Media Forensics Hub

Watt Family Innovation Center

Clemson University

Xinjiang Nylon: The anatomy of a coordinated inauthentic influence operation

Research team: Darren Linvill, Patrick Warren, Steven Sheffield, Jayson, Warren, Beau Brierre, Grant Cole, Jonathan Heijjer, Tyler Reich, Grant Saunders, Jack Taylor

1. Introduction

This report documents the presence of and tactics employed by a complex and multifaceted inauthentic social media campaign conducted over Fall of 2021 around narratives of interest to the People's Republic of China (PRC), especially those that relate to the treatment of the Uigher minority in Xinjiang province. In it, we document two large classes of social-media accounts, which differ in their origins and the role they play in this campaign, but which all have significant markers of inauthenticity. We begin by laying out narratives where this campaign was discovered, document the sorts of accounts participating in this campaign, and then turn to core tactics.

To get a sense of where these accounts were discovered, consider the dominant Hashtags about Xinjiang in October, 2021. There were about 120k tweets mentioning Xinjiang in October, 2021. In them, the most common hashtag is #Xinjiang. There were about 60k tweets in the month of October that used #Xinjiang, explicitly. As of November 15, over 20k (about 36%) of the #Xinjiang tweets came from accounts that had been suspended. But within these, there was a smaller but more suspicious narrative, characterized by the use of the #XinjiangCotton hashtag and consistent of many duplicative tweets from similar-seeming accounts. Of the 2400 tweets including this hashtag (one of the prominent hashtags in the campaign), over 1200 (51%) came from accounts that have already been suspended. That is a remarkably high rate, further indications that the accounts active on amplifying this hashtag were illegitimate and worthy of further attention.

In the influence campaign targeting the conversations around Xinjiang, the operator used a variety of inauthentic accounts. These accounts fall fairly neatly into two groups: what we call the "old" accounts" and the "new" accounts. It is difficult to know the exact number of accounts that participated in the campaign, or the total output, as many of the accounts were suspended over time and new accounts are constantly arising.

In the next section we lay out the characteristics of these accounts, before turning to the strategies they employ.

2. Accounts

For this analysis, we concentrate on a convenience sample of about 30 “old” accounts and 200 “new” that we were active in the #XinjiangCotton narrative on Twitter and that we were able to investigate in detail. There were many more “new” accounts than this (thousands) active in the narratives around #Xinjiang and #XinjiangCotton, but they were very homogeneous in character. This is also not a complete set of the “old” account group, since some were suspended before we could collect information about them and others may have escaped our collection. But it is sufficient to track their similarities and differences and how they are used. There are marked differences in the character of these two groups, along many dimensions, which we discuss in turn.

2.1 Account Profiles

Names: The accounts in the older group have a variety of international usernames, including Anglo-American, Korean, European, Indian, and traditional Chinese. These consist of a mix of Anglicanized spellings of non-English names (screenshots 1&2) and names explicitly in non-latin characters. On the other hand, the newer account names are almost exclusively traditional “American” names with “Ashley”, “Jennifer”, and “Mellissa” being common. The same pattern extends to screen names as well with the older accounts exhibiting simple hand-crafted usernames, often aligned with the usernames. The new accounts exhibit a more interesting pattern, specifically a large number of accounts, upwards of 92% of the identified accounts, having the same format of username where it is a name followed by a string of 5 or 6 random numbers, consistent with sticking with Twitter’s default recommendation. *This gives credence to the idea of a bulk creation of these newer accounts.*

Birth dates: The two groups have distinct patterns in the date of creation across accounts. The older accounts have birth dates ranging from March, 2009 at the earliest to May 2013 at the latest. The newl accounts all have birth dates in 2021. Even more specifically, handfuls of these accounts were all created within hours of each other on the same date. For example, over 60 new accounts were created between March 28 and April 1. *The extreme similarity and rapidity of the creation times of the newer accounts is further evidence of their bulk creation by a single actor. The pattern of the older accounts’ birth dates in and of itself is less suspicious as there are many Twitter users who created their accounts over 10 years ago and are still active.*

