

# El Análisis Automático de Datos, su Aporte a la Investigación Criminal

Bruno Constanzo<sup>1</sup>, Sabrina Lamperti<sup>1</sup>, Sebastián Lasia<sup>1</sup>, Ariel Podestá<sup>1</sup>,  
Pablo Cistoldi<sup>2</sup>, Ana Haydée Di Iorio<sup>1</sup>

<sup>1</sup> InFo-Lab, Laboratorio de Investigación y Desarrollo de Tecnología en Informática Forense (Universidad FASTA, Ministerio Público de la Provincia de Buenos Aires, Municipalidad de General Pueyrredon), Avellaneda 3341, Mar del Plata, Argentina

<sup>1</sup> {bconstanzo, slamperti, elsebalasia, arieluf, diana}@ufasta.edu.ar, <sup>2</sup> pcistoldi@mpba.gov.ar

<http://www.info-lab.org.ar>

**Resumen.** La aparición de la sociedad de la información facilitó el uso de las tecnologías simplificando la creación, distribución y manipulación de la información. Sin embargo, ello trajo consigo nuevos desafíos a la hora de resolver investigaciones criminales, resultando primordial el análisis de datos que pueda derivar en información potencialmente relevante para los procesos judiciales. Es por esta razón que el desarrollo de software especializado aparece como una herramienta necesaria y útil para el logro de tales objetivos. Por medio de este trabajo se pretende mostrar los avances obtenidos a través del trabajo interdisciplinario, en la búsqueda de soluciones adecuadas a las necesidades reales y locales de los Ministerios Público Fiscales de la República Argentina.

## 1 Introducción

La llamada “sociedad de la información”, potenció la capacidad de la humanidad para producir, transmitir y consumir anualmente enormes cantidades de información. A modo de ejemplo, sólo en 2007 se transmitieron 1.9 zettabytes de información en el mundo[1], siguiendo una histórica tendencia donde cada vez se producen, transmiten y almacenan más datos. Otros estudios y publicaciones indican números crecientes para distintos años sucesivos [2, 3], pero no es necesario ahondar en detalles para dejar en claro que las sociedades modernas producen enormes cantidades de información, ya sea en conjunto o de manera aislada.

En este contexto, el Ministerio Público Fiscal debe continuar con sus tareas de investigación criminal, enfrentándose no sólo a los hechos, rastros, y redes criminales, sino también al crecimiento exponencial de las fuentes de información. Cada vez con mayor frecuencia, se busca que los operadores judiciales o de organismos de investigación optimicen el tratamiento de grandes volúmenes de datos reduciendo el tiempo para descubrir puntos de interés que contribuyan a una oportuna toma de decisiones.

Dentro de esta información, proveniente de fuentes heterogéneas, se espera que los operadores judiciales sean capaces de detectar los patrones delictivos, reconstruyan secuencias temporales, descubran a los partícipes de hechos y determinen sus roles.

A modo de ejemplo, en la actualidad es posible que en una investigación se analicen diversas fuentes de información, tales como registros de llamadas entrantes y salientes, listados de transacciones bancarias, consumos de tarjeta de crédito, listados de direcciones IP (y la vinculación de estas con los abonados a un servicio de internet), registros de antenas de telefonía móvil y las conexiones de teléfonos celulares con las mismas, registros de llamadas provenientes del análisis forense de equipos móviles, conversaciones de mensajería instantánea obtenidos a través de procedimientos forenses; así como también informes registrales, notariales, bancarios, financieros, crediticios, y de organismos públicos o privados.

El desafío que se presenta es el de poder brindar a los operadores judiciales herramientas informáticas que les permitan trabajar sobre la totalidad de la información, y que estos datos, provenientes de diferentes fuentes y en distintos formatos, pueden integrarse en un único repositorio, permitiendo un análisis completo de la situación y, a su vez, la detección de patrones de comportamiento y conexiones.

Una forma de abordar esta problemática es a través del desarrollo de un software especializado que permita el análisis de datos por medio de grafos<sup>1</sup>, dado que es una de las estructuras de datos que mejor se presta a amalgamar información de fuentes dispares en un todo conexo.

