# Retweeting through the great firewall

A persistent and undeterred threat actor

Dr Jake Wallis, Tom Uren, Elise Thomas, Albert Zhang, Dr Samantha Hoffman, Lin Li, Alex Pascoe and Danielle Cave





#### About the authors

**Dr Jacob Wallis** is a Senior Analyst working with the International Cyber Policy Centre.

**Tom Uren** is a Senior Analyst working with the International Cyber Policy Centre.

**Elise Thomas** is a Researcher working with the International Cyber Policy Centre.

Albert Zhang is a Research Intern working with the International Cyber Policy Centre.

**Dr Samanthan Hoffman** is an Analyst working with the International Cyber Policy Centre.

Lin Li is a Researcher working with the International Cyber Policy Centre.

**Alex Pascoe** is a Research Intern working with the International Cyber Policy Centre.

**Danielle Cave** is Deputy Director of the International Cyber Policy Centre.

#### **Acknowledgements**

ASPI would like to thank Twitter for advanced access to the takedown dataset that formed a significant component of this investigation. The authors would also like to thank ASPI colleagues who worked on this report.

#### What is ASPI?

The Australian Strategic Policy Institute was formed in 2001 as an independent, non-partisan think tank. Its core aim is to provide the Australian Government with fresh ideas on Australia's defence, security and strategic policy choices. ASPI is responsible for informing the public on a range of strategic issues, generating new thinking for government and harnessing strategic thinking internationally.

### **ASPI International Cyber Policy Centre**

ASPI's International Cyber Policy Centre (ICPC) is a leading voice in global debates on cyber and emerging technologies and their impact on broader strategic policy. The ICPC informs public debate and supports sound public policy by producing original empirical research, bringing together researchers with diverse expertise, often working together in teams. To develop capability in Australia and our region, the ICPC has a capacity building team that conducts workshops, training programs and large-scale exercises both in Australia and overseas for both the public and private sectors. The ICPC enriches the national debate on cyber and strategic policy by running an international visits program that brings leading experts to Australia.

## Important disclaimer

This publication is designed to provide accurate and authoritative information in relation to the subject matter covered. It is provided with the understanding that the publisher is not engaged in rendering any form of professional or other advice or services. No person should rely on the contents of this publication without first obtaining advice from a qualified professional.

#### **ASPI**

Tel +61 2 6270 5100

Email enquiries@aspi.org.au

www.aspi.org.au

www.aspistrategist.org.au



■ @ASPI\_ICPC

#### © The Australian Strategic Policy Institute Limited 2020

This publication is subject to copyright. Except as permitted under the *Copyright Act* 1968, no part of it may in any form or by any means (electronic, mechanical, microcopying, photocopying, recording or otherwise) be reproduced, stored in a retrieval system or transmitted without prior written permission. Enquiries should be addressed to the publishers. Notwithstanding the above, educational institutions (including schools, independent colleges, universities and TAFEs) are granted permission to make copies of copyrighted works strictly for educational purposes without explicit permission from ASPI and free of charge.

First published June 2020.

ISSN 2209-9689 (online) ISSN 2209-9670 (print)

Cover image: Illustration by Badiucao/https://www.badiucao.com.



# Retweeting through the great firewall

A persistent and undeterred threat actor

Dr Jake Wallis, Tom Uren, Elise Thomas, Albert Zhang, Dr Samantha Hoffman, Lin Li, Alex Pascoe and Danielle Cave

# Contents

Key takeaways	03
The Chinese party-state and influence campaigns	05
Methodology	06
The Twitter takedown data	07
The size and scope of the operation	07
Self-reported account location	80
Accounts in the operation	09
Engagement	09
Retweet network analysis	10
Narrative analysis: four threats to the CCP	11
Conclusion	19
Appendix	20
Attribution and evidence	20
Tactics and techniques	20
Narratives	36
A continuing operation?	52
Notes	57
Acronyms and abbreviations	57

# Key takeaways

This report analyses a persistent, large-scale influence campaign linked to Chinese state actors on Twitter and Facebook. This activity largely targeted Chinese-speaking audiences outside of the Chinese mainland (where Twitter is blocked) with the intention of influencing perceptions on key issues, including the Hong Kong protests, exiled Chinese billionaire Guo Wengui and, to a lesser extent Covid-19 and Taiwan. Extrapolating from the takedown dataset, to which we had advanced access, given to us by Twitter, we have identified that this operation continues and has pivoted to try to weaponise the US Government's response to current domestic protests and create the perception of a moral equivalence with the suppression of protests in Hong Kong.

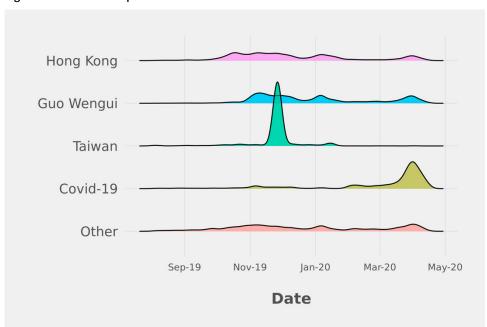


Figure 1: Normalised topic distribution over time in the Twitter dataset

Our analysis includes a dataset of 23,750 Twitter accounts and 348,608 tweets that occurred from January 2018 to 17 April 2020 (Figure 1). Twitter has attributed this dataset to Chinese state-linked actors and has recently taken the accounts contained within it offline.

In addition to the Twitter dataset, we've also found dozens of Facebook accounts that we have high confidence form part of the same state-linked information operation. We've also independently discovered—and verified through Twitter—additional Twitter accounts that also form a part of this operation. This activity appears to be a continuation of the campaign targeting the Hong Kong protests, which ASPI's International Cyber Policy Centre covered in the September 2019 report *Tweeting through the Great Firewall* and which had begun targeting critics of the Chinese regime in April 2017.

Analysing the dataset as a whole, we found that the posting patterns of tweets mapped cleanly to working hours at Beijing time (despite the fact that Twitter is blocked in mainland China). Posts spiked through 8 a.m.–5 p.m. working hours Monday to Friday and dropped off at weekends. Such a regimented posting pattern clearly suggests coordination and inauthenticity.

The main vector of dissemination was through images, many of which contained embedded Chinese-language text. The linguistic traits within the dataset suggest that audiences in Hong Kong were a primary target for this campaign, with the broader Chinese diaspora as a secondary audience. There is little effort to cultivate rich, detailed personas that might be used to influence targeted networks; in fact, 78.5% of the accounts in Twitter's takedown dataset have no followers at all. There's evidence that aged accounts—potentially purchased, hacked or stolen—are also a feature of the campaign. Here again, there's little effort to disguise the incongruous nature of accounts (from Bangladesh, for example) posting propaganda inspired by the Chinese Communist Party (CCP). While the takedown dataset contains many new and low-follower accounts, the operation targeted the aged accounts as the mechanism by which the campaign might gain traction in high-follower networks.

The operation has shown remarkable persistence to stay online in various forms since 2017, and its tenacity has allowed for shifts in tactics and the narrative focus as emerging events—including the Covid-19 pandemic and US protests in May and June 2020—have been incorporated into pro-Chinese government narratives.

Based on the data in the takedown dataset, while these efforts are sufficiently technically sophisticated to persist, they currently lack the linguistic and cultural refinement to drive engagement on Twitter through high-follower networks, and thus far have had relatively low impact on the platform. The operation's targeting of higher value aged accounts as vehicles for amplifying reach, potentially through the influence-for-hire marketplace, is likely to have been a strategy to obfuscate the campaign's state-sponsorship. This suggests that the operators lacked the confidence, capability and credibility to develop high-value personas on the platform. This mode of operation highlights the emerging nexus between state-linked propaganda and the internet's public relations shadow economy, which offers state actors opportunities for outsourcing their disinformation propagation.

Similar studies support our report's findings. In addition to our own previous work *Tweeting through the Great Firewall*, Graphika has undertaken two studies of a persistent campaign targeting the Hong Kong protests, Guo Wengui and other critics of the Chinese Government. Bellingcat has also previously reported on networks targeting Guo Wengui and the Hong Kong protest movement. Google's Threat Analysis Group noted that it had removed more than a thousand YouTube channels that were behaving in a coordinated manner and sharing content that aligned with Graphika's findings.

This large-scale pivot to Western platforms is relatively new, and we should expect continued evolution and improvement, given the enormous resourcing the Chinese party-state can bring to bear in aligning state messaging across its diplomacy, state media and covert influence operations. The coordination of diplomatic and state media messaging, the use of Western social media platforms to seed disinformation into international media coverage, the immediate mirroring and rebuttal of Western media coverage by Chinese state media, the co-option of fringe conspiracy media to target networks vulnerable to manipulation and the use of coordinated inauthentic networks and undeclared political ads to actively manipulate social media audiences have all been tactics deployed by the Chinese Government to attempt to shape the information environment to its advantage.

The disruption caused by Covid-19 has created a permissive environment for the CCP to experiment with overt manipulation of global social media audiences on Western platforms. There's much

to suggest that the CCP's propaganda apparatus has been watching the tactics and impact of Russian disinformation.

The party-state's online experiments will allow its propaganda apparatus to recalibrate efforts to influence audiences on Western platforms with growing precision. When combined with data acquisition, investments in artificial intelligence and alternative social media platforms, there is potential for the normalisation of a very different information environment from the open internet favoured by democratic societies.

This report is broken into three sections, which follow on from this brief explanation of the dataset, the context of Chinese party-state influence campaigns and the methodology. The first major section investigates the tactics, techniques and operational traits of the campaign. The second section analyses the narratives and nuances included in the campaign messaging. The third section is the appendix, which will allow interested readers to do a deep dive into the data.

ASPI's International Cyber Policy Centre received the dataset from Twitter on 2 June and produced this report in 10 days.

