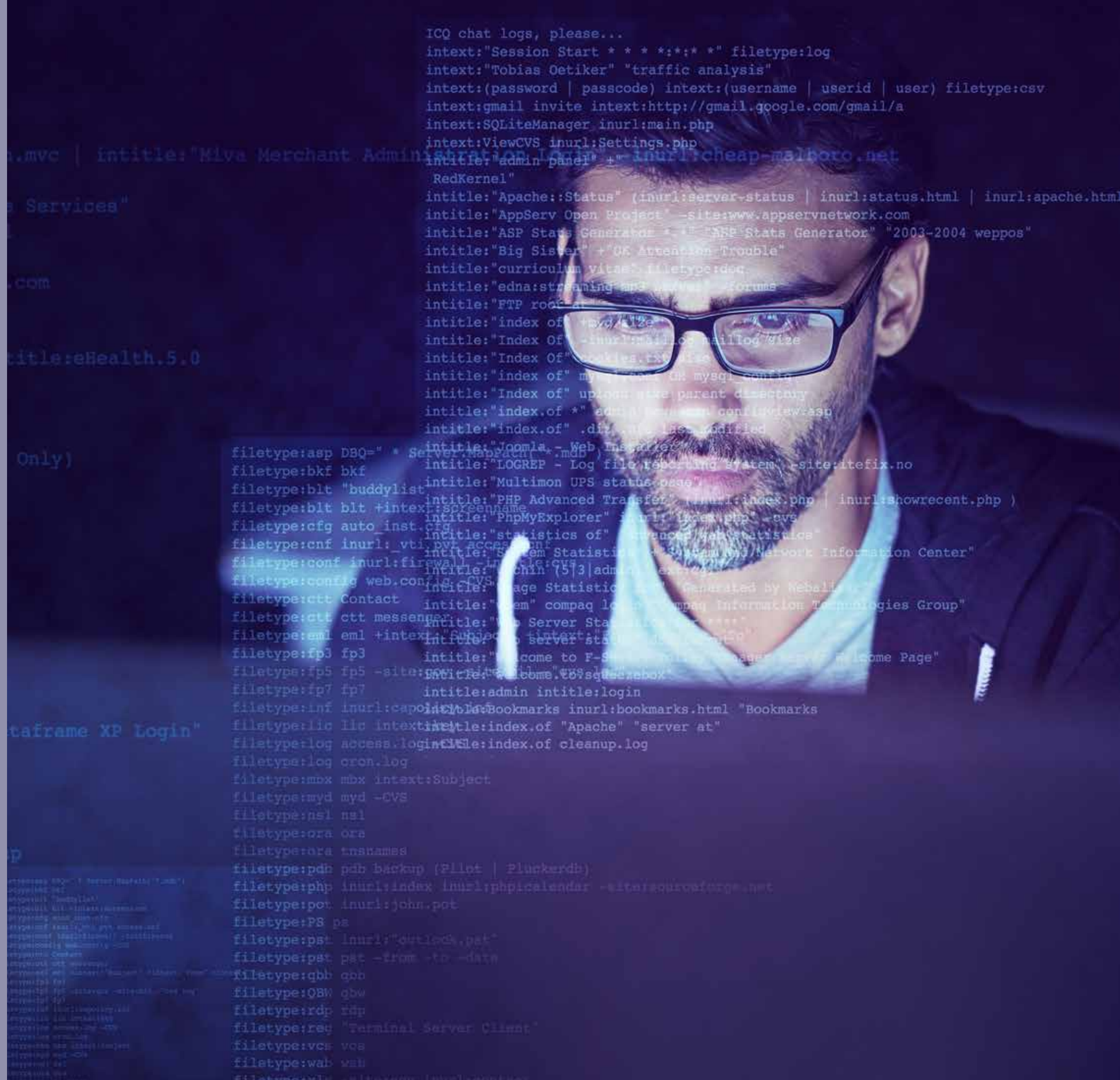
AXA XL Insurance

Deepfakes: an emerging cyber threat that combines AI, realism and social engineering



The use of deepfakes to commit cyber crime is on the rise, even though most criminals currently lack the sophistication and expertise to use the technology effectively. But as artificial intelligence evolves and becomes more user friendly, deepfakes will likely play a bigger role in cyber crime.

Cyber criminals are always looking for new methods to gain access to funds and critical information. That's largely because organizations have caught on to the phishing schemes and sneaky ransomware attacks of the past and have implemented tools and educated employees to keep cyber thieves at bay.

