Investigating Large Scale HTTPS Interception in Kazakhstan

Ram Sundara Raman
University of Michigan
ramaks@umich.edu

Leonid Evdokimov
Independent
leon@darkk.net.ru

Eric Wurstrow
University of Colorado Boulder
ewust@colorado.edu

J. Alex Halderman
University of Michigan
jhalderm@umich.edu

Roya Ensafi
University of Michigan
ensafi@umich.edu

ABSTRACT

Increased adoption of HTTPS has created a largely encrypted web, but these security gains are on a collision course with governments that desire visibility into and control over user communications. Last year, the government of Kazakhstan conducted an unprecedented large-scale HTTPS interception attack by forcing users to trust a custom root certificate. We were able to detect the interception and monitor its scale and evolution using measurements from in-country vantage points and remote measurement techniques. We find that the attack targeted connections to 37 unique domains, with a focus on social media and communication services, suggesting a surveillance motive, and that it affected a large fraction of connections passing through the country's largest ISP, Kazakhtelecom. Our continuous real-time measurements indicated that the interception system was shut down after being intermittently active for 21 days. Subsequently, supported by our findings, two major browsers (Mozilla Firefox and Google Chrome) completely blocked the use of Kazakhstan's custom root. However, the incident sets a dangerous precedent, not only for Kazakhstan but for other countries that may seek to circumvent encryption online.

CCS CONCEPTS

• **General and reference** → **Measurement**; • **Security and privacy** → **Security protocols**; **Web protocol security**; • **Social and professional topics** → **Governmental surveillance**; *Technology and censorship.*

KEYWORDS

HTTPS, Interception, Kazakhstan, MitM, Certificates

ACM Reference Format:
Ram Sundara Raman, Leonid Evdokimov, Eric Wurstrow, J. Alex Halderman, and Roya Ensafi. 2020. Investigating Large Scale HTTPS Interception in Kazakhstan. In *ACM Internet Measurement Conference (IMC '20), October 27–29, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3419394.3423665

IMC '20, October 27–29, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-8138-3/20/10.
https://doi.org/10.1145/3419394.3423665

1 INTRODUCTION

HTTPS protects billions of users: 74–95% of daily web traffic is now encrypted, providing much-needed privacy and security [1, 23]. At the same time, deep packet inspection technologies that inspect HTTPS connections have also advanced [29, 46, 50]. Although enterprise-level interception is common despite being fraught with security issues [17, 40], large-scale interception at the ISP or national level has been limited, even as increased adoption of HTTPS challenges mass surveillance and keyword-based censorship [5, 19].

Last year, in an unprecedented move, the Republic of Kazakhstan became the first country to deploy carrier-grade HTTPS interception on a national level. Starting on July 17, 2019,[1] Kazakhstan launched an HTTPS interception man-in-the-middle (MitM) attack, after instructing citizens to install a government-issued root certificate on all devices and in every browser for "security" purposes [8]. This interception, which the government described as a "pilot", covered large portions of the country's network and was active intermittently until being shut down on August 7, 2019.

While the attack was going on, we worked to understand the interception technique, measure its scope, and identify its likely targets. We first detected the interception using data from Hyperquack, a recently introduced remote technique for detecting keyword-based network interference [50]. Beginning on July 20, Hyperquack's HTTPS measurements to some (but not all) of 82 available vantage points in Kazakhstan detected rogue untrusted certificates for popular destinations such as `google.com` and `facebook.com`. The certificates were issued by the Kazakh government's custom root CA, `Qaznet Trust Network`. We later confirmed these detections with direct measurements from local virtual private servers (VPSes) and 52 in-country RIPE Atlas nodes.

We determined that the interception system would trigger on TLS connections passing through certain network locations in Kazakhstan when a targeted domain was present in the TLS Server Name Indication (SNI) header. This allowed us to probe it using connections originating from outside or inside the country destined for any HTTPS server in Kazakhstan. We used this behavior to perform comprehensive measurements from North America and two Kazakh VPSes to 6,736 TLS hosts in different parts of the country, setting the SNI header to popular domains. We also performed TTL-limited measurements to discover the location in the network where the interception was occurring. To track the attack over time, we performed measurements continuously until well after the interception system was shut down.

[1] Dates and times are in East Kazakhstan Time (UTC+6), except where noted.

Our findings show that only a fraction of the Internet traffic inside the country was subject to interception (around 7–24% of the 6,736 TLS hosts measured were affected), and that the path to all of the servers affected by the interception passed through two sets of specific hops in AS9198 (Kazakhtelecom). Of the Alexa Top 10,000 domains [4], 37 triggered interception. The majority were media and communication sites, 20 were Google services and 7 were services affiliated with Facebook. The set of targets suggests that the government's actions were motivated by surveillance, rather than increased security as was officially claimed. From our longitudinal measurements, we observed the interception being turned on and off intermittently and observed varying scale of interception, suggesting that the interception system was still being tested or tuned. Finally, the interception was turned off on August 7, with an official announcement that the system will be used again "when there is a threat [38]." We have not detected it since.

Kazakhstan's national-level HTTPS interception sets a dangerous precedent, not only for Kazakhstan—but for all governments and other powerful actors that wish to gain more control over users' Internet traffic. It also serves as an important reminder of the limits of HTTPS. Although nobody was forced to install the `Qaznet` root CA, most of the affected sites employed Strict Transport Security, so users who did not were unable to access these sites at all, even by clicking through security warnings. In the period the interception system was active, the private data of many thousands of users could have been compromised—including credentials for some of the world's most popular sites—and the security of their connections was significantly reduced.