#### The Chinese party-state and influence campaigns

The Chinese party-state has demonstrated its willingness to deploy disinformation and influence operations to achieve strategic goals. For example, the CCP has mobilised a long-running campaign of political warfare against Taiwan, incorporating the seeding of disinformation on digital platforms. And our September 2019 report—*Tweeting through the Great Firewall*—investigated state-linked information campaigns on Western social media platforms targeting the Hong Kong protests, Chinese dissidents and critics of the CCP regime.

Since *Tweeting through the Great Firewall*, we have observed a significant evolution in the CCP's efforts to shape the information environment to its advantage, particularly through the manipulation of social media. Through 2018 and 2019 we observed spikes in the creation of Twitter accounts by Chinese Ministry of Foreign Affairs spokespeople, diplomats, embassies and state media.

To deflect attention from its early mishandling of a health and economic crisis that has now gone global, the CCP has unashamedly launched waves of disinformation and influence operations intermingled with diplomatic messaging. There are prominent and consistent themes across the messaging of People's Republic of China (PRC) diplomats and state media: that the CCP's model of social governance is one that can successfully manage crises, that the PRC's economy is rapidly recovering from the period of lockdown, and that the PRC is a generous global citizen that can rapidly mobilise medical support and guide the world through the pandemic.

The trends in the PRC's coordinated diplomatic and state-media messaging are articulated as a coherent strategy by the Chinese Academy of Social Sciences, which is a prominent PRC-based think tank. The academy has recommended a range of responses to Western, particularly US-based, media criticism of the CCP's handling of the pandemic, which it suggests is designed to contain the PRC's global relationships. The think tank has offered several strategies that are being operationalised by diplomats and state media:

- the coordination of externally facing communication, including 24 x 7 foreign media monitoring and rapid response
- the promotion of diverse sources, noting that international audiences are inclined to accept independent media
- support for Chinese social media platforms such as Weibo, WeChat and Douyin
- enhanced forms of communication targeted to specific audiences
- the cultivation of foreign talent.

The party-state appears to be allowing for experimentation across the apparatus of government in how to promote the CCP's view of its place in the world. This study suggests that covert influence operations on Western social media platforms are likely to be an ongoing element of that project.

### Methodology

This analysis used a mixed-methods approach combining quantitative analysis of bulk Twitter data with qualitative analysis of tweet content. This was combined with independently identified Facebook accounts, pages and activity including identical or highly similar content to that on Twitter. We assess that this Facebook activity, while not definitively attributed by Facebook itself, is highly likely to be a part of the same operation.

The dataset for quantitative analysis was the tweets from a subset of accounts identified by Twitter as being interlinked and associated through a combination of technical signals to which Twitter has access. Accounts that appeared to be repurposed from originally legitimate users are not included in this dataset, which may potentially skew some analysis.

This dataset consisted of:

- account information for 23,750 accounts that Twitter suspended from its service
- 348,608 tweets from January 2018 to 17 April 2020
- 60,486 pieces of associated media, consisting of 55,750 images and 4,736 videos.

Many of the tweets contained images with Chinese text. They were processed by ASPI's technology partner in the application of artificial intelligence and cloud computing to cyber policy challenges, Addaxis, using a combination of internal machine-learning capabilities and Google APIs before further analysis in R. The R statistics package was used for quantitative analysis, which informed social network analysis and qualitative content analysis.

Research limitations: ASPI does not have access to the relevant data to independently verify that these accounts are linked to the Chinese Government. Twitter has access to a variety of signals that are not available to outside researchers, and this research proceeded on the assumption that Twitter's attribution is correct. It is also important to note that Twitter hasn't released the methodology by which this dataset was selected, and the dataset doesn't represent a complete picture of Chinese state-linked information operations on Twitter.

# The Twitter takedown data

### The size and scope of the operation

Figure 2: Tweets per month, by tweet language

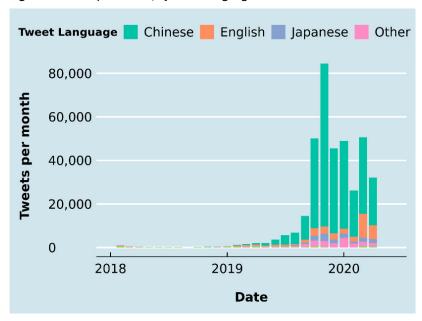


Figure 2 displays tweets per month by tweet language (as determined by Twitter) over the entire tweet dataset. The earliest tweets are from January 2018, with the latest occurring on 17 April 2020. Ninety-one per cent of all tweets were posted between October 2019 and April 2020, and 79% were in Chinese.

Figure 3: Tweets, by follower count and language, over time

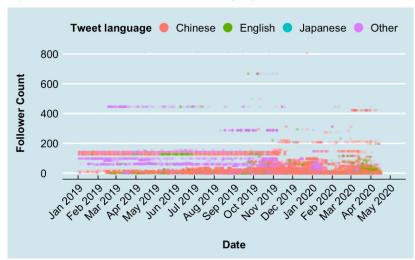


Figure 3 is another view of the dataset of tweets from January 2019 to May 2020. Individual tweets are plotted against follower count over time, coloured with the tweet language as determined by Twitter. This graph shows that the vast majority of tweeting occurs from low-follower accounts.

## Self-reported account location

Figure 4: Map of self-reported locations

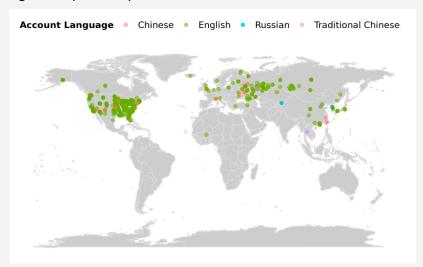


Figure 4 is a map of the self-reported locations of the accounts in the Twitter dataset, colour-coded for the language they tweeted in. Only 3% of all accounts provided location details and, being self-reported, those locations do not necessarily reflect the location of the account holder. Note that spam accounts often falsely self-report locations in the US because it's a high-value market.

Figure 5: Tweet activity, by hour of day

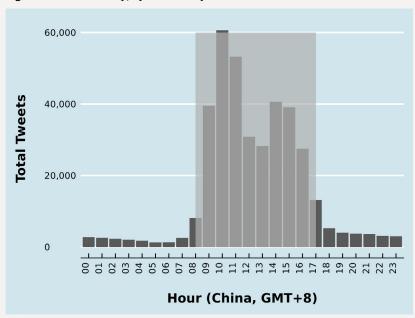


Figure 5 shows the aggregate tweet activity per hour for the entire dataset when displayed in Beijing time (GMT +8). Business hours are shaded in grey. The pattern of behaviour doesn't match the self-reported locations and is consistent with Twitter's assessment that this is a state-backed campaign.

Tweet Language

40,000

20,000

Mon Tue Wed Thu Fri Sat Sun

Day of Week

Figure 6: Tweet activity, by day of week and tweet language

Figure 6 displays the total number of tweets in the dataset by day of week, coloured by tweet language. Tweets occur mostly during weekdays, with about half as much activity during weekends.

# Accounts in the operation

Twitter released 23,750 accounts in the dataset associated with this operation.

Strikingly, the vast majority of the accounts—18,638 of 23,750—had no followers, and 95% had fewer than eight followers (Table 1).

Table 1: Follower numbers for accounts

No. of followers	Accounts	Cumulative total	Percentage of accounts
0	18,638	18,638	78.5%
1	1,605	20,243	85.2%
2	754	20,997	88.4%
3	482	21,479	90.4%
4	353	21,832	91.9%
5	307	22,139	93.2%
6	226	22,365	94.2%
7	163	22,528	94.9%
8	124	22,652	95.4%

# Engagement

Despite very low numbers of followers, some tweets received significant engagement. The high level of engagement is likely to reflect the use of commercial bot networks, as is discussed in later sections of this report. Organic engagement with the campaign appears to have been very low. One hundred and fifty-six tweets from accounts with no followers received more than 50 likes, and another 26 tweets from accounts with no followers received more than 10 retweets.

## Retweet network analysis

Figure 7: Retweet network for Hong Kong-related tweets

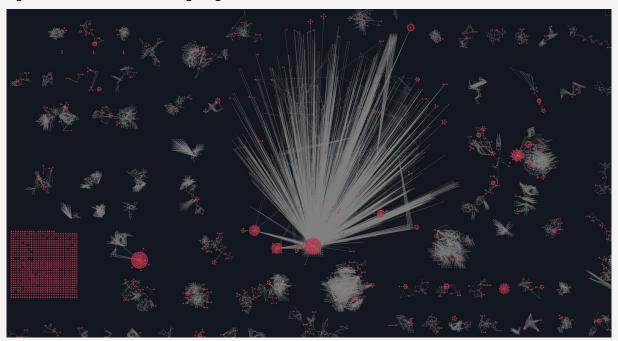


Figure 7 is a Constellation<sup>2</sup> graph of the retweet network of Hong Kong related tweets. This graph shows the disconnected nature of the tweets in this dataset—there are many isolated networks that don't engage with each other in a coordinated fashion (the red square on the left-hand side, for example, is all the accounts that are connected by retweets to only one other account, whereas the circles in the middle represent highly active nodes). The repurposed accounts with high follower counts that weren't included in this dataset may have been used to connect the graph.