User-Provided Descriptions: The older accounts have disparate characteristics of their profiles. Many of them do not have bios whereas the ones that do have things ranging from listing their gender, birthday, where they live, and how many kids they have to sharing a link and saying which high school they attended. The descriptions are in a variety of languages, including English, Korean, and Tamil. The new accounts overwhelmingly had empty descriptions. *Homogeneous and empty profiles for new accounts again support the story of cost-minimization* *batch production. The heterogeneous descriptions in the older accounts suggest heterogeneous true origins.*

Location: The patterns of user-described location are quite similar to the description. Many of the new accounts had no location specified and the handful that do located themselves in China or Taiwan. The older accounts, by contrast, are very heterogenous. They vary from “Republic of Korea” to “Glasgow” to “Polska.” None of the locations were in China, nor were they overly specific, in fact they were vague and some of them seemed to be “near” China but not inside the border with locations like “Kyrgyzstan,” “Republic of Korea,” and “South Korea.” Other locations were located in Canada and the Netherlands. When there is a old tweet history in another language it nearly always matches the location. *Only speculatively but it seems the reasoning for the sparse locations is to pass the appearance that these accounts are tweeting the pro-Xinjiang cotton propaganda from not Xinjiang.*

Follower/following structure: The following structure of the older accounts looks like that of a normal user in the country of the account’s origin, to include several prominent pop figures, politicians, etc. New accounts have very low follower and following numbers (most often zero). *The lack of investment in developing follower networks suggests that this network is not trying to influence narratives through spreading messaging to its followers. Rather they are trying to either influence search results or the algorithm or simply lending credence through likes and retweets.*

Tweet Client: The bulk of the tweet activity was done using Twitter WebApp or other online clients. This is true more so for the new accounts as they nearly universally tweeted a majority of their content using Twitter WebApp although there were several examples of Android usage. The older accounts exhibited a trend wherein before they began tweeting the cotton content they used a mix of Tweet clients from iPhone to Android to Twitter WebApp. However, once they started tweeting the propaganda the accounts switched their Tweet Clients to web-based ones such that it follows the same pattern of the newer accounts.

Emails associated with accounts: By the time they were taking part in the campaign, accounts from both groups were, largely, associated with emails from a small set of free, online email domains. The older accounts are primarily associated with accounts that appear to be hosted on usa.com, or other mail.com-affiliated domains. The new accounts are associated with emails that appear to be exclusively hosted on outlook.com or hotmail.com. *This pattern suggests that the two sets of accounts were likely inducted into the campaign by two different agents or perhaps the same agent at different times.*

Evidence of compromise: For the older accounts, the patterns of account-level characteristics, above, are consistent with them being initially created by a large heterogeneous set of agents, probably organic in nature. But the homogeneity of their behavior and their associate emails suggests that, at some point, they came under the control of a unified actor. The most likely explanation is that the original operators had their Twitter accounts compromised and repurposed to participate in this campaign. Searching for the screen names for the older accounts in Dehashed.com, a database of (mostly) publicly-disclosed breaches, revealed strongly suggestive evidence in favor of this hypothesis. More than half the screen names from older accounts appeared as a username in an email in the dehashed database, often for an email provider that primarily serves the country that appeared in user-reported location for the account. The most common of these matches (about half) were to emails hosted by the South Korean naver.com domain. Hacked in 2014, approximately 25 million Naver hosted accounts were left vulnerable to exploitation [1].

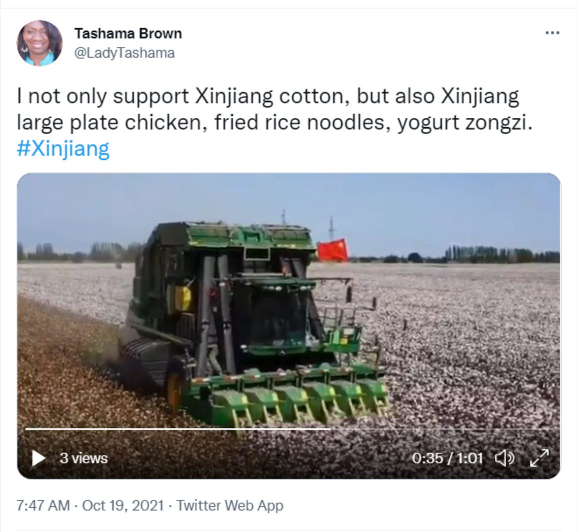
2.2 Account activity and messaging

Language: All recent tweets in both groups of accounts were either in English or Chinese, with many of English-subtitled Chinese-language videos. The language seemed to follow a pattern based on what was being tweeted: most of the propaganda about cotton in Xinjiang from "official" Chinese government accounts was in English. This was seen in both sets of accounts. Smaller tweets and other interactions were largely in Chinese, Mandarin to be specific. For some of the older accounts, there were years-old tweets that were in an entirely different language, including Korean, Czech, and French. This lends credence to many of the older accounts being compromised.

