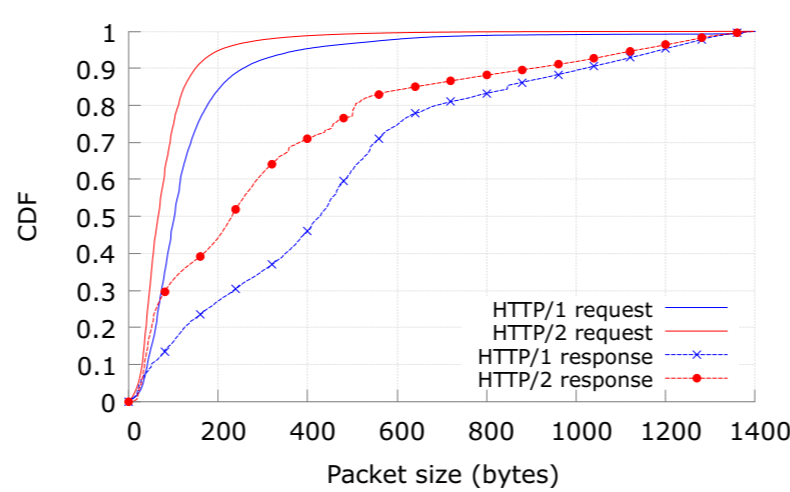
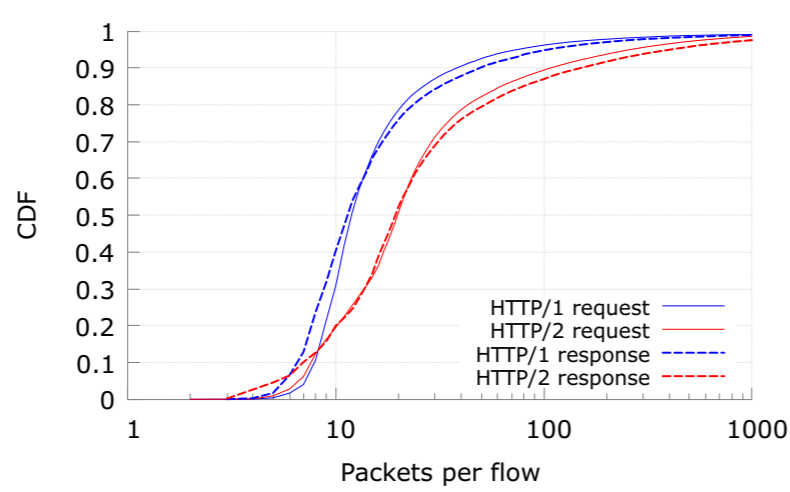
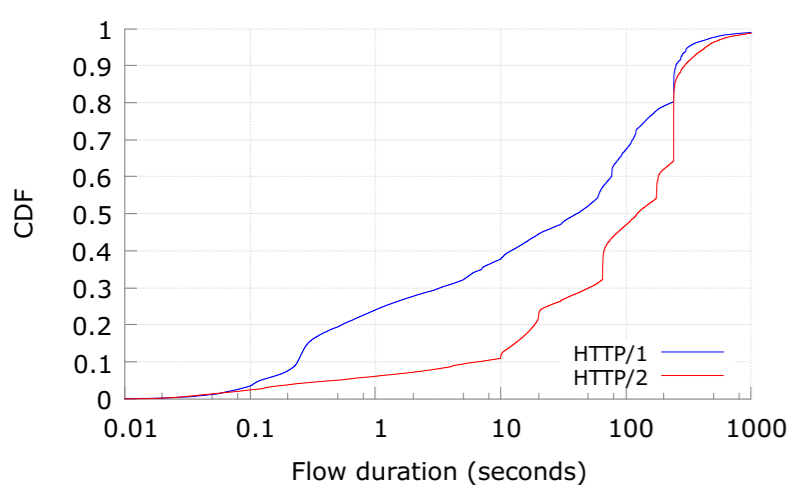
- DPI methods are expensive
- Prevalence of encryption