We hope our work will inform efforts within the HTTPS security ecosystem to plan how to respond to future incidents of national-level interception. Based in part on our findings, two major browser vendors, Mozilla Firefox and Google Chrome, completely blocked the use of the `Qaznet Trust Network` root, so that any future use will be prevented even if users manually trust the certificate [33]. We advocate similar reactions to interception events in the future, and further research into technologies that can rapidly detect and impede such attacks.

1.1 Ethics

Our measurements were guided by several ethical considerations. First, we were careful not to directly involve any human subjects in Kazakhstan, due to potential legal risks they might face. For our direct measurements using RIPE Atlas probes and VPSes in the country, we only ran preliminary tests to `google.com` and `facebook.com`, two very popular domains unlikely to draw suspicion, and did not conduct any longitudinal measurements that might overload the network.

Ethical practices for remote censorship measurement have been the subject of many papers, discussions, and workshops [13, 25, 36, 42, 56, 57]. Since IRBs have determined that work such as our study is outside of their purview, we followed community norms and the guidelines listed in the Menlo and Belmont reports [15, 37]. Specifically, our primary remote measurements to TLS hosts in Kazakhstan only used hosts that had a valid certificate, so as to exclude typical residential hosts. Moreover, we tested only domains from the Alexa Top 10,000 [4] to reduce any risk of retaliation based on visiting unusual sensitive sites. For a separate experiment, in which we tested sensitive domains from the Citizen Lab Test List [12], we limited our vantage points to servers that presented a valid EV certificate, as these are almost exclusively larger organizations.

Additionally, we followed the Internet-wide scanning best practices proposed by the ZMap Project [18]. All our measurement machines have WHOIS records and a web page served from port 80 that indicates that measurements are part of a research project and offer the option to opt-out. We did not receive any complaints during the study period.

2 BACKGROUND

In this section, we first provide background on HTTPS interception attacks, their prevalence, and efforts to detect and prevent them before providing a brief timeline of the events in Kazakhstan leading up to the large-scale interception attack.

2.1 Related Work: HTTPS Interception

To perform HTTPS interception, a network entity poses as the destination server, accepting HTTP requests from clients and transparently proxying them to the real site [9]. HTTPS is designed to prevent this by requiring the server to present a certificate, signed by a certificate authority (CA) the client trusts, that associates its public key with the requested domain. For interception to succeed, either the attacker has to cause a browser-trusted CA to falsely issue them a certificate for the target domain, or the user has to install and trust a custom CA, which the interception system can use to sign certificates that the client will accept for any site. The latter approach is commonly used in residential and enterprise settings by client-side software and middleboxes, for purposes such as malware protection and content filtering [29, 46]. However, previous work has shown that interception frequently decreases connection security due to implementation flaws and lack of support for recent standards [17, 40, 53]. Moreover, such technology provides efficient avenues for implementing censorship and surveillance [10, 22, 44, 50], since the proxy can observe or modify connection plaintext.

In contrast to the prevalence of interception within enterprises, large-scale adversarial HTTPS interception has only rarely been documented, and the few recorded instances have tended to be brief and narrowly focused. The best known incident occurred in 2011, when an attacker compromised a CA called DigiNotar and created a fake browser-trusted certificate for `*.google.com`, which an ISP in Iran used to intercept connections to Google services [5]. Large-scale interception attacks were also detected in Syria (for `facebook.com`) and China (for `github.com`) for short periods in 2011 and 2013 respectively [19, 24], both based on untrusted certificates that raised security warnings in users' browsers. Kazakhstan's 2019 attack greatly exceeded these in duration, breadth of targets, and administrative sophistication. It also represents the first time that a national government attempted to induce its citizens to install a custom CA for purposes of interception.

To defend against interception via CA compromise, researchers have proposed a variety mechanisms to complement or replace CAs [14, 30, 55] or to limit their scope of trust [26, 47], though none has seen wide adoption. The idea of certificate pinning, where

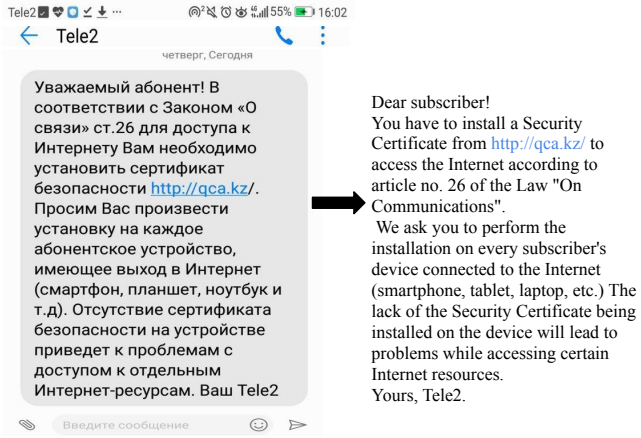

Figure 1: Kazakh users were directed to install a "security certificate"—a custom CA used to intercept HTTPS connections to popular sites. (*Image source:* [8]) ⋄

the browser remembers which certificates belong to each domain after first use, was adopted by major browsers in the past but is no longer supported [34]. Far more successful has been Certificate Transparency (CT) [28], which records certificates in a public ledger so that misissuance is at least detectable; Chrome now requires certificates from public CAs to be logged to CT. However, since the Kazakhstan attack involved users manually installing a custom CA, none of these proposals would prevent it.

