ARTICLES

Beijing’s influence operations Have Gone Algorithmic: Can India’s Answer Keep Up?

Author: Rahul Kumar Tiwari • **Published:** August 08, 2026 • **Source:** Centre for New Age Warfare Studies



EXECUTIVE BRIEF & KEY FINDINGS

China has spent two decades building a strategy for using information as a weapon, and it is now adding generative AI to that strategy. India's new counter-disinformation division is a good idea. But now its plan leans on AI tools to catch fake content as it appears. The evidence in this piece shows that AI tools regularly fail at exactly that job. The division needs a plan that still works even when the technology gets it wrong.

On 13 January 2024, Taiwan held its presidential election under a barrage of manufactured content. Fake audio of Foxconn founder Terry Gou endorsing a candidate he never endorsed. AI-generated memes counting down the days to vote out the ruling party. Synthetic news anchors reading scripts, which nobody wrote. Microsoft’s threat intelligence team identified much of it as Storm-1376, a China-linked network better known as Spamouflage, and recorded a milestone: the first time it had recorded a nation state using AI content to try to influence a foreign election.

This was not improvisation, but it was doctrine, which is China’s long-standing official playbook for using information as a weapon, now catching up with technology. The People's Liberation Army (PLA) codified its “Three Warfare” concept in early 2000s, namely directing public opinion through media, undermining an opponent's morale and decision-making, and using law and international rules as a weapon, fought continuously and mostly in peacetime. The Pentagon's 2024

report on China's military describes a newer layer, “cognitive domain operations,” aimed at winning control over how people think rather than just what they know, and says this now explicitly folds in AI and synthetic media. PLA researchers were exploring voice-cloning to imitate foreign leaders as far back as 2011; by 2020, elements of the PLA had reportedly produced a deepfake aimed at misleading the American public. In 2023, Meta shut down Spamouflage, eliminating 7,704 Facebook pages from over 50 platforms and tying it to people connected with Chinese law enforcement, the biggest known covert influence operation that has ever been shut down.

India is already in the picture. Storm 1376 delivered fabricated videos, featuring the same AI voice of the man used in Taiwan's campaign, accusing the U.S. and India of being responsible for the unrest in Myanmar. As early as 2022, Graphika had been able to piece together AI-generated Indian avatars such as a turbaned “Dr. Dass,” through their forensic investigation of Spamouflage. The theme of the evidence is clear: China sees the information environment as a permanent stage in which to operate, and it's adding generative AI to a doctrine that comes twenty years before the technology itself.

There is a single word of caution to the honour when it comes to honesty, and this is a word often left out of the propagandist’s lexicon. The machine has made volume, not victories, up till now. Taiwan Platforms removed the Gou audio, and its institutions quickly exposed it within hours of it spreading. In January 2025, Google’s threat intelligence team found that AI is not as it is often made out to be, “the game-changer,” but it enables threat actors to move faster and at higher volume. Google also pointed out that the engagement of real users for the Spamouflage AI-generated videos hasn't been significantly improved. But operations continue to roll: OpenAI blocked China-origin networks in February, June, and October of 2025, just three times.

Delhi’s Turn

India has now experienced a version that week. In Operation Sindoor of May 2025, deepfake videos were circulated against the Indian leadership with a view to being pushed by Pakistan. The Chief of Defence Staff (CDS) Gen Anil Chauhan later estimated that 15 per cent of the Armed Forces time was spent countering false narratives during the operation. Lt Gen Zubin A Minwalla, at the Jaipur press conference on the occasion of the operation’s first anniversary on 7 May 2026, announced that with the coordinated effort with the “Ministry of Information and Broadcasting,” has created an effective model of civil-military fusion, he further stated, underlining the need to build on that model as India set up a "Defence Strategic Communication Division", that leverages AI-enabled tools to counteract misinformation in real time.