Si bien actualmente hay distintos softwares existentes que permiten realizar este tipo de análisis, éstos se presentan como herramientas genéricas (como Pajek[4], Gephy[5], Cytoscape[6]) o muy enfocadas en tipo de problema en particular (como i2[7] o Maltego[8]). Es por esta razón que, al constituirse el Laboratorio de Investigación y Desarrollo en Informática Forense (InFo-Lab) en el año 2014, uno de los proyectos principales consistió en desarrollar un nuevo software especializado en la problemática expuesta, y que se dio en llamar INVESTIGA, cuya característica principal es la de encontrarse enfocado en las necesidades del Ministerio Público Fiscal de la Provincia de Buenos Aires[9].

En el desarrollo de INVESTIGA participa un equipo interdisciplinario, compuesto por ingenieros en informática, criminalistas y abogados, llevando a cabo las tareas de diseño, desarrollo, programación y test sobre el sistema desarrollado. El objetivo final, por tanto, será brindar una herramienta que permita mostrar -mediante el graficado y la visualización de diferentes resultados- las distintas informaciones requeridas a los organismos públicos y privados en el curso de una investigación criminal, de modo tal de poder establecer patrones y/o modalidades delictivas, o simplemente visualizar la evolución de la situación patrimonial de un sospechoso.

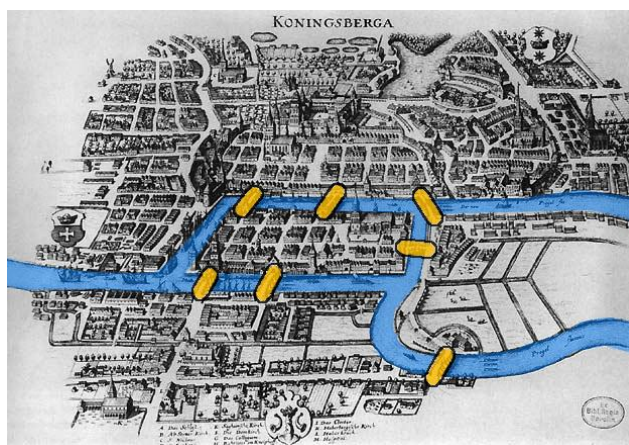
---

<sup>1</sup> En la Sociología se conoce la técnica como “análisis de redes sociales”; sin embargo, preferimos no utilizarlo en este desarrollo, dado que puede generar confusión los medios de comunicación social que permiten establecer contactos entre personas por medio de la Web.

## 2 Breve Introducción a la Teoría de Grafos

Un grafo es una estructura matemática que se compone de dos partes: nodos y vínculos. Los nodos son objetos o entidades propias que resultan de interés, y cada vínculo representa una relación o conexión entre dos objetos. En 1736 Leonhard Euler presentó una publicación sobre el problema de los siete puentes de Königsberg[10] en la que demostró que era imposible resolver el problema, y además sentó las bases para fundar la Teoría de Grafos y la Topología como ramas de la Matemática.

El problema de los puentes de Königsberg es simple: la ciudad de Königsberg es atravesada por un río, y en él hay dos grandes islas. Ambas orillas y las islas se conectan entre sí por 7 puentes, y se quiere buscar un recorrido que pasee por toda la ciudad, pero atravesase cada puente una sola vez. Este problema, en apariencia simple, ocultaba la complejidad de ser difícil de analizar sin las herramientas adecuadas.

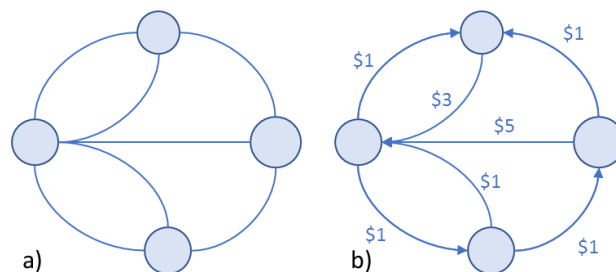


**Fig. 1.** Mapa de 1652 de la Ciudad de Königsberg<sup>2</sup>, con el agua y los puentes resaltados.

Si bien sobre el mapa de la ciudad, que se ve en la Figura 1, se pueden trazar rutas y luego evaluar si cumplen las condiciones del problema, al representarlo como un grafo se logra abstraer todo el problema a una representación mucho más clara:

---

<sup>2</sup> Imagen histórica de dominio público, disponible en [http://www.preussenchronik.de/bild\\_jsp/key=bild\\_kathe2.html](http://www.preussenchronik.de/bild_jsp/key=bild_kathe2.html). Durante la Segunda Guerra Mundial Königsberg fue bombardeada y los 7 puentes fueron destruidos. Luego de la guerra se convirtió en la ciudad soviética de Kaliningrado, y 5 de los históricos puentes fueron reconstruidos.



