RSIS S. RAJARATNAM SCHOOL OF INTERNATIONAL STUDIES
Nanyang Technological University, Singapore

RSIS COMMENTARY

www.rsis.edu.sg No. 099 – 23 July 2024

Deepfakes and the Dangers to National Security and Defence

By Benjamin Ang and Muhammad Faizal

SYNOPSIS

AI-generated disinformation in the form of deepfakes, comprising digitally manipulated video, audio, or images, has hit the headlines in Singapore. Cases from around the world demonstrate that AI-generated deepfakes combined with cyberattacks are not only a threat to the integrity of elections and scam victims but are also a threat to national security and defence.

COMMENTARY

Deepfake videos of Singapore's political leaders have been circulating since at least last year, when manipulated video and audio impersonating Senior Minister (then Prime Minister) Lee Hsien Loong circulated online, appearing to promote a cryptocurrency scheme in a TV news interview. Similar videos impersonating Prime Minister (then Deputy Prime Minister) Lawrence Wong were also circulated.

This year, more deepfake videos of Senior Minister (SM) Lee have been circulating online, this time showing him commenting on international relations. SM Lee described them as having "malicious intent" and was "dangerous and potentially harmful to our national interests". To deal with this problem of deepfake videos, the Minister for Digital Development and Information, Ms Josephine Teo, is considering ways to regulate it by proposing a labelling scheme for tools and contents and even discussing a temporary ban to counter such videos, which are anticipated ahead of future Singapore general elections.

The earlier cases were commercially motivated scams, but the recent ones have severe national security and defence implications. In SM Lee's case, the deepfake made it look as if he was commenting on foreign policy, and foreigners unfamiliar with him could be misled, thereby sowing distrust. There is evidence that hostile information

campaigns are used to weaken national cohesion or disrupt society in the lead-up to hostilities or as part of geopolitical contestation.

Cyber and AI-enabled Attacks on Defence

At the start of the Russian invasion of Ukraine, a fake video of Ukrainian President Volodymyr Zelensky appeared telling Ukrainian soldiers to surrender to their Russian opponents. Cyber attackers posted this deepfake video on a Ukrainian news website. They also circulated it on social media before it was debunked and removed.

This combination of cyberattacks and AI-enabled disinformation has continued throughout the Russia-Ukraine war. Russia is accused of using cyberattacks to disrupt critical infrastructure such as telecommunications and electricity networks in Ukraine, adversely affecting networks and air raid alert systems, ATMs, credit card payment terminals and causing a power outage during a missile strike.

AI-manipulated video has also been used against Ukraine's allies in the war. In one case, someone impersonated the Mayor of Kyiv having video calls with the mayors of Berlin, Madrid and Vienna. Another call to then UK foreign secretary, David Cameron, impersonated Petro Poroshenko, the former president of Ukraine. A deepfake video of a US government spokesperson also appeared online, discussing how Ukraine could use American weapons in limited strikes inside Russia.

In the same vein, a deepfake audio that circulated on social media in May 2024 falsely portrayed Philippines President Marcos as instructing the Filipino military to respond to Chinese attacks in the South China Sea. This deepfake could have endangered the Philippines' foreign relations and heightened tensions to the point of unintentional conflict in Southeast Asia.

In the course of the Israel-Hamas war, we also frequently see cyberattacks and information operations conducted on civilian targets to erode domestic support. These include cyberattacks on news media websites, financial institutions and government agencies. Cyberattacks have also altered digital billboards to display the Palestinian flag as well as false news about Israeli military defeats. The United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) has also been the target of intensified disinformation.

These recent cases illustrate how cyberattacks and AI-enabled disinformation can be used in situations of armed conflict. It is not difficult to imagine how the deepfake videos of our ministers could be weaponised as part of a hybrid warfare toolkit to disrupt our society or weaken our social cohesion in a time of crisis.

Need for Digital Defence Cooperation in ASEAN

Small states need help in facing this serious threat. Experts credit international cyber assistance and cooperation in enabling Ukraine to defend itself against Russia, one of the world's major cyber powers, in this domain. The US Cyber Command has supported Ukraine with offensive and defensive cyber interventions. These relationships have been built over many years.

In Southeast Asia, the spirit of relationship building is present through platforms led by ASEAN. In the defence sector, ASEAN's ADMM-Plus platform promotes building capacity and trust to address common security challenges, including those associated with cyberspace. As a region that comprises small states, which are pursuing digital economy integration, and facing rising geopolitical uncertainties, the importance of cyber defence relationships to regional security is greater than ever.

Ukraine also receives significant cyber defence assistance from the private sector, global tech firms like Microsoft, and networks of information security researchers. Private sector assistance is crucial because most essential computer systems are privately owned. These relationships also need to be built over many years.

Similarly, ASEAN countries vary in cyber capacity and the extent of their security relationships with the tech industry. Regional collaboration with the industry comes under civilian initiatives such as the ASEAN-Japan Cybersecurity Capacity Building Centre (AJCCBC) and the ASEAN-Singapore Cybersecurity Centre of Excellence (ASCCE). If the region should take a leaf out of Ukraine's playbook, ASEAN's defence sector should step up its collaboration with the industry.

Small states in Southeast Asia can also consider cybersecurity agreements similar to the 2015 China-Russian cybersecurity pact, which has two key features: mutual assurance of non-aggression in cyberspace and language advocacy for cyber sovereignty.

Such pacts are also a confidence-building measure, as they can help states prevent conflict escalation. Because cyberattacks and information operations are difficult to attribute, malicious third parties can use them to sow discord between states so that one state mistakenly blames the other. Regular dialogue between cyber defence communities can help avoid this.

Digital Defence Symposium as Enabler of Cooperation

However, such pacts may not be attainable in Southeast Asia as ASEAN is not an alliance. Furthermore, unlike China and Russia, ASEAN countries are neither major powers with a history of rivalry nor are they on the list of the leading cyber powers. Nonetheless, ASEAN's Cybersecurity Cooperation Strategy guides member countries in collaborating and promoting initiatives to create a peaceful and secure cyberspace in the region.

One such initiative is the ADMM Cybersecurity and Information Centre of Excellence (ACICE), officially launched in 2023 and based in Singapore. Although new, ACICE is well-positioned to support the ASEAN defence sector in achieving the goals of the strategy. Through its flagship conference – Digital Defence Symposium – of which the second iteration will take place on 24-25 July 2024, ACICE plants the seeds of relationship-building between ASEAN and non-ASEAN defence establishments, between ASEAN defence establishments and the tech industry, and as a node for confidence-building measures in the military cyber domain.

When the cyber defenders in ASEAN countries establish regular communications and information sharing, they will not only be able to warn one another of cyber or

disinformation attempts to sow distrust, but they can also work together to counter these cyber or information threats.

More importantly, more trust among them could also strengthen ASEAN's resilience against cyber or information threats that aim to undermine the grouping's cohesion, especially in the present climate of worsening geopolitical tensions.

Senior Fellow Benjamin Ang is the Head of the Centre of Excellence for National Security (CENS) and Head of Digital Impact Research (DIR) at the S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University (NTU), Singapore. He also oversees Future Issues and Technology (FIT) at RSIS. Muhammad Faizal is a Research Fellow in the Regional Security Architecture Programme, at the Institute of Defence and Strategic Studies (IDSS) at RSIS.

S. Rajaratnam School of International Studies, NTU Singapore
Block S4, Level B3, 50 Nanyang Avenue, Singapore 639798