URVIO, Revista Latinoamericana de Estudios de Seguridad
ISSN: 1390-3691
ISSN: 1390-4299
revistaurvio@flacso.edu.ec
Facultad Latinoamericana de Ciencias Sociales
Ecuador

Ruiz-Ruano, Ana-María; López-Puga, Jorge; Delgado-Morán, Juan-Jose
El componente social de la amenaza híbrida y su detección con modelos bayesianos
URVIO, Revista Latinoamericana de Estudios de Seguridad, núm. 25, 2019, Julio-, pp. 57-69
Facultad Latinoamericana de Ciencias Sociales
Ecuador

DOI: https://doi.org/10.17141/urvio.25.2019.3997

Disponible en: https://www.redalyc.org/articulo.oa?id=552661588004

doi.org/10.17141/urvio.25.2019.3997

URVIO, Revista Latinoamericana de Estudios de Seguridad, No. 25, diciembre 2019-mayo 2020
pp. 57-69 • RELASEDOR y FLACSO Sede Ecuador • ISSN 1390-4299 (en línea) y 1390-3691

El componente social de la amenaza híbrida y su detección con modelos bayesianos

The Social Component of the Hybrid Threat and its Detection with Bayesian Models

Ana-María Ruiz-Ruano[1], Jorge López-Puga[2]
y Juan-Jose Delgado-Morán[3]

Recibido: 2 de junio de 2019
Aceptado: 31 de julio de 2019
Publicado: 2 de diciembre de 2019

Resumen

Las sociedades contemporáneas están cada vez más condicionadas por el desarrollo de la tecnología informática. Esa tendencia deja entrever un panorama en el que cada ser humano se identifica por el binomio persona-computadora, mientras que la mayor informatización de la vida civil está generando ingentes cantidades de datos que son susceptibles de ser gestionados con fines bélicos. El objetivo de este artículo es abordar la utilidad potencial de las redes bayesianas como herramientas destinadas a la monitorización y detección temprana de ataques híbridos de carácter social a escala global. Como conclusión, planteamos que el uso de la inferencia y las redes bayesianas es útil para monitorear, detectar y supervisar el componente social de las amenazas híbridas a escala global por medio del análisis de las redes sociales.

Palabras clave: inferencia estadística; inteligencia artificial; redes informáticas; redes sociales; seguridad de los datos

Abstract

Contemporary societies are increasingly conditioned by the development of computer technology. This trend suggests a picture in which each human being is identified by the person-computer binomial while greater computerization of civil life is generating huge amounts of data that are likely to be managed for war purposes. The objective of this article is to address the potential utility of Bayesian networks aimed at monitoring and early detection of hybrid attacks of a global nature. We conclude that the use of inference and Bayesian networks is useful for monitoring, detection and supervision of the social component of hybrid threats globally through social network analysis.

Keywords: artificial intelligence; computer networks; data protection; social networks; statistical inference

1 Universidad Católica de Murcia, España, amruiz@ucam.edu, id/0000-0002-7260-0588

2 Universidad Católica de Murcia, España, jpuga@ucam.edu, orcid/0000-0003-0693-0092

3 Universidad Católica de Murcia; miembro de la Cátedra Nebrija-Santander en Gestión de Riesgos y Conflictos y miembro del Grupo Nebrija de Relaciones Internacionales, Cooperación e Integración en Áreas Regionales (RICINTAR) de la Universidad Antonio de Nebrija, España, jjdelgado@ucam.edu, orcid/0000-0002-9945-8235

Introducción

Es un hecho constatado que el mundo contemporáneo está ampliamente interconectado. El desarrollo que han experimentado las telecomunicaciones a lo largo del último siglo ha facilitado que sea cuestión de segundos ponerse en contacto con una persona que se encuentra en las antípodas. Podemos transferir texto, imágenes, sonido y video como nunca fue posible. La tecnología 5G plantea muchas más posibilidades para el presente, así como para el futuro.

Todo ese desarrollo científico-técnico ha tenido consecuencias positivas para la economía y para el desarrollo humano en general. Sin embargo, la sofisticación de las redes informáticas, así como las potencialidades que nos ofrecen coexistiría con la posibilidad de ser atacadas por agentes que pretenden desestabilizarlas. Este asunto ha sido, y sigue siendo, objeto de profundo análisis por parte de las ciencias de la computación. Sin embargo, la estabilidad y seguridad digital se enfrenta a nuevos retos.

Entre los desafíos actuales de la humanidad podríamos destacar aquellos relacionados con las amenazas **híbrida**s. Este artículo tiene como objetivo principal aproximarnos a la caracterización conceptual de lo que podríamos llamar componente social de la amenaza híbrida. Se trata de un componente mediado principalmente por el aspecto social que caracteriza a la organización humana. Por tanto, se presenta una conceptualización y se analiza **cómo puede ser** monitoreado usando la inferencia bayesiana.

La primera parte del artículo aborda la conceptualización del componente social de la amenaza híbrida utilizando una metodología basada en la revisión bibliográfica. En la segunda parte, y como objetivo secundario, presentamos una introducción a lo que denominamos inferencia bayesiana comparándola, básicamente, con la inferencia estadística clásica. La tercera parte refleja, a través de un ejemplo ilustrativo y limitado en el número de variables, el uso potencial de una red bayesiana para modelar un caso concreto de amenaza híbrida. Nos proponemos mostrar la lógica que subyace a este tipo de herramientas estadísticas en el contexto de la seguridad frente a amenazas híbridas. Por último, presentamos nuestras conclusiones y los derroteros que debería tomar el trabajo de prevención o detección del componente social de las amenazas híbridas.