That origin story cuts two ways. The institution is growing out of a coordination model that worked, not being built from scratch around a machine. What remains open is whether AI-enabled detection stays one part of that model or becomes the main way fabrications get caught as they appear, and if it becomes the main method, the doctrine will be leaning on the weakest part of the science. OpenAI built a tool to detect AI-written text and, by its own published numbers, that tool correctly caught only 26 percent of AI-written text, although wrongly flagging 9 percent of actual human writing as AI-generated; the company withdrew the tool within six months for low accuracy. A Berkeley study of deepfake detectors found their accuracy falls by 20 to 60 percentage points when tested against AI image-generation tools they were not originally trained to catch. Stanford researchers found seven widely used detectors wrongly flagged 61.3 percent of essays by non-native English writers as machine-made, while a single AI rewrite of the same text pushed detection of the real AI-generated text up to near zero. A defence spokesman leaning on these tools risks calling truth a lie and a lie truth, in public, during a crisis. Each such error feeds what legal scholars Robert Chesney and Danielle Citron call the liar's dividend: once people know convincing fakes exist, it becomes easier for anyone caught on real footage doing something wrong to simply claim the footage is fake.

A Firmer Foundation

So, build the division and build on the coordination model that values its keep, but make the detection layer more solid than real-time classification. Certify not chase: an open technical standard called Coalition for Content Provenance and Authenticity (C2PA), can be used to add a verified digital signature to official statements and footage, tracing each to its true origin. In practice, an official statement or video without this signature becomes suspect by default, and the burden shifts onto whoever is spreading the fake, not onto India's fact-checkers to disprove it afterward. Prepare audiences early: Cambridge researchers and Google Jigsaw found in their investigation that short videos teaching people to spot manipulation techniques improved recognition by about 5 percent in a YouTube field trial, roughly five percent per viewer. Modest cheap, scalable, and immune to whatever new fake-generating tool comes next, because it trains the audience rather than just the technology. And buy speed through better organisation, not better software. Taiwan is a small state, with one official language and a single government speaking to its public, conditions India's larger federal, multilingual defence-communication system does not share, so the model cannot simply be copied. But the underlying principle is worth borrowing from it: Taiwan cut official response time to a fake story from six or seven hours to under one, by pre-approving spokesperson in advance and enforcing a strict format for every rebuttal, a headline under 20 characters, a message under 200, and no more than two images. Gen Chauhan said India chose to be "measured, not reactive". Being careful is the right instinct; being slow is not the same thing, and slowness is exactly what gives a fake story its uncontested hours online.

Here is the strategic logic in plain terms. China's approach is a numbers game: flood the information space with fake content until a democracy has no choice but to chase down and disprove each one leading burning time and credibility. If India's new division relies mainly on AI tools trying to catch and label each fake as it appears, it is playing exactly the game this strategy is built to win. A different approach, certifying that is actually official, preparing citizens in advance, and moving fast enough that the true story arrives before the fake one does, it does not depend on winning that chase at all. It changes the game being played. Whether the division becomes one kind of institution or the other is still an open question, to be settled in the coming months, not on the day it was announced.

References:

1. Beauchamp-Mustafaga, N. (2019). Cognitive domain operations: The PLA's new holistic concept for influence operations. *China Brief, 19*(16). Jamestown Foundation. https://jamestown.org/cognitive-domain-operations-the-plas-new-holistic-concept-for-influence-operations/
2. Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753–1819. https://doi.org/10.15779/Z38RV0D15J
3. Coalition for Content Provenance and Authenticity. (2025). *C2PA and Content Credentials explainer (Specification 2.4)*. https://spec.c2pa.org/specifications/specifications/2.4/explainer/Explainer.html
4. Garg, A., Srivastava, G. S., & Masaki, K. (n.d.*). Beyond the benchmark: Generalization limits of deepfake detectors in the wild [Working paper]*. UC Berkeley School of Information. https://www.ischool.berkeley.edu/sites/default/files/bb_paper.pdf
5. Google Threat Intelligence Group. (2025, January). *Adversarial misuse of generative AI. Google Cloud*. https://cloud.google.com/blog/topics/threat-intelligence/adversarial-misuse-generative-ai
6. Graphika. (2023, February). *Deepfake it till you make it*. https://public-assets.graphika.com/reports/graphika-report-deepfake-it-till-you-make-it.pdf
7. Kao, S.-S. (2021, September). *Taiwan's response to disinformation: A model for coordination to counter a complicated threat* (NBR Special Report No. 93). The National Bureau of Asian Research. https://www.nbr.org/publication/taiwans-response-to-disinformation-a-model-for-coordination-to-counter-a-