#### **Tactics and techniques**

This is a persistent, large-scale operation on Twitter involving at least 23,750 accounts over a sustained period of time. During the course of this research, additional accounts on Twitter and a number of Facebook accounts were also found to be sharing identical or highly similar content. This, in combination with other indications suggesting inauthentic activity (for example, changing from a profile picture of a Bangladeshi man to a beautiful Chinese woman and abruptly switching to posting in Chinese in opposition to the Hong Kong protests) leads us to assess that these accounts are also likely part of a multiplatform campaign. Our analysis found as follows:

- The campaign used a combination of repurposed accounts (which may have been purchased in bulk from commercial resellers) and newly created accounts on both Twitter and Facebook. Some of these accounts have been active in targeting the Hong Kong protests and Guo Wengui at least as far back as early 2019.
- It is important to note that Twitter has removed repurposed (including potentially hacked) accounts from the dataset for distribution. While necessary to protect the privacy of former users, this affects the results of our research. Repurposed (and possibly hacked) accounts constitute a significant proportion of the activity on Facebook.

- At the time of writing, the campaign appears to be persisting despite takedowns by both Twitter and Facebook, including by repurposing accounts and creating new accounts to replace those removed by the platforms, and by developing new narratives in response to emerging events such as the George Floyd protests in the US in June 2020.
- Organic engagement with the campaign appears to have been very low on both platforms.
- Campaign operators on Twitter appear to have purchased engagement from a commercial provider, which helped to boost specific tweets to the top of Twitter's 'Top' search for specific hashtags. This allowed the campaign to effectively drown out organic activity on that hashtag for a period of time.
- On Twitter, 90% of tweets were made between 8 a.m. and 5 p.m. (Beijing time). Although it isn't surprising that Chinese-language tweets are made during the day in China, the relative lack of early morning or evening activity suggests that this activity isn't organic personal use of the platform.
- Fifty-two per cent of the 348,608 tweets in the dataset from Twitter are retweets. The large majority are in Chinese, and a minority in English. The English-language tweets are usually accompanied by Chinese-language hashtags and a graphic including images and several paragraphs of Chinese text.
- The campaign's content production appears to have favoured scale and speed over quality. A large proportion of the images are poorly made; translations into English and Cantonese appear to have been automatic, and many have errors; and minimal effort has been made to make repurposed accounts appear authentic.

For detailed examples that explain this information operation's tactics and techniques, see page 20 in the appendix.

## Narrative analysis: four threats to the CCP

This report focuses on the four most common themes found in the takedown dataset: Hong Kong, Guo Wengui, Covid-19 and Taiwan. Within each, we identified several recurring sub-themes of these removed tweets, which were as follows (full analysis and details are in the appendix:

- · Hong Kong
  - (1) Accusation that protesters are violent
  - (2) Demands to support Hong Kong police, and a related effort to dehumanise protesters
  - (3) Allegations of US interference
- Guo Wengui
  - (1) Focus on the relationship between Guo Wengui and Steve Bannon
  - (2) Accusations that Guo is a liar
  - (3) Accusations that Guo is immoral and a rapist
- Taiwan
  - (1) The Taiwan election
  - (2) Content on the Wang Liqiang case

#### • Covid-19

- (1) Overlap with Hong Kong and Guo Wengui content, and to a minimal extent with Taiwan content
- (2) Claim that China's pandemic management has been an exemplar, especially compared to that of the US.

In addition, our investigation found that this operation has shifted focus to incorporate current US domestic protests and civil unrest within pro-China narratives. Our assessment is that the aim of this shift is to encourage among target audiences the perception of moral equivalence between the US Government's response to its own domestic protests and the suppression of protest in Hong Kong.

The dataset is a window into what the CCP perceives as threats. The thematic issues in the takedown dataset reflect threat perceptions that the Chinese party-state would label as 'internal' affairs, but the issues play out on a global stage, engaging with international audiences and attracting international reactions. For the CCP, this means that its information operations must successfully shape, manage and control narratives outside the PRC as well as inside. This was reflected in the takedown data in terms of both messaging and the apparent intended audience.

In the takedown dataset, a notable feature of Covid-19 content was that it often intersected with narratives on Guo Wengui, Hong Kong and Taiwan. On Guo, one Tweet read that 'Since February, the coronavirus epidemic has become Guo Wengui's focus. Over the past few days, he has been trying to spread rumours and ridiculous lies with exaggerated data, while he has mobilised others to spread rumours'.

On Hong Kong, several posts suggested that pro-democracy protestors interfered with the government's ability to manage Covid-19. For instance, an image (Covid 1) of the Chairman of Wuhan General Hospital praising doctors was posted with the tweet: 'At this time of crisis, Hong Kong should be united, but the anti-epidemic period in Hong Kong is chaotic, and chaotic elements still use despicable means to gain political and monetary benefits ... The evils of chaos in HK will surely be rejected by the people...' Meanwhile, on Taiwan, attempts were made to counter the view that Taiwan's Covid-19 response was successful while the PRC's was not. One Tweet was posted with an image (Covid 2) featuring Chinese text that praised the Chinese Government's response, saying that 'China's response to fighting epidemics has been the best in the world... [while] Taiwan's response was learned from mainland China.' In a reference to the US, it also pointed to how other Western countries had performed poorly in their Covid-19 responses.

Another feature of the posts was that the intended audiences appeared varied. Some texts were composed in simplified Chinese characters, some were composed in traditional Chinese characters mixed with special Cantonese characters, and some appeared to have been first composed in simplified Chinese characters (official in mainland China) and then converted into traditional Chinese with Cantonese characters added. The images appeared to be created by both native and non-native speakers, so there were different fonts and occasional mistakes. The English-language messaging always tended to remain consistent with the Chinese language messaging. An example of English-language messaging is the suggestion that the US is interfering in Hong Kong, even promoting violence. For instance, an image (US 1) shared in an English-language version and a version with

traditional Chinese characters showed 'black hands' over a burning Hong Kong with the US Capitol building above it and the text 'Hong Kong Human Rights and Democracy Act,' and 'US moves will escalate voilence [sic] in Hong Kong'. Another (US 2) showed just the 'black hand' with a US flag and the words 'naked act of hegemony' on it (Figure 8).

Figure 8: 'Naked act of hegemony'



## **Evolving narrative targeting the US protests**

This ongoing activity follows a similar pattern to that seen in the Twitter dataset and in the independently identified Twitter and Facebook activity, in which emerging events and global news stories are incorporated into existing narratives in a way that supports the overall objectives of the Chinese state, particularly in relation to the Hong Kong protests.

Figure 9: Screenshot of Facebook post, 'American reality of blood and fire', captured 8 June 2020



This has included incorporating the evolving protest movement in the US and across the world that followed the killing of George Floyd by Minneapolis police into a broader anti-US, anti-democracy, anti-protest and pro-Hong Kong police narrative (Figure 9).

The examples in Figure 10 (with similar but not identical content, tweeted about two hours apart on 2 June) accuse the US of hypocrisy for its criticism of the response by police to protests in Hong Kong, while the US's own police and troops use violence against protests in the US, and warns Hong Kong protesters not to think they can rely on the US for support against China's national interests. Note the image on the right of the Statue of Liberty kneeling on Floyd's neck.

Figure 10: The US accused of hypocrisy, tweeted on 2 June



This activity has also included drawing on the themes of systemic racism in the US (perhaps partially as a distraction from recent scandals involving anti-black racism in China). The example in Figure 11 describes the death of George Floyd, and impacts of systemic racism and police violence against black people in the US, and Donald Trump's intention to suppress the protests with force (including driving cars into civilians) rather than talking to protesters. This image depicts Officer Chauvin holding an American flag as he kneels on Floyd's neck.

Figure 11: The US accused of racism



Activity on Facebook has also included the themes of racial discrimination as a reflection on wider social injustice in the US (note that some Chinese auto-translation software translates George Floyd's name as 'Freud', as in the comment in the example in Figure 12).

Figure 12: The US accused of racism, again



The activity has sought to convince Hong Kong protesters that the US can't be relied upon to support them; that its claims of democracy, the rule of law and freedom are 'three big lies'; and to stop being naive in trusting US politicians (Figure 13).

Figure 13: 'Don't trust the US'



As of 8 June, activity that appears highly likely to be linked to the campaign dominated the #美国骚乱 (#AmericanRiots) hashtag on Facebook (checked across multiple accounts), while active moderation on Twitter appeared to be largely, although not completely, successful in removing many of the campaign's tweets (Figure 14).

Figure 14: The #AmericanRiots hashtag



The ongoing evolution of this activity speaks to both the campaign's persistence and its resourcing. Despite the removal of tens of thousands of accounts since the activity began in 2019, the campaign operators have continued to acquire new accounts, develop new content, and adapt their existing narratives to incorporate emerging events. This is a significant investment in time, labour and budget.

Go to page 36 in the appendix for our in-depth analysis of the key narratives found in the dataset.

# Conclusion

This operation demonstrates the Chinese state's willingness to invest in ongoing efforts to manipulate social media audiences in Hong Kong and the Chinese diaspora at scale. The Hong Kong protests and Covid-19 pandemic have created a combination of crises that the party-state must contain if it's to maintain China's domestic stability and international standing. In addition, the pandemic has created an environment in which the party-state can experiment with and refine the operationalisation of propaganda and strategic disinformation in an environment with which it's unfamiliar.

There are second-order effects to this experimentation. We've reported on the phenomenon of nationalistic trolling that often aligns with the messaging of the party-state, without necessarily being explicitly state-directed. As the Chinese state leans into Western platforms, it creates opportunities to mobilise supporters from within the Chinese diaspora. That mobilisation is a fundamental tenet of the party-state's view of security and is calculated to exploit the capacity of diaspora communities to extend the party's influence overseas.

This effort to influence global public opinion is ongoing and is unlikely to relent. The Chinese party-state is invested in information management as a fundamental pillar of its global engagement. The CCP places enormous importance on propaganda domestically and is projecting this approach in its global engagement in order to secure strategic positioning on its own terms.

