Characterizing Transnational Internet Performance and the Great Bottleneck of China

PENGXIONG ZHU, University of California, Riverside, USA
KEYU MAN, University of California, Riverside, USA
ZHONGJIE WANG, University of California, Riverside, USA
ZHIYUN QIAN, University of California, Riverside, USA
ROYA ENSAFI, University of Michigan, USA
J. ALEX HALDERMAN, University of Michigan, USA
HAIXIN DUAN, Tsinghua University, China

Transnational Internet performance is an important indication of a country's level of infrastructure investment, globalization, and openness. We conduct a large-scale measurement study of transnational Internet performance in and out of 29 countries and regions, and find six countries that have surprisingly low performance. Five of them are African countries and the last is mainland China, a significant outlier with major discrepancies between downstream and upstream performance. We then conduct a comprehensive investigation of the unusual transnational Internet performance of mainland China, which we refer to as the "Great Bottleneck of China". Our results show that this bottleneck is widespread, affecting 79% of the receiver–sender pairs we measured. More than 70% of the pairs suffer from extremely slow speed (less than 1 Mbps) for more than 5 hours every day. In most tests the bottleneck appeared to be located deep inside China, suggesting poor network infrastructure to handle transnational traffic. The phenomenon has far-reaching implications for Chinese users' browsing habits as well as for the ability of foreign Internet services to reach Chinese customers.

ACM Reference Format:
Pengxiong Zhu, Keyu Man, Zhongjie Wang, Zhiyun Qian, Roya Ensafi, J. Alex Halderman, and Haixin Duan. 2020. Characterizing Transnational Internet Performance and the Great Bottleneck of China. In *Proc. ACM Meas. Anal. Comput. Syst.*, Vol. 4, 1, Article 13 (March 2020). ACM, New York, NY. 23 pages. https://doi.org/10.1145/3379479

1 INTRODUCTION

The Internet has been a driving force behind the third wave of globalization, which began after the 2000s [20, 22]. It breaks barriers between countries and allows information to flow around the world at lightning speed and extremely low cost. The performance and reliability of Internet connections crossing national borders provides an indicator of the level of infrastructure investment, globalization, and openness of the countries. Yet, although researchers have developed techniques to measure the condition of the Internet from numerous perspectives (such as congestion [36, 41,

Authors' addresses: Pengxiong Zhu, pzhu011@ucr.edu, University of California, Riverside, USA; Keyu Man, kman001@ucr.edu, University of California, Riverside, USA; Zhongjie Wang, zwang048@ucr.edu, University of California, Riverside, USA; Zhiyun Qian, zhiyunq@cs.ucr.edu, University of California, Riverside, USA; Roya Ensafi, ensafi@umich.edu, University of Michigan, USA; J. Alex Halderman, jhalderm@eecs.umich.edu, University of Michigan, USA; Haixin Duan, duanhx@tsinghua.edu.cn, Tsinghua University, China.

2476-1249/2020/3-ART13 $15.00
https://doi.org/10.1145/3379479

53, 70], outages [44], attacks [49], etc.), little attention has been paid specifically to *transnational* Internet performance—how well traffic flows across the national borders.

In this paper, we conduct a large-scale measurement study of transnational Internet performance between 29 countries and regions, with wide geographic coverage and a variety of economic development levels. We find that transnational network condition across 23 countries and regions remains stable and exhibit high throughput, while six are unstable. Of the six, five are African countries that exhibit slow speeds on both upstream and downstream links. The last is mainland China, a significant outlier with major discrepancies between downstream and upstream performance. Its downstream speed is not only unstable but even worse than that of the African countries we measured in terms of mean and median throughput, while its upstream speed is comparably high and stable to the best 23 countries and regions. So unusual is this phenomenon compared to other countries that we deem it the "Great Bottleneck of China."

This motivates us to conduct a deeper investigation of the transnational Internet performance of China. Specifically, we aim to answer the following questions:

- How widespread are slow transnational speeds within mainland China?
- Where are the network performance bottlenecks that cause these slowdowns?
- What are the possible reasons for China's unique transnational Internet performance profile?

Our results show that slow transnational Internet performance occurs all over mainland China. Interestingly, it occurs only during certain times and forms a diurnal pattern that is congestion-like (irrespective of network protocol and content). Although not all measured transnational connections suffer from slow speed, 79% do, depending on the path traversed. In pinpointing where the losses occur, we find that they happen predominantly on routers owned or operated by Chinese ISPs.

As we do not have access to the internal decision-making processes that affect Chinese ISP infrastructure nor the Chinese government's policy directives regarding transnational Internet traffic, we can only speculate about the root causes of the slowdowns. However, our observations are compatible with poor national investments in network infrastructure, or with *intentional* underprovisioning in service of broader purposes. Regardless of the motivations, the phenomenon does have at least the following impacts:

(1) The poor transnational Internet performance effectively puts any foreign business that does not have a physical presence (i.e., servers) in China at a disadvantage—anecdotal evidence suggests that servers have to be hosted in China [59] to provide a good user experience.
(2) It leads to an Internet environment where users over time will become less interested in interacting with foreign sites, resulting in a form of self-imposed isolationism.
(3) Steering user traffic and data towards domestic sites will facilitate the surveillance and law enforcement.

In all, our work makes the following contributions:

- We conduct large-scale measurements of Internet performance on transnational network paths between 29 countries and regions with various levels of economic development.
- We design experiments to investigate the scale, frequency, and causes of anomalous slowdowns on network links between China and other countries.
- We observe widespread and persistent slowdowns across more than 400 pairs of mainland China and foreign nodes over the course of more than 53 days. The only exception is Hong Kong, which can serve as a performant proxy between mainland China and the rest of the world.

- We analyze the potential root causes of the slowdown, and found that the bottlenecks are often physically located in China, raising questions on the motivations of leaving the poor transnational networks as is in the modern Internet age.

Roadmap: §2 provides background on related work. §3 measures the performance of transnational upstream and downstream connections for 29 countries and regions. We perform detailed experiments to understand mainland China's transnational network performance in §4. We discuss potential root causes of the bottleneck in §5 and ethical considerations in §6, and we conclude in §7.

2 RELATED WORK

General Internet performance measurement. End-to-end Internet dynamics have been well studied since early stages of the network's history [21, 62, 66]. Thereafter, researchers have employed end-to-end probing techniques to characterize aspects of Internet performance, including capacity and bandwidth [46, 48, 58] and throughput [32, 67]. By itself, probing is limited to finding performance issues, rather than pinpointing the exact location of a problem, but hop-by-hop probing techniques based on TTL-manipulation can be employed to derive information about where on the network path issues occur [35, 45, 53]. In this work, we combine both end-to-end and hop-by-hop approaches to study global transnational Internet throughput as well as where and how performance degradation occurs.

Internet performance troubleshooting. Performance degradation can happen anywhere along the network path, and many researchers have measured and characterized performance problems in different portions of the Internet, e.g., edge networks, non-access networks, and interconnection points. To give a few examples: Dischinger et al. [37] and Maier et al. [54] measured the broadband characteristics in edge networks (i.e., the last mile of the Internet) in North America and Europe. Tools like Netalyzr [50] and HMN [64] are browser-based applets that focus on measuring end-user network performance. Akella et al. [14] measured bandwidth and latency in non-access networks in the U.S., on both intra-ISP and inter-ISP links, to identify bottlenecks.

Inter-domain performance. Many recent measurement studies have focused on inter-domain traffic and congestion [31, 36, 41, 52, 53, 57, 70]. Luckie et al. [53] and Dhamdhere et al. [36] tackle the challenges in measuring inter-AS congestion between ISPs in the U.S. using a latency-based approach to detect congestion, i.e., TSLP. Sundaresan et al. [70] discuss the limitations of solely using end-to-end throughput measurement to infer inter-domain congestion, caused by lack of insight into network topology and path information. In our work, we focus specifically on transnational Internet performance (instead of inter-domain performance within a country), aiming to not only pinpoint the location of packet losses but also understand the potential causes (e.g., congestion or deliberate traffic throttling).

Internet performance measurement platforms. There are many distributed measurement platform that facilitate measuring network performance from diverse networks in widespread geographic areas. PlanetLab [30] provides more than 1000 servers, mostly at institutions. Unfortunately, it has been shown that its servers are unreliable due to unpredictable load issues [65].CAIDA's Archipelago (Ark) [47] has more than 180 active monitors operated from both residential network and educational network in more than 50 countries, however, it has only one monitor within the Mainland China. Other platforms, such as SamKnows [10] and BISmark [69], distribute their own hardware probes and regularly perform latency and throughput tests. The RIPE Atlas platform [68] additionally provides researchers with an API to schedule specific connectivity and reachability tests. The M-Lab platform [51] has deployed distributed servers and collects data from volunteers who run its speed and diagnostic tests toward the servers. While these platforms are valuable, our

Continent	Countries/Regions (Vantage Points)
Asia	Hong Kong (t), India (a,A,d,t), Indonesia (a), Japan (a,A,t), Korea (A,t), mainland China (a2,t2), Malaysia (a), Singapore (a,A,d,t), Thailand (t), United Arab Emirates (a)
Africa	Egypt (c), Ghana (w), Kenya (w), Nigeria (w), South Africa (s)
Europe	France (A), Germany (a,A,d,t,r3), Ireland (A), Netherlands (d), Russia (k,t), Slovakia (r), Spain (r), Sweden (A), Switzerland (r2), United Kingdom (a,A,d,r)
North America	Canada (A,d,t), United States (a,A,d,t,r2)
South America	Brazil (A)
Oceania	Australia (a,A)

a: Alibaba, A: Amazon AWS, d: Digital Ocean, k: King Server, r: Residential, s: Safe Cloud, t: Tencent Cloud, w: Web4Africa. An optional suffix indicates the number of the nodes when it's more than 1.

