UNIVERSITY OF TORONTO

MUNK SCHOOL OF GLOBAL AFFAIRS

Join the Global Conversation

The Citizen Lab

Research Brief
July 2014

Asia Chats: LINE and KakaoTalk Disruptions in China

Authors: Jakub Dalek, Philipp Winter, Andrei Dranka, Seth Hardy, Masashi Crete-Nishihata, and Adam Senft

For coverage of the report, please see Tech in Asia.

KEY FINDINGS

- Since July 2 2014, functionality of messaging applications LINE and KakaoTalk has been disrupted for users based in China as a result of DNS tampering and HTTP request filtering.
- On July 3, LINE updated the keyword list used to trigger keyword filtering for users with accounts registered to Chinese phone numbers. However, the blocking of LINE domains in China breaks the update functionality for this new keyword list.
- Photo-sharing platform Flickr and Microsoft's cloud-storage service OneDrive have also been recently disrupted in China by DNS tampering.

OVERVIEW

On July 2 2014, Tech in Asia reported that the official Web site (http://line.me) and instant messaging features of the LINE messaging application were inaccessible in mainland China. LINE Corporation acknowledged the service disruption in an announcement on its Weibo account, but did not provide details on why it was occurring. Interestingly, our monitoring of regionally-based keyword filtering in LINE for users with accounts registered to Chinese phone numbers found that the keyword list was updated on July 3 at approximately midnight EST.

Messaging application KakaoTalk experienced a similar disruption in mainland China beginning around July 2. Kakao Corporation reported that China-based users of KakaoTalk were able to send instant messages through the app, but could not access other features such as the "item store" and adding new friends.

The disruption of these chat apps occurred at the same time of reports that photo sharing platform Flickr and Microsoft's cloud-storage service OneDrive are also inaccessible in China. These service disruptions are correlated with mass democracy protests in Hong Kong on July 1 (a date that corresponds with the anniversary of the transfer of Hong Kong sovereignty from the United Kingdom to China). Some commentators have speculated that these outages may be an attempt by Chinese authorities to restrict information about the protests being transmitted through these applications and services.

In this post we examine how the Great Firewall of China (GFW) appears to be implementing DNS tampering and HTTP request filtering on KakaoTalk and LINE domains, which is disrupting service of the applications as a result. We find that Flickr and OneDrive are also blocked through DNS tampering. We also analyze the latest changes to the LINE keyword filtering list.

While there is clear technical evidence that these domains are blocked, it is unclear what the motivation behind blocking these platforms may be.

DNS TAMPERING AND TCP RESET INJECTION

Inspection of requests to LINE and KakaoTalk domains from a server based in Hangzhou, China reveal that the domains are blocked on the DNS and HTTP layers through DNS tampering and TCP Reset (RST) packet injection, respectively. Both of these filtering mechanisms are commonly used by the GFW.

DNS tampering is a practice in which deep packet inspection boxes analyze DNS traffic and send spoofed DNS responses to the client upon detecting blocked queries.

TCP RST packets are typically sent by clients or servers when there are network problems, as a response to a half-open connection, or some other network anomaly. When a RST flag is received most networking implementations in an operating system will shut down the socket used. One of the filtering mechanisms used by the GFW is inspection of HTTP traffic to check if there are any matches to keywords on block lists. If a match is present, forged TCP RST packets are injected into TCP flows which cause the endpoints to reset the connection.

DNS TAMPERING AND HTTP FILTERING OF LINE DOMAINS

The messaging application LINE works by sending HTTP requests to Naver servers within the domain space line.naver.jp (Naver is the parent company of LINE Corporation). These requests are not unlike regular HTTP requests that a browser would send when visiting web sites. All functions — from getting contact lists, changing account settings, sending messages, and visiting the in-app marketplace — are done through these HTTP requests.