Content: Both older and newer accounts posted pictures and videos with simple 1-2 sentence statements regarding the benefits of the Xinjiang region. While the videos vary slightly In content, the statements had very similar themes: the bounty of the cotton harvest, the quality of life for both the locals and workers, the quality of the cotton being produced. Often multiple accounts, often across types, shared identical or nearly identical content. Pictured below, the phrase "how can it be 'forced labor'?" is used exactly, while the links shared are all different. Sample content can be seen in Figure 1.

Networked content: Old accounts do not have much networked activity. They almost never reply and rarely retweet or like, generally 1-2 over the span of a few weeks. They also rarely post exactly duplicative content of liked or retweeted photos/videos. New accounts have a somewhat heavier retweet:tweet ratio (mostly directed at two Chinese spokesperson accounts) ; understandably so, if an account is deleted the number of retweets/likes on the original post will stand. Trends have shown that if an account plans to use a video multiple times within a pod that they will either tag another new account while quoting the tweet or will tag them in a comment.

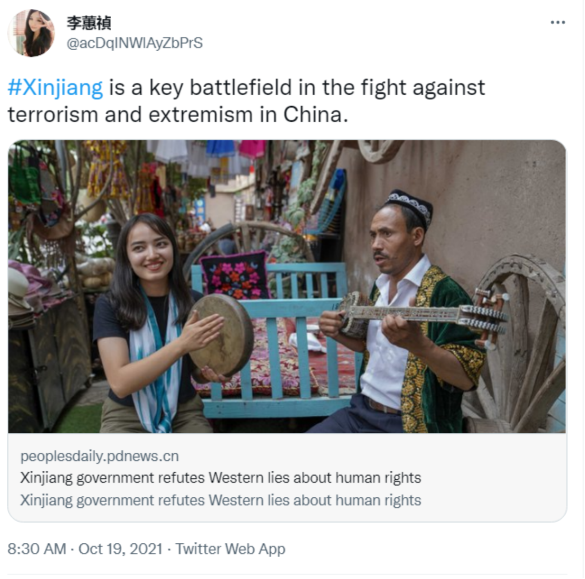
Figure 1
Sample Xinjiang Troll Accounts



Tashama Brown
@LadyTashama

I not only support Xinjiang cotton, but also Xinjiang large plate chicken, fried rice noodles, yogurt zongzi. #Xinjiang

7:47 AM · Oct 19, 2021 · Twitter Web App

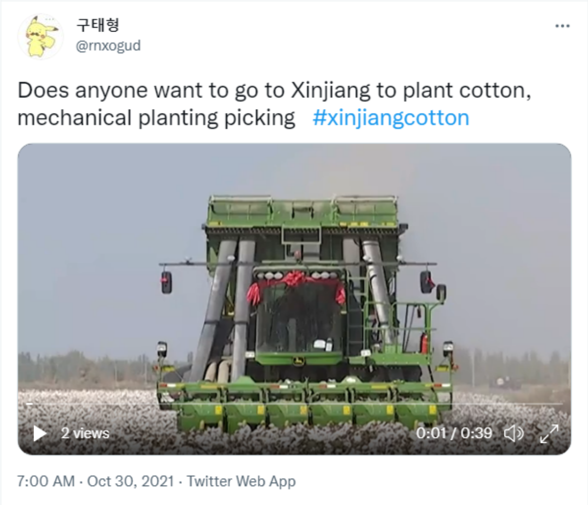


李蕙禎
@acDqINWlAyZbPrS

#Xinjiang is a key battlefield in the fight against terrorism and extremism in China.

peoplesdaily.pdnews.cn
Xinjiang government refutes Western lies about human rights
Xinjiang government refutes Western lies about human rights