Table 1. Locations of vantage points, with number of VPSes and residential devices in each country or region.

study requires performing continuous and customized tests from geographically diverse networks, hence we acquired many VPSes and residential vantage points within 26 countries.

Traffic differentiation. Due to profit motives and complex politics discussed by Claffy et al. [31], ISPs may intentionally differentiate traffic, causing performance degradation for some users. Internet censorship is another source of differential traffic treatment, resulting in prohibited content being blocked. This can be based on keywords [34] or on protocol types [40, 72]. However, we find that traffic congestion into China is not explained by differentiation based on content or protocols [38, 42, 60, 71, 75], indicating that it unlikely to be caused by active interference by the Great Firewall of China. This is interesting considering the extent of the Great Firewall of China's capabilities [55].

Nevertheless, persistent nation-wide Internet performance degradation in China could still be caused by deliberate differentiation based on domestic vs. transnational traffic, due to high international peering costs [39].

3 MEASURING GLOBAL TRANSNATIONAL THROUGHPUT

Accessing a distant server in a different country is almost certainly slower than accessing a domestic one, as the network latency increases with the distance and the network performance is determined by the weakest link (and the longer the path, the more likely a bottleneck link will be encountered). In this section, we examine the performance of both upstream and downstream connections across transnational networks among 29 countries and regions.

Experiment setup. We selected a global collection of vantage points in 29 countries and regions, using a combination of commercial VPSes and residential hosts provided by volunteers, as shown in Table 1. The list includes six continents and countries with various degrees of economic development, including ones with high GDP (e.g., the United States and China), high GDP per capita (e.g., Singapore and the United States) and ones with low GDP per capita (e.g., Kenya and Nigeria). For diversity within each country, we picked an average of two vantage points (and at most four) and ensured that they were either located in different cities or from different VPS providers. Overall, depending on the availability, the list of VPS providers we used includes Amazon AWS[3], Digital Ocean[6], Vultr[12], Alibaba Cloud[2], Tencent Cloud[11], Web4Africa[13], CityNet Host[5], SafeCloud[1], KingServer[7] (labeled for each vantage point in Table 1). A small subset of VPS providers, i.e., Amazon AWS, Alibaba and Tencent Cloud, also offer different tiers of network performance. However, as will be shown later in §4.2, they do not have an impact on the transnational network performance. In this experiment, all VPSes are 1 CPU and 1G memory,

except Sweden-AWS, which is 2 CPUs, 1G memory and network performance enhanced to up to 5 Gigabit. In total, we used 61 vantage points.

We used MaxMind[56] and our RTT-based geolocation validation method in §5 to verify the geolocation of the VPS nodes and did not find any inconsistency with their purported locations.

In order to measure throughput, we hosted an HTTP server on each VPS and used the other VPSes and the residential hosts as clients. Each client used `curl` to download data from every server for 20 seconds every 10 minutes in a round robin manner.

The web resource was a benign binary-format file with no sensitive keyword, and would be unlikely to trigger any censorship reactions. We made sure that the resource downloaded was sufficiently large so that it would not be completed within 20 seconds. During the download, we recorded a throughput data point every second. Since the goal of our experiments was to detect unstable and unexpectedly low performance, we configured `curl` to cap the client speeds at 4 Mbps to avoid using unnecessary network resources (note that the control is not perfect and small bursts may exceed the cap).

In total, we measured 2470 pairs of cross-country vantage points, 728 cross-country country-level pairs (we count receiver–sender and sender–receiver as two different pairs). The experiment ran from April 22 to April 27, 2019. Each pair was tested for at least three days. Additionally, we conducted an earlier smaller-scale pilot experiment over a shorter time period to make sure that the results of the full experiment were not anomalous.

Observation 1: Only African countries and mainland China suffer from unstable transnational network performance. We summarize the results by pairs of continents in Figure 1, which clearly illustrates the throughput differences between regions. Most continent pairs

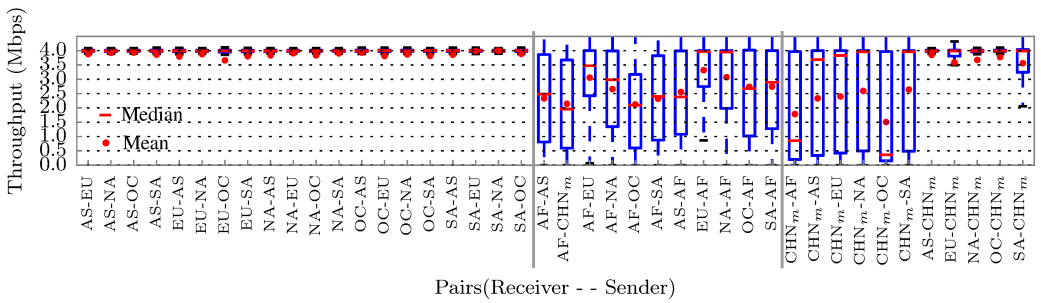
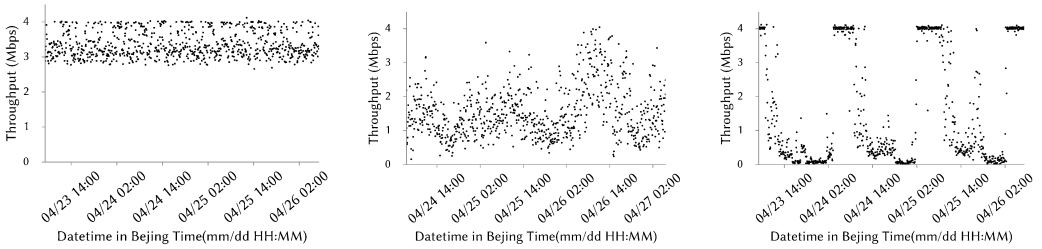


Fig. 1. Throughput of transnational links between continents (following the continent codes, e.g., NA = North America). Note that CHN$_m$ denotes mainland China, and AS denotes the rest of Asia.



(a) Lagos, Nigeria and Virginia, USA (b) Cairo, Egypt and Virginia, USA (c) Beijing, China and Virginia, USA

Fig. 2. Throughput patterns from April 23 to April 27, 2019 (Beijing Time) for three *receiver–sender* pairs (a single connection is used in each pair).

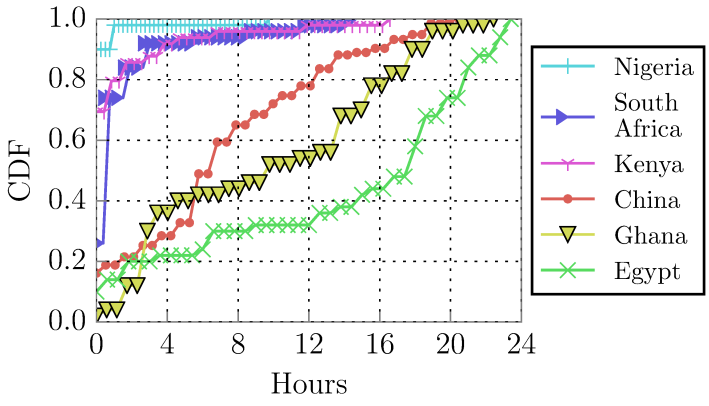


Fig. 3. CDF of slowdown hours per 24-hour period for transfers into African countries and mainland China.

have fairly stable and performant transnational networks. However, the five African countries and mainland China are obvious outliers. We isolate mainland China from the rest of the Asia in the figure, as it is the only outlier in that region.

There are several additional characteristics in our data worth noting. First, both the downstream and upstream performance of African countries' transnational networks showed significant variations. In contrast, mainland China only exhibited performance instability in the downstream direction, when data was entering China.

Second, the magnitude of the observed performance variations differs between African countries and mainland China. Overall, the African countries had better performance in terms of the standard deviation, average, and median throughput. For example, the performance of AF–EU (Africa–Europe) is noticeably better than CHN_m–AS (China mainland–Asia).

Finally, Hong Kong is an interesting data point (grouped with Asia in our figure). It is a special administrative region belonging to China, and yet we observed no discernible slowdown for its transnational networks.

Observation 2: Mainland China's transnational network performance has unique characteristics. The previous results led us to investigate Africa and China in more detail. The first question is whether all five African countries have similar results.

Slowdown window. To answer this, we define a simple metric *slowdown window* to quantify the duration where the performance suffers from significant slowdowns. Specifically, we use a sliding window of 20 minutes to identify the period during which the throughput was lower than 1 Mbps more than half the time. We then measure the total slowdown time per day for each country, aggregating over all of its corresponding senders.

Each of the five African countries exhibited a different transnational network performance profile. As shown in Figure 3, where we focus on the slowdown of downstream traffic, three of them (Nigeria, South Africa, and Kenya) rarely incurred slowdowns. On the other hand, Ghana and Egypt exhibited significant slowdowns—42% of Ghana's transnational connections and 64% of Egypt's incurred at least 12 hours of slowdown intervals.

Figure 3 also shows the downstream result for mainland China, which is somewhere in the middle compared to the two groups of African countries. For China, the majority of the transnational connections incurred between 5 and 12 hours of slowdown per day.

When we further inspect the results for upstream traffic, Ghana and Egypt also have significant slowdowns, whereas mainland China barely has any (See Figure 1). This may indicate that China's

network is better provisioned than Ghana's and Egypt's, or that its bandwidth demand is more highly asymmetric.

To gain a further understanding, in Figure 2 we graph the throughput observed by individual clients in Nigeria, Egypt, and mainland China accessing data from the same server in the U.S. Each client is located in the capital city, which ,we assume, is The client in Nigeria exhibits consistent and relatively good performance, whereas the client in Egypt has highly erratic performance with a much lower average.

