WHITE PAPER

DEEPFAKE AND ITS IMPACT ON CYBERSECURITY - A NEW FRONTIER TO ADDRESS

Abstract

Deepfake technology, which leverages advanced artificial intelligence (AI) and machine learning techniques to create hyper-realistic but fabricated media, has emerged as a significant challenge to cybersecurity. By manipulating audio & visual content to produce deceptive and convincing simulations, deepfakes have the potential to weaken trust in digital media and create a range of security risks. This whitepaper aims to explore the intersection of deepfake technology and cybersecurity, examining the threats posed by deepfakes, drivers, trends; and defense strategies. By understanding the impact of deepfakes on cybersecurity, stakeholders can develop more effective measures to protect against these emerging threats.

Infosys®
Navigate your next

Introduction

Definition of Deepfake Technology: Deepfake technology refers to the use of artificial intelligence (AI) and machine learning (ML) to create highly realistic but fake audio, video, and images. These synthetic media are generated by algorithms, particularly Generative Adversarial Networks (GANs), which can manipulate and alter existing media to produce convincing forgeries. The term "deepfake" is descended from "deep learning," a subset of AI, and "fake," representing the deceptive nature of the content.

Significance of Deepfake Technology in Cybersecurity: The advent of deepfake technology poses significant challenges to cybersecurity. Deepfakes can be used to create misleading information, impersonate individuals, conduct phishing attacks, and manipulate public opinion, thereby undermining trust and security in digital communications. The ability to produce convincing fake media has far-reaching implications for personal privacy, corporate security, and national security.

The unprecedented rise of Deepfakes - Key Drivers

Some of the key factors contributing to the rise of deepfake technologies are advancing quickly, becoming cheaper, and more accessible.

Availability of datasets & computing power

A large set of data sets with labelled visual material and many of these are freely available on the internet. Moreover, with advancement in AI and cloud computing services at low cost makes it easy for creators to make high-quality deepfakes with even a regular computer or a smartphone with internet access.

Accessibility of high-quality algorithms & pre-trained models

Openly published work or easily accessible code repositories, such as GitHub is often misused by deepfake developers. Additionally, pre-trained machine learning models only need to be trained once and can be reused, repurposed with malign intentions.

5G connectivity

Enabling higher quality videos that are portable anywhere with increased high-speed bandwidth.

Rise of 3D sensors

The latest generation of consumer electronics are equipped with 3D sensors that can be used to capture 3D information of entire scenes and scan objects.

Cat & mouse game between producers & detectors (Same authors writing as creator and detectors)

Increased image forensics & detection capabilities are driving towards increased quality. Using GAN algorithms, learning capacity has increased due to feedback loops. Further they are catalyzed by the availability of shared libraries and codes which are supplemented with products. For e.g., FaceForensics++ tool.

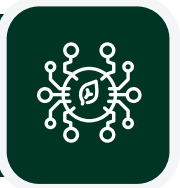
As a result, these drivers lead to a number of trends as follows:

Identity theft and financial frauds

Deepfake technology can be used to create or steal identities to impersonate legitimate individuals and commit financial fraud

Commodification of tools

Few examples include Dfaker, deepFacelab, Faceswap, FakeApp, FaceApp, Zao, DeepNude



Live real time deepfakes

Created via videoconferencing, live streaming video services, etc

Deepfake as service companies

Increased demand for deepfakes has led several companies to offer products and services as a service at low cost

Supply & demand platforms for deepfakes

Special marketplaces (dark web) are available where users or potential buyers can post requests for deepfakes

AI & 3D animation hybrids

AI-driven tech evolution aids to generate synthetic media for 3D and animation applications, but can also be manipulated for harm

"Deepfakes are a bigger problem beyond just being a Cybersecurity issue"
The harms and cascading effects of Deepfake on Individual, Organization, and Society

Psychological harm

- (S)extortion
- Defamation
- Intimidation
- Bullying
- Undermining trust

Financial harm

- Extortion
- Identity Theft
- Payment Fraud
- Stock-price manipulation
- Brand image
- Reputational damage

Societal harm

- News media manipulation
- Damage to economic stability
- Damage to justice system
- Damage to scientific system
- Erosion of Trust
- Damage to democracy
- Damage to national security

Manipulated pornographic video

Manipulated sound clip as evidence

"I'm guilty"

False statement to influence politics

"I'm against gay marriage"

INDIVIDUAL

- Reputation and psychological damage
- Someone is acquitted or made to look suspicious
- Reputation damage for politicians

1

ORGANISATION/GROUP

2

- Blackmail of family and/or company
- Court needs to adjust process to verify evidence
- Constituency loses trust in political party

SOCIETY

3

- Gender inequalities and stereotypes are intensified
- Undermining of the court system
- Loss of trust in institutions

Actors involved in Deepfake Lifecycle

The Perpetrator

The one who creates deepfake

The Victim

The targeted individual

The Technology provider

The company who developed software

The original performers

People whose figures are still used in deepfake videos

The original Authors

The company that produced original film content plays a role in legal context claiming copyrights

The Platform

An online forum or intermediary used

The Platform Users

People who upload, view or share

Managed Security Service Provider Solutions to & combat Deepfakes

Managed Security Service Providers (MSSPs) can be an integral strategy to offer a range of solutions to help organizations combat the threat of deepfakes. Here are some key areas:

Detect & Defend with AI

use of AI technologies from leading OEMs to identify changes in voice and image that may indicate deepfake.

