

24 / 07 / 2024

American (Covert) Influence



CRC

CYFLUENCE
RESEARCH
CENTER

American (Convert) Influence

For a long time, hostile digital influence campaigns were predominantly the prerogative of autocratic states. Russia, in particular, has been at the forefront of both executing such campaigns and developing new methodologies in the realm of hostile digital influence.

Recently, some Western democratic countries, such as France and the United States, have been attempting to catch up in this field, albeit arguably with less success. This effort to engage in digital influence campaigns presents unique risks for liberal democracies.

Liberal democracies inherently have more to lose by conducting such operations. Freedom of speech and the free exchange of opinions are cornerstones of any democratic society. When these societies engage in clandestine measures to distort natural discourse, they risk undermining these fundamental principles. By doing so, there is a valid argument to be made that they are eroding the very pillars on which they stand. This paradox raises critical ethical and strategic questions about balancing national security and preserving democratic values.

When it comes to influence defense teams of a country, this new reality brings up a unique set of challenges. Democracies across the world are under a constant barrage of attacks from autocratic countries. However, defense teams now need to consider a new type of threat actor—countries that are theoretically politically aligned with them or even allies.

This development introduces several complexities:

1. **Trust and Verification:** Traditional alliances are built on trust and mutual interests. When allies engage in influence operations, they strain these relationships and complicate coordinating joint efforts in other areas.
2. **Policy and Strategy:** Democracies need to develop policies and strategies that address not only threats from known adversaries but also potential subversive activities from allies. This necessitates a highly nuanced approach to intelligence gathering, threat assessment, and diplomatic engagement, underlining the complexity of the task.
3. **Legal and Ethical Dilemmas:** Engaging in counter-influence operations against allies presents legal and ethical dilemmas. Democracies must balance national security concerns with adherence to international law and the ethical standards they champion.
4. **Public Perception:** Democracies rely heavily on public trust. Revelations about influence operations against or from allies can lead to public distrust in both government institutions and international alliances.
5. **Technological Challenges:** Identifying and mitigating influence operations from allies can be technologically challenging. These actors may have access to similar or even shared technological platforms, making it difficult to detect and attribute attacks accurately.

6. **Operational Security:** The presence of allied threat actors necessitates heightened operational security within influence defense teams. Information-sharing protocols need to be carefully managed to prevent sensitive data from being exploited by those presumed to be allies.


In summary, the evolving landscape of digital influence operations demands that democracies adapt their defense strategies to address not only traditional adversaries but also their allies' complex and sometimes covert actions. This requires a delicate balance of vigilance, transparency, and strategic diplomacy to maintain the integrity of democratic institutions while safeguarding national security. By highlighting attempts by the United States of America to execute such a hostile influence campaign, we hope to add to a much-needed discussion on the use of digital hostile influence campaigns by democratic nations.

An American Hostile Influence Campaign

A Reuters investigation revealed that the US government launched a secret influence campaign to discredit China’s own Sinovac vaccine¹. The campaign targeted the Philippines’ online environment, using fake social media accounts that propagated custom-tailored propaganda. The analysis of the content led to the understanding that the campaign’s purpose was to undermine China’s efforts to improve its position in the region by providing COVID-19-related aid. However, the collateral damage in this Cold War-style smear campaign might have been the lives of innocent citizens of the Philippines.

Reuters detected at least 300 accounts on X (formerly Twitter) likely involved in this campaign targeting the Philippines. These accounts shared suspicious features such as similar creation times and a lack of identifiable information, and they all propagated the same narrative. Eventually, after Reuters approached X, the fake accounts were removed, as the company determined they were part of a coordinated campaign.

The hashtag at the center of this campaign was #Chinaangvirus, which translates from Tagalog to “China is the virus”. Additionally, the disseminated posts on X claimed that any aid coming from China, such as face masks and even the vaccine itself, should not be trusted:



This post, identified by Reuters, matched the messaging, timeframe and design of the U.S. military's anti-vax propaganda campaign in the Philippines, former and current military officials say. Social media platform X also identified the account as fake and removed it.