The client in Beijing has the worst performance of all. Interestingly, it also has a much clearer diurnal pattern than in either of the other countries. Shortly after 8 a.m. local time every day, the throughput starts to dip and stays extremely low—well below 500 kbps most of the time. It recovers swiftly at night between 1 and 2 a.m.

Our observation of poor transnational network performance in China is intriguing, given the country is a technology powerhouse and attempts to maintain tight control of many aspects of the Internet (e.g., through censorship). However, so far, our measurements of global throughput included only a small number of vantage points in China. This prompted us to perform additional experiments to demystify China's network behavior.[1]

Verification. We cross verify our results with M-Lab's NDT tests[51], which collect the China's transnational link speed since the beginning of 2019. In 64% of the 75,464 tests (with each test lasting 9 to 60 seconds), the download speed was less than 500 kbps, which generally accords with our finding of broad and severe slowdowns.

4 CHINA'S TRANSNATIONAL NETWORK PERFORMANCE

Given the unusual transnational network performance of China[2], we decided to investigate it in more detail.

First of all, China disallowed foreign ISPs to operate within the country and only recently lifted the restriction [24, 33]. According to a 2019 report [16], over many years, no foreign carriers with a presence in China's borders have been observed, and the transnational connections points are always physically located outside of mainland China. This is unique compared to many other countries where foreign carriers are allowed to have points-of-presence in-country.

Furthermore , the transnational traffic converges on the three state-own carriers, China Telecom, China Unicom and China Mobile. The three carrier giants provide 98.5% of China's total transnational bandwidth, with China Telecom taking up the proportion of 58.5% [26]. China also has a unique peering structure for transnational traffic. Only a small number of ASes in China have peering relationship with foreign ASes. Among them, China Telecom and China Unicom have the most unique peerings to foreign AS, which is 46.6% and 17.3% of the total, respectively. The following ISPs are CERNET (6.8%), an educational network, and CNNIC (6.0%), an administrative agency whose network has limited bandwidth and do not have much real traffic [74]. In addition, the recent report [16] also claims that China is connected to the rest of the Internet primarily through a limited number of connection points (most traffic passes through the United States and western Europe). In China Telecom's official website [28, 29], we can see that the number of connection points is between 18 and 28.

In this section, we hope to shed light on the following questions:

Q1. How widespread is the phenomenon throughout China? Is it happening only in specific locations, or is it universal? Does a similar slowdown occur for domestic traffic?

[1]Note that we have actually conducted some preliminary experiments even earlier than what is reported in this section, which prompted us to start the China-specific measurements (which is why the start date of the some China-specific experiments may look earlier than the ones reported in this section).

[2]For the rest of this paper, we use "China" to refer specifically to mainland China.

Fig. 4. Geolocations of vantage points within China.

Q2. What are the performance characteristics (e.g., time-of-day effects)? What factors influence the performance (e.g., in which direction are packets lost, is it path-dependent)? Does the slowdown look irregular (e.g., any traffic differentiation)?

4.1 Data and Experiments

We aim to collect a range of data to help answer the questions. We begin by first describing the vantage points from which we can collect data.

Vantage Points. Table 2 shows the vantage points which we control to facilitate the measurement. On the sender side, we have 17 VPS nodes with 1 CPU and 1GB memory in 12 different countries and regions, including 5 nodes in the US, and one node in Hong Kong (their respective VPS providers are labeled accordingly).

Our receiver nodes consist of 9 VPS nodes with 1 CPU and 1GB memory, 6 residential nodes contributed by volunteers, 1 node in the educational network (CERNET) and 1 node in a large enterprise network. We include residential vantage points as they represent real-world users, and to avoid the potential bias of VPS networks which may have better quality (we label their ISPs in the table). In addition, we have added VPSes (co-located with the above) with enhanced networking (when available) to rule out the possible explanation of local network bottlenecks. Specifically, we have chosen AWS VPSes with 10 Gbps, Alibaba Cloud VPSes with 1 Gbps, and Tencent Cloud VPSes with packet processing rate of 300,000 pps. In total, our vantage points cover 14 Chinese cities in 14 provinces, providing a wide coverage of tier-1 (Beijing, Shanghai, Shenzhen), tier-2 (Tianjin, Nanjing, Hangzhou, Chengdu, Chongqing), and other lower-tier cities as shown in Figure 4.

Data. Overall, we collect the following data to answer the Q1 and Q2 mentioned earlier.

- Downstream throughput over space and time using the vantage points under our control (Q1 & Q2). We use this metric to gain a high-level picture of the slowdown phenomenon. We test both the domestic traffic throughput as well as the transnational as a comparison.
- Downstream throughput when accessing popular foreign websites from China (Q1). The measurement represents realistic scenarios for Chinese netizens.
- End-to-end loss rate and latency collected on both ends (Q2). This can tell which direction packets lost are. It can also generally help detect if the traffic slowdown is irregular (high packet losses should typically be correlated to increased latencies in the same direction)[61].
- Routing path (Q2). This can help us determine whether the slowdowns are tied to specific paths.
- Hop-by-hop loss rate and latency (Q2). This metric is useful to locate where the packet loss or delay happens.

Role	Type	Locations
Sender	VPS	Australia(A), Brazil(A), Canada(d), Germany(d), Hong Kong(t), India(d), Japan(A), Korea(A), Russia(K), Singapore(d), Sweden(A), United Kingdom(d), United States (California(d), New York(d), Ohio(A), Oregon(A), Virginia(A))
Receiver	Educational	Beijing(edu)
	Enterprise	Beijing(com)
	Residential	Beijing (CU), Harbin(CT), Nanjing(CU), Shijiazhuang(CU), Tianjin(CU), Xiamen(CT)
	VPS	Beijing(t), Chengdu(t), Chongqing(t), Hohhot(a), Hangzhou(a), Shanghai(t), Shenzhen(a), Qingdao(a), Zhangjiakou(a)

a: Alibaba, A: Amazon AWS, com: Enterprise network, d: Digital Ocean, edu: Educational network, k: King Server, t: Tencent Cloud, CT: China Telecom, CU: China Unicom. The optional suffix indicates the number of nodes when it's more than 1.

Table 2. Vantage points we used to measure China's transnational network performance.

- Downstream throughput under different conditions (Q2). If the throughput is lower for specific types of traffic, then it is clearly not congestion-induced slowdown.

The above data are collected through three sets of experiments which we detail below.

Experiment 1: End-to-end and Hop-by-hop Tests (Transnational and Domestic). In this experiment, we perform both transnational network performance test following the sender-receiver setup in Table 2 and domestic network performance test by having Chinese vantage points download data from each other. The domestic measurement serves as a baseline and contrasts with the slower transnational network performance we observe.

End-to-end loss rate and latency are computed by comparing the raw packet traces collected on both ends.

We collected hop-by-hop metrics using a custom tool based on `mtr` [23], a utility designed to perform simple hop-by-hop measurements. We enhanced the functionality of `mtr` substantially so that it can inject TCP packets into existing flows (same four tuples), allowing us to measure the network path as close as possible to the real one used in the download. The injected packets use sequence numbers that are slightly smaller than the current values (-10) to mimic valid retransmissions. According to studies on stateful firewalls and middleboxes [63, 72], no existing firewalls will intentionally drop such packets. As a sanity check, we did two experiments between two pairs of vantage points: SF-to-Singapore for 10 hours, and SF-to-Beijing for 2 hours and found that the end-to-end loss rate of such probe packets is 0.28%. Note that the test was conducted during no-slowdown time, to avoid packet losses due to other factors.

The hop-by-hop latency (i.e., round trip time) can be measured when the router replies with an ICMP time exceeded response. We ensured that the response matched our probe packets by checking the embedded TCP sequence numbers.

Finally, the hop-by-hop loss rate can be tricky to measure, because (1) some routers are known to perform ICMP rate limiting, and (2) ICMP responses could be lost on the return path, potentially causing inflated loss rates. Borrowing a strategy proposed in prior work [75], we send only one such probe packet per second, which was shown to be slow enough to avoid triggering such limits in the wild. In case ICMP rate limits are unknowingly triggered (e.g., due to excessive background traffic), we always use the end-to-end loss rate as an upper bound to evaluate the trustworthiness of the loss rate at any particular hop.

We conducted the experiment for four times, from Mar 27 2019 to Apr 1 2019, Apr 27 2019 to May 2 2019, Jul 28 2019 to Aug 7 2019, and Sep 4 2019 to Oct 3 2019, for a total of 53 days. In the

Rank	Domain	File Domain	File Type	File Size(MB)	Origin
1	Baidu.com	downpack.baidu.com	apk	8.1	Chinese
3	Qq.com	dldir1.qq.com	apk	97	Chinese
4	Sohu.com	3g.k.sohu.com	apk	17	Chinese
5	Taobao.com	download.alicdn.com	apk	103	Chinese
8	Jd.com	storage.360buyimg.com	apk	81	Chinese
9	Sina.com.cn	downapp.sina.cn	apk	30	Chinese
31	Bing.com	www.bing.com	json	3.79	US
48	Yandex.ru	an.yandex.ru	js	1.1	Russian
50	Github.com	codeload.github.com	zip	20	US
68	Microsoft.com	download.microsoft.com	exe	15	US
76	Apple.com	www.apple.com	mp4	58	US
93	Sciencedirect.com	holdings.sciencedirect.com	zip	11	US
123	Mail.ru	rfr.agent.mail.ru	exe	56	Russian
125	Nih.gov	obssr.od.nih.gov	pdf	2.5	US
146	Ebay.com	developer.ebay.com	zip	35	US
190	Springer.com	link.springer.com	pdf	7.6	US

Table 3. Top Alexa domains used to measure the impact of the Great Bottleneck of China on Chinese users.

first experiment, we pair each receiver with 16 senders in order to exhaust all receiver-sender pairs. To avoid ICMP rate limit, each receiver only pairs with 4 senders for a single day and rotates to the next 4 senders the next day. In the following experiments, we pair each receiver with 8 randomly selected and fixed senders, each sender is also paired with 8 fixed receivers to spread the load (and to avoid ICMP rate limit). Each receiver downloads a file (through HTTP) large enough so that the connection can last for 65 seconds, which allows sufficient data points to be collected by `mtr` (we will have roughly 65 data points per hop). Each sender runs `mtr` to each receiver in a round robin manner. From Sep 26 2019 to Oct 3rd, we added the pairs with network performance enhanced VPSes. To avoid interfering with existing experiment, no hop-by-hop measurement was performed involving these nodes.