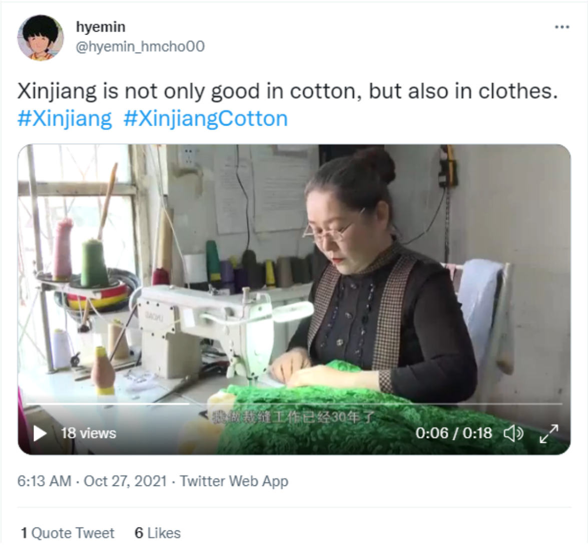
8:30 AM · Oct 19, 2021 · Twitter Web App



구태형
@rnxogud

Does anyone want to go to Xinjiang to plant cotton, mechanical planting picking #xinjiangcotton

7:00 AM · Oct 30, 2021 · Twitter Web App



hyemin
@hyemin_hmcho00

Xinjiang is not only good in cotton, but also in clothes. #Xinjiang #XinjiangCotton

6:13 AM · Oct 27, 2021 · Twitter Web App

1 Quote Tweet 6 Likes

3. Strategy

In addition to the specific behaviors for the specific accounts types laid out above, the campaign, as a whole, pursued a handful of overarching strategies to affect in information space.

3.1. Inauthentic Interactions with Official PRC Outlets

It is important to understand the cultural and political differences between how the Chinese and Americans interact with social media. The PRC banned Twitter (as well as other western social media platforms) in 2009 following riots that took place in Xinjiang province[3] and Twitter use in China is therefore very low.[2] Twitter user support for perspectives that align with the Chinese state agenda is therefore also lower than would otherwise be the case.

Interaction with tweets through likes, retweets, comments, and quote tweets is important to give a Twitter account legitimacy. A state-affiliated Twitter account with little to no interaction (likes, retweets, etc) has diminished legitimacy in the information environment. In an apparent effort to increase the face validity of their own official state accounts, Chinese information operations generate fraudulent inauthentic accounts with the aim to increase likes, retweets, and comments on a state-affiliated twitter post.

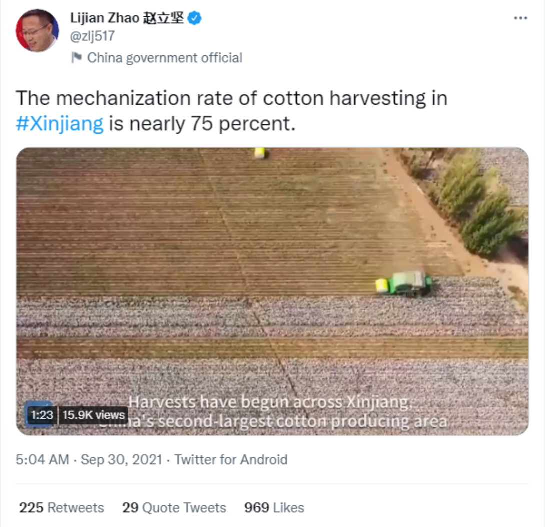
Currently they are using this tactic in part to validate messages that portray Xinjiang province in a positive light and distract the world from the atrocities that have taken place in the region in the past few years.

Inauthentic social media accounts are being employed to artificially boost the legitimacy of official Chinese state social media accounts. We can see this tactic at work in a particularly straight forward manner analyzing the account of Zhao Lijian, deputy director of the Chinese ministry of foreign affairs. This account is very active, dating back to 2010 and has 1 million followers and nearly 67 thousand tweets. Zhao posts pro-state messaging several times a day, with topics ranging from Chinese culture, to political news and economic news. This includes many posts about Xinjiang cotton. Figure 2 shows several tweets from this account, each of which shared a video. Each of these tweets shows signs of inauthentic engagement.