TRANSLATION FROM TAGALOG

#ChinaIsTheVirus

Do you want that? COVID came from China and vaccines came from China

(Beneath the message is a picture of then-Philippines President Rodrigo Duterte saying: “China! Prioritize us first please. I’ll give you more islands, POGO and black sand.” POGO refers to Philippine Offshore Gaming Operators, online gambling companies that boomed during Duterte’s administration. Black sand refers to a type of mining.)

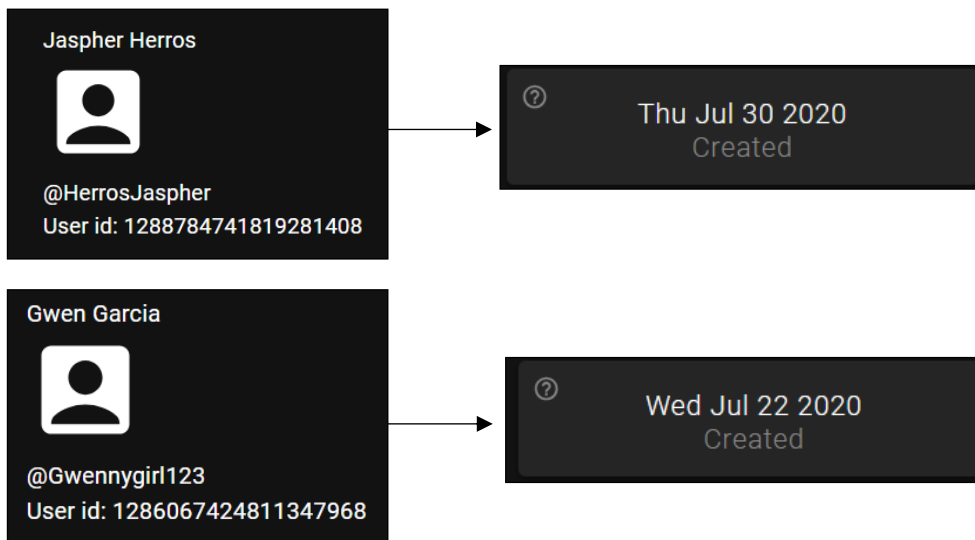
Source: <https://www.reuters.com/investigates/special-report/usa-covid-propaganda/>

¹ <https://www.reuters.com/investigates/special-report/usa-covid-propaganda/>

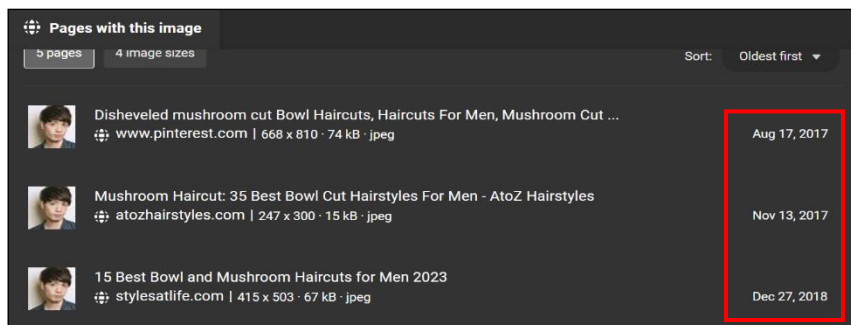
Even though most of the activity linked to this influence campaign was removed from X, we were able to identify possible online traces of it. The following 2 accounts share inauthentic features as well as activity patterns, and both propagated the same anti-China narratives over the same period.^{2 3}



Both accounts were created in July 2020, only a few days apart:

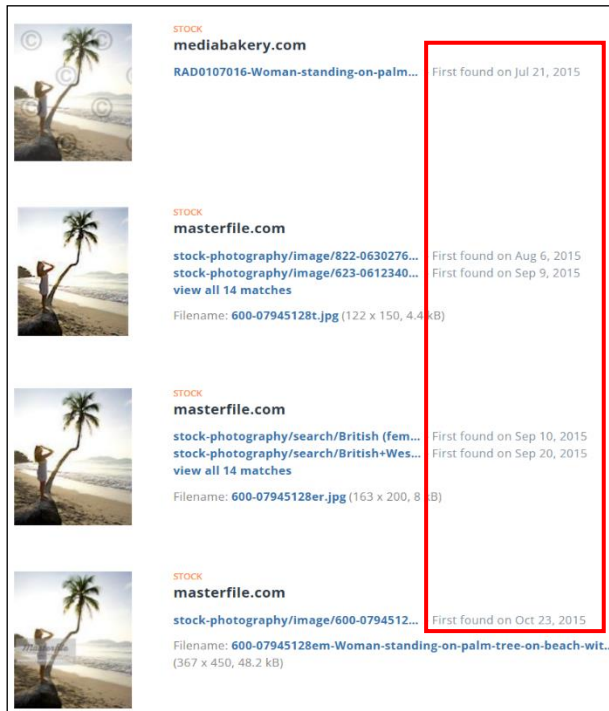


Their profile pictures can be found elsewhere online, indicating the images are not unique and not personal:

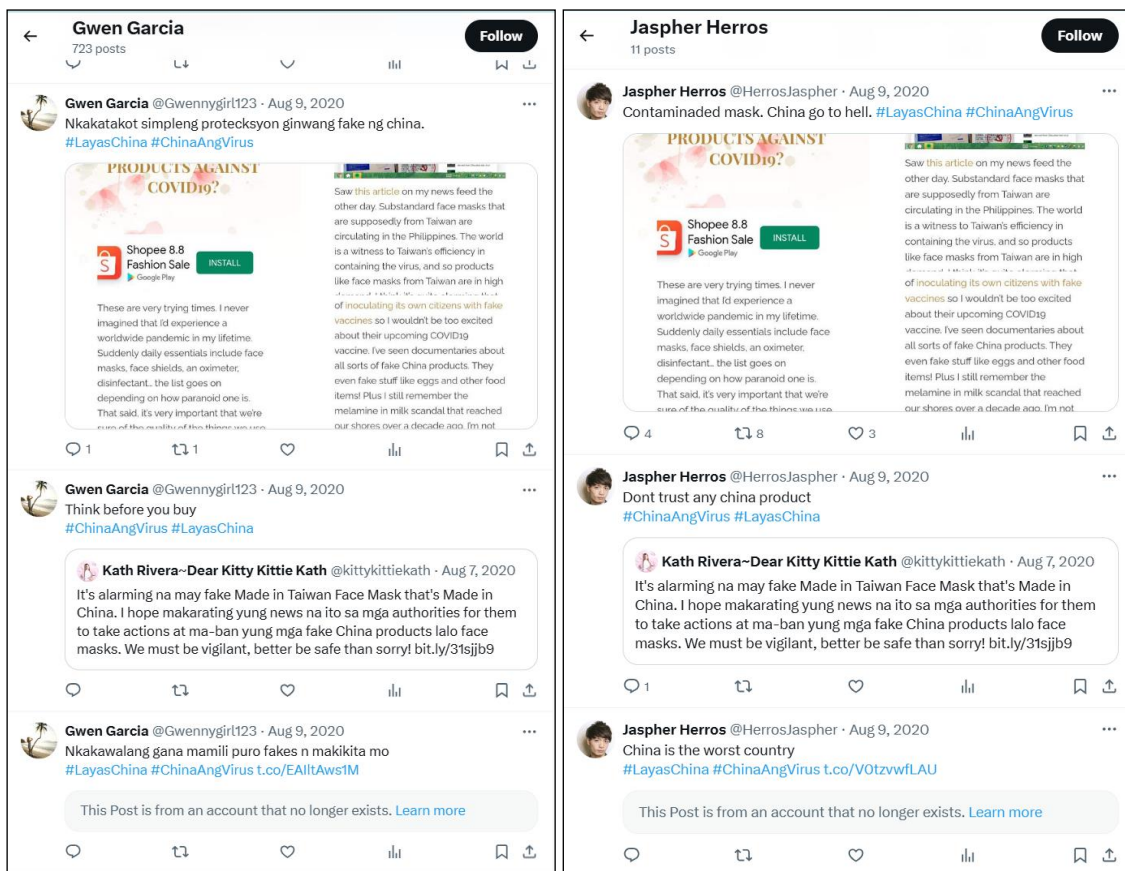


² <https://x.com/HerrosJasperher>

³ <https://x.com/Gwennygirl123>



Both accounts promoted the same hashtags and similar content in the same pattern on August 9, 2020, within the same timeframe:



These features are highly indicative of a coordinated inauthentic influence effort, similar to previously detected and reported campaigns coming from Russia and China.

According to Reuters, this influence campaign started under former president Donald Trump and continued into Biden's presidency for some time. In 2021, the Biden White House issued an edict banning this campaign, and an internal review was initiated. A Pentagon spokeswoman said that the US military "uses a variety of platforms, including social media, to counter those malign influence attacks aimed at the US, allies, and partners". She also noted that China had started a "disinformation campaign to falsely blame the United States for the spread of COVID-19".⁴

The Philippines saw a significant death toll by November 2021, caused by the pandemic and by difficulties in vaccinating the population. Based on interviews conducted by Reuters and research about the effects of skepticism towards vaccination, it is likely that the covert influence effort exploited an already vulnerable society and contributed to the high death toll in the country.

Reuters unearthed an additional influence effort that originated in the US in 2019, in which the campaign targeted China's own population⁵. According to the former US official who spoke to Reuters and had direct knowledge of this highly classified influence operation, the operation aimed to turn public opinion in China against its government.

The method used for this operation was similar to the one utilized in the Philippines, with some add-ons specifically tailored to the Chinese online audience. The operation consisted of fake online entities that were used for spreading hostile narratives about President Xi's government while leaking intelligence to overseas news sources to hurt the Chinese government's reputation. According to Reuters, some of the narratives that were disseminated by this operation included allegations that members of the ruling Communist Party were hiding ill-gotten money overseas. Another narrative aimed at China's Belt and Road Initiative, calling it corrupt and wasteful.

The uniqueness of this specific operation stems from the US officials' emphasis that these narratives were based, in fact, despite being covertly disseminated by fake online entities operated by intelligence officers. Thus, this influence effort supposedly did not conjure up fabricated pieces of information to serve as disinformation, unlike the operation targeting the Philippines, but tried to insert authentic pieces of information into the highly monitored and censored Chinese online environment. If such a negative narrative was successfully injected into China's online environment, it would likely be flagged by the monitoring Chinese authorities, and a virtual chase would start to identify the source and eliminate all online mentions of it. Indeed, two former officials told Reuters that influence efforts within China were intended to induce paranoia among the top leaders in China, so they would spend their resources chasing virtual intrusions into China's highly enclosed internet environment.

⁴ <https://www.reuters.com/investigates/special-report/usa-covid-propaganda/>

⁵ https://www.reuters.com/world/us/trump-launched-cia-covert-influence-operation-against-china-2024-03-14/?utm_source=substack&utm_medium=email

As revealing and unequivocal as these recent publications are, it is important to note that they are not the first US-led efforts to covertly influence foreign online audiences.

In 2022, Facebook's parent company, Meta, reported that people associated with the US military were behind dozens of inauthentic Facebook accounts, pages, groups, and Instagram accounts that promoted pro-US messaging while attempting to obscure their real identities⁶.