Experiment 2: Top Alexa Website Tests. To measure the impact of the Great Bottleneck of China, on average Chinese users, we use top global Alexa websites [15]. After eliminating the websites are entirely blocked in China (e.g., Google, Facebook), we pick the top 14 unblocked websites, 8 of them are foreign and hosted outside of China and 6 are domestic. All receivers resolve the domain names of the websites locally, again to represent the realistic usage scenario. The complete list of websites is shown in Table 3.

To ensure we download resources that are sufficiently large so we can conclude with confidence about the throughput results, we manually selected resources that are embedded in the pages as listed in Table 3. These selected resources are mostly hosted on sub-domains to represent the realistic usage scenario (especially for foreign websites).

Interestingly, we find that not all foreign sites are hosted physically outside of China. For example, most Chinese clients resolved www.apple.com to IPs that are physically located in China (95.59%). www.bing.com and www.microsoft.com are similar (100%). We used the Maxmind geolocation database to identify the physical location of the server IPs. The results given by Maxmind are validated using our RTT based geolocation validation method in §5.

The experiment lasted for 60 day across all nodes in China from May 31, 2019 to July 31, 2019.

Experiment 3: Traffic Differentiation Detection. To ascertain whether the observed bottleneck is an artifact of traffic differentiation policies (related to censorship or not), we performed a number of A/B tests. Specifically, we vary the traffic as follows:

- Protocol. If it is censorship-related throttling, do HTTPS, Shadowsocks, or VPN traffic experience more slowdowns than other types of traffic (note our original tests used HTTP)?
- Packet type. Does the slowdown vary depending on whether the traffic is TCP, UDP, or ICMP?
- Speed. If traffic is throttled to certain speeds, does sending packets faster result in a higher loss rate? We vary the speed from 0.1 packet-per-second (pps) to 1000 pps and back to 0.1 pps.

In each case, we ran a dedicated experiment for a minimum of one day from May 14 2019 to May 17 2019 using the same vantage points laid out in Table 2.

4.2 Results

In this section, we present our analysis of the result data and key observations.

Observation 1: Slow speeds occur extensively for transnational traffic but not for *domestic traffic*. Our data shows slow transnational network performance at all nodes inside China—every single node experienced some slowdown when downloading data from one or more servers.

With 450 pairs of vantage points that we control and 208 pairs of receiver and top Alexa websites, we are able to paint a comprehensive picture. Figure 5, in the same manner as Figure 3, shows the CDF of the number of slowdown hours (following the definition of slowdown windows earlier in §3) during the four rounds of experiments (totaling 53 days as described in §4.1).

Note that for the fourth round of experiments, we do not see significant variations between the network-enhanced VPS and otherwise, and therefore we choose not to include them in the results (in fact we found that they experienced even slightly more slowdowns). Surprisingly, the top Alexa results experience comparable or even worse slowdowns (we include only the websites that are physically hosted outside of China). The results clearly show that the slowdown is persistent and consistent over time. Roughly 70% of the pairs have a slowdown period of 5 hours or longer.

We also break down the slowdowns by receivers in China in Figure 6, where we report the average number of slowdown hours per day (aggregated over all the senders). Interestingly, almost all cities suffer from significant slowdowns across the board, except a node in Beijing (educational network) and another in Tianjin (China Unicom). The Beijing node experienced practically no slowdown in the third round of experiment (Jul 28 to Aug 7), leading us to think that the educational network may have an overall better transnational network performance. However, the number of slowdown hours increased to 6 hours averaged over 30 days in the fourth round of experiment, nullifying that hypothesis. Similarly, the Tianjin node experienced only an hour of slowdown on

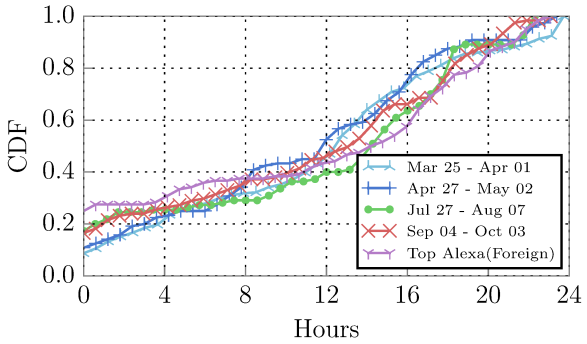


Fig. 5. CDF of hours of slowdown per day for connections from 18 Chinese to 17 foreign vantage points.

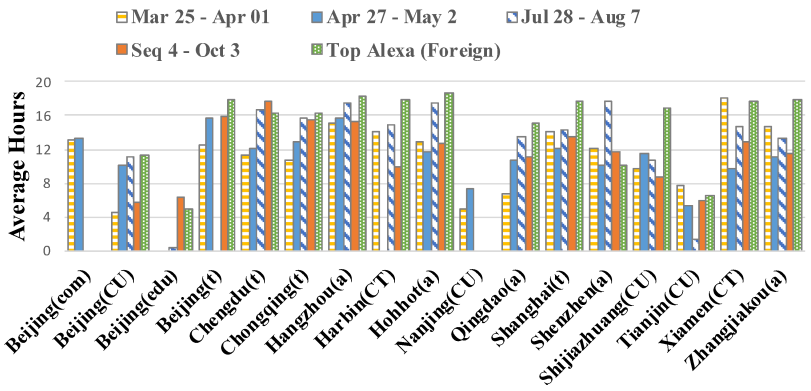


Fig. 6. Average number of slowdown hours per day by receiver. Missing bars indicate the absence of the receiver in the corresponding experiment. The ISPs labeled in the parentheses follow the convention in Table 2

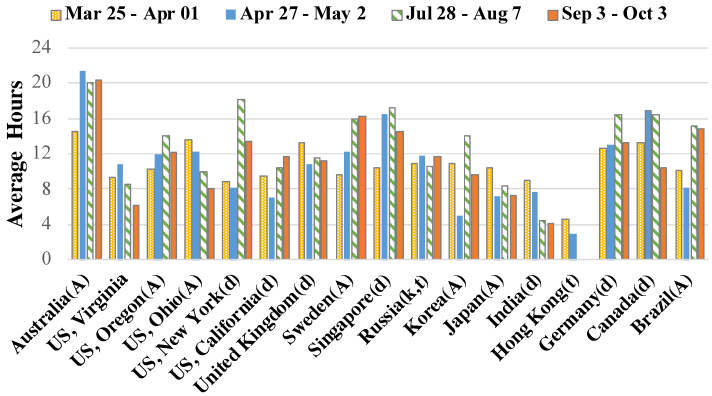


Fig. 7. Average number of slowdown hours per day by sender. Missing bars indicate the absence of the sender in the corresponding experiment.

average in the 2nd experiment but around 5-6 hours during the other two experiments. Otherwise, we see most receivers experience on average 5 to 17 hours of slowdown per day.

Conversely, in Figure 7 we show the average number of slowdown hours per receiver (aggregated over all of its corresponding senders). The slowdowns are evident for all the senders except Hong Kong, which is an outlier (with only 3 hours of slowdowns on average per day). As shown before, Hong Kong had no slowdowns when accessing data from the rest of the world (§3). Now we show that it also has much less frequent slowdowns when being accessed by nodes in mainland China. This makes Hong Kong an ideal proxy through which mainland China nodes can achieve excellent performance accessing the rest of the world. India, Japan, and Korea are the next best senders (relatively speaking), presumably because of their physical proximity to China, though they still suffer from 4 to 8 hours on average daily.

When we break down the sender by Alexa websites (again only the IPs that are physically outside of China according to Maxmind [56]), Figure 8 shows the aggregated number of slowdown hours experienced by all receivers per website. Generally the magnitude of slowdowns is high and comparable to the VPS experiments. Surprisingly, apple.com is clearly an exception. Upon checking the detailed results, this is because we found that most of the clients in China were

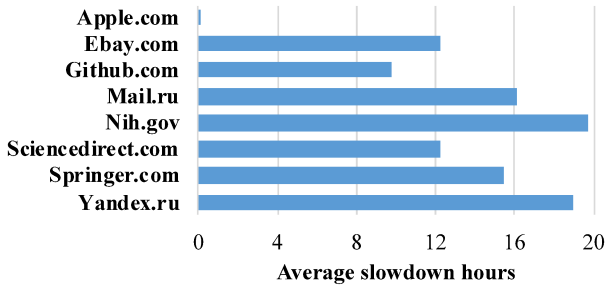


Fig. 8. Average slowdown hours per domain (with at least some resolved IPs that are physically outside of China).

