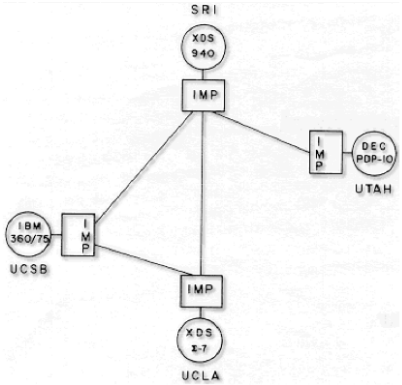
Lecture #14: wrap-up and net neutrality

WPI CS4516 Spring 2019 D term

Instructor: Lorenzo De Carli (ldecarli@wpi.edu)

A bit of history /1

- The (ancestor of) the Internet originated in the 1960s as a **military/research network**
  - **No significance** attached to uses **not related to military and research activities**
  - Among the Internet design priorities (Clark ‘88) the ability to **account for traffic** was **fairly low**
  - Not surprising: **who cares about billing if a war is going on?**



http://histoire.info.online.fr/net.html

The Internet, 1969

A bit of history /2

- The Internet started to allow non-government, non-academic users and businesses in the early ’90s
- As the Internet became a public service, **two different viewpoints** formed:
  - Internet as a **commercial/business tool**
  - Internet as tool for **knowledge sharing and societal improvement**
  - Both views coexist!

Commercial view of the Internet

- Internet is a **business tool**
- Users can purchase **goods and services** through it
- Business can **reach customers** and **advertise** through it
- **Price and functionality** of Internet service is determined by **market forces and incentives**

Examples of commercial services?

A different view

- Internet is a **public good**
- People use it to **reach and share knowledge**, **communicate** with each other, **form communities**, and enjoy **freedom of expression**
  - Also, **evade censorship**
- Internet is a public good and **access to it should be safeguarded** much like we safeguard access to drinkable water, education, etc.

The Internet has an intellectual history

- Intellectuals of the Internet age have proposed essays, arguments and manifestos on what the Internet is and what should be its governing philosophy
- John P. Barlow, ”A declaration of the Independence of Cyberspace”(1996): *<< We are creating a world that all may enter without privilege or prejudice accorded by race, economic power, military force, or station of birth. We are creating a world where anyone, anywhere may express his or her beliefs.>>*

Well, which view is right?

- Potentially, **both of them**!
- Despite claims to the contrary, the Internet **was not designed with a specific model in mind**
- The designers of the Internet were researchers and engineers, concerned with figuring out to build a large, resilient computer networks
  - Models of what the Internet should and should not be came much later!

So, in summary…

- What the Internet is and how it should be used is not a technical issue, but a **policy issue**
  - The public should decide what's the best and most useful approach to managing the Internet and controlling access to it
- That said…
  - Due to the design of the Internet and its protocols, **certain goals** (particularly those imposing very fine-grained restrictions on traffic) **are harder to implement than others**

That brings us to…

- **Net neutrality**!
- In a nutshell, “net neutrality” is the idea that **Internet service providers should not provide differential treatment to traffic originating from different sources**
- For example, this model forbids an ISP to **receive payment from a business** to **artificially prioritize their Internet service or slow down those of the competitors**

Activity

- Form **groups of ~5 people**
- Discuss and **come up with one reason in favor of net neutrality and one against**
- I'll go around the room and **ask each group to discuss what they came up with**

Arguments against NN

- The Internet is a **conduit for commercial services**
  - So **market forces** should be allowed **to decide the pricing and level of Internet connectivity**
- Certain type of services (e.g., streaming) **create disproportionate strain on ISP infrastructure**
  - So ISPs should be able to **charge those services a premium**

Arguments in favor of NN

- The Internet is an **infrastructure for accessing knowledge and public services**
- Giving ISPs the ability to limit and gate access may create **unfairness**
  - E.g., what if organizations/websites promoting particular points of view were given preferential connectivity treatment?
  - What if an ISP provides a cheap data plan which however allows access to a restricted set of websites?
    - (Basically, making access to information dependent on wealth)

