Trend Micro Stops Deepfakes and AI-Based Cyberattacks for Consumers and Enterprises

Cybersecurity leader announces capabilities to safeguard business resilience and AI adoption

DALLAS, July 30, 2024 /PRNewswire/ -- Trend Micro Incorporated (TYO: 4704; TSE: 4704), a global cybersecurity leader, today announced new innovations in its enterprise platform and consumer cybersecurity products focused on protecting all environments from the rapidly growing threat of AI-based attacks and fraud.

Trend research shows that cybercriminals are catching on to the explosion of enterprise AI use, resulting in a dramatic increase in AI-based tools available on the criminal underground. These tools are cheaper and more accessible than ever, enabling criminals at any skill level to more easily launch attacks at scale that mislead victims for purposes of extortion, identity theft, fraud, or misinformation.

Trend's new solutions launch as part of a company-wide mission to secure the AI journey for consumers and enterprises.

Kevin Simzer, COO at Trend: "Our latest research reveals several new deepfake tools that make it easy for cybercriminals at all skill levels to launch damaging scams, social engineering, and security bypass attempts. We are leading the industry in fighting back for both our enterprise and consumer customers with new capabilities to detect deepfakes and other forms of AI fraud. Like past shifts in the threat and IT landscape, we've seen the challenge of securing AI and risen to it."

Detecting and defeating these AI-based methods is central to better managing attack surface risk for enterprises and lowering overall online risk for consumers—71% of whom, in a recent Trend survey, viewed deepfakes negatively and believed that one of their top uses is for fraud.

Available soon in the Trend Vision One™ platform, new deepfake detection technology will use a variety of advanced methods to spot AI-generated content. This capability is also available for consumers today in Trend Micro Deepfake Inspector.

To learn more and download Trend Micro Deepfake Inspector, visit:https://www.trendmicro.com/deepfake-inspector

Going beyond techniques like image noise analysis and color detection, the platform also analyses user behavioral elements to provide a much stronger approach to detecting and stopping deepfakes. Upon detection, Trend immediately alerts enterprise security teams, enabling them to learn, educate, and take proactive measures to prevent future attacks.

Trend Micro Deepfake Inspector can help verify if a party on a live video conversation is using deepfake technology, alerting users that the person or persons with whom they are conversing may not be who they appear to be.

According to Gartner[1] analyst Dan Ayoub, "Readily available, high-quality GenAI applications are now capable of creating photo-realistic video content that can deceive or mislead an audience. Given the low barriers to entry in using these tools and their increasing sophistication, developing a methodological approach to detecting GenAI deepfake content has become necessary."

Deepfakes pose a significant risk to modern enterprises and individuals. An undetected deepfake can lead to financial impacts, job losses, legal challenges, reputation damage, identity theft, and potential harm to mental or physical health. In a recent Trend Micro study, 36% of consumers reported experiencing a scam attempt using a deepfake. The FBI has also previously warned of deepfake technology being used in conjunction with video calls to carry out business email compromise attacks, and to fraudulently apply for remote working positions.

This technology is not only being abused to bypass human verification but also biometric security measures like facial recognition. Trend research has also revealed recent shifts indicating a growing preference for exploiting existing LLM models through innovative jailbreaking techniques rather than developing bespoke criminal AI tools.

The launch of Trend's new solutions is part of a company-wide mission to secure customers' AI journey. Supporting a zero trust strategy, Trend also recently released new features for Trend Vision One designed to:

- Centralize management of employees' GenAI access and usage
- Inspect prompts to prevent data leaks and malicious injections
- Filter GenAI content to meet compliance requirements
- Defend against large language model (LLM) attacks

Trend Micro Deepfake Inspector is a free solution designed to alert users to potential deepfakes while they are on video call. Analysis takes place in real time and locally on the consumer device, ensuring users' data and privacy are protected at all

times.

About Trend Micro
Trend Micro, a global cybersecurity leader, helps make the world safe for exchanging digital information. Fueled by decades of security expertise, global threat research, and continuous innovation, Trend Micro's AI-powered cybersecurity platform protects hundreds of thousands of organizations and millions of individuals across clouds, networks, devices, and endpoints. As a leader in cloud and enterprise cybersecurity, Trend's platform delivers a powerful range of advanced threat defense techniques optimized for environments like AWS, Microsoft, and Google, and central visibility for better, faster detection and response. With 7,000 employees across 70 countries, Trend Micro enables organizations to simplify and secure their connected world. www.TrendMicro.com.

[1]**Gartner, Emerging Tech: Combating Deepfakes, Deception and Disinformation in Multimedia Content, Dan Ayoub, 17 April 2024**

Gartner is a registered trademark and service mark, Peer Insights and Magic Quadrant are registered trademarks, of Gartner, Inc. and/or its affiliates in the U.S. and internationally and are used herein with permission. All rights reserved. Gartner Peer Insights content consists of the opinions of individual end users based on their own experiences with the vendors listed on the platform, should not be construed as statements of fact, nor do they represent the views of Gartner or its affiliates. Gartner does not endorse any vendor, product or service depicted in this content nor makes any warranties, expressed or implied, with respect to this content, about its accuracy or completeness, including any warranties of merchantability or fitness for a particular purpose.

Gartner does not endorse any vendor, product or service depicted in its research publications and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

SOURCE Trend Micro Incorporated

For further information: Trend Micro Communications, 817-522-7911, media_relations@trendmicro.com

---

Additional assets available online: Video (1)

https://newsroom.trendmicro.com/2024-07-30-Trend-Micro-Stops-Deepfakes-and-AI-Based-Cyberattacks-for-Consumers-and-Enterprises