As the President of the US, Donald Trump, continues to engage in misinformation as a tool of domestic political influence—and with recent tweets by both the President and the White House labelled for violating Twitter's rules against glorifying violence—the US administration risks losing ground as a credible, consistent and trustworthy international voice seeking to call out disinformation and propaganda. While mitigated to an extent by the diverse voices and local, state and federal institutions in the US, this erosion in trust and credibility coincides with the Chinese party-state's pivot to using large-scale cross-platform strategic disinformation and influence operations on major Western social media networks, including Twitter, Facebook and YouTube.

In this new environment, it's essential that middle powers, civil society groups and social media operators step up. They must work more collaboratively to call out and counter state and non-state actors who are seeking, with persistence and growing sophistication, to manipulate our information environment. Knowledge of the activities and methods, as set out in this report, provides a foundation for action.

# **Appendix**

#### Attribution and evidence

The US Department of State's Global Engagement Center (GEC) has suggested that there are coordinated covert efforts to manipulate global social media in order to amplify PRC diplomats and state media. Our internal data analytics platform—ASPI's *Influence Tracker*—has noted unusual patterns in the retweet reach of particular diplomatic Twitter accounts. There's an emerging body of evidence to support the GEC's position. A study by the Digital Forensic Center identified a rapidly growing network of automated social media accounts praising the PRC's distribution of medical aid to Serbia and unfavourably contrasting the response of the EU. A similar study noted analogous bot-driven inauthentic activity attempting to drive pro-China and anti-EU sentiment on Italian social media, including artificial boosting of the Twitter account of the PRC's Italian embassy. In another investigation, Propublica analysed a network of more than 10,000 stolen and faked Twitter accounts pushing messaging strongly aligned with PRC interests. Propublica linked this network to OneSight (Beijing) Technology Ltd, an internet marketing company that's previously been contracted by state-run China News Service to boost its Twitter following.

There's been industry pushback on some of these assertions. Following its initial review of a sample, Twitter questioned the validity of the dataset from which GEC has drawn its findings. Part of the challenge is in the direct attribution of influence operations as state-directed. Making high-confidence judgements about attribution requires assessments of narrative framing, behavioural patterns, technical signals, targeting and intent. Further muddying the waters is the scale of pro-China patriotic trolling on social media platforms. Much of this activity may have some degree of authenticity while amplifying patriotic narratives at scale, harassing regime critics and demonstrating behavioural traits consistent with the inauthentic coordinated networks deployed by state-driven influence operations (for example, the use of bot networks and coordinated posting patterns). We have previously reported on a range of pro-China, anti-Taiwan patriotic trolling campaigns that have aligned with CCP interests and spiked during periods of geopolitical tension.

### **Tactics and techniques**

#### Twitter data analysis

#### The use of repurposed accounts

While the dataset received from Twitter deliberately excluded repurposed accounts, in the course of this research a number of repurposed accounts on both Facebook and Twitter were identified as being highly likely to be a part of this campaign. For the sake of accurate research and reporting, we opted to include these accounts in our research. All of the account information included in this report is or has been publicly available on Facebook's or Twitter's platforms.

The campaign operators have made significant use of repurposed accounts, which often appear to have originally belonged to users in specific locations. On Facebook, for example, there's been a very notable presence of accounts formerly belonging to users in Bangladesh, while on Twitter there have been a number of French, Russian, Indonesian and other accounts (Figure 15). This geographical distribution is likely to reflect the dynamics of the commercial account resale market; put simply, Bangladeshi or Russian accounts are comparatively cheap and widely available. It appears that these accounts may be being purchased in bulk.

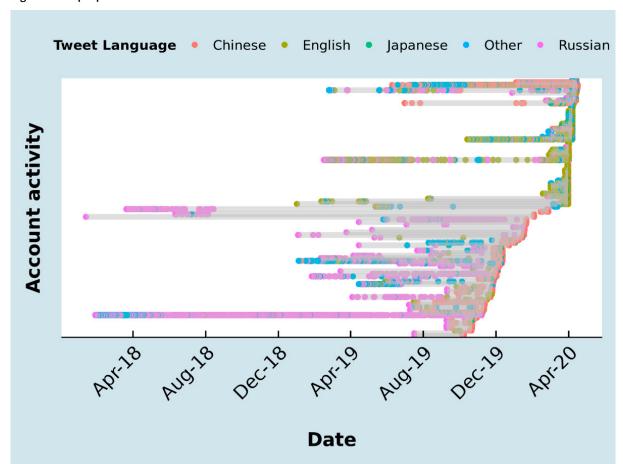


Figure 15: Repurposed Russian accounts

This graph shows repurposed Russian accounts that started tweeting in Chinese from October 2019 onwards. Each row of the plot is the tweeting activity of an individual account. Individual dots represent tweets, coloured-coded by tweet language. Grey lines represent inactive days.

Campaign operators have made minimal efforts in repurposing the accounts, taking little if any care to make the accounts' activity appear authentic. On Twitter, for example, one account abruptly shifted from tweeting in French in March 2020 to tweeting in English and Chinese in April. The account's profile picture appears to have been changed to a picture of a young woman, but the account's bio continues to link to the Facebook page of a French man (Figure 16).

Figure 16: Screenshot of account, captured 7 June 2020



In another example, the account @maerskketones appeared to still be running a script belonging to the bot account's previous owner (Figure 17). This meant that it was simultaneously being utilised in the Chinese language campaign and automatically posting content from a French-language blog. The French-language posts make use of the third-party service OverBlog Kiwi, while the Chinese-language tweets make use of the Twitter web app.

Figure 17: Screenshot of account timeline, captured 7 June 2020



On 15 March an account which had tweeted in French until January 2020, tweeted out 'Test123' in Chinese, before launching into attacks on Guo Wengui (Figure 18). The campaign operators didn't bother to delete either the old French tweets or their own test.

Figure 18: Screenshot of account timeline, captured 7 June 2020



This lazy approach to cleaning up repurposed accounts and removing old content is replicated on Facebook, where campaign operators appear to have purchased a number of accounts and pages formerly belonging to users in Bangladesh. In most cases, the campaign operators changed the account profile pictures but otherwise left the original account owner's content, including their name and personal photographs (figures 19 and 20).

Figure 19: Screenshot of Facebook account, captured 12 May 2020 and deleted following the publication of Bellingcat's analysis



Figure 20: Screenshot of Facebook account captured 12 May 2020. This Facebook account was deleted following the publication of Bellingcat's analysis



An interesting point to note is that some of these accounts have been in the hands of Chinese-language operators since at least early 2019. This includes activity directed against the Hong Kong protests dating back to July 2019; before then, the accounts posted in Chinese about fashion and travel topics. This suggests that they may have passed through the hands of Chinese-language commercial resellers before being repurposed for political content.

As an example, one account appears to have originally belonged to a young man in Bangladesh who posted frequently about soccer. The last post in Bengali was made in December 2018. In March 2019, the account, which now had a profile picture of a woman, began posting in Chinese about fashion, entertainment, lifestyle and travel (Figure 21). Some of the content appears to have come from a fashion blog, and raises the possibility that the account may have been taken over by a Chinese marketing company.

Figure 21: Screenshot of Facebook account, captured 12 May 2020 and deleted following the publication of Bellingcat's analysis



The account switched to political content in late July 2019 and has consistently posted attacks and criticisms of the Hong Kong protests and Guo Wengui since that time. It's unclear whether the account changed hands in July 2019, or whether its operators simply received a new contract from a client interested in undermining the Hong Kong protest movement. A number of other accounts examined show a similar pattern of activity, with attacks on the Hong Kong protest movement beginning on 27 July 2019. Some accounts were removed by Facebook in mid-May following the publication of reports by Bellingcat and Graphika, but others remained active or were newly activated in June 2020 (more on this below).

Although there's no attribution from Facebook at the time of writing, we've found these accounts to be sharing many identical images and media content as included in the data received from and attributed by Twitter. We have also found these accounts in some cases posting in what appears to be

a coordinated way across platforms (discussed further below). Based on circumstantial evidence, it seems highly likely that these accounts form part of the same campaign attributed by Twitter.

In addition to individual accounts, the campaign has also made use of Facebook pages. There doesn't seem to be a significant difference between the content shared on pages as opposed to content on individual accounts. Follower numbers on these pages are too low for Facebook to publicly display details about page managers under the Page Transparency tab.

#### The use of new accounts

In addition to the purchased older accounts, the campaign also appears to have either created or acquired a significant number of new accounts (Figure 22).

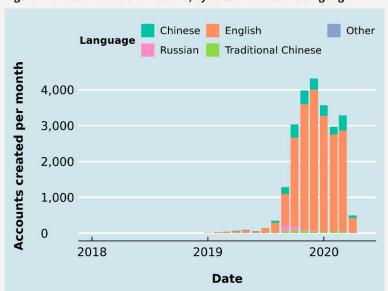
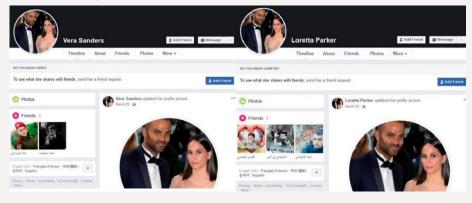


Figure 22: Account creation volume, by date and account language

On Facebook, the newer accounts and pages active in the campaign show considerable variation and may have come from a range of sources. Newer individual accounts have been primarily used to like and share content posted by older accounts. There are a number of accounts that appear to have been created in March with English-language women's names and profile pictures shared across multiple accounts—for example, 'Vera' and 'Loretta' both have a picture of NBA player Tony Parker and his partner taken at an event in Monaco in 2019 as their profile pictures (Figure 23).

Figure 23: Screenshots of 'Vera' and 'Loretta' Facebook accounts, which were used to amplify content associated with the campaign, captured 12 May



Newly created Facebook pages were also used to amplify posts from existing, repurposed accounts. For example, the page 'Banana' was used from mid-2019 onwards to share content from other accounts attacking the Hong Kong protests and Guo Wengui (Figure 24).