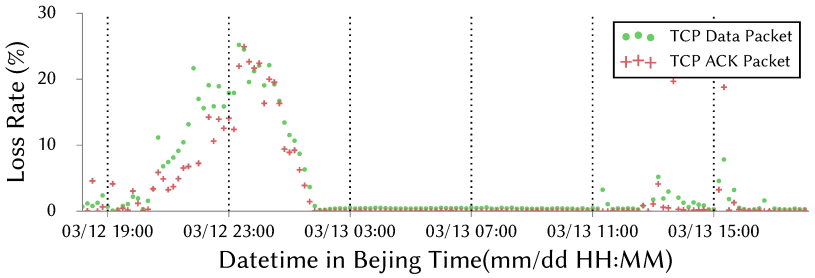
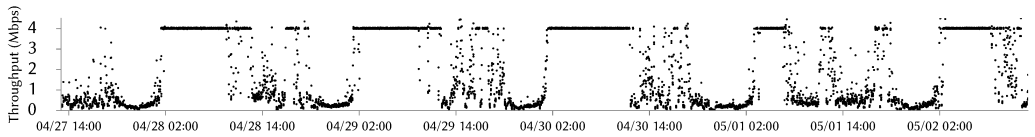


Fig. 9. End-to-end loss rate of TCP data packet of Shenzhen, China - San Francisco, USA and end-to-end loss rate of TCP ACK packet of San Francisco - Shenzhen, China.

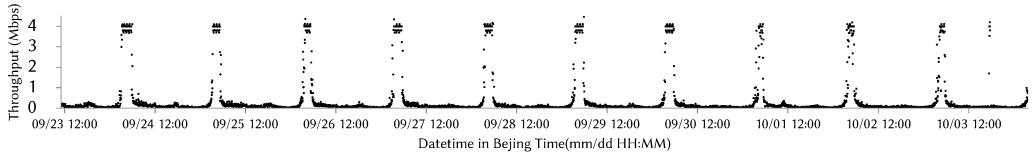
downloading data from the IPs located inside China. Together with bing.com and microsoft.com, they represent foreign websites that have a physical presence in China, all of which experienced little to no slowdowns. In addition, all domestic websites experienced virtually no slowdowns as well (which we omit in the figures). This is consistent with the end-to-end experiments where no sustained slowdowns (longer than 1 hour) were observed between any pairs of vantage points in China, indicating that the problem is indeed specific to transnational traffic.

The result clearly demonstrates that in order to do businesses with Chinese customers, companies would have no choice but to host their servers physically in China (like Apple and Microsoft did). As evident in our experiment, ebay.com had no chance of winning the competition against taobao.com with on average 12 hours of slowdown every day. On the flip side, github.com, nih.gov, and sciencedirect.com are popular websites in high demand by software developers and researchers in China. Unfortunately, the slowdowns would negatively hurt such Chinese users.

Observation 2: Packets are lost from only the direction going into China. When we inspect the end-to-end loss rate (from the raw packet traces), we find that the vast majority of packet losses occur in one direction only—from the foreign country entering China. This matches the observations we had in §3 where throughput is low only when data is sent from outside into China. The average loss rate over an entire day is typically in the range of 5% to 15%. The peak loss rate ranges from 10% to 50% (this can effectively render the network unusable). In addition, we find that the TCP ACK packets flowing in the reverse direction are almost never dropped. To confirm the results, we also repeated the experiment by reversing the sender and receiver and observe that data packets flowing from China to outside are also rarely dropped, but ACK packets flowing from outside into China experience similar high losses as incoming data packets, as illustrated in Figure 9 which captures the loss rates between Shenzhen and San Francisco as a representative



(a) Shanghai, China (VPS) from Oregon, USA, April 27 to May 2



(b) Beijing, China (residential) from Stockholm, Sweden, Sep 23 to Oct 03

Fig. 10. Diurnal patterns of transnational traffic measured at the receivers at two locations in China.

result. To confirm our intuition, we calculate the difference of ACK and data packets' loss rates over an entire day (where each data point is calculated over a 20-minute window). In the end, the mean of the difference is only 0.02% and variation is 4.58%. This indicates that the network does not treat data packets differently; rather, it is the nature that ACK packets are cumulative in TCP and therefore their losses do not impact the throughput as significantly.

Observation 3: The slowdown follows varied diurnal patterns. With regards to when and how often the slowdown occurs, we find that it typically occurs on a daily periodic basis, following certain diurnal patterns. We depict some examples in Figure 10.

Throughout our four rounds of experiments on end-to-end measurements, we observe that most locations exhibited consistent diurnal patterns each day. We sample two illustrative examples as shown in Figure 10a and Figure 10b for two different *receiver–sender* pairs. The results varied little even on holidays or weekends (May 1 and 2 are national holidays in China, and Oct 01 is the national day and China's 70th anniversary). In Figure 10a, the slowdown starts from roughly 11:30am, lasting throughout the day (with ups and downs) until 6:30pm and continues to experience consistent slowdowns through 1:30am (after midnight). In Figure 10a, the slowdown started early in the morning from 6am lasting all the way through 3:30am (after midnight) — lasting a total of more than 21 hours each day. Note that this phenomenon is consistent beyond the sampled 10-day period and is not affected by the important national day holidays or weekends.

This raises an obvious question about whether the slowdown is intentionally imposed at certain times of the day at fixed locations. To quantify the variations over time for all *receiver–sender* pairs, we calculate the standard deviations of the number of slowdown hours over the entire duration for each of the four rounds of end-to-end experiments (as well as the Alexa website experiment). As shown in Figure 11, roughly 80% to 95% of the *receiver–sender* pairs have a standard deviation of less than 3 hours, which we believe is a fairly consistent result. However, if throttling is activated and deactivated precisely according to time, then it is still difficult to explain the variation.

When we look into the cases where the standard deviation is larger than 3.5 hours. From the experiment of Apr 27 to May 02 and Sep 03 to Oct 03, the results show that the slowdown patterns do change over time — from longer slowdown in one day to much shorter ones in other days, and vice versa. We then attempt to correlate whether these slowdown pattern changes correlate with path changes. Interestingly, 11 of them did experience path changes together with the slowdown pattern changes, while 7 of them did not (and the remaining 3 had corrupted data and thus discarded).

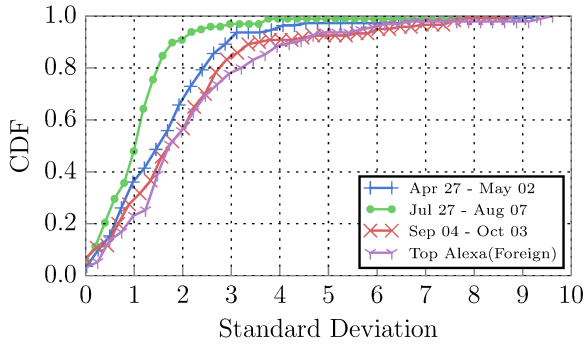


Fig. 11. CDF of standard deviation of slowdown hours per day.

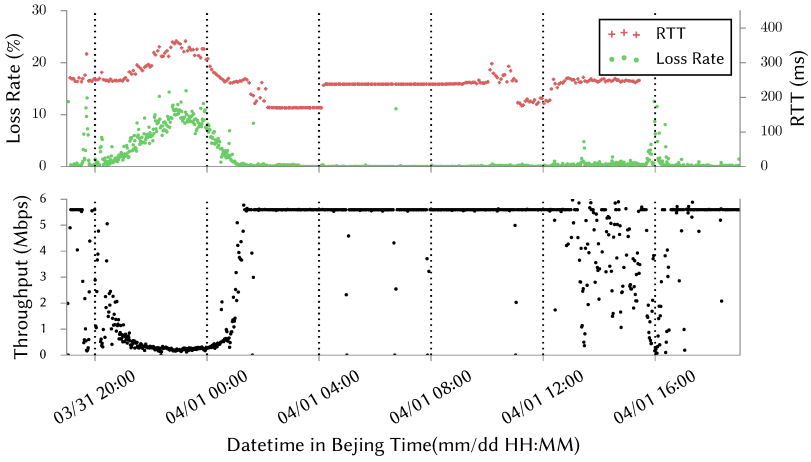


Fig. 12. Zhangjiakou, China (VPS) from London, United Kingdom, March 31

When looking at the path-changed cases in more details, we find that there are three cases involving inter-domain changes (e.g., one hop belonging to China Telecom that suddenly switches to China Unicom) while the others are intra-domain changes (e.g., load balancing).

Observation 4: No irregular traffic throttling/differentiation was observed.

From the end-to-end tests, we observed high correlation between low throughput, increased end-to-end loss rate, and increased latency. As an example, Figure 12 illustrates this for Zhangjiakou where slowdowns clearly occur when the loss rate and RTT both increase. This observation is generally in conformity with a normal congestion [61]. Of course, this is not to rule out the possibility that the congestion can be artificially imposed (e.g., by lowering the bandwidth).

Furthermore, from our A/B testing (varying the protocol, packet type, and speed), we did not detect any noticeable differences in loss rates. In particular, HTTP, HTTPS, VPN or Shadowsocks were equally affected by the poor transnational network performance. We can safely say there is no per-connection speed throttling because even if we send only one packet every 10 seconds in a flow it would experience the similar loss rates of insignificant differences.

5 ROOT CAUSE ANALYSIS OF SLOW SPEED

Clearly, the bulk of transnational traffic entering China experiences unacceptable slowdowns for a significant portion of time on a daily basis, yet we do not have the insight as to why. In this section, we try to assemble the facts from previous sections (and conduct more experiments if necessary) together to present plausible explanations of the phenomenon. We admit that the transnational network is a blackbox, and without insider knowledge, it is extremely challenging to reach a definitive conclusion, if at all possible.

To this end, we came to two classes of hypotheses as the potential root causes.

(1) Censorship-related. Intentional traffic throttling for censorship reasons, or the incapability of the Great Firewall (GFW) to inspect the large volume of traffic.
(2) Network resource provisioning. The network to handle transnational traffic maintained by Chinese ISPs has not been kept up with the demand due to policy (censorship or not), financial, or other reasons.

Note that these two hypotheses do not necessarily conflict with each other. In fact, censorship could be one of the reasons to drive the transnational network resources to be provisioned the way it is.