**Fig. 2. a)** Representación por medio de grafos del problema de los puentes. Cada nodo representa una masa de tierra, y cada vínculo representa un puente. **b)** Grafo de los puentes, con direcciones y pesos que agregan restricciones al problema original.

Tomando distancia del problema original, se propone considerar algunas situaciones adicionales para ilustrar algunos otros conceptos de grafos:

- La ciudad decide que el tránsito en algunos puentes solo podrá ir en una dirección. El grafo original ahora se convierte en un grafo dirigido, y los vínculos se representan por medio de flechas para indicar el sentido de la relación.
- Además, la ciudad decide incorporar un peaje en cada uno de los puentes para poder trabajar en obras de reparación. Dado que no todos los puentes han sufrido el mismo desgaste, y con el objetivo de incentivar el uso de unos sobre otros, hay distintas tarifas para los puentes más desgastados. En el grafo esto se representa indicando el costo junto a cada vínculo.
  - La información asociada a un vínculo se denomina “peso del vínculo”, y no se restringe únicamente a variables numéricas. Puede tratarse, por ejemplo, de fechas, horas, caudales, u otro tipo de datos, dependiendo de lo que se represente en el grafo, o conjuntos de datos.

Estas modificaciones pueden verse en la Figura 2.b), donde se ilustra cómo quedaría el puente y los posibles costos. Agregando esta nueva información, se pueden realizar nuevos análisis sobre cómo realizar un recorrido de la ciudad, por ejemplo, tratando de minimizar (o maximizar) el dinero gastado en los peajes.

El problema de los puentes se utiliza de manera ilustrativa para explicar más claramente qué es un grafo y la utilidad que presentan para modelar y analizar situaciones complejas. En otras ciencias, se pueden utilizar grafos para representar distintos tipos de problemas:

- En la Química, se puede representar moléculas por medio de grafos, en donde los nodos representan átomos y los vínculos son los enlaces entre ellos.
- En la Sociología se utilizan los grafos para representar grupos de personas y su relación (ya sea de parentesco, afectiva, laboral u otro tipo). Estos grafos de personas se conocen como “redes sociales” y su análisis permite conocer con gran profundidad cómo se maneja un grupo y quiénes son las personas centrales, más importantes, o más conectadas del mismo.

- En la Ingeniería, un proceso puede representarse por medio de grafos, donde los nodos representan estados definidos y los vínculos la transición entre ellos. Esto permite analizar cuáles son las tareas en donde deben enfocarse más recursos o esfuerzos para optimizar, ya sea costos, tiempo, aprovechamiento de recursos o alguna otra variable de interés.

### 3 Desarrollo y Análisis del Software

En la investigación del delito, sobre todo en casos complejos, se necesita acceder y trabajar con datos que no tienen una estructura en común, y que provienen de múltiples fuentes y diversas plataformas. Debido a la escala, características y forma de estas fuentes de información, resulta difícil analizarlas manualmente. Con el fin de coadyuvar en este tipo de investigaciones es que se gesta la idea de desarrollar INVESTIGA, una herramienta que permite recopilar, integrar, analizar, visualizar y distribuir grandes volúmenes de información heterogénea proveniente de múltiples fuentes de datos, para facilitar la detección de patrones y conexiones.

No solo se busca introducir un programa informático, sino mejorar el todo de la investigación y la litigación de casos complejos. Para lograr este objetivo, es necesario cambiar la forma de pensar los casos complejos, utilizar herramientas adecuadas para la tarea, y aprender nuevas habilidades. El proceso que se propone entonces consiste en:

1. **Planificar las tareas.**
2. **Trabajar en equipo**, entre los distintos involucrados, tanto en las tareas de campo, como en la obtención y análisis de los datos.
3. **Sistematizar los datos heterogéneos**, identificando en ellos la información relevante para descubrir las relaciones y patrones subyacentes en ellos.
4. **Presentar los datos:** a los investigadores (para reflexionar sobre los mismos), y ante el tribunal de forma que resulte sencilla y contundente para acompañar la estrategia de litigación.

