Iran-backed hackers interrupt UAE TV streaming services with deepfake news

Publication Date: 2024-02-08

Author: Dan Milmo

Section: Technology

Tags: Hacking, AI and the US election, Cyberwar, Deepfake, Artificial intelligence (AI), United Arab Emirates, Iran, Middle East and north Africa, news

Article URL: https://www.theguardian.com/technology/2024/feb/08/iran-backed-hackers-interrupt-uae-tv-streaming-services-with-deepfake-news

Iranian state-backed hackers interrupted TV streaming services in the United Arab Emirates to broadcast a deepfake newsreader delivering a report on the war in Gaza, according to analysts at Microsoft. The tech company said a hacking operation run by the Islamic Revolutionary Guards, a key branch of the Iranian armed forces, had disrupted streaming platforms in the UAE with an AI-generated news broadcast branded “For Humanity”. The fake news anchor introduced unverified images that claimed to show Palestinians injured and killed from Israeli military operations in Gaza. Analysts at Microsoft said the hacking group, known as Cotton Sandstorm, published videos on the Telegram messaging platform showing it hacking into three online streaming services and disrupting news channels with the fake newscaster. According to the Khaleej Times, a UAE-based news service, Dubai residents using a HK1RBOXX set-top box were interrupted in December with a message stating: “We have no choice but to hack to deliver this message to you,” followed by the AI-generated anchor introducing “graphic” footage, as well as a ticker showing the number of people killed and wounded in Gaza so far. Microsoft also cited reports of disruptions in Canada and the UK, with the channels affected including the BBC, although the BBC was not hacked directly. Microsoft said in a blogpost accompanying a report on Iranian cyber-espionage: “This marked the first Iranian influence operation Microsoft has detected where AI played a key component in its messaging and is one example of the fast and significant expansion in the scope of Iranian operations since the start of the Israel-Hamas conflict.” “The disruption reached audiences in the UAE, UK, and Canada.” Breakthroughs in generative AI – the term for technology that can swiftly produce convincing text, voice and image from simple hand-typed prompts – have triggered a rise in deepfake content online, from explicit false images of Taylor Swift to robocalls featuring Joe Biden’s AI-generated voice. Deepfake is the term for a hoax using AI to create a phoney image, most commonly fake videos of people. Experts fear AI-made material could be deployed at scale to disrupt elections this year, including the US presidential election. Iran targeted the 2020 US election with a cyber-campaign that included sending intimidating emails to voters purporting to be from members of the far-right Proud Boys group, setting up a website inciting violence against the FBI director, Christopher Wray, and others, and spreading disinformation about voting infrastructure. Microsoft said: “As we look forward to the 2024 US presidential election, Iranian activities could build on what happened in 2020 when they impersonated American extremists and incited violence against US government officials.” Microsoft said Iranian state-backed actors had launched a series of cyber-attacks and online attempts to manipulate opinion since the 7 October Hamas attacks. The tactics include exaggerating

the impact of claimed cyber-attacks, leaking personal data from an Israeli university, and attacking targets in pro-Israel Albania, Bahrain – a signatory to the Abraham accords formalising relations with Israel – and the US.