But with the rapid evolution of artificial intelligence, especially generative AI, and public demand for easy access to the new technology, cyber crime is primed to make a big leap forward.

Deepfakes – videos, pictures or audio made with AI to appear real – are the latest weapons in the cyber criminals' arsenal. There has been a notable rise in the use of deepfakes to commit cyber crime, and the only things slowing the progress is the sophistication, expertise and effort required to effectively use the technology.

But as the technology becomes more accessible and more user-friendly, deepfakes will likely play a big role in the future of cyber crime. Combatting this emerging threat will require staying one step ahead of the criminals through the use of technology and strict protocols to block access to funds or critical information and, most importantly, educating employees about how to spot a deepfake.

Why are deepfakes so effective?

Deepfakes, as we know them today, are videos, images or audio that are created to look realistic with the use of artificial intelligence In a broader sense, deepfake technology has been around for decades. Early motion pictures manipulated images long before computers and AI were available. And fake audio recordings are easy to create with or without the use of modern technology – a good celebrity impersonator can be just as effective.

Deepfake technology has been widely employed in the world of politics. Official campaigns typically avoid using such tactics, but a candidate's supporters have been known to create unflattering and realistic images of political opponents to share on social media.

The technology has some legitimate uses, particularly in the world of cinema. The motion picture, "Forrest Gump," could not have been made without the blending of Tom Hanks' main character with historical footage. More recently, AI technology was used to fascinating and realistic effect in the Netflix documentary, "Dirty Pop," about boy-band impresario Lou Pearlman. The producers combined dialogue from Pearlman's memoir with a video of him speaking to the camera and a "mouth actor" who moved his lips to match the words and some AI effects to bring it all together. If the program hadn't warned viewers beforehand that it used AI-generated trickery to make Pearlman a narrator of his own documentary, it would be hard to tell.

Deepfakes can be an effective tool for cyber crime because of ocial engineering, which is the psychological manipulation of eople into performing actions or divulging confidential information, such as passwords or access to financial accounts. Social engineering is often one of many steps in a more complex fraud scheme.

If you've ever received a realistic-looking email that appears to be from your bank or cable TV company but is actually from an unfamiliar email account, that's a scam utilizing the concept of social engineering. Deepfakes are similar to a fake email scam, but taken to a new, more sophisticated level.

Most cyberattacks that employ social engineering techniques play on the victim's emotions and create a sense of urgency, because the cyber criminals want to put the victim in an emotional state and get them to make a decision quickly before they have time to think about it critically.

Deepfake detection and protection

As quickly as deepfake technology has evolved, so have the methods to detect when an image, video or audio file is an AI-generated fake. There are several software tools available that can help detect a fake video. It's like using AI for good against those who would use it to commit crimes.

Aside from letting AI do all the detective work, humans well trained in identifying deepfakes can uncover the truth by simply analyzing the quality and consistency of the video or image. Distortions, blurriness and mismatched colors or objects can raise suspicions. Also, look for unusual behavior in those speaking in the video, like awkward motions, unnatural positions and lip movement that doesn't sync with the words being heard. Finally, it's critical to verify the source and the origin of the video or image.

To protect your organization against deepfake cyber threats, continue to follow the same tried-and-true cybersecurity protocols you have in place. Deepfakes by themselves are not a security threat, but they can be a means to an end for nefarious types to get past security protocols. Deepfakes, therefore, are really a variation of an existing threat that can make social engineering scams harder to detect.

The human element continues to be one of the biggest dangers organizations face when it comes to cybersecurity.

Organizations should continue to utilize multi-factor authentication (MFA), an electronic authentication method in which a user is granted access to a website or application only after successfully presenting two or more pieces of evidence to an authentication mechanism. Think of signing on to a website with a login, a password and, finally, a 6-digit code that was sent as a text to your smart phone.

As new threats appear on the horizon, such as deepfakes, organizations should continually review and update employee cyber training.

Deepfakes require a level of sophistication, training and effort that most cybercriminals have not yet mastered, but they are a real and emerging risk. Employees should be trained now about how to identify a deepfake threat and protect critical information from cybercriminals.

Distortions, blurriness and mismatched colors or objects can raise suspicions. Also, look for unusual behavior in those speaking in the video, like awkward motions, unnatural positions and lip movement that doesn't sync with the words being heard.

AXA XL's comprehensive cyber and technology policy offers extensive coverage and unparalleled customer service, partnering with you to prepare for and prevent cyber events, mitigate fallout during a breach, and build resilience afterward.

axaxl.com

3204_10/2024