ABRIL 2022

INFORME DETALLADO

Informe de Amenazas Adversarias

Por Ben Nimmo, líder de Inteligencia de Amenazas Global para Operaciones de Influencia y David Agranovich, director de Desarticulación de Amenazas

Meta

TABLA DE CONTENIDOS

OBJETIVO DE ESTE INFORME

Nuestros informes públicos de seguridad empezaron hace cuatro años, cuando reportamos por primera vez nuestros hallazgos sobre contenido inauténtico coordinado (CIB por sus siglas en inglés) de la Agencia de Investigación de Internet de Rusia. Desde entonces, las amenazas globales han evolucionado de forma importante y hemos ampliado nuestra capacidad para responder a una mayor variedad de comportamientos adversarios. Para ayudar a dar una mayor comprensión sobre los riesgos que vemos, ahora estamos ampliando nuestros informes para incluir ciberespionaje, contenido inauténtico y otros daños emergentes en un solo lugar, como parte de los informes trimestrales que estamos probando. También, estamos compartiendo indicadores de amenazas al final de este informe para contribuir a los esfuerzos de la comunidad de seguridad para detectar y combatir actividad maliciosa en otras partes en Internet (Ver Apéndice). Estamos abiertos a las ideas de la comunidad de seguridad para ayudar a que estos informes sean más informativos y haremos ajustes conforme tengamos retroalimentación.

RESUMEN DE NUESTROS HALLAZGOS

- Nuestro informe trimestral de amenazas piloto brinda una perspectiva general de los riesgos que detectamos a lo largo de múltiples infracciones de políticas, incluyendo Comportamiento Inauténtico Coordinado (CIB por sus siglas en inglés), ciberespionaje, y otros daños emergentes como reporte masivo.
- Tomamos acciones contra dos operaciones de **ciberespionaje** en Irán. La primera red estaba vinculada a un grupo de hackers conocidos en la industria de seguridad como UNC788. La segunda fue un grupo separado que no se había reportado previamente y que se enfocaba en industrias como la energética, telecomunicaciones, logística marítima, tecnologías de la información, entre otras. Más información aquí.
- Eliminamos una red híbrida operada por el Ministro de Interior de Azerbaiyán que **mezclaba ciberespionaje con Comportamiento Inauténtico Coordinado (CIB)** dirigido a la sociedad civil en Azerbaiyán para comprometer sus cuentas y sitios web para publicar a su nombre. Más información aquí.
- Nuestros hallazgos también incluyen **actualizaciones sobre nuestras acciones en Ucrania**, incluyendo intentos por regresar a la plataforma por parte de actores estatales y no estatales, aunado a las redes de spam que usaron tácticas engañosas para monetizar la atención del público alrededor de la guerra en curso. Más información aquí.
- Eliminamos operaciones de **Comportamiento Inauténtico Coordinado** de Brasil, Costa Rica y El Salvador, y redes de Rusia y Ucrania previamente reportadas. La red brasileña es la primera

operación principalmente enfocada en problemas medioambientales que removemos. Más información aquí.

- Como parte de la desarticulación de **amenazas nuevas y emergentes**, eliminamos una red infractora coordinada en Filipinas que afirmaba bajar sitios web y desagregarlos, principalmente aquellos de sitios de noticias. Más información aquí.
- Bajo nuestra política de **Comportamiento Inauténtico contra reportes masivos**, eliminamos una red en Rusia por abusar de nuestras herramientas para reportar de forma repetida a personas en Ucrania por supuestas violaciones de las políticas de Facebook, en un intento por silenciarlas. Más información aquí.
- También, bajo nuestras políticas de Comportamiento Inauténtico, eliminamos decenas de miles de cuentas, Páginas y Grupos alrededor del mundo por **inflar falsamente la distribución de su contenido y construir su audiencia de forma abusiva**. Lo hicimos a través de la detección automática a gran escala, complementada con investigaciones manuales. Más información aquí.

01

Eliminamos tres redes de ciberespionaje de Irán y Azerbaiyán

Actores de ciberespionaje *generalmente se dirigen a personas en Internet para recopilar datos, manipularlos para revelar información y comprometer sus dispositivos o cuentas. Cuando desarticulamos estas operaciones, eliminamos sus cuentas, bloqueamos la posibilidad de compartir sus dominios en nuestra plataforma y notificamos a las personas que creemos fueron blanco de estos grupos maliciosos. También, compartimos información con los investigadores de seguridad, gobiernos y nuestros pares en la industria para que ellos también puedan tomar acciones para frenar esta actividad. Hemos incluido indicadores de amenazas en el Apéndice de este informe.*

1. UNC788

Tomamos acciones contra un grupo de hackers en Irán, conocidos en la industria de seguridad como UNC788, que se dirigían a personas en Medio Oriente, incluyendo la milicia saudí, disidentes y activistas de derechos humanos de Israel e Irán, políticos en Estados Unidos, y académicos enfocados en Irán, activistas y periodistas alrededor del mundo. Su actividad maliciosa tenía las características de una operación persistente y con suficientes recursos, que ocultaba a sus responsables. Hemos estado monitoreando y bloqueando los esfuerzos de este grupo durante varios años, al igual que nuestros pares en otras plataformas. Esta campaña de ciberespionaje estaba activa a lo largo de Internet y se enfocaba en tácticas de phishing para robar datos de sus cuentas en línea y en compartir enlaces a sitios web maliciosos con malware.