- #1: September 30, 2021 tweet regarding the mechanization rate of cotton harvesting was retweeted or quote tweeted 260 times. Nearly 10% of these tweets came from accounts with zero followers and the median number of followers was only 91. At least Twenty of these accounts have since been suspended by Twitter.
- #2: October 25, 2021 tweet regarding Xinjiang primary students dancing was retweeted or quote tweeted 471 times. Over 12% of these tweets came from accounts with zero followers and the median number of followers was 48. Five of these retweets were the first action an account ever took. At least 21 of these accounts have since been suspended by Twitter.
- #3: October 9, 2021 tweet regarding a “dance battle” in Xinjiang was retweeted or quote tweeted 394 times. About 5% of these tweets came from accounts with zero followers and the median number of followers was 130. At least 28 of these accounts have since been suspended by Twitter.
- #4: November 8, 2021 tweet regarding Xinjiang covered in snow was retweeted or quote tweeted 238 times. Nearly 20% of these accounts have zero followers. Five of these retweets were the first action an account ever took. At least nine of the retweeting accounts have since been suspended by Twitter.

Figure 2
Sample tweets from Zhao Lijian about Xinjiang

Tweet #1



Lijian Zhao 赵立坚
@zlj517
China government official

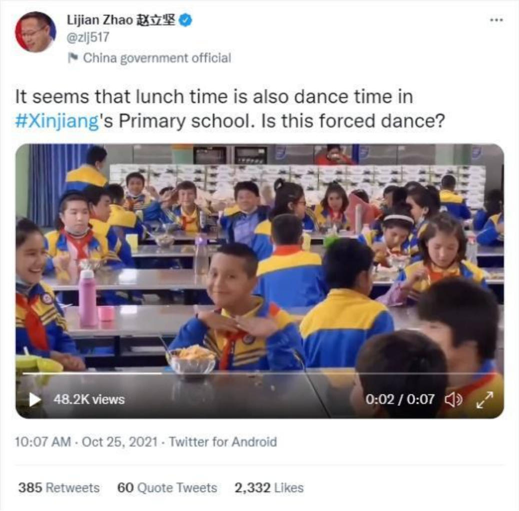
The mechanization rate of cotton harvesting in #Xinjiang is nearly 75 percent.

1:23 15.9K views
Harvests have begun across Xinjiang,
...ia's second-largest cotton producing area

5:04 AM · Sep 30, 2021 · Twitter for Android

225 Retweets 29 Quote Tweets 969 Likes

Tweet #2



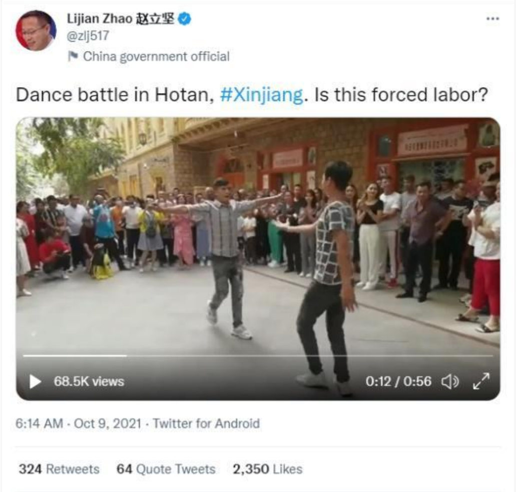
Lijian Zhao 赵立坚
@zlj517
China government official

It seems that lunch time is also dance time in #Xinjiang's Primary school. Is this forced dance?

10:07 AM · Oct 25, 2021 · Twitter for Android

385 Retweets 60 Quote Tweets 2,332 Likes

Tweet #3



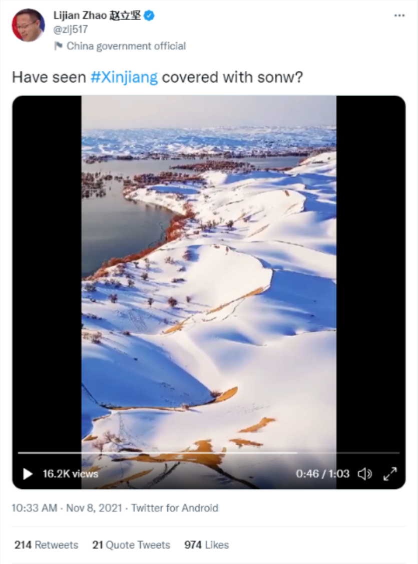
Lijian Zhao 赵立坚
@zlj517
China government official

Dance battle in Hotan, #Xinjiang. Is this forced labor?