To help us validate the hypotheses, we find it helpful to know where the bottleneck was on the path. If the bottleneck co-locates with GFW on the same hop, then it might be censorship-related. If the bottleneck is at the border of China, it may point more to the limited network resources. We propose the following approach to locate the bottlenecks.

Bottleneck Detection. We first attempt to locate the bottleneck hop — the first hop that starts to drop packets along a path.

This is actually a challenging task, as the hop-by-hop loss rate results are noisy in practice for multiple reasons: (1) Even though we know that losses occur in one direction (from outside going into China) as shown previously, the ICMP TTL-expired packets can be "lost" due to routers' ICMP rate limit mechanism [75]. (2) Sometimes a hop can drop packets (due to ICMP rate limit policy or simply noise) and yet there exists a hop afterwards that exhibit 0% loss rate. (3) Sometimes the first lossy hop can oscillate (between a few consecutive hops), again due to noise.

To address (1), we make the observation that such ICMP rate-limited hops will generate high loss rates independent of whether it is currently in slowdown time (low throughput and high end-to-end loss); while a normal hop's loss rate will correlate with the end-to-end loss rate (they go up and down together). We therefore use the following heuristics to filter out ICMP rate limiting hops. For each hop (router) h appearing in the trace results of one *receiver–sender* pair, we have loss rate h_t and end-to-end loss rate e_t for any given time t in a day T. We then compute $\sum_{t \in T}(e_t - h_t)$ as s. We also define the variable H for time series h_t and E for time series e_t to compute their Pearson correlation coefficient ρ. If both $s < 0$ (indicating the hop's loss rate is actually higher overall than the end-to-end's loss rate which is suspicious) and $|\rho| < 0.35$ (indicating that these two time series have little to no correlation), h is likely a rate limiting hop. The threshold 0.35 is empirically determined based on manual inspection of loss patterns. Note that our heuristic is based on the assumption that the loss rate of ICMP rate-limited hops is higher than the end-to-end loss rate, so our heuristic may yield false negatives (missing a rate-limiting hop) if the loss rate of a router with ICMP rate limiting triggered is in fact less than the end-to-end loss rate. Nevertheless, even if we miss rate-limiting hops, the result is that we may mistakenly think such hops are bottleneck hops (which are before the actual bottlenecks). In practice, we found that the bottleneck hops are still mostly located within China even with the conservative estimates.

To address (2) and (3), we use majority voting to eliminate the uncommon cases. For example, if a hop has experienced losses less than 50% of the time, it will not be considered a real lossy hop. Similarly, we select the first lossy hop that most frequently appeared as the real bottleneck hop.

Location of the bottleneck. After we know the bottleneck hop, the next question is whether the hop is physically located outside of China or inside. To achieve this, we rely on a simple assumption — according to [27], all transnational links have latencies larger than or equal to 24ms (already the best ultra-low latency networks China Telecom offers) except Hong Kong (which is close enough to mainland China). Therefore, we conservatively consider a hop to be located outside of China, when the RTT observed from a foreign vantage point is less than 20ms (as all foreign nodes have a higher RTT going into mainland China). For the hops whose RTTs are larger than 20ms, we then probe them using Chinese vantage points, if any of them has an RTT of less than 20ms, we consider the hop to be located within mainland China or Hong Kong. We note that it is unnecessary to differentiate a hop between mainland China and Hong Kong because we now know the links between the two are sufficiently good and unlikely to be the bottleneck. If none of the above conditions are true, we then label the hop's location as unknown. Similarly, if a hop doesn't respond to ping, we would also have to label it as unknown. However, as we work with paths instead of individual hops, we are able to extrapolate the labels of hops that are either before a foreign hop (in which case they should also be foreign hops) and after a Chinese hop (in which case they should also be Chinese hops), assuming there is no loop in the routing path.

With the above method, we are able to confirm the location (either inside of outside of China) of bottleneck hops for 72.93% of the test pairs from Jul 27 to Aug 07, and 82.52% in the test from Seq 4 to Oct 3. We also consulted the Maxmind database for the same goal but it is giving us contradictory results almost 70% of the time (e.g., a path going through US, China, US, then China which clearly doesn't match the RTT profiles). This is perhaps not surprising as it is known that Maxmind is not very accurate when it comes to the location of *routers* [43].

In the end, we categorize the bottleneck hops into outside of China (exactly one hop before entering China, or further way), inside of China (exactly the first hop entering China, or further inside). As shown in Figure 13, there is a small percentage (2%) of hop-by-hop experiments that show losses occurring completely outside of China (two or more hops away from China).

For the cases where the bottleneck hop is the last hop before entering China (over 19% of cases), it could indicate the bottleneck is related to the transnational link where the router's input queue is saturated and therefore failing to push packets fast enough onto the output queue (too much demand on the transnational link).

For the cases where the bottleneck hop is the first hop entering China (8%), it is still likely the router before entering China that has an output queue filled too fast and therefore dropped (again too much demand on the transnational link)

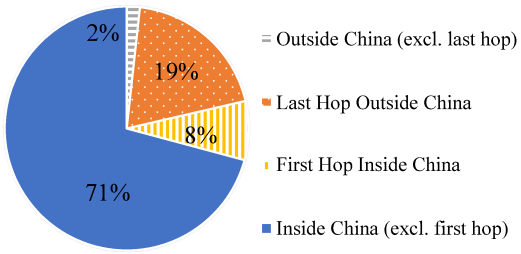


Fig. 13. The location of the bottleneck hop.

Finally, for the cases where the bottleneck hops are located deep in China (more than 70% of the time), they must not be related to the transnational links (e.g., submarine cables in the case of China and US) because the packets have already successfully reached the first router in China. This suggests that it is the Chinese ISP that is responsible for the lost packets.

Furthermore, we find that these bottleneck routers are almost always managed by the Chinese ISPs, even when the hops are physically located outside of China. We observed two common cases for the bottleneck routers located outside of China. First, cases where the IP address belongs to a Chinese AS (46.27%). Second, cases where the IP address belongs to a foreign AS, then we can look up the reverse DNS name—which is often expressive enough to explain both the owner of the IP address and the actual ISP responsible for managing the hop (38.54%). For example, Telia's IP `62.115.170.57` is located outside of China and its reverse DNS name is `chinaunicom-ic-341501-sjo-b21.c.telia.net`. After consulting a network operator, we know the naming convention of the reverse DNS name: `<Customername>-ic-<CircuitID>-<POP>-<router>.c.telia.net`. This effectively means that this is a peering link between China Unicom/Telia and China Unicom is the actual ISP managing the router on the hop. Interestingly, these two cases consist of 89.53% of the paths where the bottleneck hops are the last hop before entering China (47.58% are the first and 41.95% are the second). For the cases where the bottleneck is two or more hops away from China, the first case only takes up 31.46% while no second case was observed.

5.1 Hypothesis 1: Censorship

As China is known for its advanced censorship capabilities, it is natural to suspect it has something to do with the slowdown, especially when the slowdown patterns are so diurnally regular. This is in sharp contrast with other countries' transnational network performance, e.g., Figure 2a and Figure 2b. In addition, we observe that the changes in throughput (from slowdown to non-slowdown and vice versa) are overly sharp — country-level aggregate traffic changes are typically much more smooth. One possible explanation is that GFW is very sensitive in processing large volumes of transnational traffic and can become overwhelmed easily. However, one immediate counterargument is that GFW operates as an on-path system [72], which only processes copies of existing packets without the ability to discard existing packets. Evidently, prior work has shown that GFW fails to inject RST packets during busy hours while the packets containing sensitive keywords are still delivered successfully [34]. However, we are unable to rule out the possibility that GFW has evolved to acquire the capability to discard packets.

In fact, Great Canon (GC) [55] is such an in-path system. But it is known for intercepting a subset of traffic (based on protocol type) only. What's more, GC has been activated only twice in history (the last one in 2015 [55]). However, it might be the case that the in-path capability is re-purposed to perform general traffic throttling. If that is the case, they have done a good job because the throttling resembles natural congestion from the loss rate and latency point of view. The asymmetric performance between downstream and upstream traffic can be explained by the natural imbalance of transnational traffic (where the upstream traffic from China to outside is not significant enough to throttle).

To confirm whether GFW is now abused to slow down transnational traffic, we carefully designed a small experiment to locate the hops with GFW presence, and then try to match them with the bottleneck hops. Since the slowdown only happens in the direction from foreign into China, we use TTL-limited probes [17, 72, 74] to send probing packets from vantage points outside of China to those inside China, and record the first hop where GFW-forged RST packets are encountered. As one round of test, we perform the GFW hop measurement following the hop-by-hop packet loss measurement as described in §4. And we alternate between these two measurements which are

done close in time (in the hope that they will follow the same path). The whole experiment lasts for one day as an additional experiment. We employed 10 vantage points in China and 16 vantage points outside. Overall, we found in 34.45% of the cases, the GFW hops match the bottleneck hops. The low matching rate serves as a clue that the slowdown may not be caused by GFW. However, we only measured the GFW nodes injecting RST packets, there still could be other type of GFW devices, such as GC nodes.

5.2 Hypothesis 2: Network Resources Provisioned

According to a recent report by the China Academy of Information and Communications Technology [25], "China's international submarine cable development still lags the world's other major economies. The number of submarine cables in the U.S., Japan, the U.K., and Singapore is eight, two, five, and two times that of China, respectively, and the per capita bandwidth is 20, 10, 73, and 265 times that of China, respectively.

However, our bottleneck detection result in this section showed that this is less likely to be the primary cause of the bottleneck. In fact, most of the packet losses were observed after the traffic enters China. On the other hand, we have never observed persistent bottlenecks in domestic-to-domestic Chinese traffic, which would suggest that the bottleneck is outside the country. It is unclear to us why the network infrastructure within China is so poor compared to the transnational links (e.g., submarine cables between China and the US) which are extremely expensive to build.