Si bien podría utilizarse cualquier herramienta genérica para realizar el análisis con cualquiera de ellas que sea genérica, al utilizar las que son desarrolladas atendiendo a problemáticas de nivel local se advierte que el análisis de la información se realiza de manera más intuitiva y simple para el usuario. En este sentido, INVESTIGA ayuda a descubrir las relaciones y patrones que se encuentran presentes en los distintos conjuntos de datos, que a simple vista no podrían detectarse fácilmente, ya sea por la cantidad de información, por el formato, por la representación de los datos o porque provengan de distintas fuentes que no están relacionadas de manera directa.

En general, la mejor forma de utilizar cualquier herramienta se da cuando el usuario la conoce en profundidad, sus posibilidades y limitaciones, conoce también el caso

que está investigando, y qué puede realizar con los conjuntos de datos que tiene disponibles. Es decir, debe darse una adecuada interacción entre el conocimiento que el usuario tiene del caso y su habilidad para aprovechar las posibilidades que le ofrecen las herramientas disponibles.

Por ejemplo, ante un caso en que se investigue la fenomenología del lavado de activos<sup>3</sup>, se debe explorar en fuentes que cuenten con información de carácter patrimonial, vinculada a un conjunto de bienes y derechos de contenido económico correspondientes a una persona física o jurídica. Se debe, por tanto, recabar información sobre aspectos personales, relacionales y societarios de los distintos sujetos investigados que permitan identificar el origen de los bienes, realizando un análisis temporal, situacional, de usos y ocupaciones, incremental, entre otros[12].

La posibilidad de visualizar los datos como grafos permite que se realice el análisis de la información de manera intuitiva. Al ver cómo se organizan los nodos y las conexiones entre ellos, conociendo el significado semántico que representan, resulta muy simple interpretar el conjunto de datos y da pie a avanzar la investigación: encontrar nuevos sospechosos o partícipes, identificar domicilios y/o vehículos, por ejemplo, para solicitar secuestros de bienes, registros de llamadas o intervenciones telefónicas. Además, se pueden descubrir fácilmente patrones de actuación de una banda o asociación ilícita, descubrir “rutas del dinero”, bienes de origen ilícito, o reconstruir secuencias temporales, por citar algunos ejemplos.

Representar la información heterogénea y compleja por medio de grafos también permite realizar presentaciones más contundentes de parte del Ministerio Público Fiscal ante los tribunales, ya sea para pedir medidas de injerencia o coerción, para interrogar testigos y peritos, o para efectuar alegaciones. La presentación gráfica ayuda a que el juez y/o el jurado perciban de manera sencilla los indicios que surgen de los datos. Reformulando el viejo adagio, podríamos decir que “un grafo vale más que mil filas de datos”.

### **3.1 Formas de uso del software**

El diseño de INVESTIGA como aplicación está marcado por las necesidades del Ministerio Público y las fortalezas del equipo detrás del proyecto. En base a esa realidad, el software se ha estructurado como una aplicación web que se accede a través del navegador.

---

<sup>3</sup> Recordamos que así se denomina a aquellos hechos susceptibles de ser investigados criminalmente por tratarse de maniobras que buscan ocultar o disimular la naturaleza, origen, ubicación, propiedad o control de dinero y/o de bienes obtenidos ilegalmente a través de éste.

Dentro de INVESTIGA, algunos de los conceptos que se explicaron en la sección *Breve Introducción a la Teoría de Grafos* pueden encontrarse con otros nombres, para una comprensión más intuitiva por parte de los usuarios finales<sup>4</sup>. En particular:

- Se utiliza “entidad” como sinónimo de “nodo”. Las entidades pueden ser de distintos tipos, por ejemplo: persona, teléfono, vehículo, cuenta bancaria, etc.
- Se denomina “etiqueta del vínculo” al elemento gráfico que representa un atributo del peso del vínculo en el editor. La etiqueta muestra una parte del peso en el grafo dibujado (por ejemplo, la cantidad de llamados entre dos teléfonos), y el resto de la información que acompaña al vínculo está disponible para consulta o utilización por parte del motor de procesamiento del sistema (continuando con el ejemplo, la fecha, hora y duración de las llamadas).