(The debate is still ongoing, and probably will last for a long time)

Let's look at an analysis of real-world differential treatment

BingeOn Under the Microscope:
Understanding T-Mobile's Zero-Rating Implementation

Arash Molavi Kakhki
Northeastern University
arash@ccs.neu.edu

Fangfan Li
Northeastern University
li.fa@husky.neu.edu

David Choffnes
Northeastern University
choffnes@ccs.neu.edu

Alan Mislove
Northeastern University
amislove@ccs.neu.edu

Ethan Katz-Bassett
University of Southern California
ethan.kb@usc.edu

1. INTRODUCTION

The popularity of mobile devices for ubiquitous Internet access has led to exploding demand for relatively scarce cellular bandwidth. As a result, cellular operators increasingly manage their customers' demand using various techniques such as traffic shaping [10], transcoding [11], and *zero-rating* [9]. With zero-rating, Internet Service Providers (ISPs) do not charge users for traffic sent to/from certain services, often because those services agree to use limited bandwidth resources.

Perhaps the most well-publicized case in the U.S. is T-Mobile's "*BingeOn*" service, which zero-rates video streams from a large number of partner sites but limits the bandwidth they can use (i.e., rate limits). *BingeOn* works as follows [2,7]: First, a user must participate in *BingeOn*, which is enabled by default. If the user explicitly opts-out of *BingeOn*, traffic is not zero-rated or rate-limited by T-Mobile. If *BingeOn* is enabled, video streaming may be affected based on whether the content provider is part of *BingeOn* (and meets its technical requirements), opted out, or neither. Traffic for providers that opt-in is zero-rated and rate-limited, traffic for providers that explicitly opt-out is neither zero-rated nor rate-limited, and traffic for providers that neither opt-in nor out is rate-limited but not zero-rated. The provider opt-out option was not present in the original *BingeOn* program, but was present in March [2].

BingeOn has been highly controversial, due to concerns over network neutrality, user confusion, and technical downsides for users [12]. The resulting debate has led the EFF to call for T-Mobile to abandon *BingeOn*, and generated several responses from T-Mobile's CEO John Legere [1]. Importantly, there is little rigorous empirical data to inform the implications of zero-rating on network neutrality principles or customer-perceived Quality-of-Experience (QoE).

In this paper, we address this issue by conducting a study of T-Mobile's zero-rating policy and implementation to understand its implications for users and content providers in terms of data quota, performance, and QoE. We focus on T-Mobile and *BingeOn* due to their recent prominence, but we believe that lessons learned from this exercise will readily apply to other carriers using similar technologies to implement their policies.

We leverage our prior work on differentiation detection [11] to develop a suite of controlled tests. We deploy these tests to multiple *BingeOn* (and non-*BingeOn*) T-Mobile devices in the U.S., and correlate our tests with measurements of the billing records from these accounts. As a result, we are able to make significant headway in understanding *BingeOn*. Overall, we make five key contributions.

First, we characterize how *BingeOn* differentially impacts participating and non-participating providers. We determine that *BingeOn* is implemented solely by rate-limiting specific flows to 1.5 Mbps (consistent with the EFF's study [3]). Despite the fact that T-Mobile claims that *BingeOn* provides "optimized streaming", there is no transcoding or optimization taking place.

Second, we show how these differences translate to QoE, and how this QoE may impact users. We find that with *BingeOn* enabled, non-partner video flows see the same rate-limit (even though users are charged for these degraded flows). As a result, video services that do not support multiple quality levels could cause users to have poor video QoE.