6:14 AM · Oct 9, 2021 · Twitter for Android

324 Retweets 64 Quote Tweets 2,350 Likes

Tweet #4



Lijian Zhao 赵立坚
@zlj517
China government official

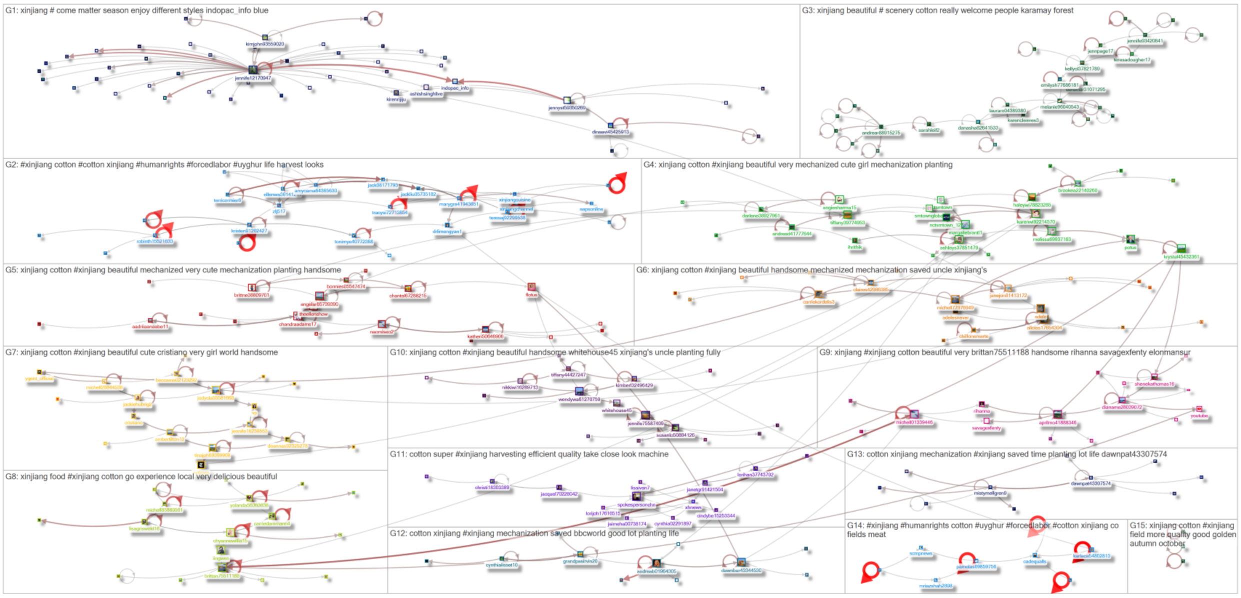
Have seen #Xinjiang covered with sonw?

10:33 AM · Nov 8, 2021 · Twitter for Android

214 Retweets 21 Quote Tweets 974 Likes

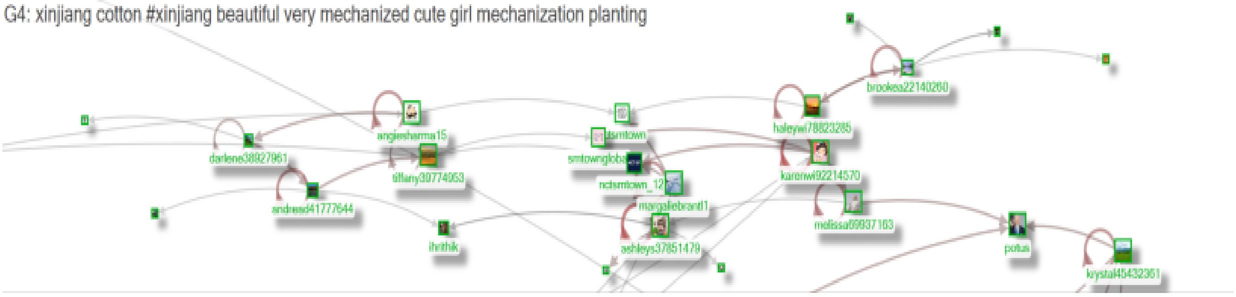
3.2 Core-Periphery Pods

NodeXL was used to analyze a sample of 230 known disinformation accounts in the Xinjiang network. Below you will see a graph that represents those accounts and who they reached out to outside the disinformation network. Generally, there are core and peripheral accounts. Most of the single link isolated accounts represent a point where the network reached out to an account outside the network via mention or reply.