State-of-the-art

- Active probing

Proposed Solution

- Flow-based classification method
- Machine learning using passive measurements

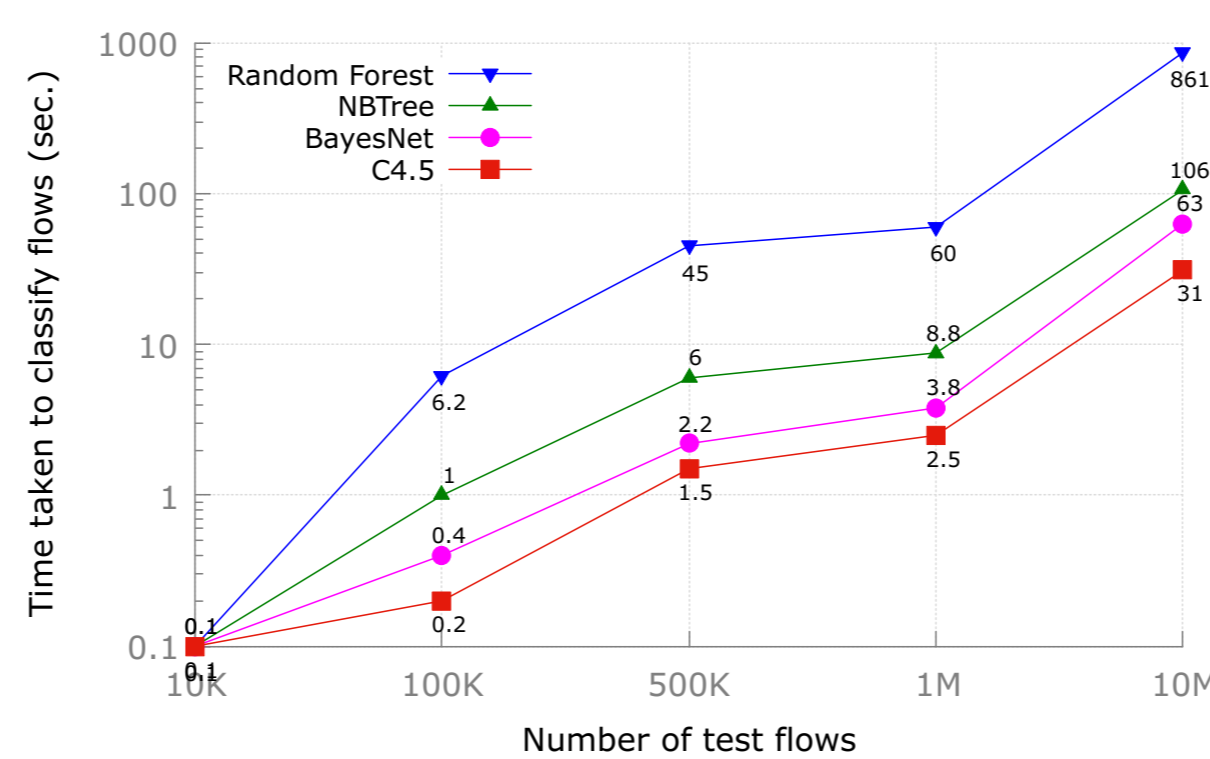
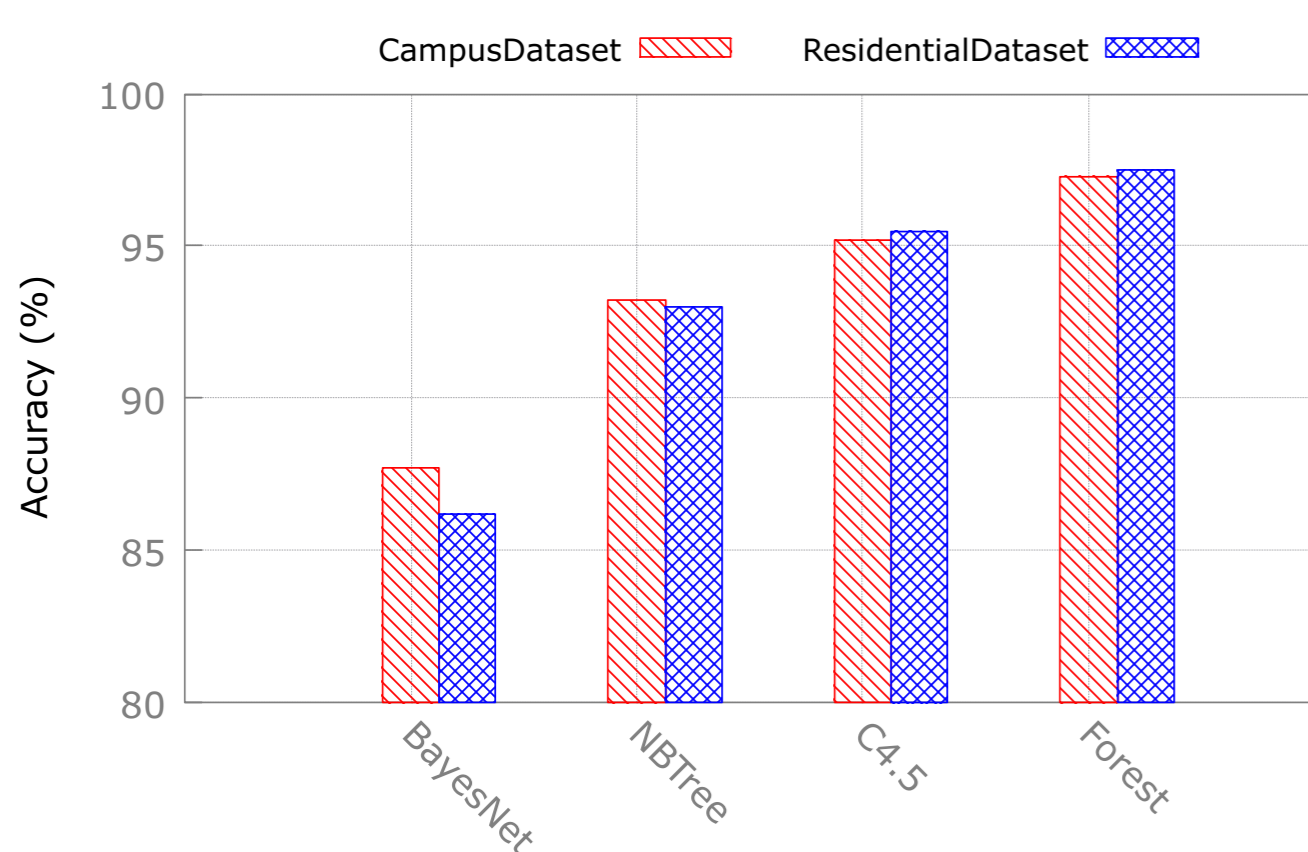


Findings

On average, HTTP/2 flows are more than 30% longer, carry more packets and have smaller packet size than HTTP/1. HTTP/2 is dominant in popular web applications

Flow-based classification

- Features: Number of client bytes, server bytes, client packets, server packets, average client bytes per packet, average server bytes per packet, flow duration.



Period(2016)	Test Instances	Accuracy(%)
Jun	800K	95.8
Jul	700K	94.8
Aug	400K	95.9
Sep	1.5M	95.5
Oct	2.1M	91
Nov	2.2M	90.4
Dec	1.4M	90.8

Findings

Decision trees are suited for classification of HTTP protocol. C4.5 has the fastest classification speed

C4.5 and Random Forest exhibit good temporal stability

HTTP/2 performance improvement using Multipath TCP

Problem

The design of a single HTTP/2 connection can lead to poor performance in certain network conditions

Objectives

- Experiment with parallel HTTP/2 connections natively in web browser.
- Explore MPTCP for parallel HTTP/2 subflows.
- Measure the effect of congestion control algorithms on performance.
- Test diverse real-world network conditions:
 - Congestion in ISP network
 - Random packet loss in WiFi
 - Bufferbloat in LTE network
- Devise application and network-aware path selection in MPTCP

