INVESTIGACIÓN

Capacitación Proactiva contra correos phishing mediante simulaciones con la herramienta GoPhish

Proactive training against phishing emails using simulations with the GoPhish tool

Ing. Rubén Olabarrieta Rivera[*], Ing. Yuneisy Quintero Prieto[2], Ing. Raúl Sanchez Arañó[3]

Recibido: 11/2025 | Aceptado: 11/2025 | Publicado: 12/2025

Resumen

El *phishing*, una amenaza centrada en la explotación del factor humano, representa un alto riesgo para las organizaciones. Frente a la insuficiencia de las capacitaciones teóricas tradicionales, este estudio implementó un programa de simulaciones controladas que utiliza la herramienta de código abierto GoPhish. A lo largo de cuatro campañas sucesivas que emplearon distintos vectores de ataque, se observó una tendencia decreciente en las tasas de interacción y una reducción significativa en la entrega de credenciales. El análisis de los usuarios participantes reveló una clara curva de aprendizaje, donde la mayoría mostró mejoras sustanciales en su capacidad de detección. Al mismo tiempo, se registró un aumento considerable en los reportes voluntarios de correos sospechosos en el entorno real, de manera especial en áreas de alta exposición. Los resultados demuestran que la combinación de entrenamiento práctico, retroalimentación inmediata y simulaciones periódicas transforma de forma efectiva al usuario

[*] TC Mariel. Cuba. rubenor@tcmariel.cu

[2] TC Mariel. Cuba. yuneisyqp@tcmariel.cu

[3] TC Mariel. Cuba. raulsa@tcmariel.cu

de un eslabón vulnerable en una barrera de defensa proactiva, que fortalece de manera sostenible la cultura de ciberseguridad organizacional.

Palabras clave: phishing, GoPhish, simulación de ciberseguridad, entrenamiento práctico

Abstract

Phishing, a threat exploiting the human factor, poses a critical risk to organizations. Confronting the inadequacy of traditional theoretical training, this study implemented a controlled simulation program using the open-source tool GoPhish. Across four successive campaigns employing distinct attack vectors, a clear decreasing trend was observed in interaction rates and credential submission was significantly reduced. Analysis of participating users revealed a strong learning curve, with most showing substantial improvements in detection capabilities. Concurrently, a considerable increase in voluntary reporting of real suspicious emails was recorded, particularly in high-exposure areas. The findings confirm that combining practical training, immediate feedback, and periodic simulations effectively transforms users from a vulnerable link into a proactive defense barrier, thereby sustainably strengthening the organizational cybersecurity culture.

Keywords: *phishing, GoPhish, cybersecurity simulations, practical training*

Introducción

El panorama de la ciberseguridad actual se caracteriza por una evolución constante y sofisticada de las amenazas, donde el correo electrónico permanece como uno de los vectores de ataque más prolíficos y efectivos. Entre estas amenazas, el *phishing* es definido como una forma de fraude en línea cuyo objetivo es robar información confidencial mediante el envío de comunicaciones que suplantan la identidad de una entidad legítima (Workman, 2008). Este se destaca no por su complejidad técnica, sino por su ingeniosa explotación de un eslabón vulnerable: el factor humano. Los atacantes han perfeccionado sus tácticas para crear mensajes fraudulentos tan convincentes

que vulneran con facilidad las defensas perimetrales mejor configuradas, dirigiéndose directamente a la psicología de los empleados.

Las consecuencias de un ataque de *phishing* exitoso son severas, y van desde el compromiso de credenciales y el robo de información sensible hasta infecciones con *ransomware* que pueden paralizar las operaciones completas de una empresa, esto genera cuantiosas pérdidas financieras y daños irreparables a la reputación corporativa. Frente a este escenario, la sola implementación de soluciones técnicas defensivas (como filtros de correo, antivirus y *firewall*) se ha revelado como insuficiente.

La defensa más crítica y subestimada radica en la capacitación continua y consciente de los usuarios finales. Para lograr esto, las empresas deben adoptar un enfoque proactivo que trascienda las capacitaciones tradicionales teóricas, mediante la simulación de ataques controlados que permitan medir y elevar el nivel de preparación real de su capital humano. En este contexto, herramientas de simulación de correos *phishing* como GoPhish emergen como recursos invaluables, lo que permite a los especialistas de ciberseguridad, ejecutar y analizar campañas de *phishing* simuladas internamente en las empresas. Esta práctica no solo identifica a los empleados más vulnerables para brindarles capacitación dirigida, sino que también sienta las bases para una cultura organizacional de ciberseguridad robusta y resistente, para transformar al eslabón más débil en una primera línea de defensa efectiva.

El objetivo principal de este trabajo es medir la susceptibilidad de la fuerza laboral, evaluar la efectividad de las herramientas de simulación para el entrenamiento proactivo en la detección de correos *phishing* e identificar áreas de mejora.

Materiales y métodos

Contexto de estudio y justificación del enfoque

El estudio se realizó en una empresa de servicios logísticos con más de 200 empleados que usan a diario correo electrónico y más del 50 % mantiene algún tipo de intercambio con clientes y proveedores internacionales lo que aumenta de manera exponencial la interacción

con correos electrónicos tipo *phishing* reales y su conocimiento de identificación es vital en la detección y minimización de incidentes. Para enfrentar esta amenaza, se implementó un programa de simulaciones controladas.

El *phishing* no es una amenaza nueva, pero su prevalencia y complejidad continúan en un crecimiento alarmante. Los reportes anuales de actores líderes en la industria proveen datos alarmantes que ilustran la magnitud del problema. El Verizon *2025* Data Breach Investigations Report (DBIR) consolida esta tendencia, situando a este patrón entre los tres principales desde 2019 (Verizon, 2025). Este informe destaca que la sofisticación ha evolucionado de forma significativa; los ataques ya no se basan en propuestas increíbles de trabajo o premios, sino en campañas prolongadas donde los atacantes construyen relaciones de confianza con las víctimas durante un periodo de tiempo.