Kazakhstan's attack was described at the time in informal online reports by our team [48] and, later, F5 Labs [54]. This paper includes significantly more detail and analysis.

2.2 Events in Kazakhstan before the Attack

Kazakhstan has a long-established, centralized policy of censorship and surveillance, and many sites have been blocked or monitored for several years [41, 43, 52]. The country is rated "Not Free" in Freedom House's "Freedom on the Net Report 2019" [20]. It regularly blocks access to political dissent, religious media, and certain social media sites [7].

In November 2015, Kazakhstan amended its communications law to require ISPs to adopt a "national security certificate" for all traffic to or from foreign destinations, with the intent of allowing the government to decrypt the communication [45]. A short time later, Kazakhtelecom, the country's largest state-owned ISP, announced plans to implement the measure [27, 39]. However, the plan was dropped following lawsuits from several organizations [3, 11]. At the same time, the Republic of Kazakhstan made a request to Mozilla to add the Root Certification Authority of Kazakhstan as a trusted CA [35]. This sparked significant discussion, but ultimately, because of incomplete audit reports and concerns that the root certificate would be used for interception, Mozilla denied the request [33, 35].

Kazakhstan's next major step towards HTTPS interception began on July 17, 2019, which we detail in this work. On that date, ISPs in Kazakhstan were instructed by the government to communicate to subscribers that they need to install (and trust) a government-issued root certificate on all devices and in every browser for "security" purposes. An SMS message sent to one ISP's subscribers is shown (with translation) in Figure 1. The certificate was not trusted by any browser by default, and needed to be manually installed by users. An initial thread about the interception was started on Bugzilla (Mozilla's bug tracker forum) on July 18 [8], which served as the starting point for our investigations.

3 TRIGGERING INTERCEPTION

The first step in investigating the large-scale HTTPS interception employed by Kazakhstan was to explore methods to trigger and detect the interference. We designed our experiments to trigger and analyze the interception based on the ethical considerations described in §1.1.

3.1 Methodology

We employed both direct measurements from inside the country and remote measurements from outside the country. For performing direct measurements, we obtained access to two VPS clients and 52 RIPE Atlas probes in the country. The two VPS clients were located in AS203087 and AS208450. We performed direct HTTPS requests to `google.com` and `facebook.com`, two domains reported in the initial Bugzilla report about the interception [8], from both the VPSes and the RIPE Atlas probes on July 20, 2019.

To increase the scale of measurements, we tested whether the interception could be triggered using remote measurement techniques from outside the country. Specifically, we used Hyperquack, a recently introduced remote measurement technique that detects network interference by sending various HTTP and HTTPS requests to thousands of infrastructural web servers around the world [50]. Hyperquack first requests several bogus (but benign) domains in the form of `<sub-domain>.example<rand>.com` from each web server. Since the web servers do not host these domains, they will likely respond with an error page. If the error response for all the requested domains are the same, Hyperquack uses this response to create a *template* that serves as the expected server response. This template includes features such as the response status code and the HTML body. In the case of HTTPS measurements, the template additionally includes the certificate, and chosen TLS version and cipher suite.

After building the template, Hyperquack requests test domains (potentially blocked domains) from each server. Since the web servers do not host these domains, the server response is expected to be the same as the template. However, if the response for the test domain differs from the template response after several retries, the measurement is marked as disrupted (for more details, refer to Sundara Raman et al. [50].)

In order to select infrastructural vantage points, we used data from Censys [16] to identify web servers that returned a valid EV certificate as these likely belong to large organizations [49]. We identified 82 such vantage points in Kazakhstan, located in 21 ASes.[2] On July 20, 2019, we performed Hyperquack HTTPS measurements to these 82 vantage points in Kazakhstan, with the input test list containing domains from the Citizen Lab Global Test List [12], a curated list of globally censored and sensitive domains, and Alexa Top 1000 popular domains [4], following the same test list selection process adopted in previous work [49, 50].

[2] AS information obtained from Maxmind [31] and Censys [16].

3.2 Results

While we did not detect any evidence of interception from our two VPSes, measurements from two of the 52 RIPE Atlas vantage points did observe the attack. The path to `google.com` and `facebook.com` from both of these probes passed through AS9198 (Kazakhtelecom). Out of the 82 Hyperquack vantage points, measurements to six had mismatching certificates between control and test measurements. Further investigation revealed that the certificate returned in these cases was signed by the Kazakhstan root CA (`Qaznet Trust Network`), the custom CA being used for interception. All six vantage points were also situated in AS 9198 (Kazakhtelecom) and geolocated to the capital city, Nur-Sultan.

From the six Hyperquack vantage points that observed the attack, connections with 27 popular social media and communication sites in the SNI header triggered interception (see Table 2). For all the other domains, the certificate was not injected, demonstrating that interception was selectively targeted. Our experiments did not indicate any change to the header or body of the response. This suggests that the system merely inspected the decrypted data, though we cannot rule out the possibility that payloads were selectively altered.