Detectamos las siguientes tácticas, técnicas y procedimientos (TTPs por sus siglas en inglés) usadas por este actor de amenaza en Internet:

- **Ingeniería social:** este grupo usaba una combinación de cuentas falsas poco sofisticadas y personas falsas más elaboradas, que posiblemente usaban para desarrollar confianza en las personas a las que se dirigían y engañarlas para dar clic en enlaces de phishing o descargar aplicaciones maliciosas. Muchas de estas cuentas se hacían pasar por activistas de derechos humanos o académicos.
- **Phishing:** esta campaña también se basaba en sitios web de phishing que alojaban sitios de eventos o archivos que pedían a las personas iniciar sesión con sus datos de Google para registrarse.
- Malware: Para comprometer las cuentas y dispositivos de las personas, este grupo copiaba y modificaba una aplicación legítima de Android, una calendario de cumpleaños, para recopilar datos de contacto y mandarlos a los servidores remotos del ciberdelincuente. También, desarrollaron un malware con capacidad de acceso remoto para Android disfrazado de Corán, una aplicación de mensajería para tener acceso a la lista de contactos de las personas, sus mensajes de texto, archivos, ubicación y activar su cámara y micrófono. A esta variedad de malware que no había sido reportada la nombramos HilalRAT (troyano de acceso remoto), tras ver “hilal” en varios ejemplares de malwares que analizamos. En el Apéndice, también compartimos una regla Yara para ayudar a que la comunidad de seguridad la identifique.

2. Grupo de hackers de Irán que no habían sido reportados anteriormente

Tomamos acciones contra un grupo de hackers en Irán que no había sido reportado anteriormente, que se dirigieron o falsificaron compañías de diferentes industrias alrededor del mundo. Esto incluyó compañías energéticas en Arabia Saudita, Canadá, Italia y Rusia; la industria de tecnologías de la información en India y los Emiratos Árabes Unidos; la industria de logística marítima en los Emiratos Árabes Unidos, Islandia, Noruega, Arabia Saudita, Estados Unidos, Israel e India; compañías de tecnología en Arabia Saudita y los Emiratos Árabes Unidos; y la industria de semiconductores en Israel, Estados Unidos y Alemania. Esta actividad tenía las características de una operación persistente y con recursos suficientes que ocultaba a sus responsables.

Este grupo usaba tácticas, técnicas y procedimientos (TTPs por sus siglas en inglés) similares a otro actor de amenazas apodado Tortoiseshell que reportamos el año pasado, pero en este caso

detectamos diferentes blancos de ataque, infraestructura técnica y malware. Identificamos los siguientes TTPs usados por este grupo en Internet:

- **Ingeniería social:** esta campaña operaba personas ficticias en diferentes redes sociales, como Instagram, LinkedIn, Facebook y Twitter para hacerlas parecer auténticas y soportar el escrutinio. Generalmente se hacían pasar por reclutadores de compañías reales y falsas en la industria o la región de la persona que era atacada, como parte de lo que parece ser un esquema de ingeniería social para engañar a las persona para que diera clic en enlaces maliciosos o instalar malware.
- **Sitios web corporativos falsos o falsificados:** esta operación incluyó una red de sitios web corporativos de reclutamiento falsos, así como dominios falsificados de compañías legítimas. También, se basaba en gran medida en tácticas de phishing por medio de correo electrónico para hacer ingeniería social a las personas, para hacerlas descargar malware, en un intento aparente por obtener información y acceso a sistemas corporativos.
- **Personalización interactiva y protección contra exploit:** este grupo tomó medidas para ocultar su actividad y proteger sus herramientas maliciosas a través de la integración de funciones interactivas que mandaban la carga maliciosa solo cuando las personas interactuaban con el atacante en tiempo real. Por ejemplo, una aplicación de entrevistas activaba una función de chat integrada para que el atacante proporcionara una contraseña para empezar la entrevista. Cuando la persona ingresaba la contraseña, se activaba la descarga de malware. Una aplicación de ajedrez también requería una contraseña, que proveían los hackers, para iniciar el juego y la descarga de malware.
- **Malware:** este grupo desarrolló aplicaciones maliciosas únicas disfrazadas de apps de VPN, una calculadora de salarios, un lector auditivo de libros o una app de mensajería. Desarrollaron malware sobre la plataforma de virtualización VMWare ThinApp, que les permitía operar en diferentes sistemas y retener la carga maliciosa hasta el último momento, haciendo la detección de malware más difícil. La carga final incluía troyanos de acceso remoto con todas las funciones incluidas, capaces de ejecutar comandos en el dispositivo del atacado, acceder y mandar archivos, tomar capturas de pantalla y descargar y ejecutar malware adicional.

3. Una red híbrida de Azerbaiyán

Desarticulamos una red compleja en Azerbaiyán vinculado con ciberespionaje y comportamiento inauténtico coordinado. Se dirigía principalmente a personas de Azerbaiyán, incluyendo defensores de la democracia, periodistas de oposición y críticos del gobierno en el extranjero. Esta campaña fue prolífica pero poco sofisticada, y fue operada por el Ministro del Interior azerí. Mezclaba una variedad de tácticas, desde phishing, ingeniería social y hacking hasta comportamiento inauténtico coordinado.

Esta operación se dirigía a sitios web y cuentas en línea de defensores de la democracia, la oposición y periodistas en Azerbaiyán, con el afán de lograr lo que parecen ser dos objetivos: obtener información personal sobre los atacados y promover ciertas narrativas acerca de ellos o a su nombre. Se enfocaba en sitios de noticias y en una serie de servicios de Internet, incluyendo Facebook, Twitter, LinkedIn, YouTube, y las plataformas rusas VK y OK. Este es otro ejemplo de una campaña híbrida de espionaje y CIB, similar a la actividad no relacionada y separada de Ghostwriter, un actor de amenaza que recientemente se dirigía a Ucrania.