Unofficial documentation by Matti Virkkunen provides details on how the LINE communications protocol works. This document gives the following example of a public message containing an image that is served through a GET request to the URL:

http://dl-obs.official.line.naver.jp/r/talk/o/u3ae3691f73c7a396fb6e5243a8718915-1379585871

When this URL is requested from a server in Canada, an image of the Nintendo character Pikachu is returned:

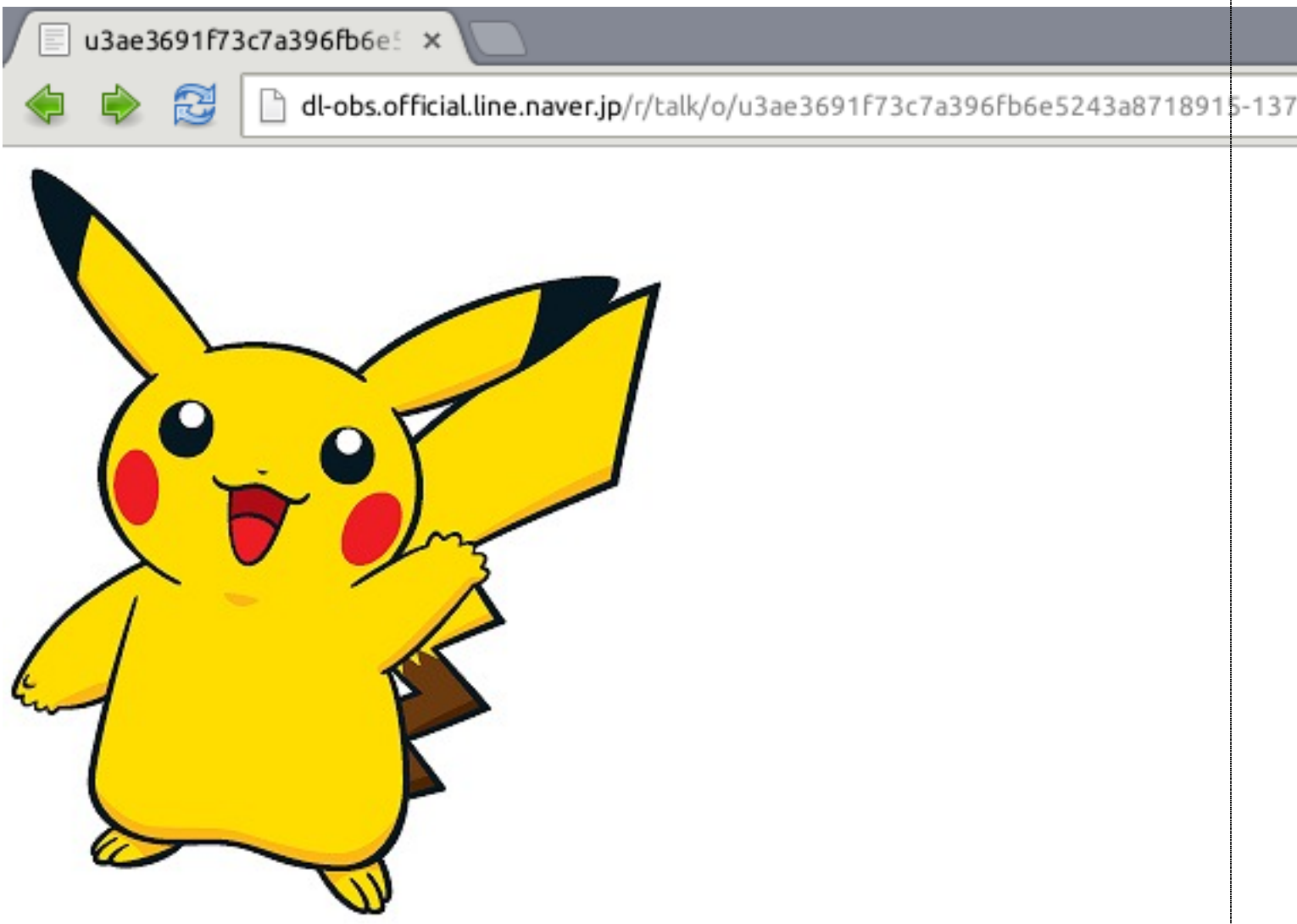

Figure 1. Content returned from http://dl-obs.official.line.naver.jp/r/talk/o/u3ae3691f73c7a396fb6e5243a8718915-1379585871 when accessed in Canada

However, when we tested this URL from a server in Hangzhou, China on July 9, the request fails. If we use the HTTP client curl on this China-based server, we find that we cannot connect, because the GFW tampered with the DNS request for the domain line.naver.jp. This tampering causes the client to connect to the bogus IP address 37.61.54.158 which does not even run a web server.