Figure 24: Screenshot of 'Banana' Facebook page, captured 7 June 2020 ı Like 🦒 Follow 🏕 Share ···

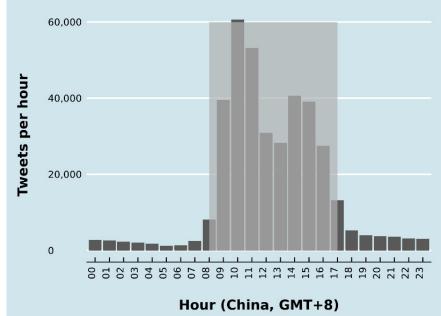


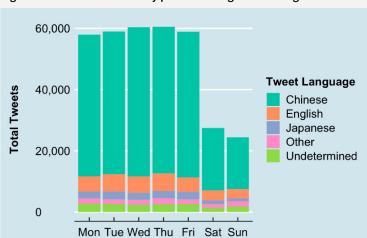
#### Tasking and operations

Ninety per cent of tweets included in the dataset received from Twitter were made from Monday to Friday, and between 8 a.m. and 5 p.m. (Beijing time), with a notable dip around lunchtime (Figures 25 and 26). While it isn't surprising that Chinese-language tweets would be made during the day in China (except for the fact, of course, that Twitter is banned in China), the fact that the activity cleaves so closely to Chinese working hours adds to the evidence that this is an organised and professional operation, as opposed to normal, organic social media activity.

60,000

Figure 25: Tweets were mainly posted between 8 a.m. and 5 p.m., China (GMT+8) time





Day of Week

Figure 26: Tweets were mainly posted during the working week

It appears likely that tasking for this operation involves some form of centralised narrative, perhaps including specific phrases or messaging in mainland Chinese, which individual operators are responsible for translating into both English and Cantonese and supplementing with visual content.

That would explain the diversity in translations of the same message (for example, due to having used different auto-translation services) and why variants on the same message are paired with different visual content within the same 24-hour period. The examples in Figure 27 show variations in English-language content, which is likely to have originated from the Chinese phrase 搬起石头砸自己的脚 (implying to hurt oneself by one's own actions), posted with different translations and different images within a short time of one another.

Figure 27: Tweeting to a shared script





Campaign operators also appear to have relied on automatic translation in places for translating phrases from standard Chinese to Cantonese. For example, the tweet in Figure 28) appears to have been initially written in standard written Chinese in simplified characters (commonly used in mainland China) and then automatically translated to Cantonese, probably by a non-native Cantonese speaker. The text included in the image says that three rioters (referring to Hong Kong protesters) are on the run. The correct expression should be 三人在逃. In the text, however, it is incorrectly written as 三人喺逃. This indicates that the author can' tell the basic usage difference in Cantonese.

Figure 28: Faulty Cantonese points to auto-translation



In the same way that campaign operators appear to be operating off the same prescribed text, they may also be being provided with images to incorporate. In the examples in Figure 29, the same photographs have been incorporated into different .jpg files, tweeted on the same day.

Figure 29: Sharing images



Content creation for this campaign appears to have often favoured speed and scale over quality, particularly as the activity expanded over the course of 2020. While some of the content, particularly that targeting Hong Kong, is of professional quality, much of the more recent content appears to have been hastily put together by campaign operators themselves. This is particularly the case for the .jpg files that combine images and paragraphs of Chinese text. Paragraphs are squished in, and images have been resized, stretched and distorted. Some images even have spell check underlines and resize anchor-points, which are usually present while composing documents (Figure 30). In these cases, the creator seems to have chosen to screenshot rather than export the images, presumably because that's quicker and easier.

Figure 30: Screenshot of campaign image files showing spellcheck underlines and graphic anchor-points



These images, which still include image positioning anchoring and spell check underlines, provide an insight into the production cycle for this content, suggesting it was rapid and lacking in quality control. As a hypothesis, this may reflect that the same individual is in charge of creating this content (based on prescribed text, images, or both) and amplifying it using their own assigned network of accounts. This would also help to explain the many relatively small and largely separate networks seen in Figure 7 on page 10.

#### Inauthentic boosting

Some tweets associated with the campaign received implausibly high, and highly specific, levels of engagement. For example, recently created accounts with almost no followers had tweets with hundreds, even thousands of likes, but barely any comments or retweets (Figure 31). Other tweets from the same account received no engagement at all. This is a pattern that's likely to reflect inauthentic activity, and in this case appears to be linked to services provided by a large commercial bot network, as previously documented on Bellingcat.



Figure 31: Screenshot of Twitter account with 1 follower and 1.1k likes on pinned tweet, captured 12 May 2020

Networks such as this can be hired, often relatively cheaply, to generate the appearance of engagement on social media platforms. An examination of the accounts boosting these tweets in the campaign shows they have also promoted everything from cryptocurrencies to online gaming and diet regimes, in languages from Arabic and Turkish to Russian and Indonesian.

There are several potential reasons why the campaign operators might use this kind of commercial boosting on Twitter. One possible reason is that it allows them to effectively dominate the hashtag, at least for a period of time, and drown out the organic conversation. For example, a search of the #香港 (Hong Kong) hashtag on Twitter (while logged out and with a clean browser installed) on 13 May 2020, found that the first 22 'Top' tweets were associated with the campaign. The use of inauthentic boosting means that, for users casually browsing Twitter, the conversation relating to Hong Kong on the 'Top' search looks markedly different from the more authentic conversation, as reflected by the difference between the 'Top' and 'Latest' searches (Figure 32).

It's important not to conflate effect with intent, however, and we can't conclude whether this was an intentional tactic or simply a convenient but unexpected side-effect for the campaign operators.

Figure 32: Screenshot comparing 'Top' and 'Latest' searches for #香港 (Hong Kong) hashtag, captured 13 May 2020

There's also a possible secondary motive for campaign operators to use paid boosting services. It's always worth remembering that campaign operators often work within similar systems and structures as other, more overt influence actors, such as traditional advertisers and marketers. Campaign operators might use paid boosting for the same reason that advertisers do: it's a quick and easy way to make the numbers look good for the client or the boss.

Latest

#### Cross-platform coordination

There's evidence of coordinated posting both within and across platforms. There's some circumstantial evidence to suspect that individual operators may be responsible for translating and generating their own content and amplifying it out through their own networks of accounts (more on this below).

In the examples in figures 33 and 34, the same content (including both the same translation and the same visual content) was tweeted first on Twitter, where it was amplified using commercial bots. It was posted a few hours later on Facebook, and that post was amplified by other campaign Facebook accounts two hours after that. While obviously not conclusive, this and similar examples may reflect a daily workflow.

Figure 33: Screenshot of Tweet, captured 7 May 2020



Figure 34: Screenshot of Facebook post and shares, captured 7 May 2020





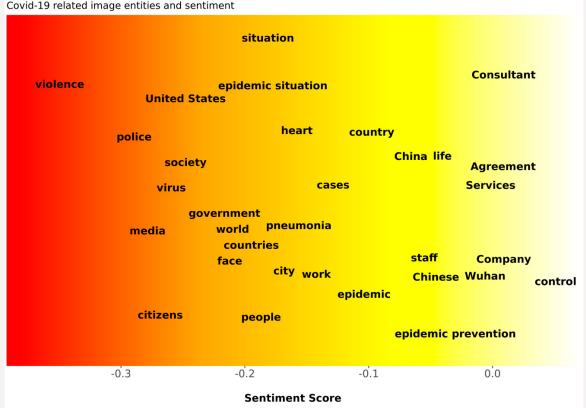
## **Narratives**

#### Covid-19

# Overall size and scope

Tweets associated with Covid-19 and coronavirus totalled to 30,478, and 4,804 images and 387 videos were associated with those tweets. Figure 35 shows a selection of entities extracted from image text and translated into English arranged horizontally by average sentiment score. Entities such as 'Wuhan', 'control' and 'China' were relatively more positive than entities such as the 'United States', 'media' and 'world'.

Figure 35: Covid-19-related entities, by the average sentiment of the translated text entities across images Covid-19 related image entities and sentiment

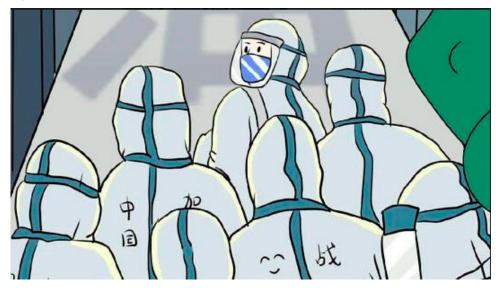


Note: The entities are arranged horizontally by the average sentiment score of the entity across all images. A vertical random jitter is added to distinguish entities. A –1 score indicates negative sentiment, 0 is neutral, and +1 is positive sentiment.

# Narrative 1: China is a leader in pandemic response

Example: In addition to the overlap with related content with Hong Kong, Taiwan and Guo Wengui, many Tweets related to Covid-19 focused on painting the PRC's response to the pandemic as successful. These images also suggested that the PRC was acting as a global leader in coming to the assistance of other countries across the globe battling the virus. Two themes pointed to the CCP's biggest fears about how it handled the outbreak. For example, one Tweet posted with an image (Covid 3) of Chinese healthcare workers (Figure 36) said: 'China has repeatedly demonstrated that "Community of Shared Destiny" is not a slogan. "Community of Shared Destiny" this time is reflected in the sending of precious testing kits and batches of medical materials to other countries; it is reflected in sending medical teams to support other countries without hesitation.'