Identificamos las siguientes tácticas, técnicas y procedimientos usadas por este actor de amenaza en Internet:

- **Sitios web comprometidos y falsificados:** este grupo operaba en Internet, con más de 70 sitios web y dominios que administraban ellos mismos o fueron comprometidos. Se dirigían a sitios en Azerbaiyán y, en menor medida, a sitios en Armenia; un número pequeño de sitios tenían dominios rusos o turcos. Una vez que comprometían estos sitios, el grupo recolectaba bases de datos con nombres de usuarios y contraseñas, posiblemente para comprometer las cuentas en línea de las personas que posiblemente reusaban sus datos de acceso en diferentes servicios de Internet. También, a veces, alojaban contenido de phishing en estos sitios web.
- **Malware y otros sitios maliciosos:** este grupo escaneaba sitios web en la región para detectar vulnerabilidades, usando herramientas como Burpsuite y Netsparker. Después, usaban técnicas públicamente conocidas para comprometer sitios vulnerables antes de subir uno de los numerosos web shells con el objetivo de mantener acceso persistente. De forma similar, para descifrar huellas digitales (hashes) obtenida de sitios comprometidos, usaban herramientas para descifrar hashes públicamente disponibles. En sus ataques a personas, este actor de amenaza es conocido por usar Windows y software de vigilancia para Android.

- **Phishing de datos de acceso:** en su actividad de phishing, este grupo se basaba en sitios web comprometidos y falsificados en los cuales pedían a las personas ingresar sus datos de acceso de redes sociales para que pudieran emitir un voto en encuestas políticas. De esta forma, un atacante podía obtener los datos de acceso de una persona para apoderarse de sus cuentas en línea. Esta operación también intentó dirigir a las personas a sus páginas web de phishing compartiéndoles enlaces en redes sociales, incluyendo a través de cuentas de figuras públicas comprometidas o cuentas que se hacían pasar por miembros del equipo de seguridad de Facebook, muchas de las cuales fueron detectadas y deshabilitadas por nuestros sistemas automáticos.
- **Reportes de industria:** nuestros hallazgos corroboran la existencia de reportes públicos sobre parte de esta actividad de OC-Media y Qurium.
- **Comportamiento Inauténtico Coordinado:** los individuos detrás de esta actividad usaban cuentas falsas y cuentas comprometidas para administrar Páginas y publicar como si fueran dueños legítimos de estas Páginas y cuentas. Generalmente publicaban en azerí, incluyendo comentarios críticos o comprometedores sobre la oposición gubernamental, activistas, periodistas y otros miembros de la sociedad civil en Azerbaiyán.

02

Actualización de seguridad de Ucrania

Desde el inicio de la invasión rusa en Ucrania, nuestros equipos han estado alertas para detectar y desarticular amenazas y abusos de la plataforma, incluyendo intentos de redes que desactivamos por reanudar operaciones. Hemos compartido nuestros hallazgos con nuestros pares en la industria de tecnología, investigadores independientes, gobiernos, autoridades e individuos atacados cuando es posible. Puedes encontrar actualizaciones de seguridad anteriores sobre Ucrania aquí.

Principales hallazgos

Actores estatales

Actores vinculados con el gobierno de Rusia y Bielorrusia estuvieron involucrados en actividades de espionaje y operaciones de influencia encubiertas en línea. Esta actividad incluía intereses en la industria de telecomunicaciones en Ucrania; los sectores de defensa y energía en Ucrania; plataformas de tecnología; así como periodistas y activistas en Ucrania, Rusia y en el extranjero.

Parece que estas operaciones se intensificaron rápidamente antes de la invasión rusa. Por ejemplo, detectamos y desarticulamos una actividad de CIB residente vinculada con el Comité para la Seguridad del Estado de Bielorrusia que repentinamente empezó a publicar en polaco e inglés sobre la rendición sin lucha de las tropas ucranianas y los líderes de la nación que huyen del país el 24 de febrero, el día que Rusia empezó la guerra. Antes de esto, este actor de amenazas en particular se enfocó en acusar a Polonia de maltratar a migrantes de Medio Oriente. El 14 de marzo regresaron a Polonia y crearon un evento en Varsovia para llamar a una protesta contra el gobierno polaco. Desactivamos esta cuenta y evento el mismo día.

Actores de amenazas persistentes conocidos

Primero, tras nuestra última actualización de seguridad sobre Ucrania, hemos visto un incremento en los intentos de ataques contra miembros de la milicia ucraniana por parte de Ghostwriter, un actor de amenazas seguido por la comunidad de seguridad. Como hemos compartido antes, Ghostwriter generalmente se dirige a personas a través de correos electrónicos fraudulentos y luego usa la información para obtener acceso a sus cuentas de redes sociales a lo largo de Internet. Desde nuestra última actualización pública este grupo ha intentado hackear decenas de cuentas de la milicia ucraniana. En algunos casos, publicaron videos pidiendo al ejército que se rindiera, como si estas publicaciones provinieran de los propietarios legítimos de la cuenta. Bloqueamos la posibilidad de compartir estos videos.

Segundo, detectamos y desactivamos un intento de regresar a la plataforma de una red que removimos en diciembre de 2020, vinculada a individuos asociados con actividades pasadas de la Agencia de Investigación de Internet de Rusia (IRA por sus siglas en inglés). Parece que su actividad fuera de la plataforma empezó el año pasado y se concentraba en un sitio web que se hacía pasar por una ONG especializada en derechos civiles en Occidente. Intentaron sin éxito crear cuentas de Facebook a finales de 2021 y en enero de 2022. A lo largo de enero y febrero de este año, el sitio web publicó acerca de violencia policiaca en Francia y en Estados Unidos, pero desde la invasión sus artículos culpaban la invasión de Rusia a la OTAN y a Occidente y acusaban a las fuerzas ucranianas de atentar contra civiles.

Actores no estatales políticamente alineados

Detectamos e impedimos un intento de volver a operar en nuestra plataforma por parte de una red que eliminamos en diciembre de 2020, vinculada a personas en la región de Lugansk en Ucrania. Esta actividad se concentraba en dos sitios web que promovían comentarios a favor de Rusia en el Cáucaso y Ucrania, y un pequeño número de cuentas en Facebook, Telegram, VK y OK. A principios de marzo, el sitio enfocado en Ucrania parece haber sido tomado para dirigir a su audiencia hacia un canal de Telegram que mostraba víctimas rusas.

También, eliminamos una red en Rusia por infringir nuestra política de Comportamiento Inauténtico contra reporte masivo. La red se coordinó para reportar falsamente a personas en Ucrania y también en Rusia por diversos tipos de violaciones, incluyendo discurso de odio, en un intento por eliminar sus cuentas y contenido de Facebook.