$ curl
http://dl-obs.official.line.naver.jp/r/talk/o/u3ae3691f73c7a396fb6e5243a8718915-1379585871
curl: (7) couldn't connect to host

Note that if we lookup the IP of dl-obs.official.line.naver.jp on a non-Chinese network connection we get the Akamai IP 184.84.243.8. There seems to be no interference in sending packets directly to this IP from the Chinese connection as both ping and traceroutes are returned correctly. Therefore, the LINE domain is blocked, but the Akamai IP is accessible.

Similarly if we visit the in-app store of LINE on an Android phone to browse stickers, the application does a GET request to the URL: http://dl.stickershop.line.naver.jp/products/0/0/1/1109/android/preview.png

This request returns the following content in Canada:

Figure 2: Content returned from http://dl.stickershop.line.naver.jp/products/0/0/1/1109/android/preview.png when accessed from Canada

Again this request fails in China:

$ curl "http://dl.stickershop.line.naver.jp/products/0/0/1/1109/android/preview.png"
curl: (7) couldn't connect to host

These results are not just limited to images. According to Virkkunen, personal chat and contact data is sent via the Apache Thrift protocol to the server gd2.line.naver.jp on port 443. In Canada these servers, without prior authentication, return a 404 response along with a distinct server response header named "X-LCR":

$ curl -I "https://gd2.line.naver.jp"
HTTP/1.1 404 Object Not Found
Content-Length: 9
X-LCR: 110

However, the same request will time out in China:

$ curl -I "https://gd2.line.naver.jp"
curl: (28) SSL connection timeout

All of these failures are due to DNS tampering of LINE domains in China. In the table below we show the differing results for requests for LINE domains from our lab network in Canada and the Hangzhou-based network we tested on:

DNS Response (IP and AS) in Canada for line.naver.jp	DNS Response (IP and AS) in China for line.naver.jp
119.235.235.44 – Line Corporation, Japan	37.61.54.158 – Baktelekom, Azerbaijan

The incorrect response of the IP "37.61.54.158" we received in China has been seen in previous reports of other domains blocked in the country, such as facebook.com.

From the China network vantage point we see that any lookup of gd2.line.naver.jp (where chat data is sent) and dl.stickershop.line.naver.jp (the in-app market place) return the same incorrect IP address.

$ dig +short gd2.line.naver.jp
37.61.54.158

$ dig +short dl.stickershop.line.naver.jp
37.61.54.158

Indeed any subdomain of line.naver.jp, whether it exists or not returns the bogus IP address as a response:

$ dig +short totallymadeup.domain.gojays.line.naver.jp
37.61.54.158

In addition to this DNS tampering any HTTP GET request that contains "line.naver.jp" in the request header will have a TCP RST returned. Therefore, there is blocking at both the DNS and HTTP level.

$ curl -H 'Host: line.naver.jp' x.x.x.x (where x = any ip)
curl: (56) Recv failure: Connection reset by peer

In media reports of the recent LINE disruption, the official website of LINE (line.me) was also reportedly inaccessible. We tested line.me on the Hangzhou-based network on July 4 and found it was being blocked through forged TCP RST packets as seen in the TCP flow graph below:

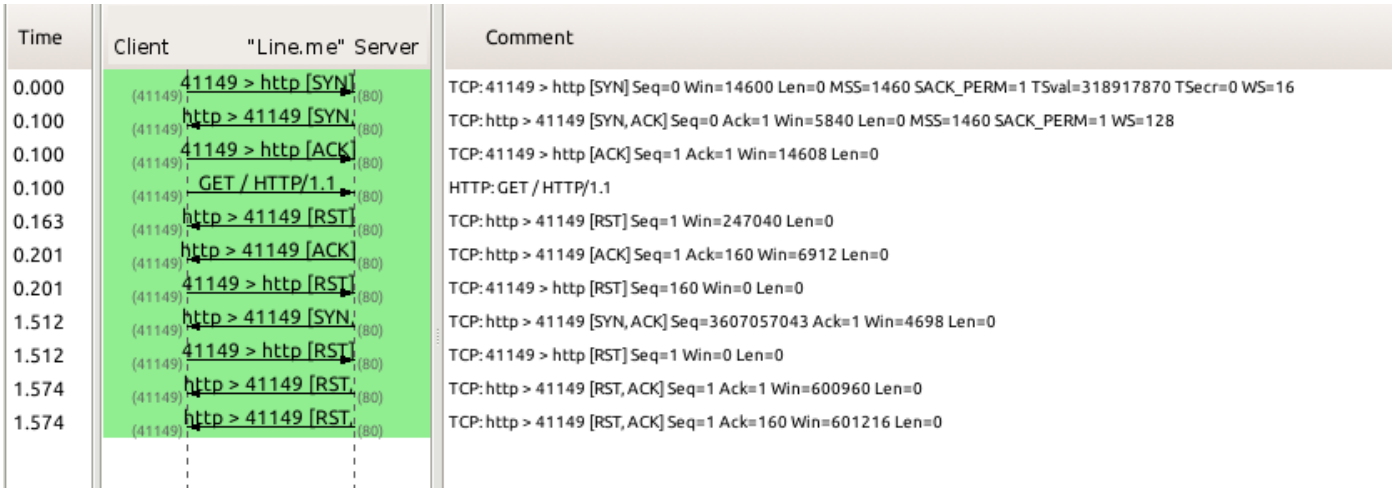

Figure 3 – TCP Flow graph of an injected RST response for line.me

However, tests we ran on July 7 from this same server showed line.me to be accessible while access to the line.naver.jp domain remained blocked. Additionally, there were also inconsistencies that were found in testing between different network vantage points in China. The testing that the findings of this report are based on was done on a server based in Hangzhou on which we consistently saw the network interference outlined. We also tested LINE domains on a commercial VPN provider based in Hefei (on Chinanet) and observed that there was no interference and the DNS requests resolved correctly in those cases. These results correspond to anecdotal reports that some users in China are able to use LINE normally while others cannot. At this time, the reason behind these inconsistencies is unknown.

LINE KEYWORD LIST UPDATE

Since the fall of 2013, we have been analyzing regional-based keyword filtering for LINE users with accounts registered to Chinese phone numbers. During our monitoring we have observed four updates to the keyword lists that the application retrieves from Naver servers. The list versions have been updated from v20, v21, v22, to the most recent update v23.

On June 3 at approximately midnight EST the keyword list was updated from v22 to v23. The only change to the list is the addition of a single keyword 买枪 “mǎi qiāng” (buy a gun). Keywords related to the purchase of illicit goods (including firearms) were prominently featured in keyword lists used to trigger censorship and surveillance features we previously discovered in the instant messaging program TOM-Skype. However, we have not found this content category in LINE keyword lists before.

This latest update is the smallest iteration to a LINE keyword list we have observed. In the previous change from list v21 to v22, 312 new keywords were added and 147 keywords were deleted. Keyword list v23 retains the other 535 keywords present on v22 and only adds the single keyword to make the total size 536 keywords.

As we have previously described, if a user is registered to a Chinese phone number, LINE will check for an updated keyword list over HTTP and HTTPS at the URLs http://line.naver.jp/app/resources/bwi and https://line.naver.jp/app/resources/bwraw. After the introduction of DNS tampering of domain names which include ‘line.naver.jp’, this mechanism for updating the keyword lists no longer functions.

Figure 4 shows what happens when LINE attempts to update its keyword list from a connection in China. The client initially tries to resolve line.naver.jp first but is redirected to the Azeri IP that does not host a web server.

