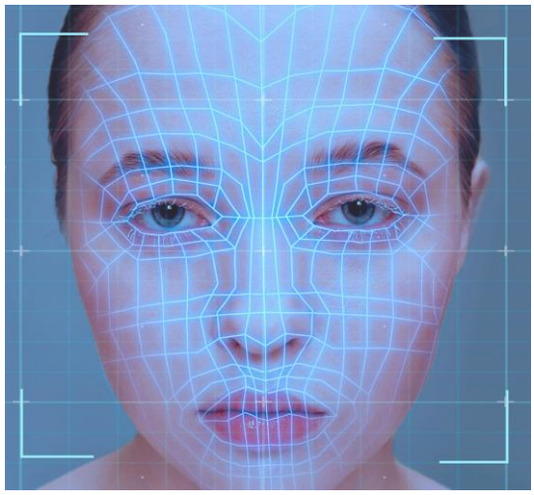
Alerta

Técnica

Sofisticación de Phishing con robo biométrico - IoCs

COLCERT – 20260203 - 092

TLP:CLEAR



Resumen Ejecutivo

El **ColCERT** ha identificado una evolución crítica en las tácticas de la campaña de phishing dirigida a entidades gubernamentales. La amenaza ha mutado de la recolección estática de fotografías ("selfies"), reportada el pasado 30 de enero, hacia un sistema avanzado de fraude biométrico dinámico.

Mediante el uso de Inteligencia Artificial ejecutada directamente en el navegador de la víctima, los atacantes ahora exigen "Pruebas de Vida" (gesticulaciones faciales) para obtener videos de alta calidad. Este material está diseñado para vulnerar sistemas modernos de autenticación y enrolamiento digital que requieren validación de identidad en tiempo real.

Detalles de la evolución técnica

❑ **Procesamiento de IA en el Cliente**

El análisis forense del artefacto malicioso (face.html) revela la integración de la librería face-api.js (basada en TensorFlow). Al acceder al sitio, se descargan automáticamente modelos de redes neuronales (pesos de IA) que permiten al navegador rastrear 68 puntos de referencia faciales en tiempo real.

❑ **Mecanismo de "Prueba de Vida" (Liveness Detection)**

A diferencia de ataques previos, la generación del archivo de video (rostro.webm) está condicionada por algoritmos de detección. El script obliga al usuario a realizar gestos específicos, como parpadear o sonreír. Solo cuando la IA local valida estas acciones físicas, se procede a la captura y preparación de los datos para su envío.

❑ **Instrumentalización de Servicios Legítimos (Webhooks)**

Para la salida de los datos (exfiltración), los atacantes están abusando de funciones de automatización conocidas como Webhooks. Se ha identificado el uso de la infraestructura de Discord como canal de recepción, permitiendo que el tráfico malicioso se mimetice con el tráfico normal de la red institucional para evadir sistemas de detección perimetral.

***Neutralidad técnica:** Se ha identificado el uso de la infraestructura de Discord como canal de recepción. Esta técnica permite que el tráfico malicioso se mimetice con el tráfico normal de la red institucional, evadiendo sistemas de detección perimetral.*

Indicadores de compromiso IoCs

Tipo de Indicador	Valor / Indicador	Observación Técnica
URL de Aterrizaje	hxxps://verificationmicrosoftsit[.]zya[.]me/face.html	Sitio de phishing con motor de captura biométrica.
Dirección IP C2	185.27.134.34	Servidor host bajo infraestructura OpenResty.
Endpoint de Salida	hxxps://discord[.]com/api/webhooks/1335759795094261882/	Canal de exfiltración (Webhook) instrumentalizado.
Hash SHA-256	cda4069024092b8f86d21ff8684a0fd106b740ad3fe901f0673f75fa708d2ec	Firma digital del script malicioso face.html.
Nombre de Archivo	rostro.webm	Video generado tras la prueba de vida.

Comparativa de evolución de IoCs

Tipo de Indicador	Reporte Inicial (2026-01-30)	Nuevos Hallazgos (Hoy)
URL de Aterrizaje	sites.google.com/view/micrologi/...	verificationmicrosoftsit[.]zya[.]me/face.html
Infraestructura C2	Servidores de Google y Replit	185.27.134.34 (Infraestructura OpenResty)
Canal de Salida	Endpoint en replit.dev	Webhook de Discord
Artefacto/Payload	Cuerpo HTML estático	Script face.html con IA de reconocimiento
Dato Exfiltrado	Fotografía estática ("selfie")	Video dinámico (rostro.webm)

Mapeo táctico mitre att&ck

Táctica	ID Técnica	Nombre de la Técnica	Aplicación en el Incidente
Acceso Inicial	T1566.002	Spearphishing Link	Uso de enlaces maliciosos bajo el señuelo de "Validación de Cuenta".
Ejecución	T1059.007	JavaScript	Ejecución de face-api.js para procesamiento biométrico en el navegador.
Defensa Evasión	T1567.001	Exfiltration Over Web Service	Abuso de Webhooks de Discord para evadir filtros perimetrales.
Defensa Evasión	T1564.008	Web Portal Messenger	Uso de plataformas legítimas de mensajería para ocultar el tráfico C2.
Recolección	T1125	Video Capture	Grabación automatizada de video tras validación de gesticulación facial.
Exfiltración	T1041	Exfiltration Over C2 Channel	Envío de los archivos capturados (rostro.webm) al servidor del atacante.