El componente social de la amenaza híbrida

No existe una definición amplia y universalmente aceptada de amenaza híbrida (Ducaru 2016). Siguiendo a Colom (2019), podemos indicar que el término "guerra híbrida" fue utilizado por primera vez en un documento oficial producido por Estados Unidos por el año 2005. En cualquier caso, existe un conjunto de elementos hostiles que pueden asociarse con los ataques híbridos. Por ejemplo, las estrategias militares o no militares que están destinadas a desestabilizar organizaciones sociales legítimamente estructuradas de manera deliberada y sincronizada pueden considerarse ataques híbridos. Los ataques o intentos de desestabilización tienen como objetivo obtener influencia política, social o económica sobre la organización social que está siendo atacada (p. e., Ducaru 2016; Hoffman 2009; Lanoszka 2016). Este tipo de ataques híbridos, que inicialmente pueden no tener un

URVIO, Revista Latinoamericana de Estudios de Seguridad No. 25 • diciembre 2019-mayo 2020 • pp. 57-69
ISSN 1390-4299 (en línea) y 1390-3691

marcado o claro componente militar, siempre han existido y la historia está plagada de ellos. El "caballo de Troya" y la Guerra Fría son claros ejemplos (uno clásico y otro sostenido en el tiempo) de ataques híbridos.

El interés particular de este artículo es un aspecto concreto de los ataques híbridos contemporáneos: el componente social-virtual. Lo que denominamos componente social del ataque híbrido contemporáneo está asociado con una idea de seguridad informática que no se circunscribe a la integridad de la información electrónica (von Solms y van Niekerk 2013). Más bien está relacionado con la veracidad de la información difundida en la red de redes con el ánimo de auspiciar, incrementar o desarrollar ataques híbridos.

Como es bien sabido, gran parte del mundo está conectado por medio de dispositivos informáticos. Cada vez es más frecuente que las personas dispongan de un móvil inteligente que es, en definitiva, una computadora. De hecho, esas "minicomputadoras personales" son claramente más potentes de lo que lo fueron sus ancestros tecnológicos hace solo dos o tres décadas. Una gran proporción de la población que reside en lo que vulgarmente se denomina "mundo desarrollado" dispone de uno o más dispositivos móviles o portátiles que superan con creces la capacidad de cómputo de la que se disponía domésticamente hace tan solo unas décadas. Según predice la Ley de Moore, la tendencia será la misma en el futuro, máquinas cada vez más baratas y más potentes.

Las computadoras actuales no son solo más potentes y tienen mayor capacidad de almacenar información, sino que cada vez están más interconectadas. Conceptos como "el internet de las cosas" o el "coche autónomo" serán pronto una realidad, según los medios de comunicación de masas y las revistas científicas especializadas. Esa conexión globalizada o conectividad globalizadora está ideada, al menos en teoría, para mejorar las vidas de las personas, pero también surgen ciertos problemas éticos que la humanidad tendrá que afrontar. Por ejemplo, ¿quién será responsable de un accidente que se produzca vinculado a la actividad de un coche autónomo?

En ese caldo de cultivo y desarrollo tecnológico, la cantidad de información que se va a producir parece abrumadora. Si las estimaciones son correctas, en pocos años la información electrónica que existirá, en términos de bytes, superará al número de estrellas que existen en el universo conocido (Butler 2016). Por ello, han sido acuñados conceptos como el de "Big Data", cuyo desarrollo es espectacular desde el punto de vista académico o científico, para poder acomodarnos a la realidad computacional que se nos avecina (Cloud Security Alliance 2012).

Todo ello plantea ciertos desafíos a las sociedades democráticas legítimamente constituidas en aras de preservar el orden social y político, en relación con posibles ataques híbridos de naturaleza informática (Lafuente 2015). Como indican algunos autores, este tipo de delitos u otros de carácter informático tenderán a ser cada vez más frecuentes, más sofisticados y más destructivos (Taddeo y Floridi 2018). En tal contexto se ubica lo que hemos denominado "componente social del ataque híbrido". Un componente que podría vincularse a lo que Rebolloso (1994) denomina clásicamente comportamiento colectivo o conducta de masas.

Es bien sabido que el comportamiento colectivo o de masas favorece la consecución de objetivos sociales deseables y positivos. Sin embargo, las masas o ciertos movimientos colectivos pueden llevar a cabo comporta-

URVIO, Revista Latinoamericana de Estudios de Seguridad No. 25 • diciembre 2019-mayo 2020 • pp. 57-69
ISSN 1390-4299 (en línea) y 1390-3691

mientos destructivos y lesivos para la propia sociedad o estructura social (Lilienfeld, Lynn, Namy y Woolf 2011). Independientemente de si los fines perseguidos son legítimos o buscan la defensa de valores positivos, la organización social tumultuosa y los amotinamientos suelen provocar consecuencias deletéreas para la propia sociedad. En el ámbito de los ataques híbridos, este tipo de comportamientos colectivos se dan cada vez más en la red, a través de, por ejemplo, sistemas de mensajería instantánea. Merecería la pena preguntarse si en esas situaciones se pone en peligro o en riesgo la seguridad ciudadana o la seguridad de la estructura social establecida.

En el ámbito de las ciencias de la computación se han desarrollado herramientas que pueden ser potencialmente utilizadas para predecir el componente social de los ataques híbridos. Por ejemplo, los modelos estocásticos altamente estructurados, como los sistemas expertos probabilísticos (Cowell et al. 1999, 37). Este tipo de modelos han surgido en el seno de la inteligencia artificial (IA) y se están postulando como herramientas eficientes para gestionar iniciativas destinadas a vulnerar la seguridad de los Estados, naciones, democracias o estructuras sociales legítimamente establecidas (Anwar y Hassan 2017; Ruiz-Ruano y Puga 2018).

Los métodos bayesianos pueden considerarse herramientas apropiadas para tomar decisiones en situaciones de incertidumbre, teniendo en cuenta el conocimiento borroso que tenemos de las variables implicadas en el problema de decisión (Edwards y Fasolo 2001). Por ello, planteamos que este tipo de herramientas estadísticas podrían ser ideales para desarrollar sistemas de monitorización del componente social en las amenazas híbridas. De hecho, la utilización de técnicas matemático-estadísticas no es nueva y algunas organizaciones gubernamentales como la CIA (Central Intelligence Agency) de los Estados Unidos de América se han apoyado en ellas para gestionar conflictos bélicos o relacionados con la seguridad (CIA 1968; Das 1999; Fisk 1994; Somiedo 2018). En cualquier caso, pese a que en este artículo defendemos que el uso de herramientas computacionales es esencial para detectar el componente social de la amenaza híbrida, tenemos que tener en cuenta sus limitaciones (Castelvecchi 2019).