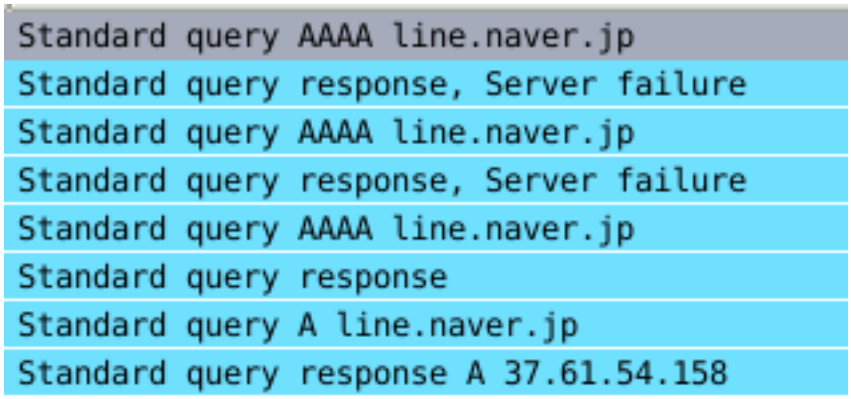

Figure 4 – Screenshot of a packet capture while attempting to update the LINE keyword list from a network in China

A potential (but speculative) explanation for the unusual addition of a single keyword to list v23 is that LINE developers may be attempting to test if the update mechanism still functions following the service disruption. However, while the keyword list update was pushed approximately one day after the service disruption was reported, it is also possible the two events are unrelated.

DNS TAMPERING AND HTTP FILTERING OF LINE DOMAINS

The KakaoTalk messaging application primarily operates through Web API calls to manage contact lists, read public messages, and adjust settings. Many of these calls rely on HTTP requests to kakao.com domains.

For example, retrieving your friends list requires the application to make a POST request to the kakao.com domain on the Android client: http://katalk.kakao.com/android/friends/update.json

Blocking a user from contacting you requires the application to make a POST request to the URL: https://katalk.kakao.com/android/friends/block.json

Account registration on KakaoTalk takes place by sending an HTTP POST request to the URL: http://ac-talk.kakao.com/android/account/validate_phone_number.json

Similar to LINE, it appears that the kakao.com domain space is being tampered with by the GFW, which is resulting in disruptions to features in KakaoTalk.

In the table below we show the differing results for requests for kakao.com domains from our lab network in Canada and the Hangzhou-based network we tested on:

DNS Response (IP and AS) in Canada for kakao.com	DNS Response (IP and AS) in China forkakao.com
110.76.141.122 – Kakao, Korea	46.82.174.68 – Deutsche Telekom AG, Germany

The bogus IP address “46.82.174.68” returned in China has been previously reported as associated with DNS tampering in the country.

In addition to blocking on the DNS layer, the GFW blocks access to kakao.com on the HTTP layer by looking at the HTTP host header. Every HTTP request with a host header that contains the string “kakao.com” is blocked by injected TCP RST packets.