Our investigation showed that connections were only intercepted if they followed a network path that passed the interception system. However, interception occurred regardless of the direction that the connection took along the path. This meant that we could trigger interception behavior from outside the country by making connections to TLS servers inside Kazakhstan and sending targeted SNI domains, allowing us to conveniently perform more detailed measurements.

Overall, we found several conditions that had to be satisfied for a certificate to be injected:

- The connection path had to pass through a particular part of AS9198 (KazTelecom), the only AS where we observed injection occur.
- The client had to send a TLS SNI header containing one of the affected domains.
- The server had to present a valid browser-trusted TLS certificate, but not necessarily a certificate for the domain provided in the SNI header.

These conditions were necessary but not sufficient. Some connections we made passed through AS9198 but did not trigger injection, despite satisfying the other conditions.

4 IN-DEPTH MEASUREMENTS

Applying our initial findings, we began more detailed, larger-scale experiments to measure additional properties of the interception system and monitor its behavior over time. Our measurement infrastructure is illustrated in Figure 2.

4.1 Methodology

4.1.1 Measurements to TLS hosts. To conduct these measurements, we needed to find TLS hosts that provided a valid browser-trusted certificate. There were over 200,000 reachable TLS hosts in 129 ASes in Kazakhstan, but only 6,736 presented a valid browser-trusted certificate according to Censys [16]. These 6,736 TLS hosts were located in 85 different ASes.

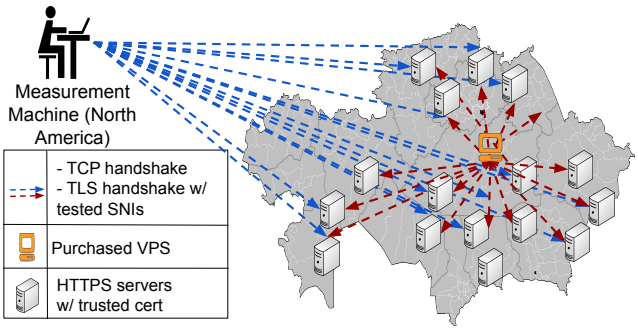

Figure 2: We performed detailed probes by connecting to TLS hosts in Kazakhstan and sending TLS connections with affected domains in the SNI header, exploiting the fact that interception could be triggered bidirectionally. ⋄

On July 22, we performed a TLS handshake from a North American client to each of these 6,736 hosts, setting the SNI to `facebook.com` and `google.com`, domains known to trigger the interception. Following that measurement, we attempted the same connections from one of our VPSes inside the country. This was to understand which networks paths were being intercepted. Additionally, we tested for interception of all domains from the Alexa Top 10,000 list from all TLS hosts where any interception was detected for `facebook.com` or `google.com`.

4.1.2 TTL-limited measurements. To locate where the interception was being performed, we employed a TTL-based technique similar to traceroute. For each TLS host where we were able to trigger interception, we made repeated connections with varying values for the IP time-to-live (TTL) field in the packet containing the SNI header, and we recorded the smallest TTL for which we received an injected certificate response. This technique allowed us to pinpoint the network location of the interception infrastructure.

On July 22, we performed this probe from a VPS in Kazakhstan to each TLS host that experienced interception. For each host we made two connections, one containing an SNI header for `facebook.com` and one for an unaffected domain, and measured the first hop for which we received a response.

4.1.3 Longitudinal measurements. In order to monitor the behavior of the interception system over time, beginning on July 23, we performed measurements from North America to the 6,736 TLS hosts every ten minutes, setting the SNI header to `google.com`, , and three other affected domains. We tested for the presence of the `Qaznet` certificate in each response.

4.2 Results

4.2.1 Extent of the Interception. Our measurements to the 6,736 TLS hosts on July 22 from North America found that only 459 servers (7.0%) had certificates injected, suggesting that HTTPS interception was occurring in only a fraction of the network in Kazakhstan. Measurements from our VPS inside the country found 1,598 (24%) TLS hosts with certificates injected. While these hosts were in different locations, the paths to all of them passed through AS9198, further confirming that this was where the HTTPS interception was taking place.

Table 1: ASes of hosts exhibiting interception were strongly biased towards AS9198, where our TTL experiments indicated the interception infrastructure was located. ⋄

AS	Name	TLS hosts
9198	JSC Kazakhtelecom	385
29555	Mobile Telecom-Service LLP	32
48502	ForteBank JSC.	23
43601	JSC BankCenterCredit	9
50482	JSC Kazakhtelecom	7
60708	KazNIC Organization	2
43934	. . .Interbank Settlement Centre. . .	1

Table 1 shows the ASes where the 459 TLS hosts that experienced interception were located. As expected, TLS hosts in AS9198 (Kazakhtelecom) experienced the largest amount of interception, since connections were more likely to pass through the intercepting hops. Kazakhtelecom is the country's largest provider, and many connections to other ISPs also passed through it.

4.2.2 Interception Location. We performed TTL-limited measurements from a VPS inside Kazakhstan to the 1,598 TLS hosts that had previously observed interception. Partway through the measurements, the interception system briefly stopped; by that point, we had performed measurements for 1,212 TLS hosts, 99.5% of which detected interception occurring at a hop earlier in the network path than the host. In the majority of cases, interception occurred only three or four network hops before the host. We confirmed similar findings from our US-based vantage point using the same technique.

