STRATOP®
RISK CONSULTING

CONFLICTOS DE LA ZONA GRIS Y SUS EFECTOS.

INFLUENCIA Y MANIPULACIÓN DE LA OPINIÓN PÚBLICA A TRAVÉS DE LAS REDES SOCIALES.

FEBRERO, 2023

Mtro. Raúl A. Álvarez Paniagua

Policy Brief febrero, 2023
Stratop Risk Consulting®

Introducción

I.Información.

A partir de la creación de herramientas como blogs, MySpace, Facebook y Twitter que pasaron de ser herramientas donde las personas tenían la posibilidad de interactuar para alcanzar un auditorio específico (por temática, geografía, área de estudio, etc.) a ser plataformas y verdaderos amplificadores para expresar, cualquier pensamiento por más radical o sencillo que sea.

Mucho se ha discutido, a últimas fechas sobre el mal uso que se les ha dado a estos servicios, pero particularmente, el que más asombra es el de la siembra de encono, rivalidades, de “ustedes-nosotros”, “nosotros-ellos”, de división. La comunicación a través de los entornos digitales.

En los últimos años, internet como un medio de diseminación de ideas políticas y como un catalizador de demandas sociales, siendo posible incluso, generar cambios en la conducción política o detener iniciativas populares de los gobiernos. Aun así, en las redes sociales se reflejan cuestiones estructurales de las sociedades, como la desigualdad, y además, manifiestan de cierto modo, nuevas preocupaciones para los ciudadanos y sus gobiernos.

Recientemente, varios episodios de la política demuestran una mayor influencia de la población en las políticas públicas de un Estado, las cuales pasan por la presión en momentos electorales, en cierta medida en la política exterior y al papel de denuncia que los medios de información han asumido.

Este ecosistema de redes sociales tiene muchas vulnerabilidades, tanto técnicas como cognitivas, que pueden explotarse utilizando técnicas de influencia, tales como pruebas sociales, identidades engañosas y bots. Las técnicas más potentes ahora se empaquetan y venden en el mercado de la manipulación de las redes sociales.

El mercado de la manipulación de los medios sociales, se divide en tres categorías superpuestas: 1) el mercado abierto de fácil acceso, 2) la web oscura y 3) el mercado de usuario a usuario sin conexión. Todos estos son canales no oficiales que brindan a los usuarios la oportunidad de comprar “Me gusta” (Likes), acciones, comentarios, suscriptores y cuentas.

En esta industria existe una contradicción con los “Términos de Servicio” de las plataformas, que no permiten la compra de los “Me gusta”, comentarios o suscriptores; sin embargo, esto sigue siendo legal en la mayoría de los países, debido a que no existe un marco legal para la regulación de sus actividades en el ciberespacio.

Actualmente, una simple búsqueda en la Dark Web producirá una lista de proveedores que ofrecen vender herramientas y servicios para la manipulación de las redes sociales. Estos servicios operan en una Zona Gris y desarrollan activamente formas de manipular las redes sociales.

Como referencia conceptual, el diplomático estadounidense George Kennan instó al uso de la guerra política mediante la aplicación lógica de la doctrina de Clausewitz en tiempo de paz para contrarrestar las actividades del adversario, definiendo la Guerra Política como "el empleo de todos los medios al mando de una nación, a excepción de la guerra"[1]. El modo de conflicto al que Kennan originalmente se refería como la Guerra Política ha sido recientemente renombrado como "Conflictos en la Zona Gris".

Ante esto, en el reporte elaborado el 3 de enero del 2017 por el Consejo Asesor de Seguridad Internacional (ISAB por sus siglas en inglés) del Departamento de Estado de los EUA, se definió la “Zona Gris”:

> *La Zona Gris se caracteriza por una intensa competencia política, económica, informativa y militar de naturaleza más ferviente que la diplomacia normal de un Estado, pero sin la guerra convencional[2]. Y que los desafíos de la Zona Gris se definen como una interacción competitiva entre actores estatales y no estatales que se encuentran entre la guerra tradicional y la dualidad de paz.*

El Del mismo modo, agrega que estos conflictos se caracterizan por la ambigüedad acerca de su naturaleza, la opacidad de las partes involucradas o la incertidumbre acerca de los marcos legales y de políticas relevantes para su encuadramiento en el arte de la guerra[3].

