Cómo reconocer llamadas falsas con IA: una amenaza creciente para las empresas

Desde ESET advierten que es cada vez es más difícil distinguir lo real de lo generado por inteligencia artificial, y que la suplantación de voz ya se usa para atacar a empresas.

Ciudad de México, México— La IA Generativa (GenAI) democratizó la creación de audio y vídeo falsos, hasta el punto de que generar un clip fabricado es tan fácil como pulsar un botón o dos. Las deepfakes pueden ser usadas de varias formas: desde eludir autenticaciones y controles, hasta infiltrarse en organizaciones creando un candidato falso y sintético para procesos de selección de personal. Sin embargo, podría decirse que la mayor amenaza que plantean es el fraude financiero/transferencias bancarias y el secuestro de cuentas de ejecutivos. ESET, compañía líder en detección proactiva de amenazas, acerca herramientas para poder identificar cuando una llamada es falsa.

El Gobierno británico afirma que el año pasado se compartieron hasta 8 millones de clips falsos, frente a los 500 mil que habían sido compartidos en 2023. La cifra real puede ser mucho mayor y, así las cosas, las organizaciones tienden a subestimar esta amenaza.

Como ha demostrado un experimento de Jake Moore, Global Security Advisor de ESET, nunca ha sido tan fácil lanzar un ataque de audio deepfake. Todo lo que se requiere es un clip corto de la víctima para ser suplantado y GenAI puede hacer el resto. Así es como podría proceder un ataque, según ESET:

1. Un atacante selecciona a la persona que va a suplantar. Puede ser un CEO, un CFO o incluso un proveedor.
2. Encuentra una muestra de audio en Internet, lo que resulta bastante fácil para ejecutivos de alto nivel que hablan en público con regularidad. Puede proceder de una cuenta en las redes sociales, de una convocatoria de beneficios, de una entrevista en vídeo o televisión o de cualquier otra fuente. Unos segundos de grabación deberían bastar.
3. Seleccionan a la persona a la que van a llamar. Para ello, puede ser necesario realizar una investigación documental, normalmente en LinkedIn, en busca de personal del servicio de asistencia informática o miembros del equipo financiero.
4. Pueden llamar directamente a la persona o enviar un correo electrónico por adelantado: por ejemplo, un director general que solicita una transferencia de dinero urgente, una solicitud de restablecimiento de contraseña o autenticación multifactor (MFA), o un proveedor que exige el pago de una factura vencida.
5. Llaman al objetivo preseleccionado, utilizando audio deepfake generado por GenAI para hacerse pasar por el CEO/proveedor. Dependiendo de la herramienta, pueden ceñirse a un discurso preestablecido o utilizar un método más sofisticado de "voz a voz" en el que la voz del atacante se traduce casi en tiempo real a la de su víctima.

"Este tipo de ataque es cada vez más barato, sencillo y convincente. Algunas herramientas son capaces incluso de insertar ruido de fondo, pausas y tartamudeos para que la voz suplantada resulte más creíble. Cada vez imitan mejor los ritmos, las inflexiones y los tics verbales propios de cada orador. Y cuando un

ataque se lanza por teléfono, los fallos relacionados con la IA pueden ser más difíciles de detectar para quien atiende", advierte Mario Micucci, Investigador de Seguridad Informática de ESET Latinoamérica.

Los atacantes también pueden utilizar tácticas de ingeniería social, como presionar a que la persona responda urgentemente a su petición, con el fin de lograr sus objetivos. Si a esto se le añade que a menudo se hacen pasar por un alto ejecutivo, es fácil ver por qué algunas víctimas son engañadas. Uno de los mayores errores se produjo en 2020, cuando se engañó a un empleado de una empresa de los Emiratos Árabes Unidos haciéndole creer que su director había llamado para solicitar una transferencia de fondos de 35 millones de dólares para una operación de fusión y adquisición.