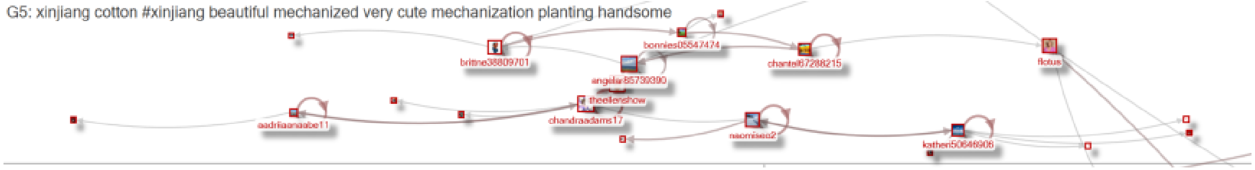

G1: xinjiang # come matter season enjoy different styles indopac_info blue
G3: xinjiang beautiful # scenery cotton really welcome people karamay forest
G2: #xinjiang cotton #cotton xinjiang #humanrights #forcedlabor #uyghur life harvest looks
G4: xinjiang cotton #xinjiang beautiful very mechanized cute girl mechanization planting
G5: xinjiang cotton #xinjiang beautiful mechanized very cute mechanization planting handsome
G6: xinjiang cotton #xinjiang beautiful handsome mechanized mechanization saved uncle xinjiang's
G7: xinjiang cotton #xinjiang beautiful cute cristiano very girl world handsome
G10: xinjiang cotton #xinjiang beautiful handsome whitehouse45 xinjiang's uncle planting fully
G9: xinjiang #xinjiang cotton beautiful very brittan75511188 handsome rihanna savagexfenty elonmansur
G8: xinjiang food #xinjiang cotton go experience local very delicious beautiful
G11: cotton super #xinjiang harvesting efficient quality take close look machine
G13: cotton xinjiang mechanization #xinjiang saved time planting lot life dawnpat43307574
G12: cotton xinjiang #xinjiang mechanization saved bbcworld good lot planting life
G14: #xinjiang #humanrights cotton #uyghur #forcedlabor #cotton xinjiang co fields meat
G15: xinjiang cotton #xinjiang field more quality good golden autumn october

The large picture image above depicts the overall network with two Chinese official accounts at its center. It depicts the massive network of signal boosters (both core and periphery) that exist purely to reply to and like the tweets of the user they fixate on. The cellular groups that we focused on were a bit more sophisticated; pictured below, they exist slightly off of the beaten path of the network, forming small pods of users that will post videos using trending hashtags revolving around major Chinese global issues; with the intent of either setting the narrative, or drowning out others who wish to speak about the atrocities in the region. These pods generally contain one or two core (hacked) accounts to add a sense of legitimacy, they are surrounded by easily identifiable, expendable, burner accounts that repost the videos and comments shared by the core accounts. Likely, erring on the side of caution, and in an attempt to maximize the lifespan of the accounts, they generally post 2-3 times on the day they choose to post, but will wait approximately five days before posting again. This tactic worked for awhile until Twitter deleted hundreds of these accounts recently.

Core accounts are those accounts in each pod that interact with at least three other disinformation accounts (usually in their own pod but not always). These accounts also conduct most of the messaging that goes out from the network mentioning or replying to non-network accounts. They make up about half of our sample network. A pod of accounts that is heavily laden with core accounts looks a lot like the graph below. Notice the interconnectivity of the accounts. This pod is nearly entirely core accounts with the isolates being accounts in the west or India that they target for replies or mentions.



Peripheral accounts tend to touch two or fewer disinformation accounts, and are rarely the source of replies or mentions to accounts outside the disinformation network (although it does happen). These accounts do produce decent numbers of tweets and are used for narrative flooding. Sometimes, core accounts retweet their content, although this is not a common tactic. These accounts seem to exist to daisy chain certain pods together. The pod below is an excellent example of this sort of structure.



Reciprocation of messaging, meaning the accounts in a pod communicated to each other in both directions in a given tweet exchange was rare. In all but one instance, reciprocation occurred between accounts in the same pod. In total there were 28 pairs of tweets that involved message reciprocity. Those 56 total messages represent just over 5% of the 920 total tweets we examined in this network.