The report stated that several clusters of inauthentic accounts posted content and tried to direct viewers to off-platform domains. The clusters focused on several geographical regions, mainly Iran, the Gulf, Central Asia, the Middle East, and North Africa. The list of specific countries included Afghanistan, Algeria, Iran, Iraq, Kazakhstan, Kyrgyzstan, Russia, Somalia, Syria, Tajikistan, Uzbekistan and Yemen.

Each cluster posted about particular themes, including sports and culture in a particular country, to appear credible; cooperation with the United States, including military cooperation; and criticism of Iran, China, or Russia. The people operating this network posed as locals in the countries they targeted and used local languages like Arabic, Farsi, and Russian to appear more credible. The main narratives propagated by this network focused on news and current events, terrorism concerns, praise of the US military, and content about COVID-19. This operation also posted content criticizing China and Russia, including Russia's invasion of Ukraine, China's treatment of the Uyghur people, Iran's influence in the Middle East, and the support of the Taliban regime by Russia and China.

Meta shared data regarding this network with independent researchers at Graphika and the Stanford Internet Observatory, who have published their findings⁷. This joint research highlighted some key features of online inauthentic coordinated behavior similar to those detected in influence efforts by Russia, China, and Iran. For example:

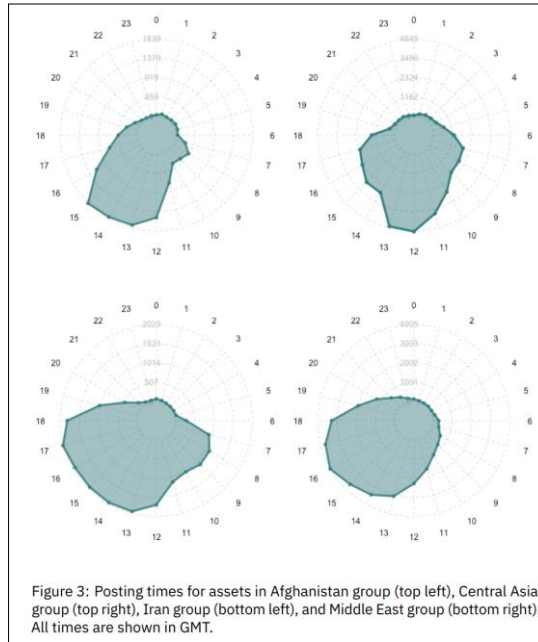
- Inauthentic accounts with GAN⁸-generated faces serving as their profile pictures.
- Inauthentic accounts posed as independent media outlets.
- Utilization of memes and short-form videos.
- Attempts to promote specific hashtags and get them viral.
- Launching online petitions.

The posting patterns also indicated coordination. For example, the daily posting times for each of the detected clusters of inauthentic accounts were quite similar, even though they were supposedly from different countries. The activity of the accounts peaked during the same period in any given 24 hours, usually between 1200 and 1800 GMT:

⁶ <https://about.fb.com/news/2022/11/metas-adversarial-threat-report-q3-2022/>

⁷ https://public-assets.graphika.com/reports/graphika_stanford_internet_observatory_report_unheard_voice.pdf

⁸ Generative adversarial networks.



Source: <https://www.reuters.com/investigates/special-report/usa-covid-propaganda/>

Moreover, accounts in some of the clusters typically posted at roughly 15-minute or 30-minute intervals in any given hour, and some almost exclusively posted in the first second of any given minute:

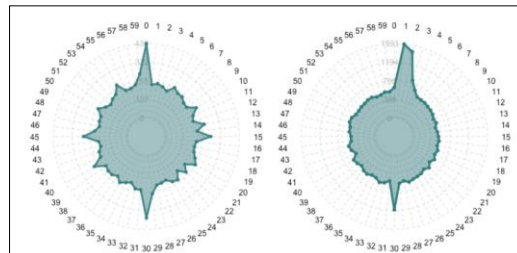


Figure 4: Radar charts showing posts by minute in any given hour for assets in the Afghanistan group (left) and Central Asia group (right).