Actores con motivaciones financieras

Como usualmente pasa con eventos globales importantes y problemas sociales relevantes, hemos visto que los estafadores alrededor del mundo recurren a la guerra en Ucrania para acumular audiencias y monetizar la atención de todos hacia esta crisis humanitaria. Sabemos que a primera vista estas actividades pueden ser malinterpretadas como operaciones de influencia respaldadas por el Estado, cuando en realidad provienen de estafadores que usan los temas sociopolíticos como una forma de señuelo de spam o clickbait.

Desde que comenzó la guerra, hemos investigado y eliminado decenas de miles de cuentas, Páginas y Grupos que usan tanto sistemas automáticos como manuales. Hemos visto a estafadores de todo el mundo usar tácticas de comportamiento inauténtico, como transmitir videos en vivo de gaming y re-compartir contenido popular como los videos de otras personas en Ucrania como una forma de simular que comparten actualizaciones en vivo. Muchos de los estafadores cambiaron de nombres de forma repetida para hacer que las personas los siguieran para intentar hacer dinero dirigiéndolas a sitios con anuncios o vendiéndoles mercancía. También eliminamos múltiples clústers de cuentas comprometidas que fueron abandonadas por mucho tiempo, que de repentinamente pasaron a ser administradas desde Rusia. Muchas de ellas compartían videos pro separatistas idénticos y amplificaban cuentas en sus propios clústers, posiblemente como parte de interacciones inauténticas pagadas.

Seguridad de la cuenta

Invitamos a las personas en Ucrania y en Rusia a incrementar la seguridad de sus cuentas en línea, incluyendo sus correos electrónicos y redes sociales. Para ayudar a mantener sus cuentas digitales seguras y proteger el acceso a redes sociales y otros sitios web bloqueados en sus países:

- Descarga una app de VPN en tus dispositivos para asegurar el acceso a sitios bloqueados, como redes sociales, a través de una conexión encriptada.
- Activa la autenticación de dos pasos usando una aplicación independiente de autenticación, como Google Authenticator o Duo.
- No reutilices tu contraseña. Las contraseñas deben ser seguras y únicas para cada una de tus cuentas.

03

Eliminamos cuatro redes por comportamiento inauténtico coordinado

Vemos CIB *como esfuerzos coordinados para manipular el debate público y lograr un objetivo estratégico, en donde las cuentas falsas son parte central de la operación. En cada caso, las personas se coordinan entre ellas para usar cuentas falsas para engañar a otros acerca de quiénes son y qué hacen. Cuando investigamos y eliminamos estas operaciones, nos enfocamos en el comportamiento en lugar del contenido, sin importar quién está detrás, qué publican o si son redes extranjeras o locales.*

Medidas continuas: *Monitoreamos los esfuerzos para restablecer la presencia en nuestras plataformas de redes que ya hemos removido previamente. Usando detección automática y manual, eliminamos constantemente cuentas y Páginas conectadas a redes que removimos en el pasado.*

1. Brasil

Eliminamos una red de 14 cuentas, nueve Páginas y 39 cuentas de Instagram por infringir nuestra política contra comportamiento inauténtico coordinado. Esta red se originó en Brasil y se dirigía a audiencias locales en este país.

Las personas detrás de esta actividad basaban su operación en cuentas falsas, muchas de las cuales fueron detectadas y desactivadas por nuestros sistemas automáticos, y se dirigían a personas a lo largo de diversas redes sociales, como Facebook, Instagram y Twitter. Esta actividad operaba en lo que parecen ser dos fases. Primero, en 2020, la operación publicaba memes sobre temas sociales y políticos, incluyendo la reforma agraria y la pandemia de COVID-19. Abandonaron esta actividad después de un par de meses, casi sin haber generado interacciones. En 2021, crearon Páginas que se hacían pasar por ONGs y activistas falsos, enfocados en temas medioambientales en la región del Amazonas en Brasil. Publicaron sobre

deforestación, incluyendo argumentos diciendo que no todo es dañino, y criticando ONGs medioambientales legítimas que se han pronunciado en contra de la deforestación en el Amazonas.

Además de publicar memes originales, también compartían contenido de medios de interés general y publicaciones de Greenpeace y fotógrafos de naturaleza, en un posible intento de hacer parecer estas cuentas más creíbles. En una instancia, vimos que esta operación usó una foto de perfil, posiblemente usando técnicas de inteligencia artificial como las Redes Generativas Antagónicas (GAN por sus siglas en inglés).

Detectamos esta red como resultado de nuestra investigación sobre contenido inauténtico coordinado en la región. A pesar de que las personas detrás de esta actividad buscaron ocultar sus identidades y coordinación, nuestra investigación encontró vínculos con individuos asociados con la milicia brasileña[1].

- *Presencia en Facebook e Instagram:* 14 cuentas de Facebook, 9 Páginas, y 39 cuentas de Instagram.
- *Seguidores:* Alrededor de 1.170 cuentas seguían una o más de estas Páginas y cerca de 23.600 seguían una o más de estas cuentas de Instagram.
- *Publicidad:* Alrededor de $34 dólares en inversión publicitaria en Facebook e Instagram, pagados en Reales brasileños.

2. Costa Rica y El Salvador

Eliminamos 233 cuentas de Facebook, 84 Páginas, dos Grupos y 27 cuentas de Instagram por infringir nuestra política contra comportamiento inauténtico coordinado. Esta red se originó en Costa Rica y El Salvador y se dirigía principalmente a Costa Rica y El Salvador.

Las personas detrás de esta actividad usaban cuentas falsas, algunas de ellas detectadas y desactivadas por nuestros sistemas automáticos, para administrar Páginas que se hacían pasar por medios de noticias, publicar memes, comentar sobre su propio contenido y el de otros y dirigir a las personas a dominios fuera de la plataforma.