Figure 36: Cartoon from Tweet about China's Covid-19 response and the 'Community of Shared Destiny'



Analysis: Amid heavy global criticism of the PRC's response to Covid-19 and accusations about the virus's origins, the Party-state has pushed back against these claims by attempting to create an image of the PRC as a responsible actor. At the first meeting of the small leading group on the epidemic, the participants called for 'doing a good job at disseminating knowledge on epidemic prevention' and 'strengthening communication and cooperation with the World Health Organization and relevant countries and regions.' The party-state has attempted to use global offerings of aid and medical supplies to deflect attention from the failures in its response, though that effort too has been met with a lot of criticism due to numerous reports of faulty equipment and subsequent cancellations of contracts and recalling of low-quality equipment.

# **Guo Wengui**

# Overall size and scope

There were 70,584 tweets, 11,832 images and 580 videos in the Twitter dataset that contained keywords associated with Guo Wengui. Figure 37 shows a selection of entities extracted from image text and translated into English.

Guo Wengui related image entities and sentiment lawenforcement SilentMajority policeforce HONGKONG ruleoflaw humanrights colorrevolutionquot eachother LegislativeCouncil ColorRevolutionquot UnitedStates bottomline epidemicsituation YuenLong **HongKongPolice** internalaffairs policeofficers LovePort policeofficer -0.6 -0.4 0.0 0.2

Figure 37: Guo Wengui related entities, by the average sentiment of the translated text entities across images

Note: The entities are arranged horizontally by the average sentiment score of the entity across all images. A vertical random jitter is added to distinguish entities. A –1 score indicates negative sentiment, 0 is neutral, and +1 is positive sentiment.

**Sentiment Score** 

# Narrative 1: Steve Bannon and Guo Wengui are colluding

Examples: One image (Guo 1) depicts Guo standing next to Bannon, who is speaking at a podium with the caption 'With justice on my [Bannon's] side, I am proud to take money from a swindler [Guo].' Another image (Guo 2) shared with high frequency, in both Chinese and English, was linked to an October 2018 Axios exclusive that revealed two contracts (Guo 3) between Guo and Bannon: 'The first contract, signed between Bannon and Guo Media, gives Bannon \$1 million for one year of consulting services beginning in August 2018' and 'The second contract, which was set to begin in August 2019 and is unsigned, offered Bannon \$1 million for consulting and set more specific expectations.' Their relationship is also discussed in the context of Covid-19. For instance, this image (Guo 4) of Guo reads 'Guo Wengui, a deceiver,' and the tweet accompanying it said 'Guo Wengui deliberately put on artificial virus labels to discredit China ... the fact is that when everyone is dedicated to prevention and control, Guo Wengui and Bannon try their best to create resistance'. Other Tweets (Guo 5) using the *Axios* image described a 'coronavirus joint performance' between Guo and Bannon (Figure 38).

Figure 38: Guo Wengui and Steve Bannon's 'coronavirus joint performance'



Analysis: Most posts related to Bannon refer to his reliance on Guo's money, suggesting that the relationship is transactional and that Bannon is happy to be Guo's microphone for a price. Given that China has become a prominent election issue against the backdrop of Covid-19, the relationship between Bannon and Guo is important because Bannon is pushing a hard line on China in a prominent way and is even using the platform to suggest regime change. On 4 June 2020, to mark the anniversary of the Tiananmen massacre, Guo Wengui and Steve Bannon ran a campaign proclaiming the 'Federal State of New China' and flew banners over New York Harbour, with Guo saying on a livestream broadcast from a boat in the harbour 'From today the Chinese Communist Party (CCP) will no longer be the lawful government of China.' On one hand, the party will see these activities driven by 'hostile forces' (Guo and Bannon) as a direct threat, and that explains why he's such a prominent target.

At the same time, however, Guo's intent and whose interests he protects are still unknown. He has claimed that other dissidents, such as Ai Weiwei and Bob Fu, are CCP spies, which posts (Guo 0) in the takedown dataset called out as fabrications. These unknown factors about Guo would shape the broader implications of the campaign against him.

# Narrative 2: Guo Wengui is a liar to be condemned

Examples: Guo is often described in the posts as a 'liar'. One image (Guo 6) used frequently in the Tweets depicts Guo as a rat and describes him as a swindler. Many posts also suggest that Guo is inviting death. For instance, this image (Guo 7) depicts Guo looking unwell and in bed and reads 'l'm afraid I'll have to take my leave first' (seeming to imply suicide), with the character for 'death' written upside down on his forehead. Names written upside down are signifiers of the condemned. During the Cultural Revolution, the same was done with the name of Liu Shaoqi, who was a key revolutionary and party leader and Mao's heir apparent turned enemy. Other posts shared images (Guo 8) of protesters outside Guo's NYC apartment, holding bilingual posters 'Guo Wengui is a big traitor', 'Guo Wengui is a rapist'. 'Guo Wengui need to pay back what he owes', 'Guo Wengui get your butt down from the 18th floor', 'Guo Wengui should get out of [the] US', 'we need to arrest Guo Wengui', and so on (Figure 39). These regular protests outside Guo's apartment are led by Liang Guanjun, who has served as a delegate to the Chinese People's Political Consultative Congress and is chairman of the US East Coast Chinese Association Federation.



Figure 39: Protesters outside Guo Wengui's New York City apartment

Analysis: Guo isn't a simple 'dissident' or 'exile'. His reasons for being in the US may be the combination of being on the losing end of a high-stakes political battle, and perhaps he continues to serve the interests of a key figure in the party-state apparatus. Whatever drives him, the party-state apparatus dedicates an enormous amount of energy to discrediting him, suggesting he's still highly relevant in PRC politics. If he is a traitor, he knows the system well enough for the CCP to be fearful of the consequences of his accusations. But he also appears to be able to exercise power within China from outside, given that his wife and daughter were given permission to leave and visit him in New York. If one assumes that Xi Jinping is in solid control of the Propaganda Department at present (and if the Twitter accounts themselves are linked to the Propaganda department), then it could mean that the decision to invest significant energy into attacking Guo is coming from the highest possible level. At the same time, it's hard to say what the implications of that investment are, because it depends entirely on who he is and whose interests he really represents (his own, or others').

# Narrative 3: Guo is immoral and a rapist

Example: Figure 40 depicts Guo sticking his hand up Lady Liberty's dress, as if he is groping her. She says 'Get your dirty hands off [me],' and he says 'I am doing this really for "freedom".' This image (Guo 10) of Guo with a woman kneeling in front of him suggests they have just engaged in a sexual act and reads 'I am Guo Lewd'. Another frequently shared image (Guo 11) appears to show an ultrasound suggesting that Guo's current assistant, Wang Yanping, is pregnant, suggesting a smear campaign against her.

Figure 40: Depiction of Guo groping the Statue of Liberty



Analysis: The CCP is known to make this accusation against its critics. For instance, *New York Times* op-ed contributor Murong Xuecun published frequently about Beijing's censorship and rights abuses. In 2014, it was reported that a cluster of over 800 Twitter accounts with few followers had started to circulate a series of articles, 'The past and present life of Murong Xuecun', attacking Murong Xuecun. Murong Xuncun wrote about the attacks against him in a 2014 op-ed, saying 'They accused me of molesting little girls. As an adult, they said, I've been a frequent visitor of prostitutes. They went on to say that I've had multiple extramarital affairs and that my infidelities led me to wife beating.' He warned that 'The online campaign against Chinese dissidents is not a problem that affects only a small group of individuals. Rather, it is seen as an attempt to manipulate opinion on a global scale.' Those accused of corruption in the PRC are quite frequently also accused of sexual impropriety, which may or may not be true but serves to paint a negative image of the person's morality. The context of these Tweets is that the party-state has accused Guo of rape, as well as fraud, money laundering, bribery and kidnapping. Additionally, Guo's former secretary has accused him of raping her in his properties in both London and New York. The allegations against Guo can't be independently verified.

# **Hong Kong**

# Overall size and scope

There were 112,280 tweets containing keywords associated with Hong Kong. Media associated with Hong Kong totalled to 16,394 images and 632 videos. Figure 41 shows a selection of entities extracted from image text and translated into English, arranged horizontally by average sentiment score (a computational categorisation of negative or positive valence derived from linguistic tone). What's notable about these entities is that, whenever 'China' was present in image text, it had a general neutral sentiment. Contrast that to 'United States', 'students', 'chaos' and 'thugs', which all generally had a more negative sentiment, noting that 'United States' and 'chaos' had similar sentiments.

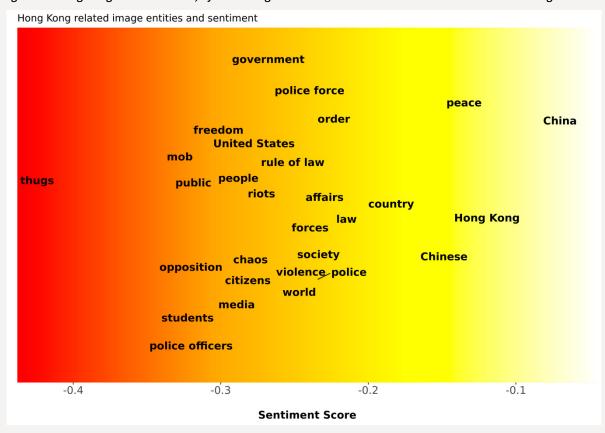


Figure 41: Hong Kong related entities, by the average sentiment of the translated text entities across images

Note: The entities are arranged horizontally by the average sentiment score of the entity across all images. A vertical random jitter is added to distinguish entities. A –1 score indicates negative sentiment, 0 is neutral, and +1 is positive sentiment.

'Police officers' also had a relatively negative sentiment. This is not to say that image text contained negative sentiments about police officers. Instead, image text might have been describing another subject negatively, but the text contained the phrase 'police officers', which the algorithm extracted from the text.

