

Tecnologías Emergentes

Técnicas de visualización analítica

Autor: David García Dolla, Nodo Gestor, SDG PLATIN.

Palabras clave: visualización interactiva, razonamiento humano, análisis de datos.

Metas Tecnológicas relacionadas: MT 6.1.2; MT 6.1.3.

La capacidad actual para recopilar y almacenar grandes cantidades de datos ha aumentado en los últimos años a mucha mayor velocidad que la capacidad para analizarlos. Si bien es cierto que se invierten grandes esfuerzos en desarrollar técnicas y algoritmos de análisis automático, lo cierto es que en la mayor parte de los casos, la complejidad inherente a la realidad hace necesario que la inteligencia humana esté presente en el ciclo de análisis. Por ello, no resulta extraño entender el auge que los métodos de visualización analítica (*visual analytics*, en inglés) han venido experimentando en las últimas décadas.

El fundamento de estos métodos consiste en complementar el proceso de toma de decisión humano con interfaces visuales interactivos, de forma que se combine la flexibilidad y creatividad humana con el conocimiento presente en las grandes cantidades de datos e información, a menudo solo identificable a través de vistas muy particulares de estos datos. Mediante estos interfaces, las personas pueden interactuar directamente con las capacidades de procesamiento de los ordenadores actuales, modificando parámetros o seleccionando métodos alternativos de análisis, y utilizando las técnicas de visualización para continuar profundizando en el proceso de razonamiento. La alternancia entre métodos de visualización y de procesamiento automático es una característica de este enfoque, permitiendo detectar desviaciones o falsas hipótesis desde etapas tempranas, lo que da lugar a resultados más fundamentados.

La figura 1 muestra el enfoque integral de estos métodos que combinan

la visualización, el análisis de datos y el razonamiento humano. Como se puede observar, existe un proceso de refinamiento paralelo entre las representaciones externas de los datos y las representaciones mentales del problema que tienen los expertos, debiendo existir coherencia entre ambas.

En definitiva, se trata de lograr que los expertos entiendan los datos en todas sus dimensiones para asegurar decisiones de mayor calidad.

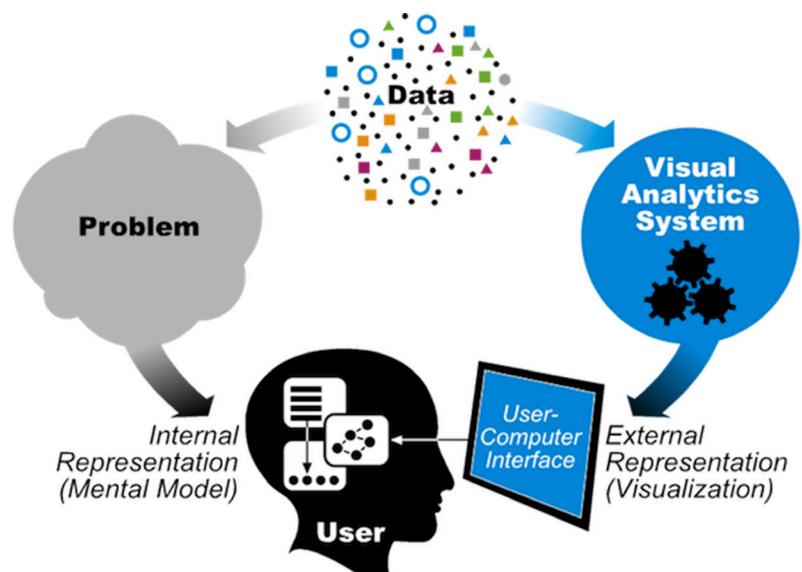
Si bien es un ámbito que se nutre de los avances de muchos otros dominios de conocimiento: los de factores humanos (cognición, percepción, etc.), visualización (gráficos por ordenador, etc.) y análisis de datos (recuperación, tratamiento y gestión de datos e información; minería de datos, etc.), todavía se considera una disciplina en desarrollo.

Asumiendo el potencial de estos métodos para abordar problemas de alta complejidad que combinen grandes volúmenes de datos procedentes de fuentes diversas y de carácter heterogéneo (asociado a diferentes dimensiones de un problema, referenciados espacial y temporalmente, con variable incertidumbre respecto a su bondad, dificultad para diferenciar entre lo esencial y lo irrelevante, etc.), es fácil entender el interés que despierta

estas disciplinas en el ámbito de defensa y seguridad.

Con la creciente disponibilidad de datos procedentes de sensores y con una capacidad de procesamiento computacional que va en aumento, pero a su vez con un creciente reto en cuanto a la complejidad de las amenazas y misiones, existe una necesidad urgente de mejorar los métodos de análisis de datos e información, pues en la actualidad constituye un cuello de botella que limita la posibilidad de tomar mejores decisiones.