Third, we reverse-engineer the classifier used for enabling zero-rating, and show that *BingeOn* as imple-

Internet-QoE, August 22-26 2016, Florianopolis , Brazil

ISBN 978-1-4503-4425-8/16/08…$15.00

DOI: http://dx.doi.org/10.1145/2940136.2940140

An analysis of *BingeOn*

- BingeOn is a **zero-rating** plan from mobile provider T-Mobile
- Zero-rating is the practice of **excluding traffic** from certain services (e.g., streaming) **from billing and monthly data caps**
  - In this case, in exchange for **accepting limits on the bandwidth** used by that traffic
- A form of differential treatment
  - Therefore, **not in accordance with the principle of net neutrality** (observation, not judgment ☺)

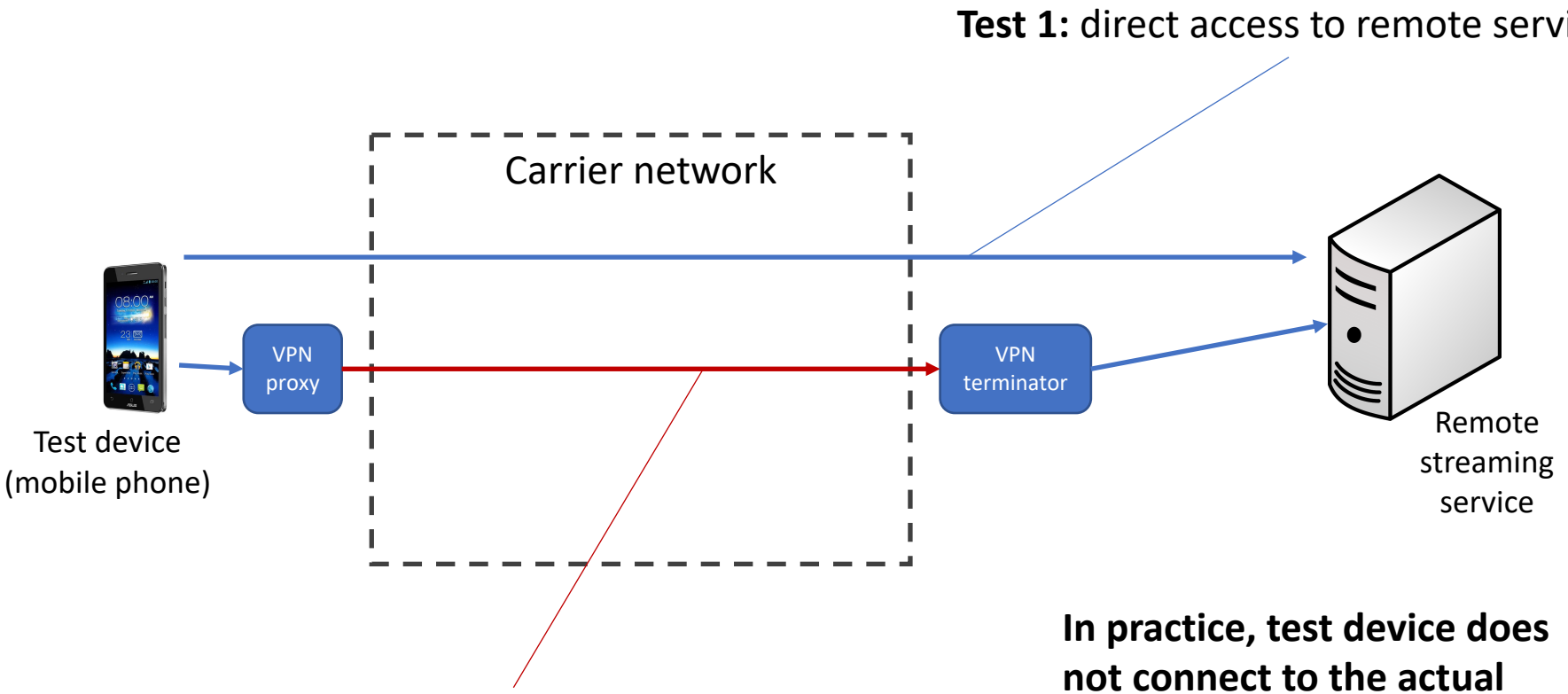
Specific promises

- Carrier states that:
  - Users enrolling in BingeOn accept that T-Mobile will "optimize quality for smartphone screens"…
  - ... and zero-rate affected streaming traffic
- Experimental questions:
  - **How does the carrier treat zero-rated traffic?**
  - **How does the carrier treats traffic not affected by BingeOn?**
    - Particularly, in regards to streaming services that are not on the list supported by BingeOn?