Retweeting of disinformation accounts in the pods by other accounts in the same pod was also a surprisingly small portion of network activity. Retweets made up 47 of 920 posts. Retweets were nearly universally from peripheral accounts. Three of our 14 pods utilized retweets. Two of the more prolific pods retweeted content from official Chinese spokespersons. A smaller pod that was generally self-amplifying retweeted a tweet from an Arkansas based cotton farmer. The tweet the disinformation accounts retweeted from Arkansas was: “Cotton and peanut harvests are in full swing.” This could be an attempt to create additional credibility in the cotton community worldwide.

Over the course of the last week it is becoming evident that the overall structure of the network is becoming more expendable in nature; the amount of time and effort to make core accounts work has become less worth it than the creation of thousands of peripheral accounts as they are taken down. Recent trends of the pods are shown to separate into groups that will focus on certain categories (music, media, government officials, etc.). They will signal boost CCP officials and media outlets in hopes of getting a verified account with a large following to interact with it. Even a single retweet or statement from a celebrity can expand the reach to millions of users who look to their idols for guidance. Most of this focus has been to western powerhouses (United States and United Kingdom), but also to dedicate a lot of manpower into the countries that have the highest regional strategic influence (South Korea and Japan).

We noted that only certain accounts in these groups, like @jennife12170947, seemed to have been used to get the message out to the broader world. Not all accounts in a given grouping were used in this manner. In the case of @jennife12170947, the account targeted prominent Hong Kong critics of the PRC, stars from India, and other pop stars (Adele, BTS, Justin Bieber, and Ed Sheeran to name a few). This targeted messaging attached to prominent verified accounts was accomplished with mentions and replies. Often the tweets would mention multiple prominent accounts in one post to maximize potential for someone looking for those accounts to see the content.

3.3 Narrative Flooding.

Most of the output from the identified accounts were original tweets that included prominent organic hashtags and hashtag that were promoting. Some output included retweets or quote-tweets of those tweets by other accounts in the network and official Chinese social media accounts. The organic hashtags they were contributing to included some of the biggest in the topic, including #Xinjiang, #Uyghar, and #UygharGenocide.

By examining key hashtags employed in conversations discussing the Uygher people, we can see clear signs of inauthentic platform manipulation of the type described above. We examined all tweets using four different hashtags in the month of October, 2021. A large portion of tweets employing three of these hashtags originated from accounts that have been suspended as of November 20, 2021. In addition, an improbable percentage of tweets using three of these hashtags were created as the first action the tweeting account performed after creation (suggesting the account was likely purpose built).

- #xinjiang:
    - 60,515 total tweets, 18,453 tweets from suspended accounts
    - 3309 tweets as first action, 2508 of these accounts suspended
- #xinjiangcotton:
    - 2374 total tweets , 1219 tweets from suspended accounts
    - 135 tweets as first action, 124 of these accounts suspended
- #uyghur:
    - 20237 total tweets , 4015 tweets from suspended accounts
    - 625 tweets as first action, 566 of these accounts suspended

We also examined a fourth hashtag, #uyghurgenocide. While troll accounts were observed employing this hashtag, they did so to less of an extent. We include data for this hashtag as a juxtaposition to the more clearly manipulated hashtags.

- #Uyghurgenocide:

- 17702 total tweets, 317 tweets from suspended accounts
- 27 tweets as first action, 1 of these accounts suspended

The underlying motivation for this flooding behavior is not entirely clear. The accounts could be trying to simply overwhelm the organic conversations in these hashtags with the PRC propagandistic narratives. That motivation would be consistent with past behavior by PRC-aligned accounts on topics like the NBA or Hong Kong.

But for the smaller hashtags, like #XinjiangCotton, this doesn't seem like it could be the whole story. On these hashtags, there is not much organic content to disrupt or overwhelm. Instead, this strategy is, perhaps, preemptive. By flooding a hashtag with high levels of low-quality content from obviously inauthentic accounts, this campaign may be able to lower the likelihood that it will organically trend. In a sense, it could take advantage of Twitter's safeguards against inauthentic hashtag promotion, by purposefully setting off the triggers, short-circuiting any push toward organic promotion.