La efectividad de esta técnica radica en su ingeniería social, un método que «influye psicológicamente en una persona para que realice una acción o revele información confidencial» (ENISA, 2022). Los actores de amenazas investigan de forma minuciosa a sus víctimas, a través de la creación de correos muy personalizados (*spear phishing*) que imitan de manera perfecta comunicaciones legítimas de superiores, colegas o servicios de confianza. El DBIR de Verizon (2025) corrobora que en la actualidad se debe ser cauteloso incluso ante los mensajes que parecen proceder de clientes o proveedores establecidos, lo que hace muy difícil su detección incluso para usuarios experimentados.

El usuario, el “eslabón débil”

La referencia tradicional en ciberseguridad ha caracterizado al usuario final como “el eslabón más débil” de la cadena de seguridad. Esta perspectiva, sin embargo, es contraproducente y se reevalua. Investigaciones en el campo de la seguridad de la información sugieren que es más efectivo ver al empleado no como un problema a ser mitigado, sino como un “factor humano” que debe ser empoderado y capacitado (Hadlington, 2017), «los empleados no son solo usuarios, sino guardianes activos de la ciberseguridad» (Hernández, 2025). Los

usuarios no fallan por negligencia inherente, sino porque la capacitación que reciben a menudo es genérica, esporádica y desconectada de las tácticas reales utilizadas por los atacantes en la actualidad.

Un estudio publicado en *Computers & Security* encontró que la formación tradicional basada en concienciación tiene un efecto limitado en el comportamiento a menos que se complemente con experiencias prácticas y retroalimentación inmediata (Pattinson et al., 2015). Es aquí donde entrenar con la realidad se vuelve crucial, la confianza cero no solo se trata de limitar accesos, sino de empoderar a las personas para que tomen decisiones informadas (Hernández, 2025). Al exponer a los usuarios a simulaciones de *phishing* realistas, se activa el aprendizaje experiencial, permitiéndoles reconocer las señales de un ataque (como remitentes extraños, URL sospechosas o urgencia injustificada) en un entorno seguro donde un error se convierte en una lección y aprendizaje, no en un desastre.

Herramienta seleccionada: GoPhish

Para implementar un programa de entrenamiento basado en simulaciones, las empresas necesitan herramientas accesibles y efectivas que evalúen la susceptibilidad de los empleados a través de ataques controlados. Si bien el mercado está dominado por soluciones comerciales reconocidas, tales como KnowBe4 y otras, líderes identificadas en el informe “Forrester Wave: Security Awareness and Training Solutions, Q1 2022” (Budge, et al., 2022), la elección de una plataforma depende a menudo del presupuesto y las necesidades específicas. Para organizaciones que priorizan la flexibilidad y la ausencia de costes de licencia, GoPhish surge como una alternativa de código abierto con una amplia aceptación.

GoPhish es una plataforma que permite a los profesionales de seguridad:

- Diseñar campañas de *phishing* altamente personalizables que replican las tácticas actuales.
- Dirigir estas campañas a grupos específicos de empleados de manera controlada.
- Medir con precisión métricas clave como la tasa de clic y la tasa de entrega de credenciales.

El valor de GoPhish, respaldado por una comunidad activa y documentación extensa, radica en la capacidad para cerrar el ciclo de aprendizaje: de la teoría a la práctica, y de la práctica a la mejora continua.

Diseño de las campañas y procedimiento experimental

Para las simulaciones se utilizaron dos vectores de ataque diferentes:

- Tipo 1: Envío de correo de solicitud de actualización de credenciales en un servicio web.
- Tipo 2: Envío de correo de solicitud de descarga de fichero.

Se ejecutaron cuatro campañas simuladas cronológicas a lo largo de dos años (2024-2025) (Tabla 1):

- Campaña 1 (Tipo 1). Actualización de credenciales en el *Webmail* de la empresa.
- Campaña 2 (Tipo 2). Solicitud de descarga de fichero «Actualización de políticas de seguridad de la empresa».
- Campaña 3 (Tipo 2). Solicitud de descarga de fichero «Listado de productos de feria».
- Campaña 4 (Tipo 1). Actualización de credenciales en el ERP (*Enterprise Resource Planning*) empresarial.

Tabla 1. Campañas en orden cronológico

	Usuarios	Clic en el enlace	Entregaron credenciales	Descarga de fichero	Reporte de usuarios
Campaña 1	176	21	10	--	0
Campaña 2	173	17	--	17	8
Campaña 3	227	40	--	40	8
Campaña 4	132	16	0	--	11

Medidas correctivas y recolección de datos

Después de cada Campaña, y en especial después de los resultados iniciales, se implementaron acciones correctivas. Tras la Campaña 1 se elaboró un Plan de Capacitación específico sobre identificación de

correos *phishing* orientado a las personas que no pasaron las simulaciones con materiales detallados y ejercicios prácticos de identificación. Además, se prepararon capacitaciones generales para todos los trabajadores sobre correos *phishing* y sesiones de concienciación general con los usuarios, enfocadas en la importancia de notificar de manera proactiva cualquier comunicación sospechosa al equipo de ciberseguridad, con el fin de agilizar la respuesta ante incidentes reales. Los datos se recopilaron de forma automatizada por la herramienta GoPhish (clic, entregas de credenciales) y con reportes manuales de los usuarios a la Dirección de IT.