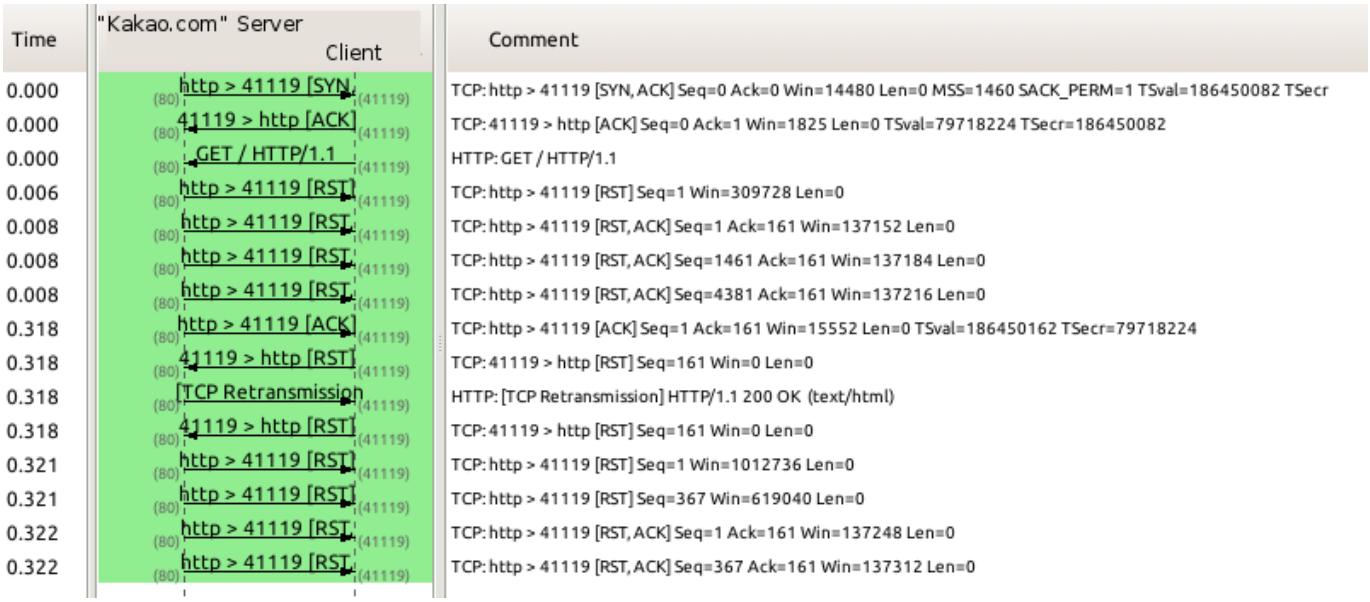

Figure 5 – TCP Flow graph of an injected RST response for kakao.com

Where KakaoTalk differs from LINE is in its handling of private chat data. This data is sent via their custom LOCO communications protocol on a different port (5228) than other functions. Brian Pak has published documentation of this protocol and implementation in a series of blog posts (here and here). According to these posts, LOCO servers are identified by IP address and not by domain. This difference may explain reports that text messaging features were accessible in China, but other features such as adding friends, and accessing the “item store”, “notices”, and “plus friend” were inaccessible.

DNS TAMPERING OF FLICKR AND ONEDRIVE

In addition to testing LINE and Kakao accessibility we tested the domains associated with the Flickr photo sharing service and Microsoft’s One Drive cloud storage service. Similar to the disruptions of LINE and KakaoTalk we find that the domain responses of Flickr and One Drive have also been tampered with.

For example, the domain flickr.com responds again with the Azerbaijan IP we saw in the LINE domain tampering:

$ dig +short flickr.com
37.61.54.158

We see the same result when attempting to access the domain onedrive.live.com :

$ dig +short onedrive.live.com
37.61.54.158

However, it is important to note that the top domain "live.com" returns the proper answer in China as it points to a Microsoft IP:

$ dig +short live.com
65.55.206.154

This result means that the live.com mail service is likely unaffected by the change but the cloud storage service is.

CONCLUSION

This recent incident is the first example of LINE and KakaoTalk domains being blocked in China of which we are aware. While the onset of the disruptions is correlated with the July 1 Hong Kong protests, this event is not necessarily the cause of the on-going outages. It is also curious that LINE, which has keyword filtering functionality of China-based users, and KakaoTalk which does not have these controls, are both being disrupted.

These disruptions are also occurring at the same time of sudden blocking of Flickr, One Drive, and more recently the removal of Instagram from Android App stores in China. Determining whether these recent disruptions and removals are related and what the underlying motivation behind them may be requires further research. We will continue to monitor the accessibility of these services and post updates as they become available.

RESOURCES

- DNS and HTTP request data for Line, KakaoTalk, Flickr, and OneDrive domains on Github
- Raw and translated LINE keyword list data on Github
- LINE Region Code Encrypter Tool for changing regions in the LINE client to disable regionally-based keyword censorship in the application

ACKNOWLEDGEMENTS

Jakub Dalek, Philipp Winter, Andrei Dranka, Seth Hardy, Masashi Crete-Nishihata, and Adam Senft undertook the research and writing of this post. Special thanks to Jason Q. Ng for translation assistance. This research is supported by the John D. and Catherine T. MacArthur Foundation.