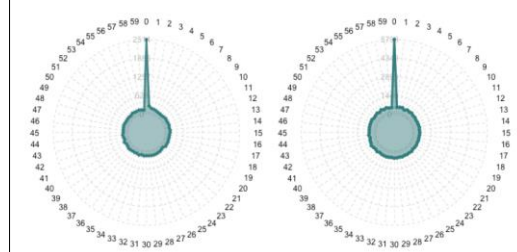


Figure 5: Radar charts showing posts by second in any given minute for assets in the Central Asia group (left) and Middle East group (right).

Source: https://public-assets.graphika.com/reports/graphika_stanford_internet_observatory_report_unheard_voice.pdf

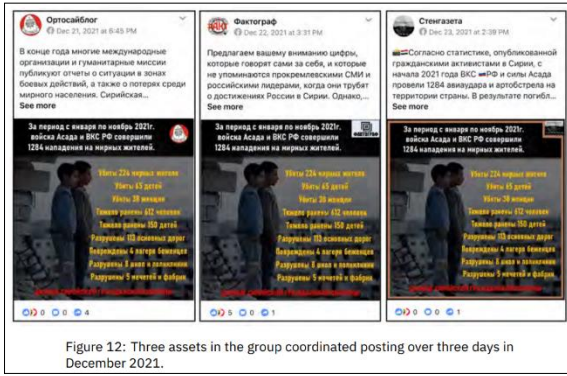


Figure 12: Three assets in the group coordinated posting over three days in December 2021.

The joint investigation by Graphika and the Stanford Internet Observatory detected coordination in the content that was also disseminated. The involved accounts posted simultaneously, used the same images, and recycled content over time across accounts.

Source: [https://public-](https://public-assets.graphika.com/reports/graphika_stanford_internet_observatory_report_unheard_voice.pdf)

[assets.graphika.com/reports/graphika_stanford_internet_observatory_report_unheard_voice.pdf](https://public-assets.graphika.com/reports/graphika_stanford_internet_observatory_report_unheard_voice.pdf)

Lastly, the narratives propagated by the clusters of inauthentic accounts focused on a range of topics relevant to the different target audiences. Still, all conveyed criticism of the activities of the governments of Russia, China, and Iran. Some were actively promoting the actions of the US as positive and beneficial:⁹



Figure 24: Posts about alleged organ harvesting of Muslims in Xinjiang (left) and blaming China for being the main sponsor of Russia's war against Ukraine (right).

Figure 19: Posts about Russia using ethnic minorities to fight in Ukraine (left) and the conscription of Central Asian migrants into the Russian military (right).

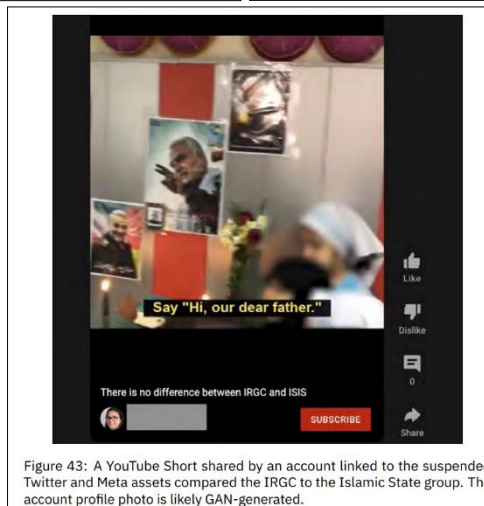


Figure 43: A YouTube Short shared by an account linked to the suspended Twitter and Meta assets compared the IRGC to the Islamic State group. The account profile photo is likely GAN-generated.

⁹ Source: https://public-assets.graphika.com/reports/graphika_stanford_internet_observatory_report_unheard_voice.pdf



CRC | CYFLUENCE
RESEARCH
CENTER