# Narrative 1: Protesters are violent

Examples: A key theme of the tweets on Hong Kong that Twitter removed is the accusation that protesters are violent. In one example (Hong Kong 1), a post written in simplified Chinese characters, with the Chinese state media *Xinhua News* logo at the bottom, reads 'Stop violence and chaos,

safeguard Hong Kong'. In another example (Hong Kong 2), an image of a vandalised China Construction Bank outlet in Hong Kong is accompanied by text written in traditional Chinese characters with Cantonese expressions: 'Destruction is easy, construction is hard. You have destroyed hard work done by generations of Hong Kongers.' The bottom of the image includes a logo with the English slogan: 'Love Hong Kong, Support Police, Oppose Violence'. In such posts, the image of violent pro-democracy protesters is contrasted with pro-CCP protesters calling for peaceful demonstrations and opposing violence. For instance, one photo (Hong Kong 3) shared with a Tweet showed a pro-CCP protester holding a sign that read 'Support ruling Hong Kong by law, protect Hong Kong, oppose violence, love the nation protect Hong Kong, peaceful demonstration.' Notably, another protester in the background of the photo appears to show someone else holding a poster that appears to show a pro-democracy protester attacking an elderly pro-CCP protester, with a headline in English 'how (rioters) treat the elder [sic]' (Figure 42).

Figure 42: Pro-CCP protester's sign



Analysis: In November 2019, Xi Jinping said, 'The most urgent task for Hong Kong at present is to end violence and chaos and restore order,' which was included in some images with Tweets in the dataset (Figure 43). Hong Kong chief executive Carrie Lam has frequently characterised the protesters as violent and destabilising. It's true that protest tactics have increasingly included violence, despite being largely peaceful early on, but the violence isn't one-sided. Police brutality and pro-CCP triad violence are also both significant forms of violence. There's also evidence that false information spread on social media has contributed to episodes of violence. By painting protesters as violent, the Party-state is justifying its own escalations—including the recent introduction of the national security law—as the only options it has left for restoring order in Hong Kong.

Figure 43: A Tweet image promoting Xi Jinping's November 2019 statement on ending violence in Hong Kong



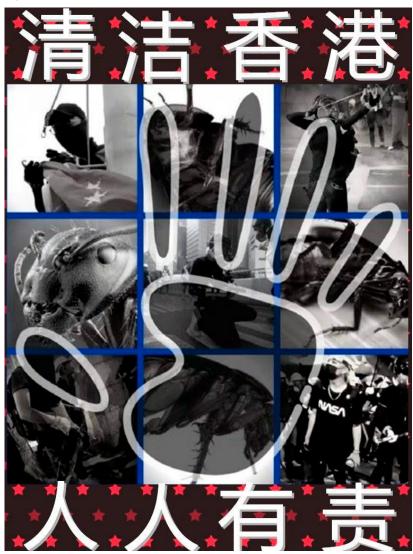
# Narrative 2: Support Hong Kong Police and dehumanise protesters

Examples: Several of the media used in these Tweets appear to have been produced in mainland China and then published for a Hong Kong audience. For instance, one propaganda video (Hong Kong 5) shared multiple times supported Hong Kong police and appears to have been made in mainland China because it included subtitles in simplified Chinese characters with a voiceover in Cantonese. Another image (Hong Kong 6) contained the logo for state media outlet *The People's* Daily. Attributed to a 'Hong Kong citizen', it said 'Hong Kong Police, add oil [jia you]!' ('add oil' is a very common expression of encouragement and support), and added, 'You protect our lives, we protect your honour.' A few of these images supporting police also suggest that mainland police are aiding Hong Kong police. One image (Hong Kong 7) shows Hong Kong police and the mainland's People's Armed Police officers standing side-by-side with the text 'Are you not battle-dressed? Let's share the coat and vest' (Figure 44). The text is a line from Shijing (Book of Odes), the oldest collection of Chinese poetry, that roughly means 'You and I are fighting for the same cause.' Another photo (Hong Kong 8), albeit unclear, appeared to show police wearing an arm patch like the People's Armed Police's new arm patch safeguarding Hong Kong, suggesting that they'll eventually replace Hong Kong police. Meanwhile, many of the images implicitly supporting police dehumanise protesters by portraying them as cockroaches. For instance, this image (Hong Kong 9) showed a Hong Kong police officer killing a cockroach [a protester] with a bottle of insect spray. Others suggested it was the responsibility of everyone to support such actions, including a faux public service announcement poster (Hong Kong 10) with an image of a cockroach and Hong Kong protesters with text (in simplified Chinese characters): 'Cleaning Hong Kong, everyone is responsible' (Figure 45)

Figure 44: 'The People's Armed Police Force and Hong Kong Police stand side-by-side'



Figure 45: Protesters depicted as cockroaches



Analysis: At the very least, the effort to support police contributes to the narrative that protesters are violent and that the police are a stabilising force. More worryingly, the narrative might serve as a pretext for increased police brutality and justification for pro-Beijing individuals and thugs assaulting protesters. Hong Kong police are already using violent tactics that were previously uncommon in Hong Kong but common in the mainland. The idea that the People's Armed Police (PAP) is present in Hong Kong is suggestive of a future increase of police brutality. Indeed, after Tiananmen in 1989, the paramilitary PAP was better trained and equipped to respond to internal security challenges such as protest and rebellion. In fact, analysts and members of the Hong Kong pro-democracy movement have alleged that mainland police have already been sent to Hong Kong as reinforcements for local police—a charge the Hong Kong Government has rejected. The narrative that it's every Hong Konger's 'responsibility' to stabilise Hong Kong mirrors the central theme of the PRC's state security strategy that 'everyone is responsible' for state security. Since this narrative is coupled with images illustrating police exterminating pests, it seems that the PRC's propaganda effort is attempting to sell to the Hong Kong public the idea that protesters deserve to be treated brutally because they're less human, and that like pests they deserve to be killed in the name of Hong Kong's security. Dehumanising the enemy or a political opponent is a common practice in China, as it has been elsewhere throughout history, and therefore that this imagery being pushed out into the public consciousness by state actors on Twitter (and elsewhere) is striking. In Xinjiang, where experts say the CCP is committing cultural genocide, the party has for years dehumanised Uyghurs by criminalising their ethnic and religious identity and simultaneously painting them in broad brush as 'terrorists' and 'separatists'.

# Narrative 3: US interference in Hong Kong

Examples: In messages presented in both Chinese and English, the Twitter accounts circulated images alleging US interference in Hong Kong. One post in English called on the US to 'Stop interfering in China's internal affairs', adding in Chinese that 'The Hong Kong Human Rights and Democracy Act is an Act that incites violence'. Many images suggest a strong relationship between prominent activists and the US Government (Hong Kong 11 and Hong Kong 12). For instance, one image (Hong Kong 13) depicted leading Hong Kong activist Joshua Wong as a puppet controlled by the US, with the logo of state media outlet *The People's Daily* logo visible at the bottom of the image. Some of these images appear in a style similar to Cultural Revolution propaganda posters. Similarly, a series of cartoons with striking similarity to Cultural Revolution style cartoons depicted prominent pro-democracy Hong Kong figures Jimmy Lai (Hong Kong 14), Anson Chan (Hong Kong 15), Albert Ho (Hong Kong 16), Martin Lee (Hong Kong 17), respectively, as a traitor, a hypocrite, an opportunist and a running dog of the coloniser (figures 46 and 47).

Figure 46: Cartoon depicting Hong Kong pro-democracy figures Martin Lee (calling him Lee the Traitor), Alvin Yeung, Jimmy Lai and Joshua Wong providing ammunition to Trump so he could destroy Hong Kong



Figure 47: Depiction of Martin Lee in Cultural Revolution-style political cartoon, where he is called a 'running dog'



Analysis: That the US Government has been behind revolutions in other countries is more than just a narrative in PRC propaganda. Chief among the CCP's descriptions of its own threat perceptions is the idea that 'internal and external hostile forces' will collude to threaten the party's leadership. The 'colour revolutions' in the early 2000s are examples, even if the reality was more complex. The Chinese Government saw the US as contributing to the destabilisation of regimes and exploiting domestic grievances and politics. In fact, the suggestion that the US is interfering in Hong Kong's internal affairs

through the Democracy Act has been combined with the claim that it's part of a 'colour revolution', and state media have accused Hong Kong protesters of colluding with 'anti-China forces' in the US. The propaganda, presented in both Chinese and English, suggests that these Twitter accounts were pushing a narrative to both Hong Kong and international audiences that the protests are neither organic nor driven by true Hong Kongers. The choice to link the imagery of the Cultural Revolution to the Hong Kong protests is a political act with which the Hong Kong audience will have direct familiarity. It suggests that the protesters are the 'enemy of the people' and should be denounced. Labelling of these figures as the 'enemy' or 'traitor' is intended to facilitate the mobilisation of patriotic support for the CCP and to consolidate the party's power in Hong Kong.

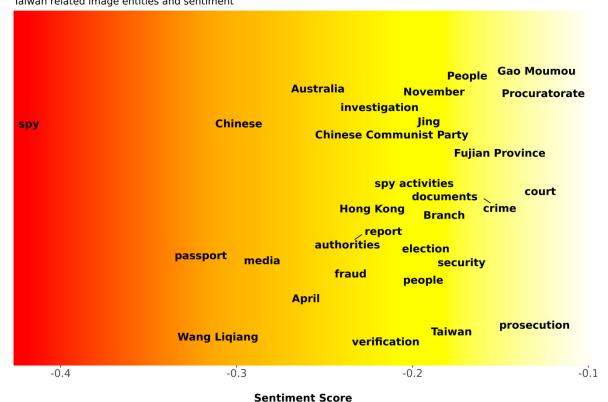
#### **Taiwan**

# Overall size and scope

The smallest subset of the total Twitter dataset consisted of 1,991 tweets that contained references to Taiwan. Those tweets were associated with 441 images and 11 videos. Figure 48 shows a selection of entities extracted from image text and translated into English, arranged horizontally by average sentiment score. 'Wang Liqiang' and the entity 'spy' were mentioned the most in these images and had generally negative sentiments, as opposed to the entities such as 'court', 'prosecution' and 'Procuratorate'.