Examining the IP addresses of the network hops in the traceroute where interception occurred, we found that 95% of the time, the last hop before the certificate was injected was 92.47.151.210 or 92.47.150.198, and the hop after injection was 95.56.243.92 or 95.59.170.59. All of these IP addresses are in AS9198 (Kazakhtelecom), suggesting a centralized design in which this AS was the only location responsible for HTTPS interception.

4.2.3 Injected Certificates. We also looked at patterns in the certificates returned by the interception system. While interception was triggered by the domain in the SNI header sent by the client, the names in the fake certificates were instead copied from those in the server's browser-trusted certificate. The fake certificates had the following properties:

- Identical Subject and Subject Alternative Name (SAN) fields to the server's real certificate.
- The public key was replaced with a host-specific 2048-bit RSA key (until July 19, 1024-bit), with exponent 3.
- The validity period (Not Before/Not After) was similar to the original certificate's but shifted six hours earlier[3].
- The serial number was similar to the original certificate's but with the last 33 bits changed randomly.
- All other x509 extensions were removed.

The use of 1024-bit RSA keys exposes users to the risk of interception by *other* governments—breaking 1024-bit RSA is likely within reach for many nation-states [2], and the CA/Browser Forum has deprecated 1024-bit RSA certificates [32]. Similarly, the use of exponent 3 in the RSA key may lead to a reduction in security that could be exploited by other malicious actors [6]. These certificates were signed by an intermediate CA (`C = KZ`, `CN = Security Certificate`) that in turn was signed by the root (`C = KZ`, `CN = Qaznet Trust Network`). The intermediate uses a 2048-bit RSA key (with more typical exponent 65,537) and is valid for three years, while the root certificate has a 4096-bit RSA key with a 30-year validity period.

[3]The validity period was updated to 24 hours on July 30, 2019 when the interception was turned back on after a four-day shutdown.

Table 2: Intercepted domains. 37 domains out of the Alexa Top 10,000 triggered interception. Most were associated with Google, Facebook, or the Russian Internet giant Mail.Ru. ⋄

Company	Domains
Google	allo.google.com, android.com, dns.google.com, docs.google.com, encrypted.google.com, goo.gl, google.com, groups.google.com, hangouts.google.com, mail.google.com, messages.android.com, news.google.com, picasa.google.com, plus.google.com, sites.google.com, translate.google.com, video.google.com, www.google.com, www.youtube.com, youtube.com
Facebook	cdninstagram.com, facebook.com, instagram.com, messenger.com, www.facebook.com, www.instagram.com, www.messenger.com
Mail.Ru	mail.ru, ok.ru, tamtam.chat, vk.com, vk.me, vkuseraudio.net, vkuservideo.net
Others	rukoeb.com, sosalkino.tv, twitter.com

4.2.4 Censor's TLS Fingerprint. Before generating a certificate, the interception system connected to the original TLS server to retrieve its real certificate for validation and replacement. We used a RIPE Atlas node in Kazakhstan to connect to a server we controlled, with the SNI header set to `facebook.com`. Instead of the expected TLS handshake from the Atlas device, our server observed a handshake from the interception system. Using TLS fingerprinting techniques from previous work [21], we generated the fingerprint (hash) of the Client Hello message. The interception system uses TLS 1.0 as the TLS record-layer version, TLS 1.2 as the ClientHello handshake version and offers 13 cipher suite options. The complete fingerprint is provided in [51]. The interception system's TLS fingerprint is virtually unseen in normal HTTPS Internet traffic (collected by [21]) and can thus be used as a unique identifier for the MitM. Sites could use this fingerprint to tell when a connection was being intercepted, and alert the user, revoke exposed credentials, or not send sensitive data. We reached out to a few affected websites, but none was able to share data about the occurrence of this fingerprint.

4.2.5 Domains Targeted. After testing affected TLS hosts with domains from the Alexa Top 10,000 [4], we found a total of 37 domains that triggered interception. These domains are mostly social media and communication sites, and are listed in Table 2. When ISPs instructed users to install the Kazakhstan root certificate, they claimed that its purpose was to protect against fraud, hacking, and illegal

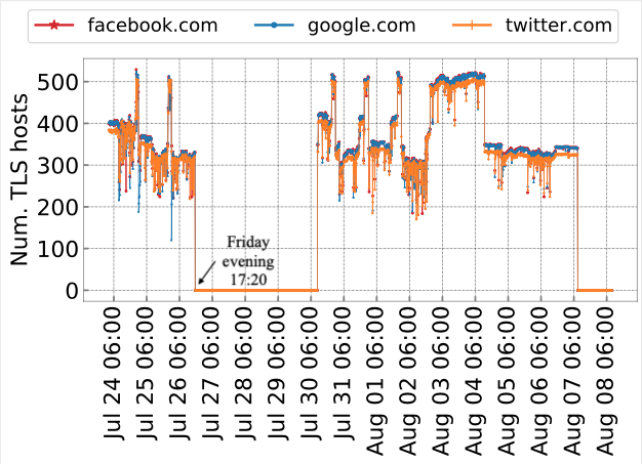

Figure 3: Longitudinal counts of TLS hosts (out of 6,736 hosts measured) exhibiting interception show daily patterns (likely due to routing changes) and an extended outage, during which the system was tuned. ⋄

content. However, the set of targets suggests that the actual intention may have been to surveil users' online interactions and communications.

