Date : 15-10-2024

Volume 1, Issue 3

Cybersecurity Digest

Bi-Weekly Update by the CISO Section of Meghalaya Rural Bank

"Never trust anything you see on the Internet"

The Kerala Deepfake scam serves as a reminder of the importance of being vigilant when using technology. People should be careful about the information they share online and should be cautious about the people they interact with online.

The Evolution of Generative AI: Deepfakes

Generative AI refers to the subset of AI technologies capable of creating new content that resembles human-generated outputs, text, images, videos or even code. This capability is built upon machine learning models, particularly generative adversarial networks (GANs), variational autoencoders (VAEs), and, more recently, large language models (LLMs) like OpenAI's GPT series. These models are trained on vast datasets, learning to replicate and innovate on the data patterns they're exposed to.

On the other hand, deepfakes, a portmanteau of "deep learning" and "fake", are a specific application of generative AI focused on creating hyper-realistic video and audio recordings. Leveraging techniques such as GANs and deepfakes can manipulate existing media to make it appear that individuals are saying or doing things they never did. Initially gaining notoriety in misinformation and digital forgery, deepfakes have also found legitimate applications in filmmaking, gaming, and virtual reality, demonstrating the technology's ambivalent potential.

In the meantime, deepfakes, which are high-quality AI-generated fake videos, have been circulating online. Synthetically generated deepfake videos have exceeded acceptable limits in terms of reality distortion. This disruptive technological development significantly impacts the truth. Due to the easy accessibility of generative AI models, the probability of misuse of this technology aggressively increases. Though deepfake technology came into this game in 2017, large language models like ChatGPT, Bing, Bard, GPT-4, etc. add a whole new dimension to this scenario. In particular, political figures are targeted in deepfake videos, contributing to the erosion of media credibility.

This technology has spread information, rumors, and propaganda, provoked political strife, blackmailed individuals, and threatened democracy. The capabilities of this technology to manipulate reality have gone too far. This paradigm-shifting technological revolution is altering the truth and creates a zero-trust society insidiously.

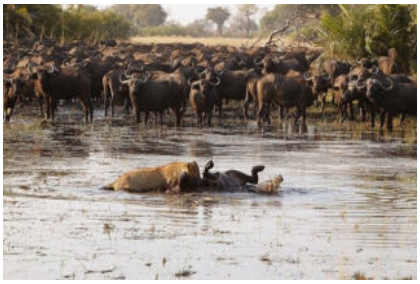
Case Study: Kerala's First Deepfake Fraud

In July 2022,Kerala witnessed its first reported deepfake fraud case. A 73-year-old-man, Radhakrishnan, fell victim to a sophisticated deepfake scam that resulted in the loss of ₹ 40,000.

The victim received a Whatsapp call from someone who appeared to be his former colleague, Venu Kumar. The caller, using deepfake technology, perfectly mimicked Venu Kumar's voice and appearance. The caller asked Radhakrishnan for a loan of ₹ 40000, claiming that he was in urgent need of money. Believing that the caller was indeed his former colleague, Radhakrishnan transferred the money without hesitation.

After realizing that he had been scammed, Radhakrishnan filed a police complaint. The police investigation revealed that the deepfake call was made by a fraudster using sophisticated AI software.

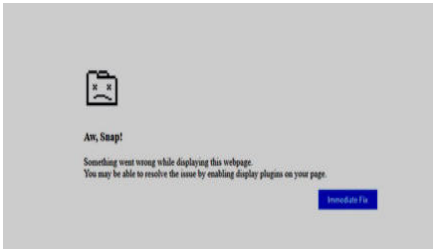
Hunting in Cyberspace: Tracking down Watering Hole Attacks



The term **Watering Hole Attack** comes from the hunting. Rather than tracking its prey over a long distance, the hunter instead determines where the prey is likely to go, most commonly to a body of water (the watering hole) and the hunter awaits there. When the prey comes of its own will, often with its guard down, the hunter attacks.

A watering hole is a cyberattack technique used by attacker to identify websites that are frequently visited by their target victims and then infecting these websites by injecting malicious code (often JavaScript/ HTML) into web pages to gain access to visitors' devices and information or access assets on network.

The payload may be automatic, or the attacker may cause a bogus prompt to appear or fake error page telling the user to take an additional action that will download malicious code.



Fake error page to lure victim to click

How a Watering Hole Attack works?

- **Target Identification**: Attackers identify websites frequently used by their target victims. These websites could be industry forums, professional association websites, or popular news portals.
- **Website Compromise**: Attackers exploit vulnerabilities in the target websites to inject malicious code. This could involve exploiting unpatched software or weak security configurations.
- **Victim Infection**: When victims visit the compromised websites, the malicious code executes on their devices, infecting them with malware. The malware exploits weaknesses in the target's cybersecurity. The attacker delivers and installs malware without the target realizing.

Cyber Crime News

- Cyber attack cripples **Uttarakhand Government IT System**, halts entire digital operations–***The420.in***
- Cyberattack group '**Awaken Likho**' targets Russian Government with advanced tools—***TheHackerNews***
- **American Water**, the largest water utility in US, is targeted by a cyberattack—***APNews***

Cyber Security Awareness Month -October 2024

NABARD, DoS CSITE-Cell is commemorating Cybersecurity Awareness Month in October 2024.

Unleash your inner cyber superhero and join the movement to secure our world.