[1] El proceso de atribuir una actividad infractora a ciertos actores de amenaza ha sido ampliamente debatido en la comunidad de seguridad. El enfoque de Meta sobre atribución se basa en las señales técnicas y de investigación disponibles. Cuando, con base en la evidencia disponible, nuestros equipos de investigación especializados no detectan evidencia clara de comando y control, pero sí detectan un número de individuos asociados con la entidad detrás de la operación, Meta atribuye la actividad a “individuos vinculados a la entidad”.

Esta red también amplificó contenido de Páginas de políticos y negocios locales. Usualmente publicaba sobre ambos bandos del espectro político, incluyendo apoyo a los candidatos políticos que competían entre sí. Algunas de estas cuentas tenían fotos de perfil, posiblemente generadas usando técnicas de inteligencia artificial como Redes Generativas Antagónicas (GAN por sus siglas en inglés).

Las personas detrás de esta actividad publicaban principalmente en español acerca de noticias y eventos actuales en Centroamérica. También, publicaban mensajes de apoyo a una compañía de telecomunicaciones en Costa Rica y criticaban a sus competidores.

Detectamos esta red tras revisar reportes públicos sobre una parte de su actividad fuera de la plataforma y tomamos acciones antes de las elecciones en Costa Rica. Aunque las personas detrás de la operación intentaron ocultar su actividad y coordinación, nuestra investigación encontró vínculos con individuos asociados con Noelix Media, una empresa de Relaciones Públicas con oficinas en Costa Rica y El Salvador. Noelix está ahora vetada de nuestra plataforma.

- *Presencia en Facebook e Instagram:* 233 cuentas de Facebook, 84 Páginas, dos Grupos y 27 cuentas de Instagram.
- *Seguidores:* Alrededor de 212.000 cuentas seguían una o más de estas Páginas, cerca de 10 cuentas se unieron a uno o más de estos Grupos, y cerca de 2.000 cuentas seguían una o más de estas cuentas de Instagram.
- *Publicidad:* Cerca de $128.000 dólares en inversión publicitaria en Facebook e Instagram, pagados principalmente en dólares estadounidenses y colones costarricenses.

3. Rusia y Ucrania

Reportamos esta aplicación de la política como parte de nuestra actualización de seguridad el 27 de febrero de 2022.

Eliminamos una pequeña red de 27 cuentas de Facebook, dos Páginas, tres Grupos y cuatro cuentas de Instagram por infringir nuestra política contra comportamiento inauténtico coordinado. Esta red operaba desde Rusia y Ucrania y se dirigía principalmente a Ucrania.

Descubrimos una red relativamente pequeña de 27 cuentas de Facebook, dos Páginas, tres Grupos y cuatro cuentas de Instagram que se dirigían a personas en Ucrania en diferentes redes sociales y a través de sus propios sitios web. Esta red usó cuentas falsas y administró personajes y marcas falsos a lo largo de Internet, incluyendo en Facebook, Instagram, Twitter, YouTube, Telegram, Odnoklassniki y VK, para parecer más auténticos en un aparente intento de soportar el escrutinio de las plataformas e investigadores. Estos personajes falsos usaron fotos de perfil posiblemente generadas usando técnicas de inteligencia artificial como las Redes Generativas Antagónicas (GAN por sus siglas en inglés). Eliminamos esta operación, bloqueamos la posibilidad de compartir sus dominios en nuestra plataforma y compartimos información con otras empresas de tecnología, investigadores y gobiernos.

Las cuentas falsas afirmaban estar basadas en Kiev y se hacían pasar por editores, un ex ingeniero de aviación, y un autor de una publicación de hidrografía, la ciencia de mapear el agua. Esta operación operó una serie de sitios web disfrazados de sitios de noticias independientes, publicando afirmaciones sobre una supuesta traición de Occidente a Ucrania y Ucrania como un Estado fallido.

Nuestra investigación encontró vínculos con esta red y otra operación que eliminamos en abril de 2020, vinculada con individuos en Rusia, la región de Donbás en Ucrania y dos organizaciones de medios en Crimea, NewsFront y SouthFront, ahora sancionadas por el gobierno estadounidense.

- *Presencia en Facebook:* 27 cuentas de Facebook, dos Páginas, tres grupos y cuatro cuentas de Instagram.
- *Seguidores: Alrededor de* 3.450 cuentas seguían una o más de estas Páginas y cerca de 415 cuentas seguían una o más de estas cuentas de Instagram.
- *Publicidad:* Alrededor de $200 dólares en inversión publicitaria en Facebook e Instagram pagada principalmente en rublos rusos y dólares estadounidenses.

4. Rusia

Reportamos esta aplicación de nuestras políticas como parte de nuestra actualización de CIB el 16 de febrero de 2022.

Eliminamos una pequeña red de tres cuentas de Facebook por infringir nuestra política en contra de comportamiento inauténtico coordinado. Esta red se originó en San Petersburgo, Rusia, y se dirigió principalmente a Nigeria, Camerún, Gambia, Zimbabwe y Congo.

Las personas detrás de esta actividad usaron cuentas falsas para crear personas falsas que se hacían pasar por editores o un ejecutivo de Relaciones Públicas basado en Europa de habla árabe. Estas cuentas tenían fotos de perfil, posiblemente generadas usando técnicas de inteligencia artificial como las Redes Generativas Antagónicas). Vimos dos periodos de actividades, la mayor parte de estos poco exitosos. Primero, esta operación trató de solicitar apoyo de periodistas independientes para escribir artículos sobre Siria a través de Grupos de periodistas de habla árabe. Después de un periodo de inactividad, parece ser que se enfocaron en África, en un intento para cooptar medios de comunicación para que publicaran historias a su nombre sobre política africana, incluyendo críticas a la influencia francesa en África. Estamos notificando a las personas que creemos fueron contactadas por esta red.