Resultados y Discusión

Resultados por campaña y análisis de la evolución

La Campaña 1 consistió en la simulación de un correo electrónico que, a través de un enlace fraudulento, solicitaba a los usuarios la actualización de sus credenciales (Figura 1) en un portal que emulaba la interfaz del Webmail corporativo (Figura 2). Tras la introducción de los datos por parte del usuario, la solicitud era redirigida al portal oficial de la empresa, registrándose el evento como un intento de *phishing* exitoso. Los resultados de esta campaña revelaron una tasa de clic del 20 % y, de estos, una tasa de entrega de credenciales del 48 %. Estas cifras, alarmantes al inicio, evidenciaron una clara vulnerabilidad del personal ante este tipo de ataques, a pesar de la existencia de capacitaciones periódicas en la empresa. Un hallazgo adicional crítico fue la casi nula notificación de los correos fraudulentos por parte de los empleados.

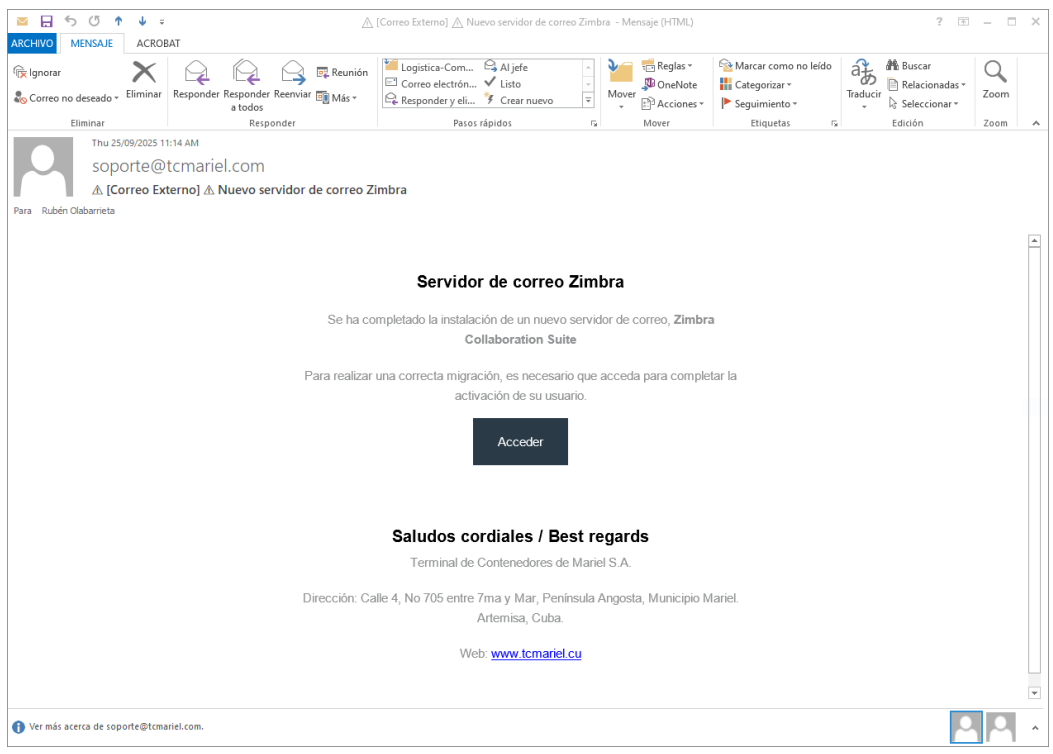

Figura 1. Correo de solicitud de actualización de credenciales en el Webmail de la empresa