4.2.6 Longitudinal Analysis. The results from our longitudinal measurements are shown in Figure 3. Interception was paused the evening of Friday, July 26, and resumed four days later, on the morning of July 30, with some changes to the logic for setting certificate validity periods. This suggests that the system was under active testing and development. Overall, we saw a median of 340 TLS hosts observing the interception when it was active. We noticed some periodic trends, such as a daily increase from 9 P.M. to midnight. Further investigation revealed that TLS hosts from four subnets belonging to mobile providers observed the interception only during this period, indicating a possible routing change.

5 DISCUSSION

Kazakhstan's HTTPS interception attacks represent an escalation in efforts by certain governments to gain access to encrypted communications. Unlike previous state-sponsored interception attacks, which were limited in scope and sophistication [19, 24], it covered a wide range of popular sites and lasted several weeks, potentially allowing the government to capture data and credentials for many thousands of users.

Significantly, Kazakhstan was able to intercept HTTPS without compromising a browser-trusted CA, as in some previous incidents [5]. Instead, users were forced to trust the government's custom root CA (and allow interception), or access to many of the targeted sites would be effectively blocked. Because of misleading communication from ISPs that suggested the certificate was intended to protect users' security, many users may have installed it without knowing its adverse effects. Browser security indicators would then give them a false sense of security, since the lock icon would be displayed even when the custom certificate was in use. We tried contacting some targeted services for information about how many users were affected, but none were able to share their data.

Informed by our findings, two major browser vendors, Mozilla Firefox and Google Chrome, responded on August 21, 2019, shipping changes that completely blocked use of the `Qaznet` root, even if manually installed [33]. Although this step was taken after the interception system was shut down, it prevents the system from being used again without users having to install a different certificate. We advocate an even quicker response if there are similar incidents in the future. Because of the prevalence of network security products that require users to install custom certificates, the option to add trusted certificates is necessary. However, we recommend that browsers add non-intrusive visual indicators to alert users about possible security risks each time a custom root is being used.

Additionally, we recommend further research into and higher adoption of defense mechanisms against large-scale MitM attacks in the HTTPS ecosystem [28]. We also encourage content providers to employ techniques to detect and share information regarding large-scale HTTPS interception attacks from particular countries or networks. As described in §4.2.4, interception systems may have unique TLS fingerprints, which would allow content providers to alert users whose connections are intercepted or take other protective actions.

Kazakhstan's interception system has not been active since being shut down on August 7, 2019, but, having showcased its capabilities, the government has stated its intention to turn the system on again "when required." The international community should prepare for that possibility—and for the event that another government conducts the same style of attack. Future measurement research can help by continuously monitoring for large-scale interception events such as Kazakhstan's.

6 CONCLUSION

With countries such as China and Russia practicing extensive censorship and moving closer to a controlled and balkanized Internet, end-to-end encryption is more important than ever for keeping users safe. In this paper, we explored Kazakhstan's government-sanctioned HTTPS interception attack in detail using direct and remote measurements. Such attacks threaten the protection offered by HTTPS and weaken security and privacy for the country's Internet users. It appears that the Kazakh government is willing to conduct further interception in the future, and other governments may adopt similar techniques. We urge the Internet security community to prepare for such events, by performing closer monitoring and by instituting policies for how to respond. If such interception attacks become normalized, decades of progress towards an end-to-end encrypted web will be lost for many of the Internet's most vulnerable users.

7 ACKNOWLEDGMENTS

The authors thank the shepherd Alan Mislove and the anonymous reviewers for their helpful feedback. We are also grateful to Wayne Thayer, Dana Keeler, and J.C. Jones from Mozilla for their help and prompt response in blocking the use of the MitM certificate. This work was supported in part by the U.S. National Science Foundation Award CNS-1518888 and a Google Faculty Research Award.

REFERENCES

[1] J. Aas, R. Barnes, B. Case, Z. Durumeric, P. Eckersley, A. Flores-López, J. A. Halderman, J. Hoffman-Andrews, J. Kasten, E. Rescorla, S. Schoen, and B. Warren. Let's Encrypt: An automated certificate authority to encrypt the entire web. In *ACM Conference on Computer and Communications Security (CCS)*, 2019.

[2] D. Adrian, K. Bhargavan, Z. Durumeric, P. Gaudry, M. Green, J. A. Halderman, N. Heninger, D. Springall, E. Thomé, L. Valenta, B. VanderSloot, E. Wustrow, S. Zanella-Béguelin, and P. Zimmermann. Imperfect forward secrecy: How Diffie-Hellman fails in practice. In *ACM Conference on Computer and Communications Security (CCS)*, 2015.

[3] K. Afifi-Sabet. Kazakh government will intercept the nation's HTTPS traffic. ITPro, July 19, 2019. https://www.itpro.co.uk/network-internet/34051/kazakh-government-will-intercept-the-nation-s-https-traffic.

[4] Alexa. Top 1,000,000 sites, July 2019. http://s3.amazonaws.com/alexa-static/top-1m.csv.zip.

[5] S. Bhat. Gmail users in Iran hit by MITM attacks. Techie Buzz, August 30, 2011. http://techie-buzz.com/tech-news/gmail-iran-hit-mitm.html.

[6] D. Bleichenbacher. Forging some RSA signatures with pencil and paper. Presentation in the rump session, CRYPTO, 2006.