En el epígrafe siguiente presentamos una escueta descripción de algunos de los elementos de los métodos bayesianos que creemos útiles para detectar el componente social de la amenaza híbrida. Presentamos nuestras conclusiones respecto a la utilidad y viabilidad de este tipo de herramientas para detectar y monitorizar amenazas de seguridad a gran escala.

Inferencia bayesiana

Los métodos de inferencia bayesiana son usualmente contrastados o confrontados con la inferencia clásica o frecuentista (Alonso y Tubau 2002; Cowell et al. 1999; O'Hagan y Luce 2003; Serrano 2003). Se considera que la estadística bayesiana tiene su origen en un trabajo atribuido al reverendo Thomas Bayes (1763), publicado a título póstumo. En esta sección presentaremos algunas de las características más sobresalientes de la estadística bayesiana, siguiendo algunos puntos de O'Hagan y Luce (2003).

En primer lugar, los métodos bayesianos o incluso los híbridos-bayesianos asumen que existe un componente subjetivo relevante en los problemas que han de ser resueltos. Desde esas ópticas se asume como normal la subjeti-

URVIO, Revista Latinoamericana de Estudios de Seguridad No. 25 • diciembre 2019-mayo 2020 • pp. 57-69
ISSN 1390-4299 (en línea) y 1390-3691

vidad y se trata de modelarla estadísticamente. Dicho de otro modo, se considera que existen incertidumbres en los planteamientos de los problemas que han de resolverse y ello es modelado estadísticamente por medio de, por ejemplo, distribuciones de probabilidad. Hacer esto supone una perspectiva humilde frente a la resolución de cualquier problema, dado que se está asumiendo que existe conocimiento incompleto sobre el aspecto de la naturaleza que se está estudiando. El conocimiento incierto sobre la realidad de un problema es, entonces, modelado o incluido en este, utilizando el conocimiento subjetivo que se tiene. Por ello, los métodos bayesianos se caracterizan por ser técnicas que asumen cierta subjetividad en la resolución de los problemas. Es más, se suele decir que asumen que la probabilidad en sí misma es un elemento más subjetivo que objetivo. Así, en vez de considerar que la probabilidad de ocurrencia de un fenómeno es algo real que existe como tal, se considera que las estimaciones de probabilidad son producto de la naturaleza cognitiva humana. Es decir, que las estimaciones de probabilidad son, en cualquier caso, subjetivas.

Como señala Dixon (1970), la probabilidad de ocurrencia de un fenómeno no sería una propiedad que pertenece al sistema o al evento observado, sino que, más bien, sería una propiedad o una característica que depende del observador del sistema. Creemos conveniente hacer aquí una sublime apreciación que ya sugería Jeffreys (1931), quien ha sido considerado uno de los máximos exponentes de la inferencia bayesiana. El hecho de que las estimaciones de probabilidad puedan ser consideradas algo subjetivo no implicaría necesariamente que la mente humana sea la que crea la realidad (Berger y Luckmann 1968). Más bien, se diría desde un enfoque bayesiano, que se acepta la idea de que habría ciertas estimaciones de probabilidad que encajan mejor o peor con los datos experimentales. Por tanto, la misión de las personas que resuelven problemas de probabilidad consiste en ir acomodando creencias y experiencias con base en los datos empíricos u observados.

Podríamos decir que los métodos bayesianos combinan información previa (denominada en el contexto estadístico "distribuciones *a priori*") con información observada (también denominada datos o distribución de verosimilitud), para producir distribuciones posteriores o actualizadas (también denominadas "distribuciones *a posteriori*") del problema estudiado. Sobre las distribuciones posteriores se suelen llevar a cabo procesos de inferencia estudiando, por ejemplo, datos estadísticos de tendencia central (media, mediana o moda, por ejemplo) o identificando intervalos de credibilidad bayesianos.

Los métodos bayesianos son técnicas que se asemejan mucho al razonamiento natural humano (Anscombe 1961; Bolstad 2007; Puga, Krzywinski y Altman 2015). Por ejemplo, en el ámbito de estudio de la psicología y la neurociencia, estudios de principios de este siglo sugieren que el razonamiento humano, principalmente el causal, se asemeja mucho a lo que cabría esperar si se estuviese utilizando inferencia bayesiana (véanse, por ejemplo, los trabajos de Glymour 2001, 2003; Gopnik et al. 2004, 2001, experimento 3; Gopnik y Schulz 2004; Sobel, Tenenbaum y Gopnik 2004, experimento 3). Dicho de otro modo: cuando observamos el aprendizaje humano y tratamos de modelarlo formalmente, se hacen predicciones que son muy congruentes con lo que postula la estadística bayesiana.

Mientras que la estadística frecuentista se centra, desde un punto de vista probabilísti-

URVIO, Revista Latinoamericana de Estudios de Seguridad No. 25 • diciembre 2019-mayo 2020 • pp. 57-69
ISSN 1390-4299 (en línea) y 1390-3691

co, en estimar la probabilidad de que unos datos muestrales (*D*) provengan de cierta distribución hipotética (*H*), es decir, *P*(*D*|*H*), la estadística bayesiana está más interesada en conocer la credibilidad o verosimilitud de las hipótesis planteadas en función de los datos empíricos observados, *P*(*H*|*D*). De ese modo, usando el teorema de Bayes y considerando que los parámetros poblacionales son entes que se distribuyen aleatoriamente (no como la estadística frecuentista o clásica, que considera a estos como entes fijos), los métodos bayesianos permiten actualizar los modelos probabilísticos a medida que se van recogiendo u observando datos muestrales frente a la resolución de problemas.