Experimental goals

- **Measure and compare** characteristics of traffic affected by BingeOn, and traffic not affected by it
- **Reverse-engineer** methodology carrier uses to identify traffic that should be affected by BingeOn, and identify **possible weaknesses**

Experimental setup

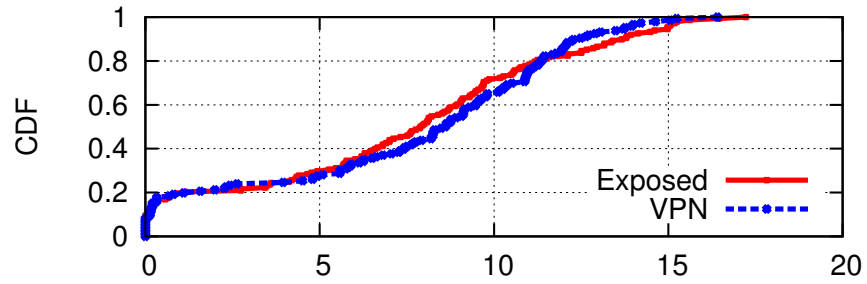


In practice, test device does not connect to the actual streaming service but to a replay server that plays back traffic captures

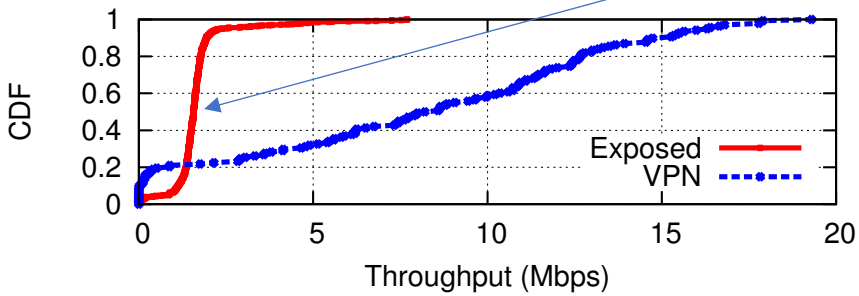
More on experimental setup

- Two different test devices are used, one with BingeOn enabled and one with BingeOn disabled
- **Advantage #1:** allows **identification** of how the same traffic is treated by carrier **when analysis is possible and when it is not**
- **Advantage #2:** allows **traffic modifications** to understand **how affected traffic is fingerprinted**

Main results /1



(a) *BingeOn* Disabled



(b) *BingeOn* Enabled

Netflix streaming traffic capped to ~1.5Mb/s

Main results /2

	***BingeOn* disabled**		***BingeOn* enabled**	
	Exposed	Random	Exposed	Random
Throughput (Mbps)	7.707 (0.404)	7.231 (1.238)	**1.540 (0.003)**	7.205 (2.264)
RTT (ms)	206 (33)	201 (31)	**126 (2)**	141 (48)
Retransmission rate (%)	0.099 (0.093)	0.121 (0.1217)	**18.158 (0.732)**	0.272 (0.135)

Quality	**Time to Start (sec)**		**video loaded in 1 min (%)**		**#rebuffers**		**Buffer/Play Time (%)**	
	BingeOn On	*BingeOn* Off	*BingeOn* On	*BingeOn* Off	*BingeOn* On	*BingeOn* Off	*BingeOn* On	*BingeOn* Off
auto*	1.54	1.32	4	3	0.0	0.0	2.48	2.15
hd1080	4.64	2.6	1	3	7.22	0.11	43.63	4.2
hd720	2.76	1.64	2	5	2.5	0.0	7.44	2.65
large	1.76	1.32	3	8	0.0	0.0	2.82	2.19
medium	1.58	1.08	4	12	0.0	0.0	2.56	1.76
small	1.26	1.0	7	15	0.0	0.0	2.07	1.63