En cuanto a la interfaz para el usuario final, el sistema cuenta con dos grandes partes:

- En primer lugar, la interfaz de menús e importación, similar a un conjunto de *wizards* o ayudantes. Desde esta interfaz el usuario selecciona sobre qué quiere trabajar, ya sea un nuevo grafo o uno preexistente, y si desea realizar una importación de datos al sistema.
- En segundo lugar, la interfaz del editor, donde el usuario puede crear, explorar y modificar grafos, está pensado en las necesidades reales de las investigaciones en el país. Se la considera como el área principal de trabajo para el usuario, y es donde podrá explotar el potencial de la herramienta, de acuerdo a cada caso en el que se requiera su empleo. Esto marca una diferencia con respecto a otras posibilidades de análisis y procesamiento de datos.

Las formas básicas en que el usuario puede trabajar sobre INVESTIGA son:

- El trabajo en forma libre, en donde el usuario ingresa directamente al editor con un grafo en blanco y va colocando entidades y vínculos sobre el grafo, realizando las conexiones en forma manual. Para casos muy simples, o en donde ya cuenta con un diagrama (en papel, u otro formato de “difícil acceso” para el sistema), resulta una forma adecuada de trabajo, aunque no se aprovechan completamente las capacidades del motor de relaciones del sistema.
- La importación de datos permite al usuario la carga de gran cantidad de datos estructurados de manera rápida. La interfaz de menús guiará el proceso de importación, permitiendo que el usuario seleccione un modelo de datos (o “plantilla de importación”) -adecuado a la información que sus datos representan-, o cree un modelo nuevo si no cuenta con uno adecuado. Como parte del proceso, se crean todas las entidades representadas en los datos, y luego los vínculos entre ellas por medio del motor de relaciones de INVESTIGA. Una vez finaliza-

---

<sup>4</sup> En su desarrollo se tuvo en cuenta que las investigaciones están a cargo de instructores del Ministerio Público, y que en ocasiones se deriva su análisis a miembros de fuerzas de seguridad; por lo que no es requerido conocimientos avanzados de informática para su utilización.

do el proceso de importación, el sistema muestra el grafo resultante en el editor, listo para que el usuario comience a explorarlo, analizarlo y modificarlo.

- Este modo permite trabajar con datos que sean remitidos por organismos, o bien a través de un listado que haya confeccionado el propio investigador. Actualmente, se permite importar planillas de cálculo, en formato CSV<sup>5</sup>. Los datos que se quieran importar deben ser convertidos a dicho formato independientemente de su origen.

Además de estos casos de uso “básicos”, se puede trabajar de manera mixta con la herramienta, importando datos, modificando el gráfico de manera manual, y luego volviendo a importar datos para incorporar nueva información. En la siguiente sección se ilustran casos de ejemplo sobre cómo utilizar INVESTIGA donde se verán reflejadas estas formas de trabajo.

## 4 Aplicaciones Prácticas

A continuación, se presentan casos reales, adaptados para su publicación, en donde se ha utilizado el software de análisis de grafos. Se busca resaltar cómo el uso de una herramienta como INVESTIGA aporta a una mejor visualización de los datos provistos por distintos organismos o empresas, convirtiéndolos fácilmente en información sustancial para una mejor resolución de las investigaciones.

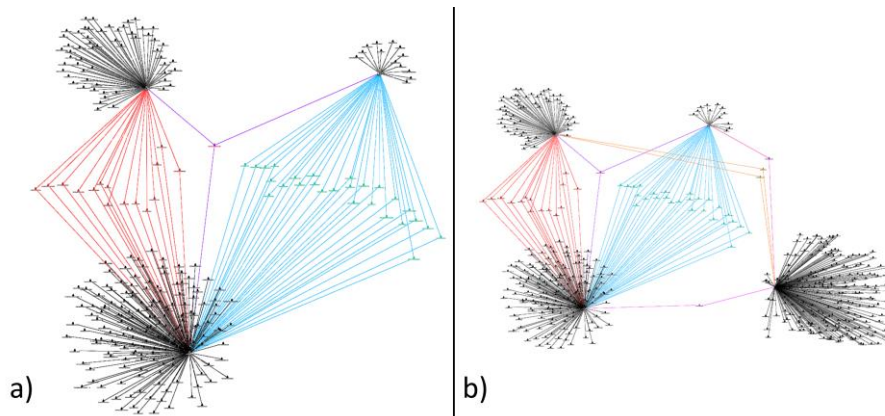
### 4.1 Caso: Choferes “fantasma” en empresas de transporte