El Factor de Bayes o *FB* está recibiendo un considerable interés en los últimos tiempos en un amplio abanico de disciplinas científicas tanto teóricas como aplicadas (Held y Ott 2018; Hoijtink, van Kooten y Hulsker 2016; Jarosz y Wiley 2014; Jeon y De Boeck 2017; Morey y Rouder 2011; Morey, Wagenmakers y Rouder 2016). Es un estadístico que estima la medida en que una hipótesis es más probable que otra, teniendo en cuenta los datos muestrales o las evidencias disponibles (Kass y Raftery 1995). Matemáticamente, el Factor de Bayes (FB_{AB}) resulta de dividir la probabilidad de que ocurra una hipótesis (pongamos, *A*) entre la probabilidad de que ocurra otra (digamos, *B*), teniendo en cuenta cuán probables son a la luz de los resultados observados empíricamente. Dado que este factor es una fracción o proporción, puede ser interpretado de forma comparativa atendiendo a la verosimilitud de una hipótesis respecto a otra, habiéndolas condicionado a los datos observados. Por tanto, cuando el Factor de Bayes es igual a uno, podríamos decir que ambas hipótesis son igualmente probables a la luz de los datos muestrales observados. Por su parte, cuando $FB_{AB} > 1$ diríamos que la hipótesis *A* es más probable que la hipótesis *B,* teniendo en cuenta los datos observados empíricamente. La magnitud de ese estadístico indicaría cuánto más probable es la hipótesis *A* frente a la *B*. Por ejemplo, un factor de Bayes igual a 20 indicaría que la hipótesis *A* es 20 veces más probable que la *B*. Por el contrario, y de manera análoga, cuando $FB_{AB} < 1$, la conclusión a la que podemos llegar es que la hipótesis más probable teniendo en cuenta los datos observados es la *B*.

Los Factores de Bayes son estadísticos muy útiles en la toma de decisiones porque alumbran o clarifican los caminos a seguir frente a situaciones de incertidumbre. Dado que su naturaleza es comparativa y confrontan la verosimilitud de unas hipótesis frente a otras, pueden ser utilizados para dar más o menos relevancia a las creencias que tenemos sobre las soluciones hipotéticas que planteamos ante los problemas. Aunque existen varias formas de interpretar los Factores de Bayes (Kass y Raftery 1995), vamos a presentar, con propósitos ilustrativos, los puntos de cortes que propuso Jeffreys (1948). Según este autor, existe una evidencia anecdótica en los datos en favor de una hipótesis frente a la otra cuando el Factor de Bayes está comprendido entre 1 y 3. Si, por su parte, la fracción de verosimilitud se encuentra entre 3 y 10, se podría considerar que la evidencia es substancial. Si el valor del factor está entre 10 y 30, diríamos que la evidencia registrada frente a la primera de las hipótesis es fuerte. Por último, si el factor está comprendido entre 30 y 100, diríamos que la evidencia empírica que apoya a una hipótesis frente a la otra es muy fuerte. Cuando estuviese por encima de 100, podríamos decir que es decisiva.

URVIO, Revista Latinoamericana de Estudios de Seguridad No. 25 • diciembre 2019-mayo 2020 • pp. 57-69
ISSN 1390-4299 (en línea) y 1390-3691

Un ejemplo de la aplicación de métodos bayesianos

En esta sección vamos a presentar un ejemplo muy sencillo que ilustra cómo podrían utilizarse los métodos bayesianos para detectar o monitorizar el componente social de las amenazas híbridas. Utilizamos un modelo de red bayesiana porque permite ilustrar fácilmente cómo se puede organizar la información existente sobre un problema y cómo se puede utilizar de cara a su resolución estratégica. Estas herramientas han sido utilizadas anteriormente tanto con fines militares como en el ámbito de la detección inteligente de delitos (Das 1999; Garbolino y Taroni 2002; Oatley y Ewart 2003). Un ejemplo mucho más profundo y elaborado que este es el modelo de Análisis de Riesgos Adversarios (Ríos, Ríos y Banks 2012), sobre cómo los sistemas expertos probabilísticos y los diagramas de influencia pueden ser utilizados para la modelización de toma de decisiones bajo incertidumbre. Ese modelo, que podría perfectamente acomodarse al estudio de la amenaza híbrida, representa un paradigma que combina simultáneamente la teoría de juegos y el análisis de riesgos, frente a la toma de decisiones bajo incertidumbre.

Las redes bayesianas son modelos multivariados que permiten representar tanto la dimensión cuantitativa como cualitativa de un problema (Cowell et al. 1999). La dimensión cualitativa es lo que se conoce como Gráfico Dirigido Acíclico o GDA. Esto es, una estructura gráfica que conecta variables (representadas por nodos o vértices) por medio de aristas dirigidas (representadas por flechas). Como su nombre indica, los ciclos o *loops* no están permitidos en los grafos de una red bayesiana. Las flechas que conectan las variables tienen sentido estadístico dado que representan o indican que una variable depende de otra o que una variable está influenciada por otra u otras. Por su parte, la dimensión cuantitativa del modelo quedaría determinada por un conjunto de funciones de probabilidad condicional que especifican las relaciones probabilísticas definidas por los enlaces presentes en el grafo. Tanto las características gráficas del modelo como las funciones de probabilidad son usadas por el teorema de Bayes para actualizar la probabilidad de ocurrencia de los eventos del modelo y, por tanto, para tomar decisiones respecto a la resolución de problemas.