Summary of main results

- “Optimized streaming” really means **capping streaming traffic** to 1.5 Mb/s with a **token bucket**
- **No optimization is performed** – the mechanism **relies on the streaming server adapting the quality** to the bandwidth limitations
- Streaming services with limited ability to adapt quality **may suffer from poor quality of experience**

About traffic identification

- One of the fundamental results of accountability not being a core Internet goal is that **there is not built-in mechanism for traffic identification and classification**
- But zero-rating requires distinguishing streaming traffic to apply bandwidth caps
- Experimental question: **how is streaming traffic fingerprinted by carrier?**

Analysis of traffic fingerprinting

Application	Detection criteria		How to evade detection?	
	BingeOn*/Music Freedom	**Video**	***BingeOn***	**Video**
Netflix	Specific GET arguments and the term "Netflix"	Same as *BingeOn*	Randomize GET argument and "Netflix" in response	Same as *BingeOn*
HBOGo	`Host` header ends with "hbogo.com"	`Content-Type` header (video/mp2t)	Randomize `Host` header	Randomize `Content-Type` header
ShowTime	`Host` header ends with "showvodhls.edgesuite.net**"	`Content-Type` header (video/mp2t)	Randomize `Host` header	Randomize `Content-Type` header
Hulu	`Host` header ends with "hulu.com"	`Content-Type` header (video/mp2t)	Randomize `Host` header	Randomize `Content-Type` header
Amazon Video	`Host` header ends with "amazonvod.loris.llnwd.net**"	`Content-Type` header (video/mp2t)	Randomize `Host` header	Randomize `Content-Type` header
Veoh	Not part of *BingeOn*	`Content-Type` header (video/mp4)	N/A	Randomize `Content-Type` header
Vimeo	Not part of *BingeOn*	Unknown***	N/A	N/A
YouTube (HTTP)	`Host` header ends with "googlevideo.com"	`Content-Type` header application/octet-stream	Randomize `Host` header	Randomize `Content-Type` header
YouTube (HTTPS)	Server name in the SNI ends with "googlevideo.com"	Same as *BingeOn*	Randomize "googlevideo" in the SNI	Same as *BingeOn*
Spotify*	"Spotify" in `Host` and `User-Agent` headers	N/A	Randomize the term "Spotify"	N/A
Pandora*	`Host` header ends with "p-cdn.com"	N/A	Randomize `Host` header	N/A

Analysis of traffic fingerprinting /2

- Identification based on **heuristic rules**
- Regular expression matching on hostname, content type
- **Fragile, easily circumvented**!
- Also, creates opportunity for **free-riding**

Possible zero-rating-enabled attack

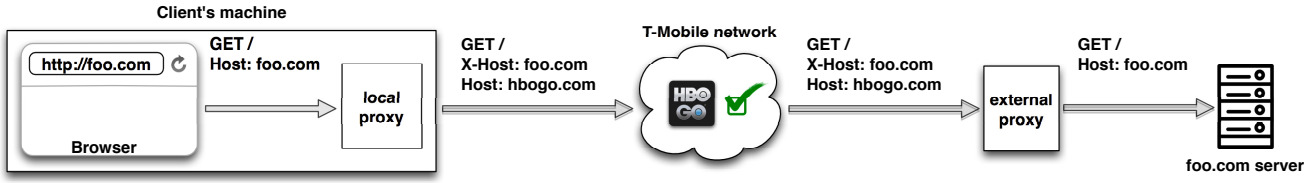


Figure 2: Subverting *BingeOn*: a) a local proxy copies the `Host` header into an optional parameter and overrides it with a *BingeOn*-enabled domain b) T-Mobile incorrectly classifies the traffic and zero-rates it c) an external proxy reverts the changes and forwards the request to destination.

Take-away points

- Whether net neutrality should be implemented is an issue of **policy** and **public opinion**
- Implementing differential treatment of traffic is **technically complex** and **inherently fragile**, regardless of whether net neutrality is considered positively or negatively
- Traffic identification mechanisms can be **evaded**, and keeping up with attacks may create a **significant overhead for carrier**

The End is Near