A more plausible reason is financial related. In the early years, Chinese ISPs do not have a good reputation of making international peering easy [73], primarily because they wanted to grow their own transit business and make themselves to be top tier ISPs, since they control traffic flows in and out of China. According to a recent report by China Academy of Information and Communications Technology [33], all the three state-own ISPs have set up a premium transnational network (primarily for business uses) to maximize their profit. For example, China Telecom Global's official website[4] explicitly claims four tiers of services to connect to Chinese users. (1) China Access, (2) ChinaNet Paid-Peer, (3) Global Transit (GT), (4) Global Internet Access (GIA). Basically, the first three share the same point-of-presence or international gateway and therefore similar potential bottleneck, while Global Internet Access has a different dedicated CN2 international gateway.

To verify the existence of tiered services, we found an exotic VPS provider [19] reselling these tiers of networks through hosting, and they provide test IPs that allow customers to check the loss rate and latency to these nodes from within China [19].

We performed a day-long ping test (against the 7 test IPs) from five Chinese nodes, and did find that GIA gives the best performance. All five nodes experienced on average 3% loss rate, with the maximum loss rate of 7% observed in a node in Shenzhen. GT offers the second best network performance with an average loss rate of 4% (and a maximum loss rate of 14% observed in the same Shenzhen node). Finally, China Paid-Peer and China Access offer an average loss rate of 5% (and a maximum of 15% observed in Shenzhen as well).

This illustrates the severity of transnational bottlenecks by China Telecom, even when considering the top tier of its service. Unfortunately, we do not have enough information to accurately determine which paths traversed which tier of services in our earlier experiments. However, the average daily loss rates of our earlier experiments are roughly on par, starting from 5%, all the way to 15% in some receiver–sender pairs.

From the hosting company, we know that the GIA network include specific IP ranges `59.43.x.x` [18]. Unfortunately, from cross checking the paths we collected from foreign VPS nodes, we are unable to find any that traverses the GIA network, which explains why the slowdown was so evident.

The hosting company also offered one test IP in Hong Kong connected to mainland China through the major Hong Kong telecommunication provider, HKT, which is marketed as the best performing network to mainland China. We also tested this IP through one day of ping experiments. Indeed, the loss rate ranged from only 0.1–1%. This supports our observation earlier that Hong Kong (in §3 is the best proxy to reach mainland China.

In [39], the authors points out the cost of peering directly with China Telecom is expensive. Buying access from a US ISP that peers with China Telecom is way cheaper, which will most likely go through the lowest tier.

If the financial motivation is indeed the main factor, then clearly the Chinese ISPs are not yet successful in attracting big foreign companies to pay for the higher tiers of transnational links (we saw most Alexa websites still have poor network performance). In addition, we also tested a few major Chinese applications and websites from foreign vantage points. The idea is that if users outside of China want to access these services (we identify servers that are physically located in China), then they would also need to suffer from slowdowns. This is especially problematic if the services offer real-time video or voice streaming services such as QQ and Wechat. In both cases, we found that packets from foreign countries actually went through Hong Kong's major ISP PCCW [9]. This prompted us to test the foreign-equivalent version, Skype, and found that packets also go through Hong Kong. Finally, we surveyed a list of popular VPN providers specifically targeting Chinese users, and all of them have nodes in Hong Kong, giving further evidence that Hong Kong is a performant proxy. The list includes ExpressVPN, NordVPN VyprVPN, PureVPN, Surfshark VPN, VPN.ac, NeVPN, IronSocket Network, Buffered VPN, Astrill VPN, PrivateVPN, SwitchVPN, TunnelBear, Windscribe, and Netease UU. Interestingly, when we checked the Chinese Top Alexa websites in Table 3, we find that none of them went through the best GIA network of China Telecom's. This is likely be cause the GIA network is still not as performant compared to ISPs in Hong Kong (3% loss rate vs. 0.1 to 1% in our tests) and therefore those who really need performance would prefer Hong Kong ISPs.

Summary. All in all, we believe the slowdown would either (1) come from some sort of government policy — including but not limited to, discouraging Chinese users to access foreign services, setting a higher bar for foreign companies to do business with Chinese users, or (2) driven by financial motivations by Chinese ISPs. At the end of the day, we admit that this analysis is largely a best effort and we believe pinpointing the root causes further would require potentially insider knowledge about the government policy and ISP's inner-workings.

6 ETHICAL CONSIDERATIONS

Our investigation used 77 vantage points in different geographic regions, some from volunteers and some from VPS providers. In all cases, we directly communicated with the volunteers or VPS providers and used the author's normal email address. We did not collect any personally identifiable information from network traffic, nor did we attempt to access sensitive sites. We provided detailed and accessible instructions to ensure that our data collection was inline with volunteer expectations. We also minimized the potential performance burden on our volunteers by rate-limiting our measurements to make sure our tests did not use more than 4 Mbps of their bandwidth.

7 CONCLUSION

In this paper, we examined transnational network performance across multiple countries and conducted an in-depth investigation of the Great Bottleneck of China. Although we are not the first to recognize that foreign download speeds in China are poor, our work illuminates the dynamics and potential causes of China's unique transnational bottleneck. China's anomalously slow inbound

network performance appears to be the result of congestion, not only at the international border but also within China. It severely affects users in a large variety of geographic locations across the country, and results in a more isolated Chinese Internet.

ACKNOWLEDGMENTS

We would like to thank our sheperd, Michael Sirivianos, for his guidance on improving our work, as well as the anonymous reviewers for their insightful comments that helped improve the quality of the paper. We also want to thank anonymous volunteers, users from NANOG. Shitong Zhu and Vagelis Papalexakis for their assistance and support. This work was in part supported in part by the National Science Foundation (CNS-1518888).

REFERENCES

[1] 2017. SafeCloud. https://www.safecloudonline.com.
[2] 2019. Alibaba Cloud. https://us.alibabacloud.com/.
[3] 2019. Amazon AWS. https://aws.amazon.com.
[4] 2019. China Telecom Global. https://www.chinatelecomglobal.com.
[5] 2019. CityNet Cloud. https://citynet.net.
[6] 2019. Digital Ocean. https://www.digitalocean.com/.
[7] 2019. King Servers. https://www.king-servers.com/.
[8] 2019. NetRadar. https://www.netradar.com/.
[9] 2019. Pacific Century CyberWorks (PCCW). https://www.pccwglobal.com/.
[10] 2019. SamsKnow. https://samknows.com.
[11] 2019. Tecent Cloud. https://cloud.tencent.com/.
[12] 2019. Vultr. https://www.vultr.com/.
[13] 2019. Web4Africa. https://web4africa.com/.
[14] Aditya Akella, Srinivasan Seshan, and Anees Shaikh. 2003. An Empirical Evaluation of Wide-area Internet Bottlenecks. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC)*.
[15] Alexa Internet, Inc. 2019. Alexa Top 500 Sites. https://www.alexa.com/topsites.
[16] Dave Allen. 2019. Analysis by Oracle Internet Intelligence Highlights China's Unique Approach to Connecting to the Global Internet. https://blogs.oracle.com/internetintelligence/analysis-by-oracle-internet-intelligence-highlights-china%e2%80%99s-unique-approach-to-connecting-to-the-global-internet.
[17] Anonymous. 2012. The Collateral Damage of Internet Censorship by DNS Injection. *ACM SIGCOMM Computer Communication Review* (2012). https://doi.org/10.1145/2317307.2317311.
[18] BandwagonHost. 2018. DC1/DC2/DC4/QNET/CN2/Direct Route via China Telecom and China Unicom and Their Relationships. https://www.bandwagonhost.net/858.html.
[19] BandwagonHost. 2019. BandwagonHost Data Center Test IPs. https://www.bandwagonhost.net/test-ip.
[20] Luc Berlin. 2011. The Internet and Globalization: Ten Tips to Building an Effective Digital Strategy for Global Success. Retrieved March 23, 2019 from https://gbr.pepperdine.edu/2011/12/the-internet-and-globalization-ten-tips-to-building-an-effective-digital-strategy-for-global-success/
[21] Jean-Chrysostome Bolot. 1993. Characterizing end-to-end packet delay and loss in the internet. *Journal of High Speed Networks* 2, 3 (1993), 305–323.
[22] Artur Borcuch, Magdalena Piłat-Borcuch, and Urszula Świerczyńska-Kaczor. 2014. The Influence of the Internet on globalization process. *Journal of Economics and Business Research* (2014).
[23] BitWizard B.V. 2019. *MTR*. http://www.bitwizard.nl/mtr.
[24] CGTN. 2019. BT becomes first foreign telecoms firm to secure Chinese license. http://www.chinadaily.com.cn/a/201901/29/WS5c4fbdfca3106c65c34e70b2.html.
[25] China Academy of Information and Communications Technology. 2018. White Paper on China International Optical Cable Interconnection. http://www.caict.ac.cn/english/yjcg/bps/201808/P020180829385778461678.pdf.
[26] China Internet Network Information Centre. 2018. Statistical Report on Internet Development in China. https://cnnic.com.cn/IDR/ReportDownloads/201807/P020180711391069195909.pdf.
[27] China Telecom Americas. 2019. Ultra-Low Latency Connectivity for Financial & Capital Markets. https://www.ctamericas.com/wp-content/uploads/2019/07/China-Telecom-Americas-Ultra-Low-Latency-Financial-Services-Presentation.pdf.
[28] China Telecom Global. 2019. China Access. https://www.chinatelecomglobal.com/products/ChinaAccess/.
[29] China Telecom Global. 2019. GIA. https://www.chinatelecomglobal.com/products/GIA/.