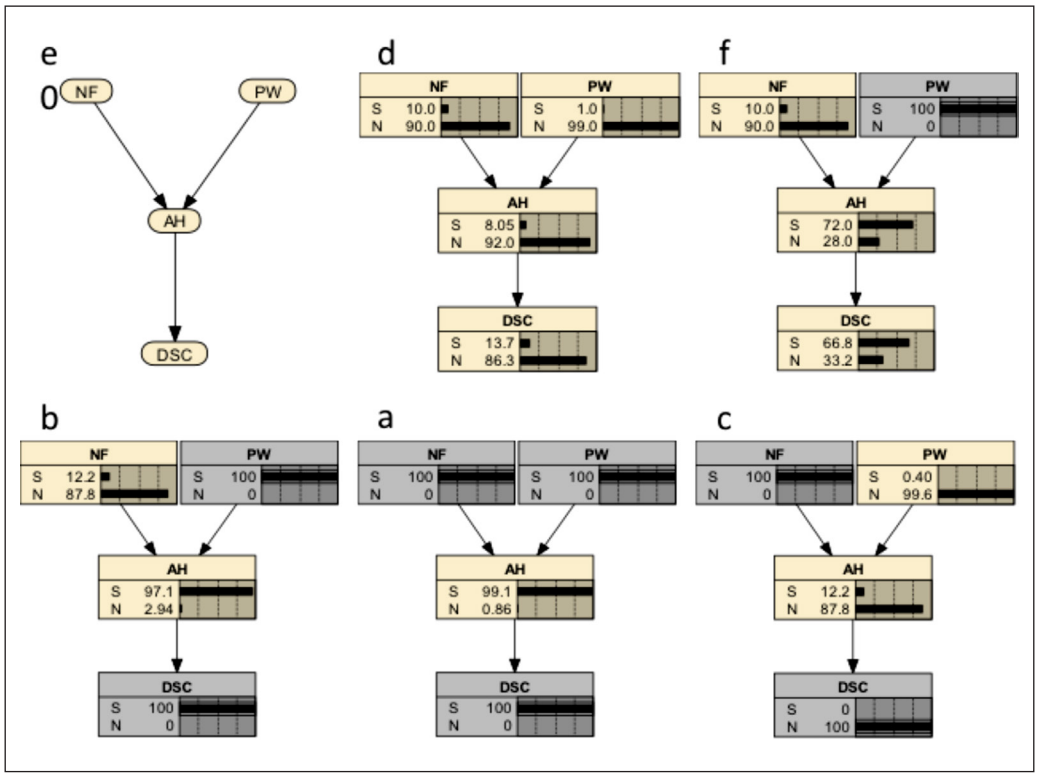
Supongamos que estamos monitoreando la aparición del componente social de la amenaza híbrida. Para referirnos al evento "se está perpetrando algún tipo de hostilidad vinculada con el componente social de la amenaza híbrida" vamos a utilizar las letras *AH*. Es decir, vamos a considerar que la amenaza híbrida es una variable que puede estar presente con cierta probabilidad. Si atendemos a los planteamientos de Lanoszka (2016), podríamos identificar algunas señales que se vinculan con el inicio de la ocurrencia de amenazas híbridas. Por ejemplo, la presencia de noticias falsas que pueden radicalizar o polarizar a la sociedad, el asalto de páginas webs oficiales o institucionales y la debilitación de la sociedad civil son eventos que se han asociado tradicionalmente con estadios iniciales de ataques híbridos (Grinberg et al. 2019). Por tanto, vamos a considerar que todos esos eventos son variables estadísticas y las vamos a modelar como variables dicotómicas que asumirán dos valores posibles: "se está produciendo el evento en cuestión" y "no se está produciendo el evento en cuestión". Además, vamos a representar las noticias falsas con las letras *NF*, el asalto o pirateo a páginas webs con las letras *PW* y la debilitación de la sociedad civil con la expresión *DSC*.

URVIO, Revista Latinoamericana de Estudios de Seguridad No. 25 • diciembre 2019-mayo 2020 • pp. 57-69
ISSN 1390-4299 (en línea) y 1390-3691

Para ilustrar cómo se podría modelar este problema vamos a utilizar el software Netica en su versión 6.05. Aunque Netica es un software comercial, la compañía proporciona el acceso a una versión demo totalmente funcional que está limitada en el número de variables (López 2012). Como se puede apreciar en el panel *a* de la figura 1, un gráfico hipotético que podría ilustrar la relación entre todas esas variables podría ser aquel que especifica que la amenaza híbrida dependería de la presencia de noticias falsas y del pirateo de páginas webs. El grafo del panel *a* también representa que la debilitación de la sociedad civil dependería de la amenaza híbrida. Este grafo, por tanto, ilustra una situación ficticia e hipotética que solo tiene sentido con fines ilustrativos, en la que la sospecha de un ataque híbrido inminente estaría condicionada por la proliferación de noticias falsas y por el ataque sistemático de webs oficiales o gubernamentales. Por su parte, el gráfico modela la situación que indicaría que la amenaza híbrida genera cierto grado de debilitación cívica en la sociedad objeto de un ataque híbrido.

Vamos a considerar cada una de las variables mencionadas, que se relacionan con la amenaza híbrida, como variables dicotómicas

Figura 1. Representación de la estructura y el funcionamiento de una red bayesiana



Leyenda: AH: amenaza híbrida, NF: noticias falsas, PW: pirateo web, DSC: debilitación de la sociedad civil, S: sí, N: no. Las probabilidades están expresadas en términos porcentuales.

Fuente: elaboración propia.

en las que serán posibles valores afirmativos (*S*) o negativos (*N*). Además, consideraremos que cada una de estas variables está definida paramétricamente por una función de probabilidad condicional en función de las variables de las que depende. De esta manera, el modelo que hemos presentado en el panel *a* de la figura 1 quedaría especificado o definido por ocho parámetros u ocho estimaciones de probabilidad.

Supondremos que las estimaciones de probabilidad (expresadas en forma porcentual) son fruto del análisis exhaustivo que se ha llevado a cabo en una agencia de inteligencia gubernamental. Según este, la probabilidad basal de ataques piratas contra webs oficiales es del 1 % y la probabilidad de difusión de noticias falsas es de un 10 %.

Estos parámetros podrían tener sentido considerando que es más fácil que se difunda un bulo falso que el asalto organizado a una web oficial o gubernamental. Las dos variables no dependen de ninguna otra (técnicamente se dice que son variables madre). No obstante, la variable que representa la amenaza híbrida sí depende de las dos anteriores. Por ello, quedaría definida por cuatro parámetros de probabilidad condicional, que se relacionarían con el resultado de combinar cartesianamente las dos variables previas. Una posible parametrización para esta variable podría ser la siguiente, expresada en notación matemática:

$P(\text{AH} = \text{S} \mid \text{NF} = \text{S}, \text{PW} = \text{S}) = 90\ \%$,
$P(\text{AH} = \text{S} \mid \text{NF} = \text{N}, \text{PW} = \text{N}) = 5\ \%$,
$P(\text{AH} = \text{S} \mid \text{NF} = \text{S}, \text{PW} = \text{N}) = 56\ \%$ y
$P(\text{AH} = \text{S} \mid \text{NF} = \text{N}, \text{PW} = \text{S}) = 70\ \%$.