Por lo tanto, se puede afirmar que aquellas actividades encubiertas o ilegales de arte no tradicional que están por debajo del umbral de la violencia armada organizada; incluida la interrupción del orden, la subversión política en contra del gobierno o un empleo intensivo de las organizaciones no gubernamentales para la movilización social, la manipulación de las redes sociales y la corrupción financiera como parte de un diseño integrado para lograr una ventaja estratégica, son las tácticas implementadas actualmente en la Zona Gris.

[1] Max Boot and Michael Doran, “Political Warfare”, Council on Foreign Relations (28 de Junio de 2013), disponible en https://cfrd8-files.cfr.org/sites/default/files/pdf/2013/06/Policy_Innovation_Memorandum_33_Boot.pdf
[2] Declaración del General Joseph L. Votel comandante del USSOCOM ante el Subcomité del Comité de Servicios Armados de la Cámara de Representantes sobre Amenazas y Capacidades Emergentes (18 de marzo de 2015), Washington, D.C. EUA; disponible en https://docs.house.gov/meetings/AS/AS26/20150318/103157/HMTG-114-AS26-Wstate-VotelUSAJ-20150318.pdf
[3] Véase “The Gray Zone”, documento técnico elaborado por el Capitán Philip Kapusta, del Comando de Operaciones Especiales de los EUA. (USSOCOM), “Definición de los desafíos de la Zona Gris”, (9 de septiembre de 2015), pág. 1, disponible en: https://info.publicintelligence.net/USSOCOM-GrayZones.pdf

Análisis

El Comando de Operaciones Especiales de los EUA (USSOCOM), ha examinado varios estudios de casos para comprender mejor cómo conceptualizar este conjunto de problemas y responder en consecuencia. Derivado de lo anterior, considera que las técnicas en la Zona Gris incluyen[4]:

a. Uso intensivo del ciberespacio, operaciones de información, esfuerzos para socavar la resistencia pública / aliada / local / regional, y la información / propaganda en apoyo de otros instrumentos híbridos[5];
b. Operaciones encubiertas, espionaje, infiltración y subversión;
c. Operaciones de Fuerzas Especiales y otras unidades armadas controladas por un Estado, y personal militar no reconocido (Fuerzas Irregulares);
d. Apoyo logístico, político y financiero para movimientos insurgentes y terroristas;
e. Alistamiento de actores no gubernamentales, incluidos organizaciones políticas, religiosas, étnicas o sectarias extremistas;
f. Asistencia a fuerzas militares y paramilitares irregulares en otros países;
g. Las presiones económicas que van más allá de la competencia económica normal;
h. Manipulación y desacreditación de las instituciones democráticas de un Estado, incluido el sistema electoral y el poder judicial;
i. Ambigüedad calculada, uso de operaciones encubiertas no reconocidas, engaño y negación plausible; y
j. Uso de amenazas explícitas o implícitas, amenazas de uso de la fuerza armada, terrorismo y el abuso de algunos sectores de la población civil.

El término “Conflictos en la Zona Gris” puede ser nuevo; sin embargo, el fenómeno no lo es. Si bien muchas de las técnicas utilizadas ahora están basadas en tecnología moderna, en particular las comunicaciones cibernéticas y el uso intensivo de las redes sociales, muchas son tan antiguas como la historia. El Dr. Josep Baqués[6] explica que, a diferencia de las campañas militares, los Conflictos en la Zona Gris plantean alcanzar sus metas a más largo plazo, por lo tanto, sin prejuicio de cuál sea el objetivo final de la Zona Gris, suele desgastar a los actores afectados, deteriorando su legitimidad, su *modus vivendi*, su cohesión social, su economía o todos esos factores a la vez, en un proceso que suele conllevar varios años, pero que es estratégicamente rentable, considerando que la relación costo-beneficio de esta estrategia es exponencial.