Son muchos los dominios en los que se están utilizando estas técnicas (lucha contra IEDs, inteligencia, ciberdefensa, análisis de redes de comunicaciones, gestión de emergencias, etc.). Como ejemplo de ello, a continuación se apuntan algunos los grupos de estudios que la Organización de Ciencia y Tecnología de la OTAN (STO) ha puesto en marcha en los últimos años que las utilizan: IST-116 (*“Visual Analytics”*), IST-110 (*“Visualization for Analysis”*), IST-122 (*“Cyber Security Science and Engineering”*), SAS-IST-102 (*“Intelligence Exploitation of Social Media”*), IST-110 (*“Interactive visualization of dynamic networks”*), IST-133 (*“Visual analytics - cyber security”*), IST-131 (*“Distributed Data Analytics for Combating Weapons of*



Esquema del proceso paralelo de visualización analítica (fuente VA community)

tecnologías emergentes

Mass Destruction”, SAS-ET-CZ (“*Visual Analytics for Communicating Defense Investment Uncertainty and Risk*”), IST-SAS-139 (“*Visual Analytics for Exploring, Analysing and Understanding Vast, Complex and Dynamic Data*”), HFM-248 (“*Social Media and Information Technology for Disaster and Crisis Response*”), etc.

Resulta también interesante ver como la STO, a través de algunas de las actividades anteriores, ha establecido contacto con una de las principales comunidades internacionales dedicadas a esta tecnología, *Visual Analytics Community* (<http://vacommunity.org>), la cual está apoyada por los principales organismos del Gobierno de EE.UU vinculados a Defensa y Seguridad (*Department of Defense, Department of Homeland Security*).

Se trata de una comunidad de empresas y organismos de investigación dedicada a promover el desarrollo de esta tecnología a través de, entre otras actividades, concursos competitivos anuales en el que los participantes deben resolver situaciones ficticias normalmente vinculadas a situaciones de crisis o de riesgo para la seguridad en las que es necesario manejar grandes cantidades de datos muy heterogéneos.

Así, en el reto de 2015, actualmente en ejecución (*VAST Challenge 2015*), se recrea una situación ficticia en la

que, en el marco de un homenaje a una antigua estrella del fútbol que tiene lugar en un parque de atracciones, se produce un crimen perpetrado por una persona desequilibrada. Si bien el autor del crimen es capturado, la policía está interesada en recrear los detalles de lo ocurrido para lo que se plantean retos asociados a determinar el movimiento de individuos y grupos en el parque (obtenidos a partir de los datos de pago de todos los visitantes al parque durante un fin de semana), así como de las comunicaciones entre los visitantes y entre estos y los servicios del parque (usando los datos de la app que proporciona el parque).

Por su parte, en el reto de 2014, se planteaba una reunión de altos directivos de una empresa gasística en la isla en la que tienen sus principales explotaciones, durante la cual se produce un secuestro de sus directivos, siendo sospechosos los miembros de una organización local que luchaba activamente contra estas explotaciones. La información aportada para resolver la situación de crisis consistía tanto en informes de texto (noticias aparecidas en prensa durante los últimos años sobre las actividades de la compañía y de los activistas locales, informes sobre los líderes de esta organización, cabeceras de correos electrónicos entre empleados de la empresa), datos de movimiento du-

rante el fin de semana del secuestro (posición GPS de sus vehículos, transacciones de tarjetas de crédito, etc.) así como las comunicaciones durante la situación de crisis (llamadas al servicio de emergencias, etc.).

En 2012, el reto se centró en una situación de ciberseguridad, en la que un banco imaginario necesita herramientas para aumentar la consciencia situacional de su red de casi un millón de ordenadores de forma que pueda percibir su estado y detectar problemas, para lo que se planteaban dos retos, el primero vinculado a la visualización de enormes cantidades de datos y el segundo a temas de seguridad en redes de ordenadores.

En estos concursos no solo premia a los participantes que sean capaces de encontrar la solución a los retos planteados, sino que también se valora la originalidad y efectividad de las soluciones de visualización interactivas empleadas. Al finalizar los concursos, se desarrolla un simposio en el que se ponen en común las diferentes aproximaciones realizadas.

Constituye, en definitiva, una manera muy interesante de promover el desarrollo de una tecnología que está llamada a tomar cada vez una mayor relevancia en las soluciones tecnológicas de gestión de datos e información, en particular en el ámbito de defensa y seguridad.



Fig. 2. Imagen de la actividad IST-133 sobre analítica visual aplicada a la ciberdefensa. (Fuente: STO).