[7] bne IntelliNews. Kazakhstan blocks Tumblr for promoting terrorism, porn, April 12, 2016. https://www.intellinews.com/kazakhstan-blocks-tumblr-for-promoting-terrorism-porn-94928/.

[8] Bugzilla. MITM on all HTTPS traffic in Kazakhstan, 2019. https://bugzilla.mozilla.org/show_bug.cgi?id=1567114.

[9] F. Callegati, W. Cerroni, and M. Ramilli. Man-in-the-middle attack to the HTTPS protocol. *IEEE Security & Privacy*, 2009.

[10] Z. Chai, A. Ghafari, and A. Houmansadr. On the importance of encrypted-SNI (ESNI) to censorship circumvention. In *USENIX Workshop on Free and Open Communications on the Internet (FOCI)*, 2019.

[11] C. Cimpanu. Kazakhstan government is now intercepting all HTTPS traffic. ZDNet, July 18, 2019. https://www.zdnet.com/article/kazakhstan-government-is-now-intercepting-all-https-traffic/.

[12] Citizen Lab. Block test list. https://github.com/citizenlab/test-lists.

[13] J. R. Crandall, M. Crete-Nishihata, and J. Knockel. Forgive us our SYNs: Technical and ethical considerations for measuring internet filtering. In *ACM SIGCOMM Workshop on Ethics in Networked Systems Research*, 2015.

[14] I. Dacosta, M. Ahamad, and P. Traynor. Trust no one else: Detecting MITM attacks against SSL/TLS without third-parties. In *European Symposium on Research in Computer Security*. Springer, 2012.

[15] D. Dittrich, E. Kenneally, et al. The Menlo Report: Ethical principles guiding information and communication technology research. Technical report, U.S. Department of Homeland Security, 2012.

[16] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman. Censys: A search engine backed by Internet-wide scanning. In *ACM Conference on Computer and Communications Security (CCS)*, 2015.

[17] Z. Durumeric, Z. Ma, D. Springall, R. Barnes, N. Sullivan, E. Bursztein, M. Bailey, J. A. Halderman, and V. Paxson. The security impact of HTTPS interception. In *Network and Distributed Systems Symposium (NDSS)*, 2017.

[18] Z. Durumeric, E. Wustrow, and J. A. Halderman. ZMap: Fast internet-wide scanning and its security applications. In *USENIX Security Symposium*, 2013.

[19] P. Eckersley. A Syrian man-in-the-middle attack against Facebook. EFF Deeplinks Blog, May 5, 2011. https://www.eff.org/deeplinks/2011/05/syrian-man-middle-against-facebook.

[20] Freedom House. Freedom on the net report, 2019. https://freedomhouse.org/countries/freedom-world/scores.

[21] S. Frolov and E. Wustrow. The use of TLS in censorship circumvention. In *Network and Distributed Systems Symposium (NDSS)*, 2019.

[22] S. Gatlan. South Korea is censoring the Internet by snooping on SNI traffic. Bleeping Computer, February 13, 2019. https://www.bleepingcomputer.com/news/security/south-korea-is-censoring-the-internet-by-snooping-on-sni-traffic/.

[23] Google. Transparency report: HTTPS encryption on the web, 2020. https://transparencyreport.google.com/https/overview.

[24] M. Johnson. China, GitHub and the man-in-the-middle. GreatFire.org, January 30, 2013. https://en.greatfire.org/blog/2013/jan/china-github-and-man-middle.

[25] B. Jones, R. Ensafi, N. Feamster, V. Paxson, and N. Weaver. Ethical concerns for censorship measurement. In *NS Ethics '15: Proceedings of the 2015 ACM SIGCOMM Workshop on Ethics in Networked Systems Research*, 2015.

[26] J. Kasten, E. Wustrow, and J. A. Halderman. Cage: Taming certificate authorities by inferring restricted scopes. In *Intl. Conference on Financial Cryptography and Data Security (FC)*, 2013.

[27] Kazakhtelecom JSC. Kazakhtelecom JSC notifies on introduction of national security certificate from 1 January 2016, 2015. Archived at https://web.archive.org/web/20151202203337/http://telecom.kz/en/news/view/18729.

[28] B. Laurie, A. Langley, and E. Kasper. Certificate transparency. *ACM Queue*, 2014.

[29] H. Lee, Z. Smith, J. Lim, G. Choi, S. Chun, T. Chung, and T. T. Kwon. maTLS: How to make TLS middlebox-aware? In *26th Network and Distributed Systems Symposium (NDSS)*, 2019.

[30] M. Marlinspike. Convergence, 2011. Archived at https://web.archive.org/web/20160803195327/http://convergence.io/.

[31] MaxMind. https://www.maxmind.com/.

[32] Mozilla. Phasing out certificates with 1024-bit RSA keys. The Mozilla Blog, September 08, 2014. https://blog.mozilla.org/security/2014/09/08/phasing-out-certificates-with-1024-bit-rsa-keys/.

[33] Mozilla. Mozilla takes action to protect users in Kazakhstan. The Mozilla Blog, August 21, 2019. https://blog.mozilla.org/blog/2019/08/21/mozilla-takes-action-to-protect-users-in-kazakhstan/.

[34] Mozilla. HTTP Public Key Pinning (HPKP), 2020. https://developer.mozilla.org/en-US/docs/Web/HTTP/Public_Key_Pinning.