[4] No todas las técnicas de uso de estas técnicas pueden considerarse operaciones de Zona Gris. Por ejemplo, las Operaciones de Fuerzas Especiales (SOF por sus siglas en inglés) tienen muchas funciones, como el compromiso y el mantenimiento de la paz, que no son esfuerzos de la Zona Gris, aunque solo sea porque carecen del objetivo de obtener una ventaja competitiva sobre un adversario.
[5] El ciberespacio representa un instrumento de ataques de la Zona Gris que es particularmente desafiante, no solo por su novedad, sino, en la mayoría de los contextos, por la mayor dependencia y vulnerabilidad de los EUA que por los posibles adversarios. En consecuencia, la preparación tanto organizacional como operacional para enfrentar los desafíos de la Zona Gris requiere atención en los problemas cibernéticos.
[6] Investigador del Instituto Español de Estudios Estratégicos quien aborda la conceptualización de los conflictos en la Zona Gris en su Documento de Investigación 02/2017 “Hacia una definición del concepto Gray Zone (GZ), IEEE, Granada, España (2017).

Los medios empleados para alimentar esta lógica son variados, entre los más usuales y citados está la propaganda política, entendida como una información políticamente orientada y además distorsionada. Los medios de difusión empleados son de lo más diversos, desde la utilización de periodistas afines, desplazados al epicentro de la Zona Gris, hasta el empleo de blogs y redes sociales de amplio espectro, aprovechando las ventajas del ciberespacio.

Pero también incluye, la posibilidad de planear operaciones de información y/o desinformación[7] que incluyan expertos en ciberseguridad, a miembros de los servicios de inteligencia e incluso componentes de operaciones psicológicas[8] apoyados de sólidos marcos teóricos de corte social/constructivista (ingeniería social), para la elaboración de narrativas en el adecuado contexto ya sea geográfico, sociológico/humano e histórico y desde el cual se genera la Zona Gris[9].

El hecho de utilizar la información como arma, lleva a confrontar, a causar confusión, eliminar la confianza en las instituciones e incluso entre las personas. Son operaciones claramente dirigidas con el objetivo de la transformación del tejido social y el debilitamiento de las instituciones.

Dichas narrativas, buscan lograr la intensificación y el impacto del debate político, mediante la asociación de grupos insatisfechos con la situación social y económica imperante, así como la inserción de movimientos opositores a través de la empatía, permitiendo al individuo tener la sensación de identificación y pertenencia a estos grupos; con el fin de amplificar la discordia en el entorno, así como el malestar social generalizado,