El análisis basado en el marco de trabajo MITRE ATT&CK revela una operación que prioriza la evasión de defensas mediante el uso de infraestructura legítima. El **ataque comienza con un acceso Inicial basado en ingeniería social**, **pero escala rápidamente en complejidad técnica al utilizar JavaScript avanzado** para realizar un procesamiento de IA en el dispositivo de la víctima (Edge Computing).

El punto más crítico se observa en la recolección, donde el adversario no solo captura datos, sino que valida la presencia física del usuario mediante video. Finalmente, la exfiltración se camufla utilizando servicios de mensajería de confianza, lo que permite que el robo de identidad biométrica pase inadvertido para los sistemas de monitoreo convencionales al no generar alertas de conexiones a dominios sospechosos conocidos.

Recomendaciones técnicas y preventivas

A. Para Equipos de TI y seguridad (Administradores)

Bloqueo de Infraestructura Maliciosa:

- ❑ Implementar el bloqueo inmediato en Gateways de enlace y Firewalls del dominio raíz zya.me y sus subdominios asociados.
- ❑ Denegar conexiones hacia la dirección IP 185.27.134.34.

Monitoreo de Exfiltración (Webhooks):

- ❑ Configurar reglas de detección para peticiones POST hacia la API de Discord (discord.com/api/webhooks/) que no correspondan a herramientas de integración autorizadas por la entidad.
- ❑ Auditar el tráfico saliente en busca de archivos con extensión .webm o .mp4 hacia servicios de mensajería o almacenamiento en la nube no corporativos.

Endurecimiento de Navegadores (Hardening):

- ❑ Aplicar políticas de grupo (GPO) para restringir el acceso automático a periféricos de video y audio. La cámara debe permanecer bloqueada por defecto y requerir autorización explícita para sitios de confianza exclusivamente.
- ❑ Bloquear la ejecución de scripts provenientes de dominios de baja reputación o creados recientemente.

Detección de Modelos de IA:

- ❑ Monitorear la descarga de archivos .weights, .shard o archivos JSON de gran tamaño asociados a librerías de visión artificial (face-api.js, tensorflow.js) en portales no relacionados con el desarrollo técnico de la entidad.

B. Para funcionarios y usuarios finales

- ❑ *Cero confianza en biometría por enlace:* Informar a todo el personal que ninguna entidad gubernamental solicita pruebas de vida (parpadear, sonreír o mover el rostro) a través de un navegador web para validar cuentas de correo.

- ❑ *Reporte de "Cámara Activa":* Instruir a los usuarios para que cierren inmediatamente cualquier pestaña y reporten al área de seguridad si observan que el indicador LED de su cámara se activa sin su consentimiento en un sitio externo.

- ❑ *Validación de dominios:* Antes de ingresar cualquier dato o permitir acceso a periféricos, verificar que la URL pertenezca estrictamente al dominio oficial .gov.co.