[35] Multiple authors. Nation state MITM CAs? (thread). mozilla.dev.security.policy mailing list, 2016. https://groups.google.com/forum/#!msg/mozilla.dev.security.policy/wnuKAhACo3E/cpsvHgcuDwAJ.

[36] A. Narayanan and B. Zevenbergen. No encore for Encore? Ethical questions for web-based censorship measurement, 2015. Available at SSRN: https://ssrn.com/abstract=2665148 or http://dx.doi.org/10.2139/ssrn.2665148.

[37] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research, 1978.

[38] National Security Committee of the Republic of Kazakhstan. В отношении сертификата безопасности, 2019. http://knb.gov.kz/ru/news/v-otnoshenii-sertifikata-bezopasnosti.

[39] S. Nichols. Is Kazakhstan about to man-in-the-middle diddle all of its Internet traffic with dodgy root certs? The Register, December 3, 2015. https://www.theregister.co.uk/2015/12/03/kazakhstan_to_maninthemiddle_all_internet_traffic/.

[40] M. O'Neill, S. Ruoti, K. Seamons, and D. Zappala. TLS proxies: Friend or foe? In *ACM Internet Measurement Conference (IMC)*, 2016.

[41] OpenNet Initiative. Country profile: Kazakhstan, 2010. https://opennet.net/research/profiles/kazakhstan.

[42] C. Partridge and M. Allman. Addressing ethical considerations in network measurement papers. In *NS Ethics '15: Proceedings of the 2015 ACM SIGCOMM Workshop on Ethics in Networked Systems Research*, 2015.

[43] P. Pearce, B. Jones, F. Li, R. Ensafi, N. Feamster, N. Weaver, and V. Paxson. Global measurement of DNS censorship. In *USENIX Security Symposium*, 2017.

[44] R. Ramesh, R. Sundara Raman, M. Bernhard, V. Ongkowijaya, L. Evdokimov, A. Edmundson, S. Sprecher, M. Ikram, and R. Ensafi. Decentralized control: A case study of russia. In *Network and Distributed Systems Security Symposium (NDSS)*, 2020.

[45] N. Shapovalova. Security certificate of the Republic of Kazakhstan: the state will be able to control the encrypted Internet traffic of users. Dentons, Dec. 2015. https://www.dentons.com/en/insights/alerts/2015/december/30/security-certificate-of-the-republic-of-kazakhstan.

[46] J. Sherry, C. Lan, R. A. Popa, and S. Ratnasamy. Blindbox: Deep packet inspection over encrypted traffic. In *ACM SIGCOMM*, 2015.

[47] C. Soghoian and S. Stamm. Certified lies: Detecting and defeating government interception attacks against SSL. In *21st ACM Symposium on Operating Systems Principles (SOSP)*, 2010.

[48] R. Sundara Raman, L. Evdokimov, E. Wustrow, A. Halderman, and R. Ensafi. Kazakhstan's HTTPS Interception, 2019. https://censoredplanet.org/kazakhstan.

[49] R. Sundara Raman, P. Shenoy, K. Kohls, and R. Ensafi. Censored Planet: An Internet-wide, Longitudinal Censorship Observatory. In *ACM Conference on Computer and Communications Security (CCS)*, 2020.

[50] R. Sundara Raman, A. Stoll, J. Dalek, A. Sarabi, R. Ramesh, W. Scott, and R. Ensafi. Measuring the deployment of network censorship filters at global scale. In *Network and Distributed System Security Symposium (NDSS)*, 2020.

[51] TLSFingerprint.io. The Kazakhstan interception system's TLS fingerprint (f09427b5aaf9304b), 2019. https://tlsfingerprint.io/id/f09427b5aaf9304b.

[52] B. VanderSloot, A. McDonald, W. Scott, J. A. Halderman, and R. Ensafi. Quack: Scalable remote measurement of application-layer censorship. In *USENIX Security Symposium*, 2018.

[53] L. Waked, M. Mannan, and A. Youssef. The sorry state of TLS security in enterprise interception appliances. *Digit. Threat. Res. Pract.*, 1(1), 2019.

[54] D. Warburton. Kazakhstan attempts to MITM its citizens. F5 Labs Blog, August 1, 2019. https://www.f5.com/labs/articles/threat-intelligence/kazakhstan-attempts-to-mitm-itscitizens.

[55] D. Wendlandt, D. G. Andersen, and A. Perrig. Perspectives: Improving SSH-style host authentication with multi-path probing. In *USENIX Annual Technical Conference (ATC)*, 2008.

[56] B. Zevenbergen et al. *NS Ethics '15: Proceedings of the 2015 ACM SIGCOMM Workshop on Ethics in Networked Systems Research*. ACM, 2015.

[57] B. Zevenbergen, B. Mittelstadt, C. Véliz, C. Detweiler, C. Cath, J. Savulescu, and M. Whittaker. Philosophy meets Internet engineering: Ethics in networked systems research. GTC Workshop Outcomes Paper, 2015. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2666934.

Figure 4: The certificate chain of Kazakhstan's custom root. ⋄

A APPENDIX

Certificate Chain. Figure 4 shows the parsed certificate chain from our measurements. The root certificate (top left) with subject `Qaznet Trust Network` has a validity period of 30 years. The intermediate `Security Certificate` (bottom left) has a three year validity period and the leaf certificate (right) has the same validity period as the original certificate (but shifted by six hours).