Detectamos esta actividad como parte de nuestra investigación interna sobre sospechas de comportamiento inauténtico coordinado con vínculos a la actividad que desarticulamos en agosto de 2020. Aunque las personas detrás de esta operación intentaron ocultar sus identidades y coordinación, nuestra investigación encontró vínculos a individuos asociados con actividad pasada de la Agencia de Investigación de Internet de Rusia.

- *Presencia en Facebook:* tres cuentas de Facebook.

04

Eliminamos una red de reportes masivos en Rusia

Bajo nuestra política de Comportamiento Inauténtico, eliminamos actividad de reportes masivos *cuando detectamos redes adversarias que se coordinan para abusar de nuestros sistemas de reporte para eliminar cuentas o contenido erróneamente de nuestra plataforma, generalmente con el objetivo de silenciar a otros.*

Rusia

Eliminamos una red de 200 cuentas que operaban desde Rusia. Los individuos detrás de la actividad se coordinaban para reportar a las personas por supuestas infracciones a las políticas de la plataforma, incluyendo discurso de odio, acoso e inautenticidad, en un intento por eliminar a estas personas y su contenido de Facebook. La mayoría de estos reportes falsos se enfocaron en personas en Ucrania y Rusia, pero la red también reportó a usuarios en Israel, Estados Unidos y Polonia.

Las personas detrás de esta actividad se basaron en cuentas falsas, auténticas y duplicadas para presentar cientos, en algunos casos miles, de reportes contra sus objetivos, abusando de las herramientas de reporte. Muchas de las cuentas de esta red fueron detectadas y deshabilitadas por nuestros sistemas automáticos. Sus reportes coordinados incrementaron a mediados de febrero, justo antes de la invasión a Ucrania. Posiblemente en un intento para no ser detectados, las personas responsables de esta actividad coordinaron los reportes masivos en su Grupo sobre cocina, que tenía alrededor de 50 miembros cuando lo eliminamos.

Detectamos esta red como resultado de nuestra investigación interna sobre sospechas de comportamiento inauténtico en la región. Nuestro análisis identificó vínculos limitados entre esta actividad y una red rusa que eliminamos por CIB en 2019.

05

Eliminamos una red infractora coordinada en Filipinas

Eliminamos redes infractoras coordinadas *cuando detectamos a personas, ya sea usando cuentas auténticas o falsas, trabajando juntas para infringir o evadir nuestras Normas Comunitarias.*

Esta aplicación de nuestras políticas, principalmente basada en el comportamiento, complementa nuestras políticas de contenido existentes, bajo las cuales ya eliminamos contenido y cuentas que infringen nuestras Normas Comunitarias, incluyendo incitación a la violencia, acoso y bullying o desinformación de salud dañina. Reconocemos que, en algunos casos, estas infracciones de contenido son perpetuadas por un grupo organizado trabajando en conjunto para amplificar el comportamiento dañino de sus miembros y repetidamente violar nuestras políticas de contenido. En algunos casos, el riesgo de daño potencial causado por la actividad total de la red supera el impacto de las publicaciones individuales o cuenta. Para abordar estos esfuerzos coordinados de forma más efectiva, desarrollamos protocolos para la aplicación de nuestras políticas que nos permiten tomar acciones contra la red central de cuentas, Páginas y Grupos involucrados en este comportamiento. Como parte de este marco, podríamos tomar una serie de acciones, como reducir el alcance del contenido y deshabilitar cuentas, Páginas y Grupos.

Filipinas

Eliminamos una red de más de 400 cuentas, Páginas y Grupos en Filipinas que trabajaron juntas para infringir nuestras políticas contra daño coordinado, bullying y acoso, discurso de odio, desinformación, incitación a la violencia y evasión de las normas de forma sistemática.

Las personas detrás de esta actividad afirmaron ser hacktivistas y se basaron principalmente en cuentas auténticas y duplicadas para publicar y amplificar contenido sobre ataques de

Denegación de Servicio (DDoS), recuperación de cuentas, así como la desagregación y vulneración de sitios web en Filipinas. Comentaron acerca de ataques DDoS contra los sitios del Premio Nobel en diciembre de 2021 que no fue exitoso, y acusaban a figuras públicas en Filipinas de ser comunistas (una táctica conocida como “red-tagging”).

Esta red afirmaba bajar sitios y desagregarlos, incluyendo aquellos de entidades de noticias, negocios y escuelas. También ofrecían servicios de ciberseguridad para proteger a sitios web de ataques, como los que afirmaban haber atacado. Finalmente, este grupo invitó públicamente a estos nuevos miembros a unirse y llevar a cabo ataques DDoS.

06

Eliminamos comportamiento inauténtico

¿Qué es el Comportamiento Inauténtico (IB)? *Mientras CIB está orientado principalmente a engañar a las personas acerca de quién está detrás de una operación para manipular el debate público para un objetivo estratégico, el Comportamiento Inauténtico se enfoca principalmente a amplificar e incrementar la distribución de contenido. Generalmente, pero no de forma exclusiva, tiene motivaciones financieras y mezcla muchas tácticas con spam y actividades fraudulentas.*

¿Cómo aplicamos las políticas contra Comportamiento Inauténtico?

Este año, eliminamos decenas de miles de cuentas, Páginas y Grupos alrededor del mundo por inflar falsamente la distribución de contenido y construir audiencias de forma abusiva. Nos apoyamos en una variedad de medidas para aplicar nuestras políticas contra IB, desde las alertas, la reducción en la distribución de contenido, hasta la remoción de actores de IB y clústers de actividad de nuestra plataforma. Más información de Comportamiento Inauténtico aquí.

Los operadores de IB generalmente se enfocan en la cantidad, en lugar de la calidad de interacciones. Por ejemplo, pueden usar un gran número de cuentas falsas poco sofisticadas para publicar masivamente su contenido o darle “Me gusta”. También, podrían intentar monetizar la atención de las personas, ya sea dirigiéndolas a sitios web con anuncios fuera de la plataforma o vendiéndoles playeras u otros productos. Como respuesta a las detecciones y remoción, generalmente tratan de reconstruir su actividad de manera agresiva. Debido a que por lo general trabajan a escala, combatimos IB a través de la aplicación de nuestras políticas y detección automática a escala, complementada por investigaciones manuales para ayudarnos a identificar nuevas tácticas o vacíos en la detección.