Dicho esto, hay formas de detectar a un impostor. Desde ESET sostienen que dependiendo de lo sofisticada que sea la GenAI que están utilizando, puede ser posible discernir:

- Un ritmo antinatural en el discurso del orador
- Un tono emocional antinaturalmente plano en la voz del orador
- Respiración antinatural o incluso frases sin respiración
- Un sonido inusualmente robótico (cuando se utilizan herramientas menos avanzadas)
- Ruido de fondo extrañamente ausente o demasiado uniforme

Además, en términos corporativos se recomienda empezar por la formación y concienciación de los empleados. Estos programas, según ESET, deben actualizarse para incluir simulaciones de audio deepfake que garanticen que el personal sepa qué esperar, qué está en juego y cómo actuar. Se les debe enseñar a detectar los signos reveladores de la ingeniería social y los escenarios típicos de deepfake. Deben realizarse ejercicios de red team para comprobar si los empleados asimilan bien el proceso correcto:

- Verificación fuera de banda de cualquier solicitud telefónica, es decir, utilizar cuentas de mensajería corporativas para comprobar con el remitente de forma independiente
- Dos personas que firmen las transferencias financieras importantes o los cambios en los datos bancarios de los proveedores
- Contraseñas o preguntas acordadas previamente que los ejecutivos deban responder para demostrar que son quienes dicen ser por teléfono

"Las falsificaciones son sencillas y su producción cuesta poco. Dadas las enormes sumas que pueden obtener los estafadores, es poco probable que veamos pronto el final de las estafas de clonación de voz. Por lo tanto, la mejor opción que tiene una organización para mitigar el riesgo es un triple enfoque basado en las personas, los procesos y la tecnología. Para que se adapte a medida que avanza la innovación en IA, es importante que sea revisado periódicamente. El nuevo panorama del ciberfraude exige una atención constante", concluye Micucci de ESET.

Para saber más sobre seguridad informática visite el portal corporativo de ESET: https://www.welivesecurity.com/es/seguridad-corporativa/como-reconocer-llamadas-falsas-con-ia-una-amenaza-creciente-para-las-empresas/

Por otro lado, ESET invita a conocer **Conexión Segura**, su podcast para saber qué está ocurriendo en el mundo de la seguridad informática. Para escucharlo ingrese a: https://open.spotify.com/show/0Q32tisjNy7eCYwUNHphcw

Visítanos en:

@ESETLA /company/eset-latinoamerica /esetla /ESETLA /@esetla

Acerca de ESET

ESET® proporciona seguridad digital de vanguardia para prevenir ataques antes de que ocurran. Al combinar el poder de la IA y la experiencia humana, ESET® se anticipa a las ciberamenazas conocidas y emergentes, asegurando empresas, infraestructuras críticas e individuos. Ya sea protección de endpoints, nube o dispositivos móviles, sus soluciones y servicios nativos de IA y basados en la nube son altamente efectivos y fáciles de usar. La tecnología de ESET incluye detección y respuesta sólidas, cifrado ultraseguro y autenticación multifactor. Con defensa en tiempo real las 24 horas, los 7 días de la semana y un sólido soporte local, mantiene a los usuarios seguros y a las empresas funcionando sin interrupciones. Un panorama digital en constante evolución exige un enfoque progresivo de la seguridad: ESET® está comprometido con una investigación de clase mundial y una potente inteligencia sobre amenazas, respaldada por centros de I+D y una sólida red global de socios. Para obtener más información, visite https://www.eset.com/latam o síganos en LinkedIn, Facebook y Twitter.

Datos de Contacto de Comunicación y Prensa

*Para más información se puede poner en contacto a través de **prensa@eset-la.com** o al +54 11 2150-3700.*

Contacto de prensa Serna PR

Gesuri Lazcano |gesuri.lazcano@sernagrp.com | Cel. 55 9106 3371
Alma Sánchez | alma.sanchez@sernagrp.com | Cel. 55 2748 7884