Zero Trust Maturity Model

The Cybersecurity and Infrastructure Security Agency (CISA) proposes the Zero Trust Maturity Model to help in mitigating deepfake attacks. This can help in detecting and preventing deepfake attacks before they cause harm.

Continuous Monitoring and Threat Intelligence

Equip Security Operations Centers with monitoring and detection capabilities 24X7.

Digitally watermarking audio visual materials

User Engagement & Open Communication

Define policy at top level. This entails making it easier for people to know where a specific piece of information came from, how it was created, and whether it is reliable.

IAM solutions & Multifactor Authentication

Implement robust IAM practices including MFA challenge response system to provide extra layer of security.

Red Teaming for AI systems

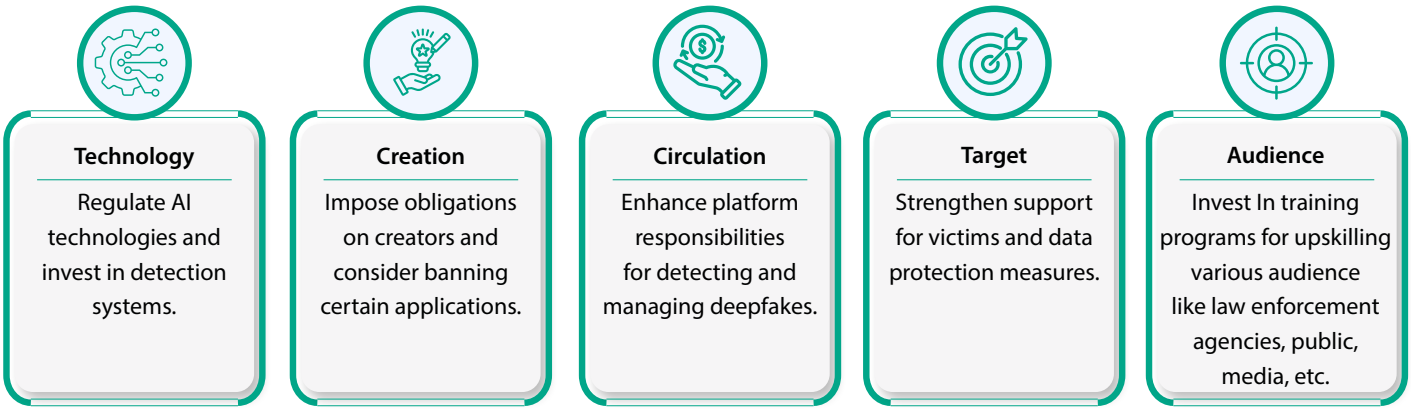
Systematic testing of AI systems to identify vulnerabilities and prevent misuse for creating deepfake content.

Invest in awareness and training

Cyber Security Awareness, campaigns, Training software; Support people who spot deepfakes.

Call to action

A multi-dimensional level approach is required to counter deepfake threat, with the government playing a crucial role in policy making. Addressing deepfakes requires a multi-dimensional approach with Policy recommendation at each level. Below are the five dimensions:



Collaboration across various bodies like Law enforcement agencies, Industry partners and academia can raise awareness around deepfake can create a robust defense against deepfake threats, thereby protecting individuals, organizations, and society as a whole.

Conclusion

With the rise of deepfakes it has shown both the creative and destructive potential of synthetic media. In particular to Cyber space, it has added another layer to cybersecurity threats. From misinformation campaigns to identity theft, the outcomes of this innovative technology are far-reaching. It impacts not only the enterprises but also a threat to national security for example in geo-political, and cyberwarfare situations misleading society. Apart from the malicious use of deepfake; then beneficial application of the technology has made rapid advancements and more convincing than ever; thus, making easy picking for fraudsters to use the technology for malicious intent.

However, as daunting as the threat of deepfakes may appear, it is not undefeatable. We can reduce the risks posed by deepfakes significantly - by raising awareness, executing defensive strategies, leveraging the power of artificial intelligence, and promoting open communication.

This would mean the importance of maintaining constant surveillance and taking proactive cybersecurity measures. As we progress into the new digital age, it becomes clear that our cybersecurity strategy must be as dynamic and adaptable as the threats we face.

References:

https://cybervie.com/blog/deepfake-impact-on-cybersecurity/

https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2021)690039

Deepfakes and LLMs: The Cybersecurity Frontier We Can't Ignore

https://www.gartner.com/en/documents/5358363

About the Author

Varun Soni

Principal Consultant

Varun has 20 years of cybersecurity experience and technical expertise in multiple cyber domains. His specialization includes security strategy planning and implementation, consulting, Center of Excellence, and Managed Security operations. He is an accomplished leader with demonstrated hands-on-experience in establishing and implementing large-scale cybersecurity transformations for Fortune 500 companies globally, across industry verticals. In his current role, he leads the strategic partnership with Zscaler, managing the joint go-to-market cybersecurity offerings. He has completed professional security certifications viz. CISSP, CISM, CCSK, AWS-SA, and many more.

For more information, contact askus@infosys.com

Infosys®
Navigate your next