Este enfoque nos ayuda a aprender y mejorar nuestras defensas en respuesta a la adaptación de los agresores, mientras eliminamos estos clústers de actividad a escala, sin importar si buscan promover rumores sobre celebridades o clickbait sociopolítico, con la intención de acomular audiencias.

A continuación algunas de las estrategias engañosas que hemos visto en las operaciones de IB, usadas para impulsar sus interacciones:

Cambio de contexto

Los operadores de IB a menudo buscan engañar y aumentar su audiencia afirmando que se dedican a un tema popular y luego cambian a otro no relacionado cuando este se vuelve viral. Conocen bien a sus audiencias y cambian rápidamente su enfoque para publicar sobre las últimas noticias o escándalos para engañar a las personas y orillarlas a dar clic en los enlaces a sus sitios. Como era de esperar, la política también se ha convertido en un tema de interés común para generar spam. A primera vista, estas actividades pueden confundirse, y a menudo se confunden, con operaciones de influencia con motivaciones políticas, cuando en realidad utilizan temas políticos como otra forma de clickbait, de manera similar a los memes de cachorros o celebridades.

Por ejemplo, en las primeras etapas de la guerra en Ucrania, investigamos un dato de un periodista y eliminamos una red de cuentas de Instagram administradas desde Estados Unidos, algunas de las cuales afirmaban estar reportando en vivo desde Ucrania. Intentaron monetizar, incluso vendiendo productos temáticos a través de su sitio web. Después de que comenzara la invasión rusa, este grupo pasó rápidamente de publicar sobre "conducción aterradora" y "videos de Airsoft" a temas militares. En otro caso, eliminamos una página de Vietnam que pasó de publicar videos sobre "tips de vida" y joyería a publicar sobre equipos militares y el conflicto de Ucrania, todo para llevar a las personas a un sitio web fuera de la plataforma.

Hacerse pasar por comunidades auténticas en diferentes países

Las redes de IB a menudo pretenden tener su sede en un país, cuando en realidad operan desde otro completamente diferente. Esta táctica a menudo va de la mano con el "cambio de contexto" que describimos anteriormente. Incluye *spammers* y estafadores extranjeros que actúan en masa alrededor de cualquier tema de interés en un país o región en particular, como una elección, una crisis sociopolítica o un desastre natural, para acumular audiencias y monetizar su atención.

Por ejemplo, varios grupos de spam con sede en Vietnam y Bangladesh se hicieron pasar por partidarios del Canadian Trucker Convoy para sacar provecho del interés de la gente en esta protesta. En un caso, crearon grupos para los partidarios del convoy o páginas de Facebook diseñadas para que pareciera que estaban dando actualizaciones sobre el convoy, y luego

publicaron enlaces a sitios web de comercio electrónico o enlaces a sitios de marketing de terceros.

Reporte masivo, “me gusta” y compartir contenido para hacerlo parecer más popular

Por lo general, los actores de IB se basan en cuentas falsas para dar “Me gusta”, comentar o compartir contenido en un esfuerzo por crear la falsa percepción de que el contenido es orgánicamente popular. En un caso que investigamos el último trimestre, un grupo de cuentas falsas creadas en Bangladesh parecía haber cambiado de administrador y se usaba para comentar, dar me gusta y compartir contenido en la Página de un excandidato al Senado de Estados Unidos, proveniente de Arizona.

Este comportamiento puede aumentar la cantidad de “Me gusta” e interacciones en las publicaciones, pero nuestras investigaciones a lo largo de los años han encontrado que rara vez, si es que alguna, resulta en interacciones entre personas reales. Estos grupos de amplificación a menudo se manifiestan como una "burbuja" de interacciones o "click clique" donde solo a sus propios miembros interactúan con las publicaciones de los demás, en lugar de personas reales fuera de esa burbuja.

A detalle: Las elecciones en Filipinas

Al igual que con cualquier evento cívico importante, hemos visto a los operadores de IB de varios países, incluidos Filipinas, Vietnam, Estados Unidos y Tailandia, participar activamente alrededor de las próximas elecciones de Filipinas en mayo, utilizando las tácticas comunes de IB que describimos anteriormente. Parecían enfocarse en aumentar su audiencia para una eventual monetización o para hacer que el contenido de sus clientes pareciera más popular de lo que realmente es.

Cambio de contexto

Eliminamos varios grupos que cambiaron el enfoque de sus Páginas y Grupos para hablar de las elecciones mientras intentaban aumentar su número de seguidores. En un caso, una red incluía una serie de Páginas que cambiaban de temas no relacionados a política a temas políticos y, en algunos casos, regresaban a otros temas no relacionados. Una Página que compartía principalmente videos de baile se renombró para convertirse en

"Noticias de Bongbong Marcos", mientras que otra Página que comenzó apoyando a un político luego cambió su nombre a "Tu respuesta financiera" y comenzó a publicar consejos sobre préstamos. Entre estos Grupos, vimos intentos adversarios de evadir nuestra detección automática de cambio de nombre, incluida una Página que cambió su nombre ocho veces durante ocho años, cambiando gradualmente de "Historia y curiosidades de Filipinas" a centrarse en candidatos políticos particulares.

Esfuerzos engañosos para hacerse pasar por comunidades auténticas en diferentes países

Eliminamos varios grupos en Vietnam, Tailandia y EE.UU. que se hacían pasar por miembros de comunidades locales en Filipinas en un aparente intento de monetizar la atención de las personas sobre las elecciones. En febrero, identificamos un grupo de Páginas operadas por spammers en Vietnam que usaban un VPN para aparentar que tenían su sede en Filipinas. Se hicieron pasar por simpatizantes de campañas políticas o entidades de noticias locales y usaron nombres como Noticias de Tendencia en Filipinas, Duterte Live, Relacionada con Francis Leo Marcos y Noticias Pinas. Afirmaron compartir imágenes en vivo mientras pretendían ser fuentes de noticias locales en un intento de llevar a las personas a sus sitios web llenos de anuncios.