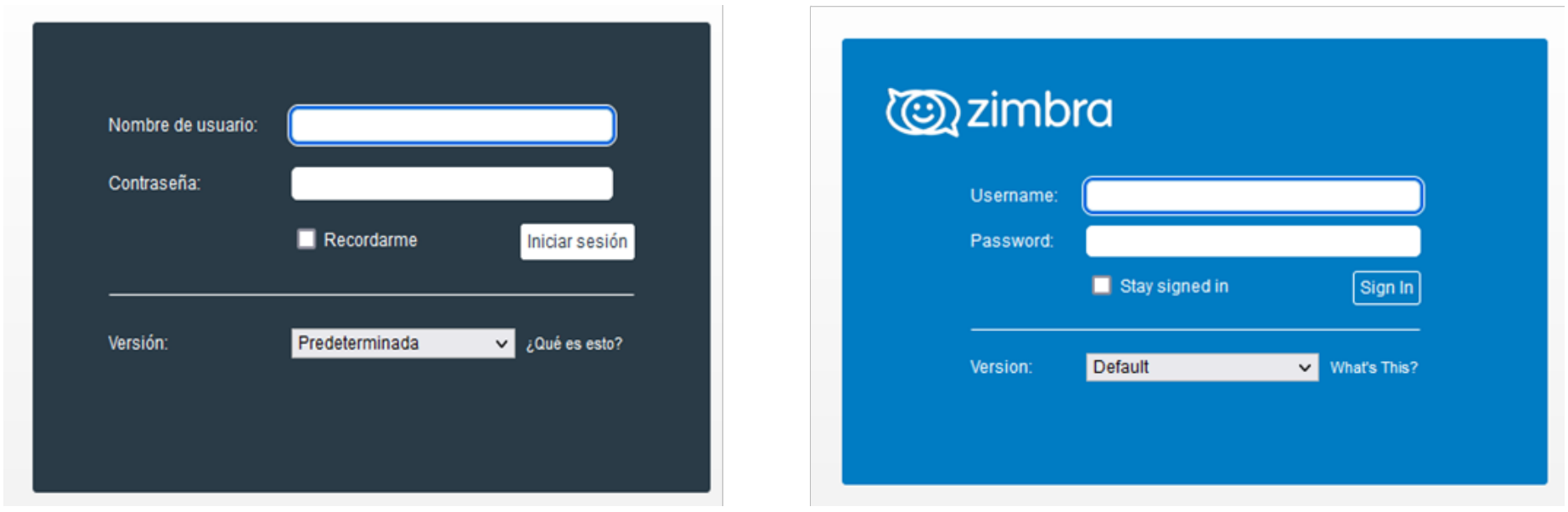

Figura 2. Interfaz falsa (izquierda) y real (derecha) del Webmail corporativo

La Campaña 2 se diseñó para evaluar la efectividad de las medidas de concientización implementadas, con el uso de un vector de ataque distinto: la descarga de un archivo malicioso. En esta simulación, el correo electrónico instaba a los usuarios a descargar un archivo bajo la denominación «Actualización de políticas de seguridad de la empresa». El evento se registró como un intento de *phishing* exitoso si el usuario procedía a descargar el fichero. Los resultados mostraron una mejora sustancial: la tasa de clic y descarga del archivo se redujo al 10 %. De manera aún más significativa, se registró un incremento en las notificaciones por parte de los empleados, pasando de un 0 % en la campaña inicial a un 4 %. Este aumento en los reportes, aunque modesto, sugiere una mayor concientización y una respuesta proactiva del personal, lo que evidencia la efectividad preliminar de las capacitaciones específicas y generales impartidas con posterioridad a la primera simulación.

La Campaña 3 incrementó de manera significativa el nivel de sofisticación mediante la aplicación de técnicas avanzadas de ingeniería social. El correo, diseñado con un mayor apego a la realidad operativa de la empresa, solicitaba la descarga de un archivo denominado «Listado de productos de feria». Al ejecutarse, este archivo iniciaba una secuencia automatizada que simulaba comportamientos maliciosos, como la ejecución consecutiva de aplicaciones predeterminadas del sistema Windows©, culminando con la activación de un segundo *script* (Figura 3). Dicho *script* desplegaba en el Escritorio de las estaciones de trabajo, mediante una ventana de comandos (CMD), una representación gráfica en forma de carabela (Figura 4), con el objetivo

de emular de manera más verosímil las consecuencias observables de un ataque real.

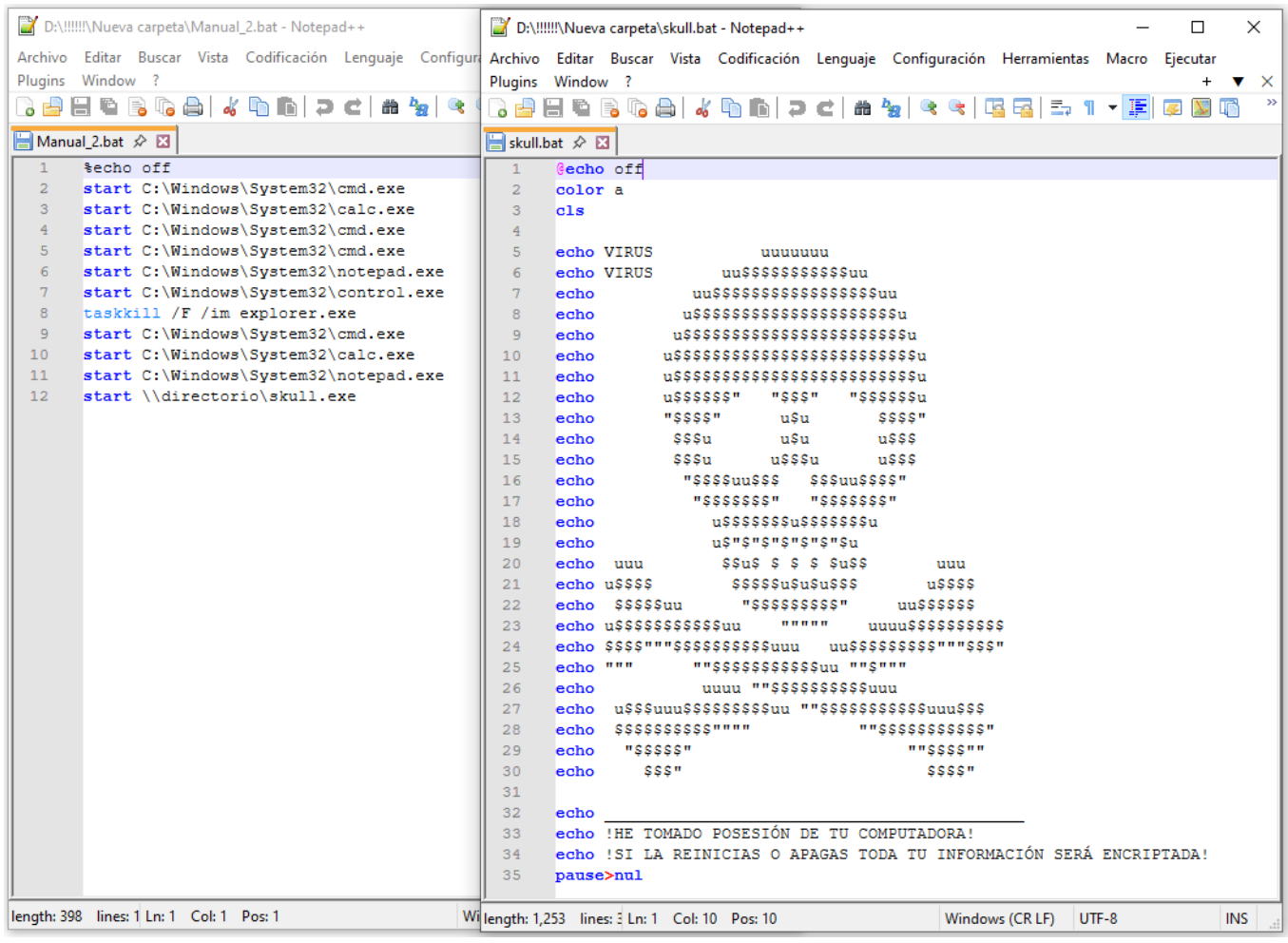

Figura 3. Primer (izquierda) y segundo (derecha) script que se ejecutaban al descargar el fichero