Vector de Ataque : Phishing y robo biométrico

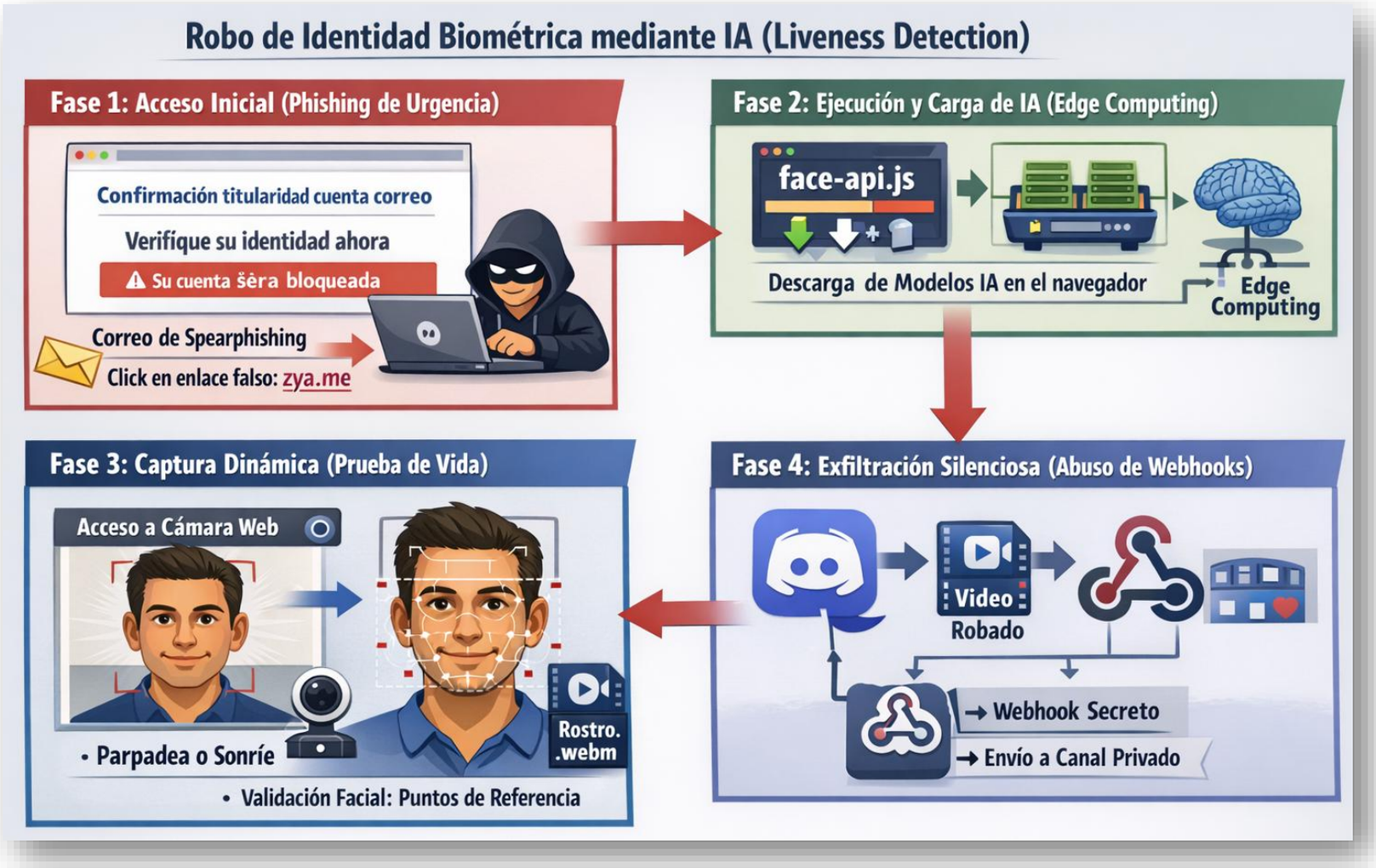


Figura 1. *Flujograma del vector de ataque para el robo de identidad biométrica dinámica. Se observa la transición desde el acceso inicial vía phishing hasta la exfiltración de video dinámico mediante el abuso de servicios legítimos de mensajería (Webhooks de Discord).*

Glosario técnico

- **Biometría Dinámica:** Método de autenticación que utiliza características físicas del usuario que están en movimiento o cambio (como el parpadeo o el movimiento de los labios) para verificar la identidad. Es más segura que la biometría estática (fotos) porque es más difícil de duplicar artificialmente.

- **Prueba de Vida (Liveness Detection):** Técnica de seguridad utilizada para determinar si una muestra biométrica (rostro) proviene de un ser humano real y presente en el momento de la captura, y no de una fotografía, pantalla o máscara.

- **Edge Computing (IA en el Cliente):** Se refiere al procesamiento de datos (en este caso, el análisis facial por Inteligencia Artificial) que ocurre directamente en el dispositivo del usuario (computador o celular) en lugar de enviarse a un servidor central. Esto permite que el ataque sea más rápido y difícil de detectar por firewalls.

Glosario técnico

- **Webhook:** Es una dirección URL que actúa como un "buzón automático". Permite que una aplicación envíe información a otra de forma instantánea. En este incidente, se utilizó para enviar el video robado directamente a un canal privado de los atacantes.

- **OpenResty:** Plataforma web basada en Nginx que permite integrar scripts complejos. Los atacantes la utilizan en su infraestructura para gestionar el tráfico malicioso y ocultar la ubicación real de sus servidores de control.

- **Formatos de Video .WebM:** Formato de archivo multimedia diseñado para la web. Su uso en este ataque confirma que el objetivo no era una foto fija, sino un registro de video dinámico con capacidades de alta compresión.

Fuentes

Google Threat Intelligence: Telemetría global de VirusTotal, Mandiant y Google Safe Browsing para la validación de reputación de infraestructura y atribución de tácticas de actores de amenaza.

MITRE ATT&CK®: Framework internacional de tácticas, técnicas y procedimientos (TTPs) utilizado para el mapeo del comportamiento del adversario.

FIRST (Forum of Incident Response and Security Teams): Estándares globales para la coordinación y respuesta a incidentes de ciberseguridad.

Face-api.js Documentation: Especificaciones técnicas de redes neuronales aplicadas a la detección de puntos de referencia faciales en entornos web.

ColCERT: Repositorio histórico de alertas y actualizaciones AL-20260130-091 sobre campañas de phishing dirigidas a entidades de gobierno.