LITORALPRESS

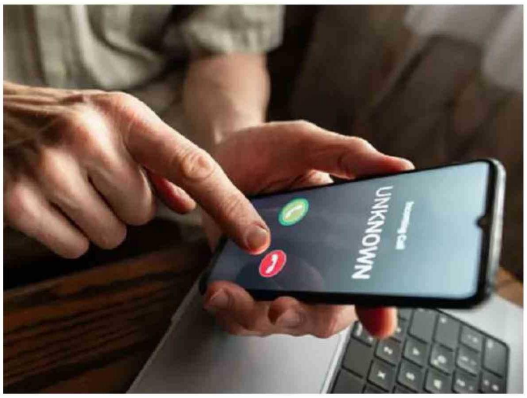
Fecha: 18-01-2025
Medio: El Divisadero
Supl. : El Divisadero
Tipo: Noticia general
Título: Nuevo tipo de estafas mediante uso de IA

Pág. : 3
Cm2: 593,9

Tiraje: 2.600
Lectoría: 7.800
Favorabilidad: No Definida

Del fraude por e-mail a falsas voces y rostros:

Nuevo tipo de estafas mediante uso de IA



El avance de la Inteligencia Artificial (IA) ha transformado para siempre nuestra vida digital. Aunque los chatbots y algoritmos hacen más fácil y eficiente nuestra experiencia en línea, también han creado nuevas amenazas de ingeniería social, que al igual que las estafas tradicionales, buscan robar datos personales, información bancaria y detalles sensibles, tanto de personas como de empresas.

Isabel Manjarrez, investigadora de seguridad del Equipo Global de Investigación y Análisis de Kaspersky, señala que con la IA, los ciberdelincuentes han perfeccionado tácticas como las estafas de phishing, cuyos ataques en Chile aumentaron 125% en 2024, en comparación con el año anterior, según el Panorama de Amenazas de Kaspersky.

Agrega que a través de mensajes falsos, los estafadores engañan a sus víctimas para que revelen inadvertidamente sus correos, contraseñas o datos de sus tarjetas bancarias. Se dirigen tanto a usuarios como a empresas, de forma masiva o personalizada, disfrazados de notificaciones de bancos, proveedores de servicios, sistemas de pago electrónico u otras organizaciones, incluso puede parecer que provienen de alguien que conocemos.

"Es común que estos mensajes de estafa estén mal redactados y tengan errores ortográficos, no obstante, los modelos de aprendizaje automático (LLM's) han mejorado la capacidad de los delincuentes para que los mensajes tengan una gramática impecable y sean mucho más convincentes; incluso logrando imitar páginas auténticas de diferentes compañías aumentando así el éxito de sus ataques. Con esto, ahora pueden dirigirse a sus víctimas en distintos idiomas y replicar estilos de comunicación de una persona específica, como un amigo o un socio, pues los modelos de aprendizaje también son capaces de analizar publicaciones en redes sociales u otra información pública de esa persona", indica la experta.

Otro ejemplo que detalla Isabel Manjarrez son las ultra falsificaciones, también conocidas como *deepfakes*: contenidos multimedia manipulados con Inteligencia Artificial. "Con apenas unos segundos de una grabación de voz, la IA puede generar clips de audio; además, también facilita la alteración de imágenes y videos modificando rostros y sus expresiones. Se estima que, en promedio, ocurre un intento de *deepfake* cada cinco minutos y que, para 2026, hasta el 90% del contenido en línea podría generarse de esta forma. Esto es preocupante porque, dado que los *deepfakes* pueden replicar de manera convincente la imagen o la voz de una persona, el riesgo de suplantación de identidad aumenta", advierte la investigadora.

Explica que los ciberdelincuentes han adoptado estas herramientas avanzadas para emplear tácticas más complejas. Por ejemplo, roban cuentas de aplicaciones de mensajería, como Telegram o WhatsApp; con los mensajes de voz en los chats crean grabaciones que imitan a los dueños de esas cuentas y las envían a contactos de confianza de las víctimas, como amigos, familiares o clientes, para estafarlos. Además, al poder manipular imágenes y videos, realizan videollamadas suplantando la identidad de personas conocidas.

"Con estas técnicas, los delincuentes solicitan transferencias bancarias urgentes o información confidencial, aprovechando la confianza de otros para cometer fraude, tanto a nivel personal como empresarial. Es alarmante, pero posible", señala.

Añade que los ataques con *deepfakes* han evolucionado de tal forma que los estafadores también recurren a imágenes y voces de celebridades o figuras públicas para difundir sus trampas. Estas incluyen desde anuncios falsos que promueven hacer inversiones en plataformas fraudulentas o inexistentes hasta estafas de tipo romántico en las que engañan a las víctimas para ganar su confianza y pedirles dinero. También han utilizado audios falsos en ataques dirigidos a individuos y sistemas bancarios con autenticación por voz.

"En resumen, con la Inteligencia Artificial los criminales pueden automatizar la producción masiva de contenido fraudulento, haciendo sus ataques más sofisticados y difíciles de detectar. Por eso, a medida que esta tecnología avanza, nuestra defensa debe enfocarse en dos frentes: técnico y educativo", sostiene Isabel Manjarrez.

Medidas preventivas

La experta indica que, a nivel técnico, existen soluciones prometedoras que se pueden adoptar, como las marcas de agua, para etiquetar contenido generado por IA; detectores de *deepfakes*, para identificar características específicas de contenido manipulado; y firmas digitales, las cuales ya se utilizan en transacciones bancarias y comunicaciones importantes, para verificar la autenticidad de imágenes, videos y audios alterados. Sostiene que el principal desafío de estas medidas es evolucionar a la misma velocidad que los grandes modelos de lenguaje y la IA generativa, exigiendo una constante actualización para que sean efectivas.

Isabel Manjarrez expresa que a nivel educativo, hay una brecha crítica; un desconocimiento de lo fácil que es explotar la Inteligencia Artificial.

"La ciberdelincuencia aprovecha esta falta de conocimiento, lo que resalta la necesidad de un diálogo abierto y de campañas educativas sobre los riesgos. Si bien los *deepfakes* y las estafas impulsadas por la IA presentan retos cada vez mayores, comprender estas amenazas es el primer paso para enfrentarlas. No hay por qué tener miedo, si combinamos soluciones y mejores prácticas de seguridad con una adecuada alfabetización cibernética, tanto usuarios como organizaciones podemos reducir los riesgos y contribuir a la construcción de un entorno digital más seguro y resiliente", concluye.