Interacciones inauténticas

Identificamos varios esfuerzos para hacer publicaciones constantes, con tasas similares al spam, para dirigir a las personas a Páginas específicas o sitios web fuera de la plataforma. En un caso, encontramos una agencia de administración de redes sociales que usaba cuentas falsas y Páginas duplicadas para amplificar contenido político y de entretenimiento de manera inauténtica. La agencia usó una red de más de 700 cuentas para publicar, compartir y comentar publicaciones y compartir contenido a través de Grupos grandes.

En otros casos, detectamos y eliminamos actividad de interacciones inauténtica operada por las mismas personas que apoyaban a varios candidatos en la misma elección al. En el periodo previo a las elecciones, eliminamos alrededor de una docena de grupos centrados en interacciones falsas.

Continuamos monitoreando de cerca la situación en el período previo a las elecciones de mayo en Filipinas y tomaremos medidas si encontramos actividades infractoras que

intenten aprovechar el interés de las personas en las elecciones utilizando tácticas de IB.

c Apéndice: Indicadores de Amenazas

1. UNC788, Irán

Domains y C2s

- bnt2[.]live
- archery.dedyn[.]io
- market.vinam[.]me
- signin.dedyn[.]io
- Market.dedyn[.]io

Hashes

- 43535540e94b39279af925e9548dce7f
- 9b91427d195b8b7e75fbbc29a798bede
- aaa55f1e48aba8856661fedc0074e81a
- 6e0ec6bd0bef489c83c2dce4876de5c8
- 70875705e8bc3887cec4ef1873cdb152
- aa7330d2d360cac61394843d8af730bb
- ab533be4ff9c99e8a03bc4cd413badb6

Regla Yara

```
rule hilal_rat_dex: {
  meta:
    source = "Facebook"
    date = "2022-04-07"
    description = "Detects custom android rat impersonating various
applications that siphons phone details to a C2."
  strings:
    $class0 = "Lcom/hilal/SysUpdater/MainActivity;"
    $class1 = "Lcom/hilal/adm/R;"
    $file0 = "cacaca.dat"
    $file1 = "ccc.dat"
    $file2 = "fifi.dat"
    $file3 = "smr.dat"
```

```
    $file4 = "smse.dat"
    $typo0 = "Erron in Decryption"
    $typo1 = "GetDevcie"
    $sec1 = "6123cc12ef9bd0bf1592c69bf769853fb0a00084" // AES key
    $cmd0 = "/Aud"
    $cmd1 = "/Cam"
    $cmd2 = "/Upd"
    $cmd3 = "/Con"
    $func1 = "CamStart"
    $func2 = "AudStop"
    $func3 = "AudStart"
    $func4 = "DownFi"
    $func5 = "ScrSht"
    $func6 = "CamList"
    $func7 = "CamStop"
    $func8 = "ListExplore"
    $interesting_string0 = "isMyServiceRunning?"
    $interesting_string1 = "Checking new version... Please wait..."
    $notification_service0 = "**********  onNotificationPosted"
    $notification_service1 = "********** onNOtificationRemoved"
    $phnum = "PhNumber"
  condition:
    filetype_dex and
    10 of them
    or all of ($file*)
    or all of ($func*)
    or all of ($interesting_*) and all of ($notification_*)
}
```

2. Grupo no reportando anteriormente (Irán)

Dominios y C2s

- alharbitelecom[.]co
- apply-jobs[.]com
- applytalents[.]com
- appslocallogin[.]online
- careers-finder[.]com
- cloudgoogle[.]co
- cortanaservice[.]com
- cortanaupdate[.]co
- defenderupdate[.]ddns[.]net
- edge-cloudservices[.]com
- elecresearch[.]org
- enerflex[.]ddns[.]net
- enerflex[.]org
- etisalatonline[.]com
- exprogroup[.]org
- freechess[.]live
- funnychess[.]online
- getadobe[.]ddns[.]net
- getadobe[.]net
- globaltalent[.]in
- googleservices[.]co
- googleupdate[.]co
- helpdesk-product[.]com
- khaleejtimes[.]co
- librarycollection[.]org
- linkedinz[.]me
- listen-books[.]com
- lukoil[.]in
- mastergatevpn[.]com
- microsoftcdn[.]co
- microsoftdefender[.]info
- microsoftedgesh[.]info
- mideasthiring[.]com
- office-shop[.]me
- onedrivelive[.]me
- onedriveupdate[.]net
- online-audible[.]com
- online-chess[.]live
- outlookde[.]live
- outlookdelivery[.]com
- remgrogroup[.]com
- saipem[.]org
- sauditourismguide[.]com
- savemoneytrick[.]com
- sharepointnotify[.]com
- sparrowsgroup[.]org
- supportskype[.]com
- talent-recruitment[.]org
- talktalky[.]azurewebsites[.]net
- thefreemovies[.]net
- updateddns[.]ddns[.]net
- updatedefender[.]net
- updatedns[.]ddns[.]net
- updateservices[.]co

3. Azerbayián

Dominos

- localadmin[.]online
- localadmin[.]ru
- analyzeryandex[.]000webhostapp[.]com
- vote2021[.]w3spaces[.]com

URLs para hacer phishing de datos de acceso

- localadmin[.]online/votes/security
- localadmin[.]online/vote
- localadmin[.]online/vote/fb/login.html

4. CIB: Costa Rica, El Salvador

Dominios

- latinoamericareporta[.]com
- revistadcr[.]com

5. CIB: Rusia, Ucrania

Dominos

- kavkazru[.]press
- politica[.]in[.]ua

6. CIB: Rusia, Ucrania

Dominios

- monitor-ua[.]com
- ukraine2day[.]com