Se denuncia que existen personas registradas como choferes de distintas empresas de transporte público de pasajeros, que cumplirían horas de trabajo dentro de cada una de ellas. Sin embargo, el testimonio los choferes indicaba que nunca habían trabajado en estas empresas, o bien que sólo prestaban servicios en una. Asimismo, uno de ellos detectó que estaba registrado como en relación de dependencia al momento de solicitar un beneficio social ante la ANSES, cuando en realidad estaba desempleado. Se hizo evidente que existía un aumento excesivo de la cantidad de personal en determinados meses del año. El objetivo de incorporar de personal de forma ficticia se realizaba con el fin de petitionar subsidios estatales por parte de las empresas, lucrando a su favor.

---

<sup>5</sup> *Comma-Separated Values*, un formato de archivo simple que permite representar tablas de datos. Es posible convertir los archivos a dicho formato a través de un editor de hoja de cálculo, o mediante el uso de un script ya programado.





**Fig. 3. a)** Gráfico inicial con tres empresas, sus empleados y los vínculos que denotan la relación de dependencia. Algunos choferes están registrados en más de una empresa. **b)** Gráfico resultante al integrar información de una cuarta empresa.

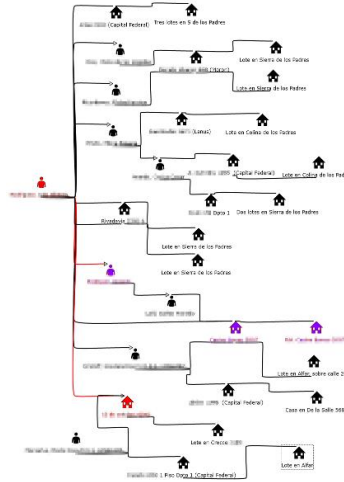
Al realizar un entrecruzamiento de los listados de choferes registrados ante AFIP por cada una de las empresas de transporte, se determinan las personas que cumplen rol de chofer simultáneamente en varias empresas (Figura 3a). Al comienzo de la investigación se contaba con tres listados de personas provisto por la AFIP relacionadas a tres empresas de transporte público de pasajeros. Posteriormente se requirió información de una cuarta, agregándose los datos de esta empresa al gráfico ya existente. El software analiza los nodos en común y los integra al grafo existente (ver Figura 3b).

Se puede apreciar cómo al agregar una cuarta empresa se realiza la vinculación de forma automática a través del propio motor del software, lo que nos permite visualizar rápidamente los nuevos casos de personas que figuran registradas con más de un empleador.

## 4.2 Caso: Defraudaciones inmobiliarias

Se trata de una veintena de defraudaciones inmobiliarias cometidas en Mar del Plata y zona. Cuatro de estos imputados, dos gestores y dos abogadas, conformaban una asociación ilícita que llevaba adelante las maniobras delictivas tendientes a lograr la escrituración de inmuebles ante la Justicia Civil.

En el gráfico siguiente se muestra cómo uno de los integrantes de la asociación ilícita, se encuentra vinculado con cada uno de los hechos, a partir de distintas conexiones (domicilios o personas en común). Este sujeto principal, era quien conectaba de distintas formas con las víctimas, y a su vez las derivaba a las abogadas que luego iniciaban el proceso de escrituración.



**Fig. 5.** El sistema permitió graficar de forma sencilla, las interconexiones existentes entre personas y domicilios que se repetían entre la veintena de casos.