[7] Los medios para el desarrollo de este tipo de operaciones son: A). Manipulando las métricas sociales. Mediante la manipulación de métricas sociales replicando los "Me gusta", comentarios, acciones, vistas, seguidores, etc., en todas las plataformas de redes sociales. La manipulación de las métricas sociales se realiza mediante una gama de herramientas: a). Empleo de cuentas automáticas falsas. b). Con plataformas "freelance", usualmente empleando personas (operadores o suckpuppets) para su desarrollo en las llamadas "granjas de bots", quienes coordinan las interacciones mediante el uso de software. c). A través del intercambio de un "Me gusta" o el reenvío del contenido o publicación de Influencers, a cambio de la misma acción interactuando con otros usuarios y el contenido compartido a través de su cuenta en las redes sociales. Así como la manipulación de objetivos específicos como foros web (blogs), encuestas públicas en redes sociales, sitios web para manipular el resultado de una encuesta política o difamar a un competidor comercial, etc. B). Empleo de cuentas falsas adquiridas en el mercado negro. Las cuentas falsas o comprometidas sirven como la herramienta básica para obtener acceso a una plataforma de redes sociales, un requisito para poder manipular las plataformas con falsos "Me gusta", vistas, comentarios, etc. Es importante señalar que este tipo de cuentas serán empleadas a fin de evitar el rastreo de los usuarios; por lo que el precio de la compra de dichas cuentas depende en gran medida del nivel de seguridad de la red para una plataforma determinada; cuanto más difícil sea registrarse y mantener una cuenta, mayor será el precio. Las siguientes características afectan el precio de una cuenta falsa: a). Tipo de cuenta. Como en cualquier otro mercado, el costo de crear / mantener un servicio influye en su precio. Las cuentas que se pueden registrar automáticamente son las más baratas, pero también las menos confiables, mientras que las que se registran manualmente son más confiables, pero también más caras, al igual que las cuentas genuinas que han sido hackeadas. Existen tres tipos de cuentas falsas: 1). Cuentas registradas automáticamente. Estas cuentas se registran utilizando programas diseñados para este fin y, por lo general, no son muy sofisticadas. 2). Cuentas registradas manualmente, estas cuentas están registradas por operadores humanos y, a menudo, se pueden rellenar con contenido personalizado. 3). Cuentas pirateadas (hackeadas). Son cuentas que pertenecen a personas reales cuyas credenciales han sido pirateadas o robadas. Debido a sus registros históricos genuinos, las cuentas pirateadas son mejores para eludir las salvaguardas de la plataforma. En ocasiones, las cuentas pirateadas se consideran consumibles para recopilar información sobre los usuarios o como parte de un esquema de orientación a corto plazo, porque el propietario legítimo puede retomar el control de la cuenta. Sin embargo, dado que los propietarios reales a veces carecen de la voluntad y la capacidad de recuperar el acceso, dichas cuentas a veces permanecen bajo el control de otros usuarios para fines de manipulación. b). Verificación de la cuenta. Las cuentas de las redes sociales se pueden verificar por teléfono y por correo electrónico. Las cuentas verificadas son más caras, pero es menos probable que se bloqueen. La calidad del proceso de verificación también es un factor del precio: la verificación manual es más confiable, mientras que las cuentas verificadas automáticamente tienen un mayor riesgo de ser bloqueadas y el trabajo realizado a través de estas cuentas (me gusta, comentarios, etc.) se elimina. c). Contenido. Cuanto más contenido se proporcione para una cuenta creada con fines de manipulación, más realista se verá. Dependiendo de la tarea que se requiere que realice una cuenta, un nivel más alto de autenticidad percibida puede ser más o menos deseable. Una cuenta básica se vende generalmente en cuatro categorías: 1). Cuentas sin contenido, 2). Cuentas con una foto de perfil, 3). Cuentas con una foto de perfil y unas cuantas fotos, 4). Cuentas con una foto de perfil, fotos, y una gama de publicaciones. Además de estas categorías generales, se podrán desarrollar aquellas que acorde a la experiencia y habilidad de usuarios más avanzados con cuentas personalizadas y altamente desarrolladas que son casi imposibles de diferenciar de las cuentas reales. d). Edad de la cuenta. La historia de una cuenta (su edad), es Importante para la confianza y credibilidad. El mercado negro ofrece cuentas con una amplia gama de edades, desde unos cuantos días hasta más de 7 años. Cuanto más antigua es la cuenta, más cara es, ya que es menos probable que una cuenta antigua sea detectada como maliciosa.

[8] Los seres humanos construyen sus relaciones, tanto dentro como fuera de las redes sociales, lo que resulta más efectivo es influir a partir de proponer posturas que resultan discordantes. Los seres humanos siempre buscan reafirmación a partir de tener razón. Es por lo que a estas posturas se les conoce como "riesgos cognitivos" los cuales generan un ciclo de polarización que es aprovechado para ganar simpatías con tinte político y social. Esto puede ser aprovechado para generar afinidades a partir de la discordia. Una manera, relativamente sencilla, de distinguirlo es por el lenguaje. No se dan datos sino más bien abundan los adjetivos y se parte de los pronombres "nosotros" y "ustedes", para que, mediante interacciones coercitivas, los usuarios con quienes interactúan tomen postura, además, sustentándose profundamente en filias y fobias que por definición son irracionales. Existen 4 etapas plenamente identificadas para lograr dinamizar los "riesgos cognitivos" y posicionar una narrativa, estas etapas son:

A) En la primera etapa del ciclo a partir de conocerse un hecho o evento, surge un pico de interacción en el cual debe buscarse atraer la mayor atención posible con la premisa de asignar responsabilidades, difundir sospechas, mostrar información sin sustento, señalar "presuntos culpables" y exigir respuesta de las autoridades. B) A partir de generar elementos discordantes con la misma intensidad y alcance, en la segunda etapa se mantienen las interacciones de la primera etapa, pero, además, se empieza a mostrar solidaridad con las víctimas y quienes pueden resultar afectados a consecuencia del hecho o evento. Así, habiendo marcado posturas y generando controversia, se refuerzan los elementos discordantes entre los extremos ideológicos y de simpatía. C) Para la tercera etapa, el ciclo disminuye en su volumen de interacción e inicia la campaña sostenida a largo plazo con un discurso que se torna totalmente político y social. Los extremos ideológicos ya han tomado postura de manera prácticamente irreconciliable. D) Finalmente, como cuarta etapa, las ideas, supuestos, información descontextualizada, rumores, y todo lo que sirva como sustento para una u otra postura, se irá tornando más oscuro, más difuso, pero será difundido tanto como sea necesario para mantener la confrontación entre ambos extremos. Para este punto la influencia ya salió del espacio virtual y corrió de boca en boca volviéndose un tema de dominio público "bien conocido".

[9] Es importante señalar en este aspecto que no deben confundirse con las metas de las campañas de comunicación, aunque en muchos medios se les reconozca como comunicación estratégica.

“Sacar de contexto datos, recurrir a fuentes dudosas, y además de todo, el catálogo de falacias argumentativas, forman parte del modelo de interacción coercitiva y coacción que facilita la manipulación en las redes sociales.”

incrementando la confusión disminuyendo la capacidad de manejo de controversias, dinamizando las expresiones de descontento individual, lo cual puede ser un factor determinante para lograr el efecto exponencial deseado.

Esto es simple, se trata de mantener a la gente enojada y confundida. La ira es una emoción fácilmente manipulable y la confusión dificulta tomar decisiones. Este modelo funciona para incentivar la acción, o para inhibirla, no solo se trata de propaganda para invitar a realizar una determinada acción, también sirve para promover la idea de no hacer o dejar de hacer una acción dada.

Las interacciones predominantes, para amplificar la manipulación y en consecuencia la división social (por ejemplo: cuestiones raciales, creencias religiosas, clases sociales, división económica, movimientos migratorios, entro otros) sirven para construir las narrativas dirigidas a una audiencia llena de confusión, pero lista a recibir su contenido sin ser cuestionado, aprovechando el clima social, no con el fin de favorecer a uno u otro actor político o social. Como se mencionó anteriormente, el método pasa por las actividades de desinformación, causar confusión y reducir la confiabilidad en los medios de comunicación tradicionales. La duda promovida en los medios tradicionales potenciará el manejo de las redes sociales.

Sacar de contexto datos, recurrir a fuentes dudosas, y además de todo, el catálogo de falacias argumentativas, forman parte del modelo de interacción coercitiva y coacción que facilita la manipulación en las redes sociales. Al respecto, dos palabras simples pero poderosas describen la forma en que se logra el impacto: Coerción y Coacción, como parte además de una campaña de desinformación.

Coacción: Imponer una forma de presión verbal, en forma casi siempre de amenazas veladas o explicitas, a quien difiere en posturas o ideas. Es llevar a alguien a tomar una postura, mediante presión y amenazas, en respuesta a un libre ejercicio de expresión de opinión o ideas.

Coerción: La coerción digital es una forma de censura. Viene en forma de ataques infundados y reiterados, descalificaciones, groserías, violencia. No atiende ninguna forma de razonamiento. Es linchamiento digital.

Conclusiones

Por lo anterior, las redes sociales se han convertido en el escenario donde se busca influir en el prestigio, imagen pública y credibilidad de actores políticos, líderes sociales, sindicales, instituciones y tomadores de decisión, combinando el uso de la presión social y la manipulación con narrativas convenientes para este propósito. De ahí que, el entorno social digital en las redes sociales provee del campo de acción y materia para trabajar con gran precisión los objetivos que interesen a un poder político, social o fáctico determinado. La ciencia de datos (Big Data) permite analizar las interacciones de usuarios de redes sociales logrando la descripción detallada de un perfil altamente georreferenciado de cada uno de estos[10].