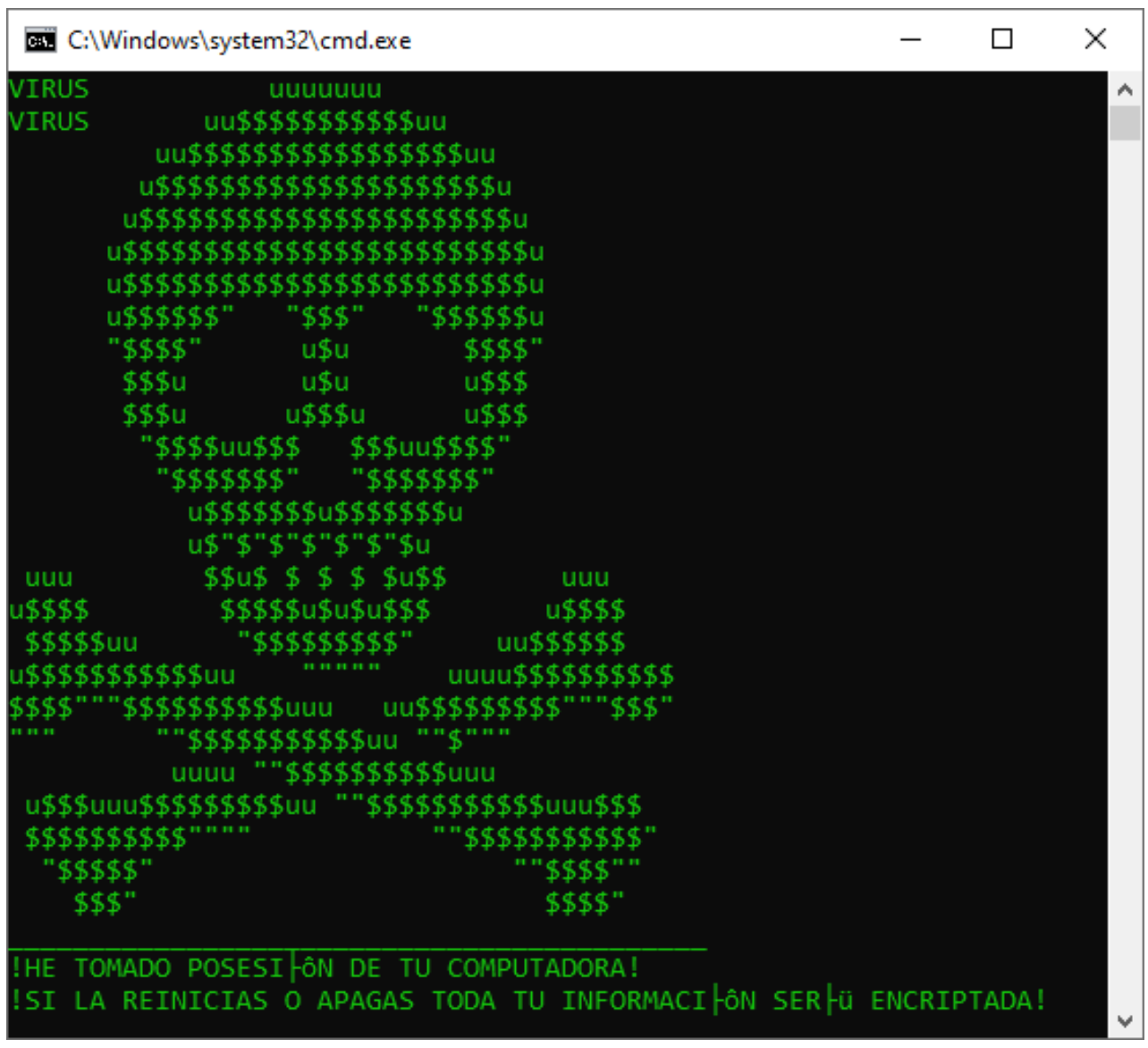

Figura 4. Representación gráfica en forma de calavera en ventana CMD

Los resultados de esta campaña fueron reveladores: si bien se registró un incremento considerable en las interacciones con el correo fraudulento (Figura 5), el número de reportes se mantuvo en ocho usuarios. Las entrevistas de seguimiento realizadas en las diferentes direcciones de la empresa identificaron la causa principal de esta

aparente discrepancia. El notable impacto visual de la simulación anterior había generado conversaciones internas entre los trabajadores, lo que resultó en un conocimiento generalizado de la prueba. En consecuencia, los usuarios que identificaron el correo optaron por no reportarlo, al reconocerlo como un elemento interno de capacitación y no como una amenaza externa genuina. Este fenómeno subraya la importancia de variar los vectores de ataque en simulaciones sucesivas para evitar la contaminación de los resultados y mantener la validez del ejercicio.