Así, según lo expresado en a), diríamos que la probabilidad de que se estén produciendo amenazas híbridas dado que se han producido noticias falsas y ataques webs sería del 90 %. Por su parte, si nos fijamos en b), diríamos que la probabilidad de que se esté produciendo una amenaza híbrida en ausencia de noticias falsas y de ataques de webs oficiales o gubernamentales sería del 5 %. Los parámetros c) y d) se interpretarían análogamente.

Los parámetros asociados con la variable DSC tendrían que estar expresados en términos de los valores de la variable AH. Supongamos que sus parámetros expresados en notación probabilística son los siguientes: $P(\text{DSC} = \text{S} \mid \text{AH} = \text{S}) = 90\ \%$ y $P(\text{DSC} = \text{S} \mid \text{AH} = \text{N}) = 7\ \%$. O expresado en palabras, que la probabilidad de que la sociedad civil se debilite cuando se ha producido una amenaza híbrida es del 90 %, mientras que la probabilidad de que esta misma sociedad se debilite es del 7 % cuando la amenaza híbrida no ha tenido lugar. Los valores complementarios de todas esas probabilidades se obtienen como la diferencia respecto a la unidad. Así, por ejemplo, $P(\text{AH} = \text{N} \mid \text{NF} = \text{S}, \text{PW} = \text{S}) = 1 - P(\text{AH} = \text{S} \mid \text{NF} = \text{S}, \text{PW} = \text{S}) = 10\ \%$. Sobre todos estos valores de probabilidad son sobre los que el teorema de Bayes opera para hacer estimaciones del estado de cada una de las variables, teniendo en cuenta observaciones o evidencias sobre la situación evaluada.

Cuando todos los valores de probabilidad son tenidos en cuenta por la red bayesiana, se genera lo que se conoce como distribución previa del modelo. Esta aparece reflejada en el panel *b* de la figura 1. Representa las probabilidades vinculadas a cada uno de los estados de la variable cuando no se tiene información sobre la situación o cuando no se ha observado ningún dato sobre el problema. Por ejemplo, en el panel *b* se puede apreciar que, en principio, la probabilidad de ataque híbrido es relativamente pequeña, del orden del 8,5 %. Sin embargo, con este tipo de modelos podemos valorar el impacto que tendría la observación de alguno o algunos de los estados de estas

URVIO, Revista Latinoamericana de Estudios de Seguridad No. 25 • diciembre 2019-mayo 2020 • pp. 57-69
ISSN 1390-4299 (en línea) y 1390-3691

variables. De esa manera se puede evaluar el impacto que tendrían diferentes escenarios sobre la probabilidad de ocurrencia de un ataque híbrido y podrían llevarse a cabo acciones destinadas a minimizarlo. Por ejemplo, imaginemos que supiésemos que se ha perpetrado un ataque pirata a una web del gobierno. Como se aprecia en el panel *c* de la figura 1, podríamos concluir que la probabilidad (o nuestra creencia) de que se esté produciendo un ataque híbrido aumenta hasta el 72 %. Si seguimos utilizando el teorema de Bayes y nos informan que se ha observado una debilitación de la sociedad civil (panel *d* de la figura 1), nuestra estimación de probabilidad para el ataque híbrido aumentaría hasta el 97,1 %.

Obsérvese que la probabilidad de que se esté produciendo un ataque híbrido aumenta hasta el 99,1 % (panel *e* de la) si, además, se detecta que se han propagado noticias falsas. Por último, tal y como se ilustra en el panel *f*, un escenario en el que se producen amenazas falsas, pero donde la sociedad civil está fortalecida, nos llevaría a pensar que la probabilidad de amenaza híbrida es tan solo del 12,2 %. Este tipo de valoración de escenarios es de vital utilidad en contextos en los que se han de llevar a cabo acciones destinadas a solventar misiones estratégicamente relevantes. Por ello, los modelos bayesianos podrían ser de utilidad para identificar y gestionar los componentes sociales de las amenazas híbridas a escala internacional (Ducaru 2016).

El ejemplo que hemos presentado es muy simple y, probablemente, poco realista, pero se ha concebido para tratar de ilustrar la utilidad de estos modelos estadísticos como aliados en la toma de decisiones estratégicas en situaciones bajo incertidumbre. Además, el modelo es estático en el tiempo y podría no reflejar las verdaderas relaciones que se establecen entre las variables. Afortunadamente, diferentes técnicas surgidas en el seno de la Inteligencia Artificial han sido desarrolladas para identificar y encontrar modelos estadísticos con los datos de los que se dispone. Algunos autores sugieren que esa estrategia de trabajo puede ser fructífera (Anwar y Hassan 2017). Además, existen herramientas informáticas que permiten encontrar las estructuras de red bayesiana más plausibles, a partir de conjuntos de datos (por ejemplo, Heckerman 1995; Ruiz-Ruano 2015; Scutari 2010).

Conclusiones

La humanidad parece aproximarse a una situación en la que el binomio persona-computadora tiende a parecerse a lo que se nos ha presentado repetidamente en las películas o relatos de ciencia-ficción. Da la sensación de que, como sugieren algunos vaticinios quizá no extremos, llegará el momento en el que las computadoras podrán estar conectadas directamente a nuestros cerebros. Quizá esas proyecciones futuristas nunca lleguen a materializarse, pero lo que es cierto es que, cada vez más, las personas parecemos depender más de las computadoras para comunicarnos e interactuar con nuestro entorno social. Independientemente de si valoramos eso como algo positivo o negativo, hay que reconocer que existen ciertos riesgos relacionados con el inicio o desarrollo de ataques híbridos en las democracias legítimamente constituidas.