Dado que, la mayoría de las conversaciones en línea se han trasladado a las plataformas de redes sociales, la cantidad de Ataques de Denegación Distribuida de Servicio (DDoS, por sus siglas en inglés) tradicionales ha disminuido. En la actualidad, es difícil montar un ataque DDoS exitoso en Facebook o Google debido a la gran cantidad de tráfico ordinario. Como resultado, ha surgido una nueva forma de ataques DDoS, mediante los ataques coordinados que se realizan en páginas dentro de las plataformas de redes sociales utilizando cuentas manuales o automáticas que informan falsamente las cuentas o publican activaciones de las plataformas para eliminar contenido o prohibir cuentas por moderación. Aunque el efecto suele ser temporal, esta nueva forma de ataque dirigido logra el mismo propósito que los ataques DDoS tradicionales.

[1] Para entender las interacciones de usuarios y su impacto con la manipulación de las redes sociales en necesario considerar lo siguiente:
En primer lugar, las redes sociales aumentan la velocidad y la difusión de la información, por lo que deben entenderse que las narrativas e incentivos subyacentes que se emplean en este espacio digital no son estáticos y responden a estímulos sin precedentes, con el potencial de generar nuevas emociones e ideas, permitiendo un mapeo de las tendencias que se generan en un entorno geográfico determinado.
En segundo lugar, lo primordial es centrarse en la interacción de usuarios de las redes sociales y en la capacidad de adaptación de sus narrativas mediante la asociación de bloques de datos, la segmentación de grupos comunes, identificando sus narrativas y los efectos en la comunicación donde se dinamizan los riesgos cognitivos.
Finalmente, las redes sociales generan nuevos datos e información crucial, es decir, ¿qué tan popular es un Tweet en particular, ¿cuántas veces se comparte una publicación en particular, ¿cuántas visitas a la página o cuántos seguidores tiene una cuenta en particular? Estos datos no solo proporcionan nueva información, sino que el suministro de esta información altera fundamentalmente el entorno estratégico.

Este tipo de ataque, se usa para restringir el trabajo de activistas de derechos humanos y civiles, periodistas independientes, etc., atacando sus cuentas de redes sociales para asegurarles que dejen de usar las plataformas.

La suma de todos los elementos antes citados, facilitará la movilización de la población; en muchas ocasiones, la clave del éxito de los Conflictos en la Zona Gris será esa implicación, la cual se sustanciará a través de los mecanismos de protesta que podríamos calificar de ordinarios y perfectamente legales como los son las manifestaciones o huelgas en forma pacífica, y en mayor magnitud la toma de instalaciones, edificios de gobierno o plazas públicas e inclusive de la infraestructura crítica.

Con esta evolución de los Conflictos en la Zona Gris, mediante la manipulación de las redes sociales se propician las condiciones para generar una escalada en los eventos, desde estrategias no violentas de movilización social hasta la resistencia civil en función del carácter más o menos proactivo o reactivo, dentro de la propia Zona Gris y del tipo de narrativas que impulsen a la acción o motiven a la inacción, así como del tipo de narrativa empleada en cada caso.

A consecuencia de lo anterior, la combinación de herramientas empleadas en la Zona Gris pueden llegar a provocar el colapso de un Estado y de sus instituciones sin necesidad de emplear la violencia de forma directa por parte de otro Estado u organización antagonista, buscando la sobrerreacción de la sociedad por medio de diversas presiones mediáticas, a fin de deslegitimizarlo ante la opinión pública local en internacional, o incluso, que busque justificar, mediante argumentos jurídicos, un incremento de medidas para intentar reestablecer el orden y paz social.

Este documento documento se distribuye de forma gratuita sin fines de lucro y ha sido elaborado por el personal de:
Stratop Risk Consulting®
Para más información, otros materiales, consultas, productos o posibles ideas de análisis de factores de riesgo asociados, por favor contáctenos en:

stratoprisk.com contact@stratoprisk.com @StratopRisk StratopRisk STRATOP RISK CONSULTING