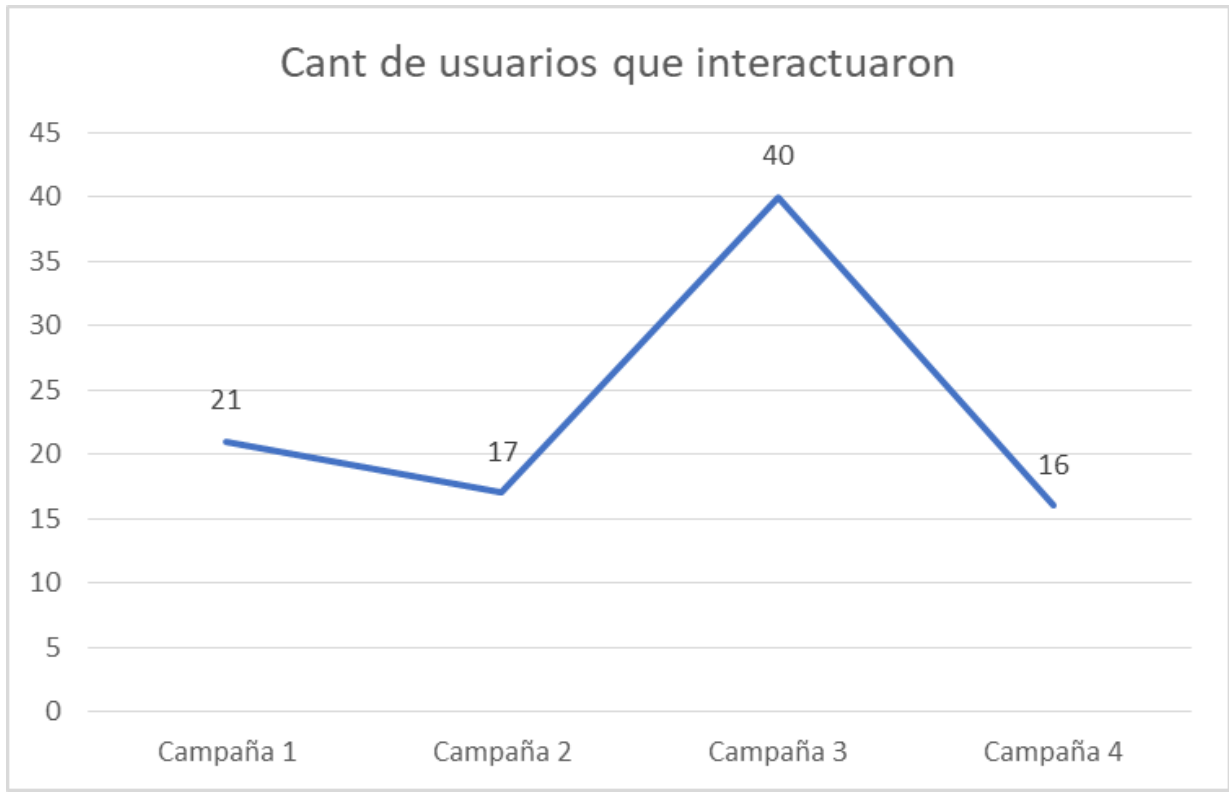

Figura 5. Cantidad de interacciones por Campañas

La Campaña 4 retomó el vector de ataque de la Campaña 1 (la solicitud de actualización de credenciales), pero aplicándolo esta vez al acceso del sistema ERP corporativo, con una plataforma de *phishing* rediseñada para esta simulación. Los resultados demostraron una mejora significativa en la resiliencia de los usuarios. Esta campaña registró el menor nivel de interacción, con 16 usuarios haciendo clic en el enlace fraudulento. De manera aún más destacable, la tasa de entrega de credenciales fue nula, ya que ninguno de los usuarios que accedió a la página falsa introdujo sus datos. Al mismo tiempo, se observó un aumento considerable en el número de reportes por parte de los empleados (Figura 6). El análisis comparativo de las cuatro campañas evidencia una clara tendencia de mejora progresiva: la reducción drástica de las interacciones y la tasa de éxito del *phishing*, junto con el incremento sostenido de las notificaciones, confirman la efectividad del programa de capacitación continuo.