En este artículo hemos mostrado cómo podrían modelarse estadísticamente ciertos aspectos de la dimensión social de la guerra híbrida. Hemos utilizado una herramienta estadística perteneciente al ámbito de la inferencia bayesiana, las redes bayesianas, y hemos visto

URVIO, Revista Latinoamericana de Estudios de Seguridad No. 25 • diciembre 2019-mayo 2020 • pp. 57-69
ISSN 1390-4299 (en línea) y 1390-3691

cómo sus cómputos pueden ser utilizados para gestionar o enfrentar la toma de decisiones. Los modelos probabilísticos multivariados pueden servir para generar diferentes escenarios que producen distribuciones de probabilidad condicionada a diferentes observaciones. Así, este tipo de modelos bayesianos u otros que pudiesen desarrollarse serían apropiados para monitorear la amenaza híbrida en su vertiente social, dada su habilidad para racionalizar las situaciones de alta incertidumbre.

Los modelos estadísticos serán eficientes en la medida en que las definiciones de las variables estén claramente definidas. Por tanto, dado que la definición de guerra híbrida no es todavía algo universalmente aceptado, podríamos incurrir en errores si tratásemos de diseñar modelos destinados a predecirla. La propuesta que hemos utilizado en este trabajo alude a una definición de guerra híbrida caracterizada por agresiones graduadas, que incrementan en grado de hostilidad. Puede ser un buen punto de partida, pero sería necesario seguir trabajando en este campo. Como sucede en muchas áreas de estudio del ámbito social, hay que tener en cuenta que, así como la guerra híbrida es algo naturalmente dinámico y complejo, la dimensión social de los ataques híbridos no deja de ser un concepto escurridizo y difícilmente manejable desde un punto de vista científico. La agitación social siempre ha existido, existe y existirá; reconocer acciones deliberadas que se encuadren en el marco del concepto de guerra híbrida es una tarea de complejidad abismal. Las máquinas y los algoritmos informáticos que podemos utilizar para detectar fenómenos sociales como el que analizamos en este artículo tampoco están libres de error, de un tipo o de otro, y quizá no sea lo más apropiado confiar en ellas para la resolución de cierto tipo de problemas.

Bibliografía

Alonso, Diego, y Elisabet Tubau. 2002. "Inferencias bayesianas: una revisión". *Anuario de Psicología* 33: 25-47.

Anscombe, Francis John. 1961. "Bayesian statistics". *The American Statistician* 15: 21-24. dx.doi.org/10.2307/2682504

Anwar, Amaan, y Syed Imtiyaz Hassan. 2017. "Applying Artificial Intelligence Techniques to Prevent Cyber Assaults". *International Journal of Computational Intelligence Research* 13: 883-889.

Bayes, Thomas. 1763. "An essay towards solving a problem in the doctrine of chances". *Philosophical Transactions* 53, 370-418. dx.doi.org/10.1098/rstl.1763.0053

Berger, Peter Ludwig, y Thomas Luckmann. 1968. *La construcción social de la realidad*. Buenos Aires: Amorrortu.

Bolstad, William. 2007. *Introduction to Bayesian Statistics*. Hoboken: Wiley.

Butler, Declan. 2016. "A World Where Everyone Has a Robot: Why 2040 Could Blow Your Mind". *Nature* 530: 398-401. dx.doi.org/10.1038/530398a

Castelvecchi, Davide. 2019. "Machine Learning Comes Up Against Unsolvable Problem". *Nature* 565: 277. dx.doi.org/10.1038/d41586-019-00083-3

CIA (Central Intelligence Agency). 1968. "Bayes' theorem in the Korean war". Intelligence Report No. 0605/68, Directorate of Intelligence.

Cloud Security Alliance. 2012. "Top ten big data security and privacy challenges", https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Big_Data_Top_Ten_v1.pdf

Colom, Guillem. 2019. "La amenaza híbrida: mitos, leyendas y realidades". *Instituto Español de Estudios Estratégicos* 24. http://www.ieee.es/Galerias/fichero/docs_opinion/2019/DIEEEO24_2019GUICOL-hibrida.pdf

Cowell, Robert, Philip Dawid, Steffen Lauritzen, y David Spiegelhalter. 1999. *Probabilistic networks and expert systems*. Harrisonburg: Springer.

Das, Balaram. 1999. *Representing uncertainties using bayesian networks*. Australia: Department of Defence/Defence Science and Technology Organization.

Dixon, John. 1970. *Introducción a la probabilidad. Texto programado*. México: Limusa-Wilely.

Ducaru, Sorin Dumitru. 2016. "The Cyber Dimension of Modern Hybrid Warfare and Its Relevance for NATO". *Europolity* 10: 7-23.

Edwards, Ward, y Barbara Fasolo. 2001. "Decision Technology". *Annual Review of Psychology* 52: 581-606.

Fisk, Charles. 1994. "The sino-soviet border dispute: a comparison of the conventional and Bayesian methods for intelligence warning", https://www.cia.gov/library

Garbolino, Paolo, y Franco Taroni. 2002. "Evaluation of Scientific Evidence Using Bayesian Networks". *Forensic Science International* 125: 149-155.

Glymour, Clark. 2001. *The Mind's Arrows. Bayes Nets and Graphical Causal Models in Psychology*. Cambridge: MIT Press.

Glymour, Clark. 2003. "Learning, prediction and causal Bayes nets". *Trends in Cognitives Sciences* 7: 43-48.

Gopnik, Alison, Glymour, Clark, Sobel, David, Schulz, Laura, Kushnir, Tamar, y Danks, David. 2004. "A Theory of Causal Learning in Children: Causal and Bayes Nets". *Psychological Review* 111: 3-32.

Gopnik, Alison, y Laura Schulz. 2004. "Mechanisms of Theory Formation in Young Children". *Trends in Cognitives Sciences* 8: 371-377.