Figure 48: Taiwan-related entities, by the average sentiment of the translated text entities across images

Taiwan related image entities and sentiment

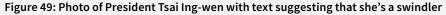


Note: The entities are arranged horizontally by the average sentiment score of the entity across all images. A vertical random jitter is added to distinguish entities. A –1 score indicates negative sentiment, 0 is neutral and +1 is positive sentiment.

#### Narrative 1: The Taiwan election

Many Taiwan-relevant posts engaged Taiwanese politics and the recent presidential election. For instance, one shared an image from an article (Taiwan 0) about conflict within the ruling Democratic Progressive Party. The image says the article, published in simplified Chinese characters (Taiwan uses traditional script) was published by *China Taiwan* (中国台湾网, taiwan.cn). The China Taiwan website is supervised by the Taiwan Affairs Office of the State Council of the PRC. According to *Apple Daily*, taiwan.cn is one of the 'content farms' on the Taiwan Government's watchlist. The Kuomingtang's unsuccessful candidate in the January 2020 presidential election, Han Kuo-yu, featured frequently. For instance, one post included an image (Taiwan 1) of an article published by *China Taiwan* saying that Han's policy slogan, 'Safe Taiwan, prosperous people' can cure the 'Tsai Ing-wen disease'. Many posts

(Taiwan 2) also referred to President Tsai Ing-wen as a swindler. For example, one image of Tsai Ing-wen laughing is captioned (in traditional Chinese characters) with the text: 'Taiwan[ese] people are so easy to cheat' (Figure 49).





Significance: According to a 2019 report on the PRC's information warfare against Taiwan, Taiwan-based disinformation outlets linked to pro-Beijing groups and individuals reposted articles produced by *China Taiwan*. Another Taiwan newspaper, *The Liberty Times*, said in 2019 that 23 Taiwan online media outlets (18 of them are based in Taichung) posted full-text *China Taiwan* articles. At the beginning of 2019, data suggested that President Tsai Ing-wen and the Democratic Progressive Party (DPP) would lose both the presidency and the legislature in the January 2020 elections. In January 2019, after the DPP's poor showing in local elections, DPP party leaders called for Tsai not to run for a second term. After the protests began in Hong Kong in mid-2019, public opinion began to shift in favour of Tsai and the DPP, who oppose 'reunification' with the PRC. The Kuomintang's candidate, Han Kuo-yu, was widely considered during the election to have the CCP's backing, while the CCP has openly and vocally opposed President Tsai. Though Han was defeated in the presidential election, he remained as the southern city of Kaohsiung's mayor. On 6 June, however, Han was defeated in a recall vote, becoming the first municipal mayor to be recalled since martial law ended in Taiwan.

# **Narrative 2: Wang Liqiang**

Examples: Numerous posts linked to content on Taiwan covered the story of Wang Liqiang, who defected to Australia, claiming to have been a PRC spy. Some Tweets shared images of Tsai Ing-wen and Wang Liqiang placed together with no text provided, and another with the same images (Taiwan 3) overlaid with the text 'I will let you hype this once!', suggesting that Wang said to Tsai 'I will let you use me for this campaign' (Figure 50). Many other posts described the case, including the images of a man said to be Wang in a PRC courtroom and images of his fake passports (Taiwan 4).

Figure 50: Side-by-side photos of Wang and Tsai, suggesting collusion between them



Analysis: Wang's defection and ensuing events became a major issue in the Taiwanese presidential election in January 2020. He claimed to have been involved in influence operations against Taiwan, including through Facebook, that were designed to support opposition presidential candidate Han Kuo-yu. Wang said he had left China shortly before he was supposed to travel to Taiwan for operations targeting the presidential election. After his allegations were aired, Taiwanese authorities stopped two individuals whom Wang accused of being intelligence officers from leaving the country and placed exit bans on them. According to *The Age*, a senior official from Taiwan's opposition Kuomintang party contacted Wang ahead of the election and attempted to induce him to say that the incumbent DPP had paid him to make up his claims about Chinese intelligence activity.<sup>3</sup> A Facebook page supporting Han Kuo-yu that Wang claimed had been set up as part of an influence operation was taken down by Facebook shortly before the election. The connection between Wang's case and the Taiwanese election would have made it an extremely sensitive issue for the CCP, which has been hostile towards the DPP. The posts appear designed to discredit Wang and suggest links between Wang and the DPP's President Tsai Ing-wen.

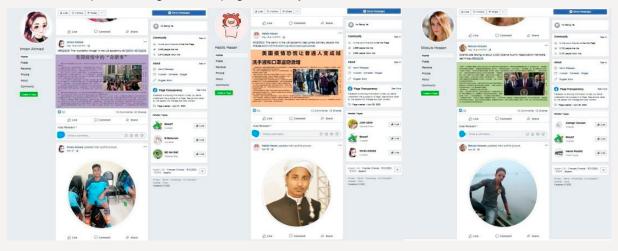
# A continuing operation?

One of the key features of this campaign is persistence in the face of repeated discovery and takedowns. That includes the September 2019 takedowns by Twitter that led to *Tweeting through the Firewall* and numerous smaller removals in the months since. During the weeks in which this research was conducted, dozens and possibly hundreds of accounts linked to the campaign were removed from Twitter and Facebook.

The campaign operators appear to have sourced, created or activated new accounts within a matter of days. For example, after Bellingcat published an investigation into what it characterised as pro-Chinese bot activity on Twitter and Facebook on 5 May 2020, a number of Facebook accounts identified in that research were removed.

The rebuilding process appears to have begun almost immediately. This included the creation of new pages and the activation of existing accounts. Like the previous, deleted Facebook accounts, a number of the recently activated or repurposed accounts appear to have initially been identifying as Bangladeshi men and created on 27 or 28 April 2020 (Figure 51). It isn't clear whether these individuals are real people; the accounts don't appear to have ever been used for personal social media use. On 18 May, these accounts made their first posts connected to the campaign and changed their profile pictures (although not their names).

Figure 51: Screenshot of three Facebook accounts created on 27–28 April, which changed their profile pictures and made their first posts relating to the campaign on 18 May 18



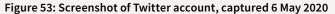
Some of the accounts recycled previous profile pictures. For example, the same picture of a blonde woman was used for the account Md Joy Jnr, which was deleted in mid-May, as for the account MD Alom, which was created 19 May (Figure 52). This adds to circumstantial evidence suggesting that these accounts may have been purchased in bulk from the same reseller.

| Mode | Process
| Mod

Figure 52: Screenshots of Facebook accounts showing same profile image on a deleted Facebook account and a new Facebook page

Similarly, despite the takedown of at least 23,750 accounts by Twitter, which forms the bulk of the dataset for this report, activity that appears to be part of the same campaign has continued into June 2020.

On 2 June, for example, a Twitter account that appears to have previously belonged to an India-based consultant and tweeted in English primarily about Indian politics up until 27 June 2019, began tweeting content that appears to be related to this campaign, targeted at the protests in the US (Figure 53).





Other recently created accounts also resumed tweeting and retweeting content that appears likely to have been part of the campaign (Figure 54).

Figure 54: Screenshots of Twitter accounts captured 5 June 2020



In a further piece of circumstantial evidence, some of the newly created Twitter accounts use the same profile pictures as recently created Facebook accounts that also appear to be active in the campaign (Figure 55).

Figure 55: Screenshot showing Twitter and Facebook accounts, captured 6 June 2020



Although these accounts have been active only for a relatively short period of time, there's already evidence of cross-posting of content across platforms (Figure 56).

Figure 56: Screenshot showing identical content posted on Twitter and Facebook accounts on 2 June, captured 8 June 2020



# **Notes**

- 1 As determined by Twitter.
- 2 Constellation is data analytics and visualisation software available at https://www.constellation-app.com/
- 3 Note that ASPI International Cyber Policy analyst Alex Joske was a co-author of this story published in Australia's *The Age* newspaper.

# **Acronyms and abbreviations**

CCP Chinese Communist Party

DPP Democratic Progressive Party (Taiwan)

EU European Union

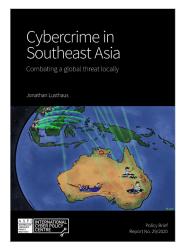
GEC Global Engagement Center (US)

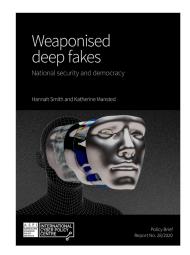
PAP People's Armed Police

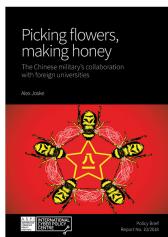
PRC People's Republic of China

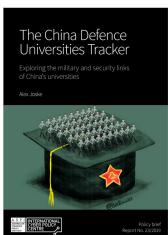
# Some previous ICPC publications



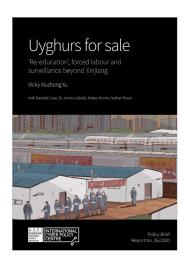


















# Stay informed via the field's leading think tank, the Australian Strategic Policy Institute.

The Strategist, ASPI's commentary and analysis website, delivers fresh ideas on Australia's defence and strategic policy choices as well as encouraging discussion and debate among interested stakeholders in the online strategy community. Visit and subscribe to an email digest at www.aspistrategist. org.au.