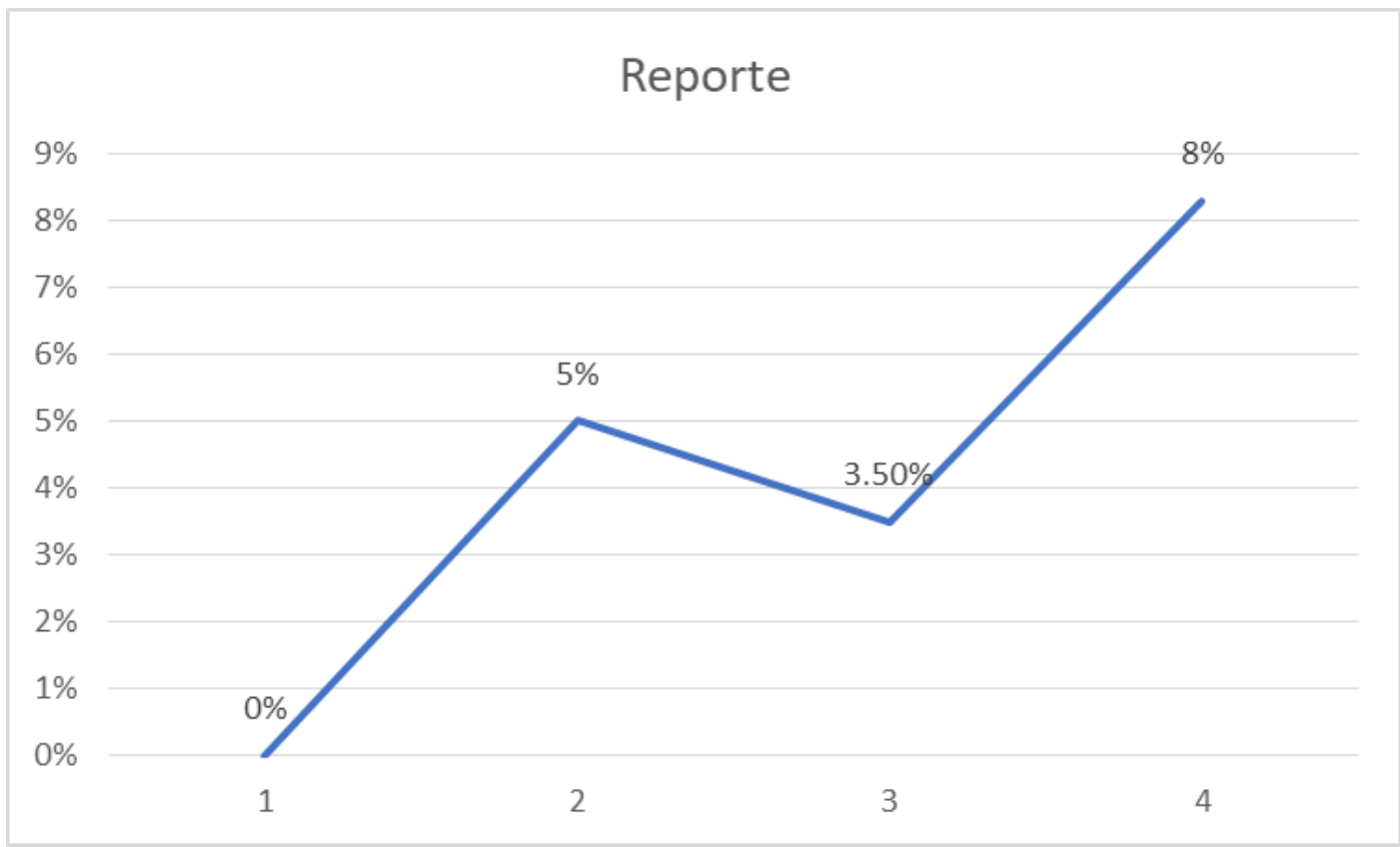

Figura 6. Reporte de recepción de los usuarios en cada Campaña

La tendencia decreciente en las tasas de interacción y entrega de credenciales a lo largo de las sucesivas campañas demuestra de manera fehaciente la efectividad del programa de capacitación implementado. Un análisis detallado de los 227 usuarios participantes en total en todas las simulaciones revela que la mayoría mostró una clara curva de aprendizaje: mientras 46 usuarios (20.3 %) solo interactuaron en una simulación, solo 21 (9.3 %) lo hicieron en dos ocasiones y solo 2 usuarios (0.9 %) mantuvieron un comportamiento vulnerable en tres de las cuatro campañas (Figura 7). Esta distribución evidencia que el entrenamiento práctico y la retroalimentación inmediata resultaron efectivos para la inmensa mayoría del personal.