complicated-threat/

8. Kumar, A. (2026, May 7). *After Operation Sindoor, India builds powerful new air defence shield against drones and missiles*. Asianet Newsable. https://newsable.asianetnews.com/india/operation-sindoor-one-year-india-new-air-defence-plan-aims-to-stop-drones-missiles-and-future-threats-articleshow-vnd9o5r
9. Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns, 4*(7), Article 100779. https://doi.org/10.1016/j.patter.2023.100779
10. Martin, A. (2023, August 29). *Chinese law enforcement linked to largest covert influence operation ever discovered.* The Record. https://therecord.media/spamouflage-china-accused-largest-covert-influence-operation-meta
11. Microsoft Threat Analysis Center. (2024, April 4). *China tests U.S. voter fault lines and ramps AI content to boost its geopolitical interests* [East Asia report summary]. Microsoft On the Issues. https://blogs.microsoft.com/on-the-issues/2024/04/04/china-ai-influence-elections-mtac-cybersecurity/
12. Microsoft Threat Analysis Center. (2024, April 4). *Same targets, new playbooks: East Asia threat actors employ unique methods* [Threat intelligence report]. Microsoft. https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/MTAC-East-Asia-Report.pdf
13. OpenAI. (2023, January 31). *New AI classifier for indicating AI-written text* [withdrawal notice added July 20, 2023]. https://openai.com/index/new-ai-classifier-for-indicating-ai-written-text/
14. OpenAI. (2025, February). *Disrupting malicious uses of our models: An update*. https://openai.com/index/disrupting-malicious-ai-uses/
15. OpenAI. (2025, June). *Disrupting malicious uses of AI: June 2025*. https://openai.com/global-affairs/disrupting-malicious-uses-of-ai-june-2025/
16. OpenAI. (2025, October). *Disrupting malicious uses of AI: October 2025*. https://openai.com/global-affairs/disrupting-malicious-uses-of-ai-october-2025/
17. Patil, S. (2026). *Operation Sindoor at one: Taking stock of India's cyber preparedness*. Observer Research Foundation. https://www.orfonline.org/expert-speak/operation-sindoor-at-one-taking-stock-of-india-s-cyber-preparedness
18. Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., & Lewandowsky, S. (2022). Psychological inoculation improves resilience against misinformation on social media. *Science Advances, 8*(34), Article eabo6254. https://doi.org/10.1126/sciadv.abo6254
19. Rosen, G. (2023, August 29). *Raising online defenses through transparency and collaboration*. Meta Newsroom. https://about.fb.com/news/2023/08/raising-online-defenses/
20. Shah, D. P. (2024). Understanding the PRC's information operations against India. *Indian Public Policy Review, 5*(3), 1–37. https://doi.org/10.55763/ippr.2024.05.03.001
21. Singh, R. (2025, May 31). 15% of Army time lost to fake news during Op Sindoor: CDS Gen Anil Chauhan. *Business Standard*. https://www.business-standard.com/external-affairs-defence-security/news/operation-sindoor-india-pakistan-war-fake-news-cds-anil-chauhan-125053100353_1.html
22. U.S. Department of Defense. (2024). *Military and security developments involving the People's Republic of China: Annual report to Congress*. https://media.defense.gov/2024/Dec/18/2003615520/-1/-1/0/MILITARY-AND-SECURITY-DEVELOPMENTS-INVOLVING-THE-PEOPLES-REPUBLIC-OF-CHINA-2024.PDF

The views expressed are those of the authors and do not represent the views of CNAWS.

Image Source: ChatGPT