[30] Brent Chun, David Culler, Timothy Roscoe, Andy Bavier, Larry Peterson, Mike Wawrzoniak, and Mic Bowman. 2003. Planetlab: an overlay testbed for broad-coverage services. *ACM SIGCOMM Computer Communication Review* (2003).
[31] KC Claffy, David D Clark, Steven Bauer, and Amogh D Dhamdhere. 2016. Policy Challenges in Mapping Internet Interdomain Congestion. *Policy Challenges in Mapping Internet Interdomain Congestion (August 24, 2016). TPRC* (2016).
[32] Collin Anderson 2015. *New Opportunities for Test Deployment and Continued Analysis of Interconnection Performance.* https://www.measurementlab.net/blog/interconnection_and_measurement_update.
[33] State Council. 2016. Telecommunications Regulations of the People's Republic of China, Order No.666. http://www.china.org.cn/business/laws_regulations/2010-01/20/content_19273945.htm.
[34] Jedidiah R Crandall, Daniel Zinn, Michael Byrd, Earl T Barr, and Rich East. 2007. ConceptDoppler: A weather tracker for Internet censorship. In *ACM Conference on Computer and Communications Security.*
[35] Leiwen Deng and Aleksandar Kuzmanovic. 2008. Monitoring persistently congested Internet links. In *IEEE International Conference on Network Protocols.*
[36] Amogh Dhamdhere, David D Clark, Alexander Gamero-Garrido, Matthew Luckie, Ricky KP Mok, Gautam Akiwate, Kabir Gogia, Vaibhav Bajpai, Alex C Snoeren, and Kc Claffy. 2018. Inferring Persistent Interdomain Congestion. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM).*
[37] Marcel Dischinger, Andreas Haeberlen, Krishna P. Gummadi, and Stefan Saroiu. 2007. Characterizing Residential Broadband Networks. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).* http://doi.acm.org/10.1145/1298306.1298313.
[38] Marcel Dischinger, Alan Mislove, Andreas Haeberlen, and Krishna P Gummadi. [n.d.]. Detecting bittorrent blocking. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).*
[39] DrPeering. 2014. What up w/Peering in China? http://drpeering.net/AskDrPeering/blog/articles/Ask_DrPeering/Entries/2012/4/25_What_up_w_Peering_in_China.html.
[40] Roya Ensafi, David Fifield, Philipp Winter, Nick Feamster, Nicholas Weaver, and Vern Paxson. 2015. Examining how the Great Firewall discovers hidden circumvention servers. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).*
[41] Rodérick Fanou, Francisco Valera, and Amogh Dhamdhere. 2017. Investigating the Causes of Congestion on the African IXP substrate. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).*
[42] Tobias Flach, Pavlos Papageorge, Andreas Terzis, Luis Pedrosa, Yuchung Cheng, Tayeb Karim, Ethan Katz-Bassett, and Ramesh Govindan. [n.d.]. An Internet-Wide Analysis of Traffic Policing. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM).*
[43] Manaf Gharaibeh, Anant Shah, Bradley Huffaker, Han Zhang, Roya Ensafi, and Christos Papadopoulos. 2017. A look at router geolocation in public and commercial databases. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).* ACM.
[44] Vasileios Giotsas, Christoph Dietzel, Georgios Smaragdakis, Anja Feldmann, Arthur Berger, and Emile Aben. 2017. Detecting peering infrastructure outages in the wild. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM).*
[45] Ningning Hu, Li Erran Li, Zhuoqing Morley Mao, Peter Steenkiste, and Jia Wang. 2004. Locating Internet bottlenecks: Algorithms, measurements, and implications. In *ACM SIGCOMM Computer Communication Review.* ACM.
[46] Ningning Hu and Peter Steenkiste. 2003. Evaluation and characterization of available bandwidth probing techniques. *IEEE journal on Selected Areas in Communications* (2003).
[47] Young Hyun and Kc Claffy. 2015. Archipelago measurement infrastructure. http://www.caida.org/projects/ark/.
[48] Manish Jain and Constantinos Dovrolis. 2003. End-to-end available bandwidth: measurement methodology, dynamics, and relation with TCP throughput. *IEEE/ACM Transactions on Networking (TON)* (2003).
[49] Mattijs Jonker, Aiko Pras, Alberto Dainotti, and Anna Sperotto. 2018. A First Joint Look at DoS Attacks and BGP Blackholing in the Wild. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).*
[50] Christian Kreibich, Nicholas Weaver, Boris Nechaev, and Vern Paxson. 2010. Netalyzr: illuminating the edge network. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).* ACM, 246–259.
[51] Measurement Lab. 2019. Network Diagnostic Test (NDT). http://www.measurementlab.net/tools/ndt.
[52] Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide, and Farnam Jahanian. 2010. Internet interdomain traffic. *ACM SIGCOMM Computer Communication Review* 40, 4 (2010), 75–86.
[53] Matthew Luckie, Amogh Dhamdhere, David Clark, Bradley Huffaker, et al. 2014. Challenges in Inferring Internet Interdomain Congestion. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).*
[54] Gregor Maier, Anja Feldmann, Vern Paxson, and Mark Allman. 2009. On dominant characteristics of residential broadband internet traffic. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC).* ACM, 90–102.
[55] Bill Marczak, Nicholas Weaver, Jakub Dalek, Roya Ensafi, David Fifield, Sarah McKune, Arn Rey, John Scott-Railton, Ron Deibert, and Vern Paxson. 2015. An analysis of china's "great cannon". In {*USENIX*} *Workshop on Free and Open Communications on the Internet.*

[56] MaxMind Inc 2016. *Maxmind GeoLite databases.* Retrieved May 4, 2019 from http://dev.maxmind.com/geoip/legacy/geolite/
[57] Measurement Lab Consortium and others. 2014. ISP interconnection and its impact on consumer internet performance. https://www.measurementlab.net/publications/isp-interconnection-impact.pdf.
[58] Bob Melander, Mats Bjorkman, and Per Gunningberg. 2000. A new end-to-end probing and analysis method for estimating bandwidth bottlenecks. In *Globecom'00-IEEE. Global Telecommunications Conference. Conference Record (Cat. No. 00CH37137)*, Vol. 1. IEEE, 415–420.
[59] mlytics. 2018. Why your website is slow in China (and how to fix it). https://mlytics.com/blog/why-your-website-is-slow-in-china/.
[60] Arash Molavi Kakhki, Abbas Razaghpanah, Anke Li, Hyungjoon Koo, Rajesh Golani, David Choffnes, Phillipa Gill, and Alan Mislove. 2015. Identifying traffic differentiation in mobile networks. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC)*. ACM, 239–251.
[61] SB Moon, J Kurose, P Skelly, and D Towsley. 1998. Correlation of Packet Delay and Loss in the Internet TITLE2. (1998).
[62] Vern Edward Paxson. 1998. *Measurements and Analysis of End-to-end Internet Dynamics.* Ph.D. Dissertation. Berkeley, CA, USA. UMI Order No. GAX98-03325.
[63] Zhiyun Qian and Z Morley Mao. 2012. Off-path TCP sequence number inference attack-how firewall middleboxes reduce security. In *IEEE Symposium on Security and Privacy*. IEEE, 347–361.
[64] Alan Ritacco, Craig Wills, and Mark Claypool. 2009. How's My Network? A Java Approach to Home Network Measurement. In *Proceedings of International Conference on Computer Communications and Networks*. IEEE, 1–7.
[65] Marcelo Santos, Stenio Fernandes, and Carlos Kamienski. 2014. Conducting network research in large-scale platforms: Avoiding Pitfalls in PlanetLab. In *IEEE International Conference on Advanced Information Networking and Applications*.
[66] Stefan Savage, Andy Collins, Eric Hoffman, John Snell, and Thomas Anderson. 1999. The end-to-end effects of Internet path selection. In *ACM SIGCOMM Computer Communication Review*, Vol. 29. ACM, 289–299.
[67] Speedtest 2012. *Speedtest by Ookla - The Global Broadband Speed Test.* Retrieved April 20, 2019 from https://www.speedtest.net/
[68] RN Staff. 2015. RIPE Atlas: A global internet measurement network. *Internet Protocol Journal* (2015).
[69] Srikanth Sundaresan, Sam Burnett, Nick Feamster, and Walter de Donato. 2014. BISmark: A Testbed for Deploying Measurements and Applications in Broadband Access Networks. In *USENIX Annual Technical Conference*.
[70] Srikanth Sundaresan, Xiaohong Deng, Yun Feng, Danny Lee, and Amogh Dhamdhere. 2017. Challenges in Inferring Internet Congestion Using Throughput Measurements. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC)*.
[71] Mukarram Bin Tariq, Murtaza Motiwala, Nick Feamster, and Mostafa Ammar. 2009. Detecting network neutrality violations with causal inference. In *Proceedings of the international conference on Emerging networking experiments and technologies*. ACM, 289–300.
[72] Zhongjie Wang, Yue Cao, Zhiyun Qian, Chengyu Song, and Srikanth V Krishnamurthy. 2017. Your state is not mine: a closer look at evading stateful internet censorship. In *ACM SIGCOMM Conference on Internet Measurement Conference (IMC)*. ACM, 114–127.
[73] Winther, Mark. 2006. Tier-1 ISPs: What They Are and Why They Are Important, NTT White Paper. https://www.us.ntt.net/downloads/papers/IDC_Tier1_ISPs.pdf.
[74] Xueyang Xu, Z Morley Mao, and J Alex Halderman. 2011. Internet censorship in China: Where does the filtering occur?. In *International Conference on Passive and Active Network Measurement*. Springer, 133–142.
[75] Ying Zhang, Zhuoqing Morley Mao, and Ming Zhang. 2009. Detecting traffic differentiation in backbone ISPs with NetPolice. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*.

Received October 2019; revised December 2019; accepted January 2020