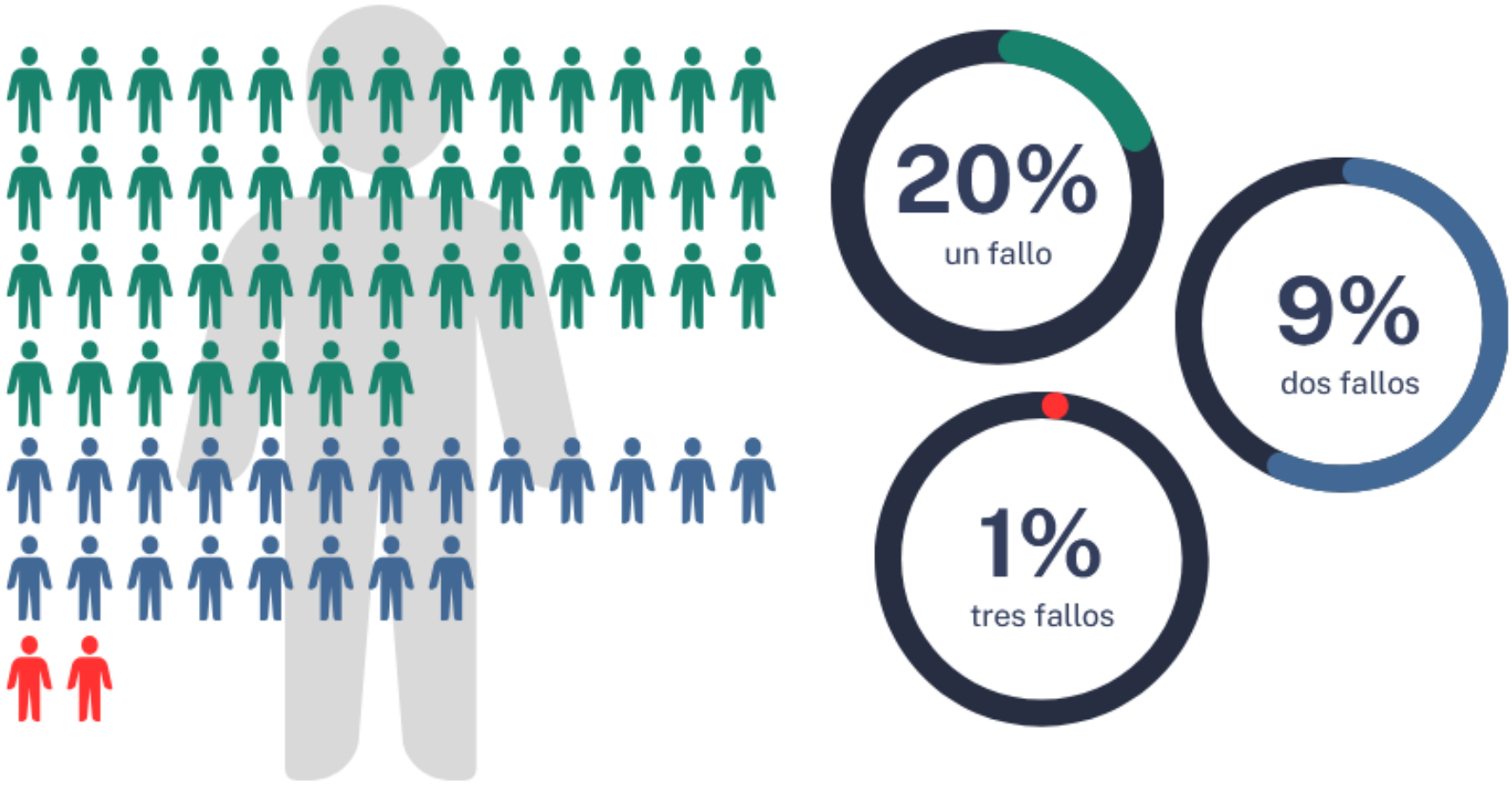

Figura 7. Cantidad de fallos reiterados por los usuarios

El éxito se ve reforzado por un indicador conductual clave: la Dirección de IT reportó un aumento sustancial en los reportes voluntarios de correos sospechosos por parte de los empleados en el entorno productivo, que trasciende el ámbito de las simulaciones. Un análisis más detallado revela que las direcciones con mayor volumen de reportes de correos *phishing* reales fueron las de Logística y Comercial, lo cual es coherente con su alta exposición a comunicaciones internacionales. Este hallazgo no solo valida la eficacia de la concientización, sino que apunta a una asimilación palpable de los protocolos de seguridad por parte de los usuarios y se traduce en un fortalecimiento de la cultura de ciberseguridad organizacional. Como resultado directo de esta mayor vigilancia, se ha incrementado la efectividad en el bloqueo proactivo de dominios maliciosos, cerrando así un ciclo de mejora continua en la postura de seguridad de la empresa.

Conclusiones

Los hallazgos de este caso de estudio, en línea con las estadísticas globales, demuestran que los enfoques de capacitación teórica tradicional resultan insuficientes para mitigar de manera eficaz el riesgo asociado al *phishing*.

La implementación de herramientas de simulación como GoPhish se revela como una estrategia crucial, pues permite cuantificar de forma objetiva el riesgo al transformar vulnerabilidades abstractas en métricas tangibles. Esta capacidad de presentar datos concretos (como el porcentaje específico de empleados que interactúan con un simulacro) facilita la comunicación efectiva con la alta dirección y la necesidad de priorizar recursos de seguridad.

Además, estas herramientas posibilitan una capacitación altamente dirigida y una optimización de recursos, al identificar con precisión a los usuarios más vulnerables. Este enfoque permite diseñar programas proactivos de concientización personalizados que maximiza el impacto de las iniciativas de entrenamiento. De manera paralela, las simulaciones continuas fomentan un estado de alerta constante y una cultura de seguridad, lo que integra la ciberseguridad en la rutina

laboral y transforma al capital humano de un eslabón pasivo en un componente activo de la defensa organizacional.

El costo de implementar una solución como GoPhish es mínimo comparado con el costo potencial de pérdidas financieras y de reputación de una sola brecha de datos exitosa. Invertir en la preparación del factor humano es, por tanto, una de las estrategias de mitigación de riesgo más efectivas disponibles.

La fortaleza cibernética de una organización ya no puede depender únicamente de *firewall* y *software*. La defensa debe ser basada en las personas también. Las herramientas de capacitación son fundamentales para lograr la operatividad de esta visión y esto facilita a las empresas transformar a su fuerza laboral de un objetivo potencial en una barrera de defensa robusta y consciente, elevando así su postura de seguridad digital de manera significativa y sostenible.

Referencias bibliográficas

Budge, J., Blankenship, J., Sjoblom, S., & Nagel, B. (2022). The Forrester Wave™: Security Awareness and Training Solutions, Q1 2022. *Forrester Research*. Disponible en https://www.livingsecurity.com/hubfs/Forrester%20Wave%20Report%202022%20-%20Security%20Awareness%20and%20Training%20-%20Q1%202022.pdf

ENISA. (2022). ENISA Threat Landscape 2022. *European Union Agency for Cybersecurity*. Disponible en https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022

Hadlington, L. (2017). Human Factors in Cybersecurity; Examining the Link Between Internet Addiction, Impulsivity, Attitudes Towards Cybersecurity, and Risky Cybersecurity Behaviours. *Heliyon, 3 (2017)*, Artículo No-e00346. https://doi.org/10.1016/j.heliyon.2017.e00346

Hernández Domínguez, A. (2025). *Más allá de la tecnología: Apostar por las personas en la estrategia de confianza cero*. Disponible en http://www.cubadebate.cu/especiales/2025/02/21/mas-alla-de-la-tecnologia-apostar-por-las-personas-en-la-estrategia-de-confianza-cero/

Pattinson, M., Jerram, C., Parsons, K., McCormac, A., & Butavicius, M. (2015). Why do some people manage phishing e-mails better than others? *Information Management & Computer Security*, *23*(3), 270-281.

Verizon (2025). 2025 Data Breach Investigations Report. V*erizon Business*. Disponible en https://www.verizon.com/business/resources/T163/reports/2025-dbir-data-breach-investigations-report.pdf

Workman, M. (2008). Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security. *Journal of the American Society for Information Science and Technology, Volumen* 59, Issue 4, pages 662-674. Disponible en https://sci-hub.se/10.1002/asi.20779