Gopnik, Alison, David Sobel, Laura Schulz, y Clark Glymour. 2001. "Causal Learning Mechanisms in Very Young Children: Two, Three, and Four-Years-Olds Infer Causal Relations from Patterns of Variation and Covariation". *Developmental Psychology* 37: 620-629.

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, y David Lazer. 2019. "Fake News On Twitter During The 2016 U.S. Presidential Election". *Science* 363: 374-378. 10.1126/science.aau2706

Heckerman, David. 1995. *A Tutorial On Learning with Bayesian*. Redmon: Microsoft Research.

Held, Leonhard, y Manuela Ott. 2018. "On P-values and Bayes Factors". *Annual Review of Statistics and its Application* 5: 393-419. dx.doi.org/10.1146/annurev-statistics-031017-100307

Hoffman, Frank. 2009. "Hybrid Warfare and Challenges". *Joint Force Quarterly* 52: 34-39.

Hoijtink, Herbert, Pascal van Kooten, y Hulsker, Koenraad. 2016. "Bayes Factors Have Frequency Properties-This Should Not Be Ignored: A Rejoinder to Morey, Wagenmakers, and Rouder". *Multivariate Behavioral Research* 51: 20-22. 10.1080/00273171.2015.1071705

Jarosz, Andrew, y Jennifer Wiley. 2014. "What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors". *Journal of Problem Solving* 7: 2-9. dx.doi.org/10.7771/1932-6246.1167

Jeffreys, Harold. 1931. *Scientific Inference*. Cambridge: Cambridge University Press.

Jeffreys, Harold. 1948. *Theory of Probability*. Oxford: Oxford University Press.

Jeon, Minjeong, y Paul De Boeck. 2017. "Decision Qualities of Bayes Factor and P Value-Based Hypothesis Testing". *Psychological Methods* 22: 340-360. dx.doi.org/10.1037/met0000140

Kass, Robert, y Adrian Raftery. 1995. "Bayes Factors". *Journal of the American Statistical Association* 90: 773-795. dx.doi.org/10.1080/01621459.1995.10476572

Lafuente, Guillermo. 2015. "The Big Data Security Challenge". *Network Security* 2015: 12-14. 10.1016/S1353-4858(15)70009-7

Lanoszka, Alexander. 2016. "Russian Hybrid Warfare and Extended Deterrence in Eastern Europe". *International Affairs* 92: 175-195.

Lilienfeld, Scott, Steven Jay Lynn, Laura Namy, y Nancy Woolf. 2011. *Psicología. Una introducción*. Madrid: Pearson.

López, Jorge. 2012. "Cómo construir y validar redes bayesianas con Netica". *Revista Electrónica de Metodología Aplicada* 17: 1-17.

Morey, Richard, y Jeffrey Rouder. 2011. "Bayes Factor Approaches for Testing Interval Null Hypothesis". *Psychological Methods* 16: 406-419. dx.doi.org/10.1037/a0024377

Morey, Richard Donald, Eric-Jan Wagenmakers, y Jeffrey Rouder. 2016. "Calibrated Bayes Factors Should Not Be Used: A reply to Hoijtink, van Kooten, and Hulsker". *Multivariate Behavioral Research* 51: 11-19. dx.doi.org/10.1080/00273171.2015.1052710

Oatley, Giles, y Brian Ewart. 2003. "Crimes Analysis Software: 'Pins in Maps', Clustering and Bayes Net Prediction". *Expert Systems with Applications* 25: 569-588.

O'Hagan, Anthony, y Bryan Luce. 2003. *A premier on Bayesian statistics in health economics and outcome research*. Sheffield: MEDTAP International.

Puga, Jorge, Krzywinski, Martin, y Naomi Altman. 2015. "Points of Significance: Bayesian statistics". *Nature Methods* 12: 377-378. dx.doi.org/10.1038/nmeth.3368

Rebolloso, Enrique. 1994. "Conducta colectiva y movimientos colectivos". En *Psicología social,* coordinado por José Francisco Morales, 763-800. Madrid: McGraw Hill.

Ríos, David, Jesús Ríos, y David Banks. 2012. "Adversarial Risk Analysis". *Journal of the American Journal Association* 104: 841-854. dx.doi.org/10.1198/jasa.2009.0155

Ruiz-Ruano, Ana María. 2015. "Aprendizaje estructural de redes bayesianas para modelar el emprendimiento académico de base sostenible y tecnológica". Tesis doctoral, Facultad de Ciencias de la Salud, Universidad Católica San Antonio de Murcia. http://hdl.handle.net/10952/1556

Ruiz-Ruano, Ana María, y Jorge Puga. 2018. "Seguridad informática e inteligencia artificial en la era de la información masiva". En *Conflictos y diplomacia, desarrollo y paz, colaboración y medio ambiente*, dirigido por César Augusto Giner y Juan José Delgado, 711-724. Navarra: Aranzadi.

Scutari, Marco. 2010. "Learning Bayesian Networks with the bnlearn R Package". *Journal of Statistical Software* 35 (3): 1-22. dx.doi.org/10.18637/jss.v035.i03

Serrano, José. 2003. *Iniciación a la estadística bayesiana*. Madrid: Muralla/Hespérides.

Sobel, David, Joshua Tenenbaum, y Alison Gopnik. 2004. "Children's Causal Inferences from Indirect Evidence: Backwards Blocking and Bayesian Reasoning in Preschoolers". *Cognitive Science* 28: 303-333.

Somiedo, Juan Pablo. 2018. "El análisis bayesiano como piedra angular de la inteligencia de alertas estratégicas". *Revista de Estudios en Seguridad Internacional* 4 (1): 161-176. dx.doi.org/10.18847/1.7.10

Taddeo, Mariarosaria, Luciano y Floridi. 2018. "Regulate Artificial Intelligence to Avert Cyber Arms Race". *Nature* 556: 296-298. doi.org/10.1038/d41586-018-04602-6

Von Solms, Rossouw, y Johan van Niekerk. 2013. "From Information Security to Cyber Security". *Computers and Security* 38: 97-102. doi.org/10.1016/j.cose.2013.04.004