### 4.3 Casos de entrecruzamiento de llamadas

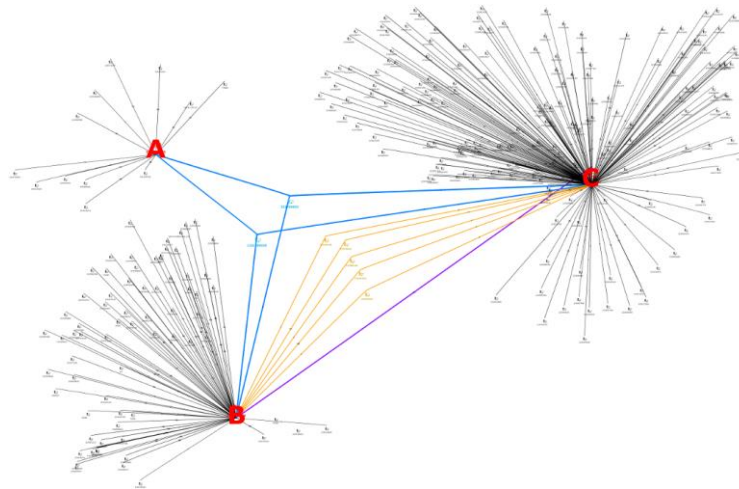
Cuando se piden a las empresas prestadoras de servicio telefónico los registros de llamadas entrantes y salientes, es común que respondan enviando un documento de texto con la salida de pantalla de un sistema propio (ver la Figura 6). Puede verse en él un encabezado (que se repite en cada “hoja”) que indica de qué línea se trata, cuando se emitió el reporte, qué rangos de fecha cubre, y el encabezado de las columnas de datos: Fecha y Hora, número pagador (el cliente de la prestadora de servicio), dirección de la llamada (Entrante, Saliente o Transferencia), el número destino y la duración de la llamada.

El análisis clásico que realizaban los investigadores con este tipo de información, es imprimiendo el listado y realizando un trabajo “manual” de entrecruzamiento, verificando línea por línea y tomando nota, a modo de detectar las vinculaciones entre teléfonos, o su ubicación en una determinada zona. Esto suele conllevar esfuerzo mental, concentración y realización de algún tipo de bosquejo manual por parte de quien realiza la labor, pudiendo incurrirse en errores involuntarios, o en un déficit en la explotación total de los datos.

1	JUD011L			
2	Detalle de Llamadas por Nro. Pagador			
3	Telefono	[REDACTED]	Hasta	[31/12/2016-99:99]
4	Fecha desde	[01/12/2015-00:00]		
5	Ejecucion	[12/Abr/2017-11:33]		
6	-----			
7	Fecha y Hora	Nro. Pagador	Dir. Nro. Destino	Duracion
8	-----			
10	01/12/2015 10:45:29	[REDACTED]	S [REDACTED]	00:00:02
11	01/12/2015 10:47:00	[REDACTED]	E [REDACTED]	00:05:21
12	01/12/2015 11:14:29	[REDACTED]	E [REDACTED]	00:00:15
13	01/12/2015 11:15:11	[REDACTED]	S [REDACTED]	00:00:42
14	01/12/2015 11:20:34	[REDACTED]	S [REDACTED]	00:00:01
15	01/12/2015 11:21:34	[REDACTED]	S [REDACTED]	00:01:10
16	01/12/2015 11:23:20	[REDACTED]	S [REDACTED]	00:02:44
17	01/12/2015 13:18:37	[REDACTED]	S [REDACTED]	00:01:51
18	01/12/2015 13:53:13	[REDACTED]	E [REDACTED]	00:00:07

**Fig. 6.** Ejemplo de listado de llamadas en el formato enviado por la prestadora de servicios. Se han pixelado datos sensibles para el ejemplo.

Para analizar esta información con INVESTIGA, en primer lugar, debe convertirse a formato CSV. Esta transformación de los datos se puede realizar con herramientas como Microsoft Excel, LibreOffice Calc o *scripts* de programación. Luego se procede a realizar un gráfico por medio del importador de datos, y se utiliza una plantilla<sup>6</sup>. Una vez finalizado el proceso de importación, se obtiene un grafo “de estrella”, donde el número pagador (la línea de interés) se encuentra en el centro, rodeada por los números con los que se comunicó en el período de tiempo analizado. Cada vínculo entre el número pagador y los otros números lleva la información de fecha, hora, duración y dirección de la llamada.



**Fig. 7.** Grafo de cruzamiento de llamadas, resultado de combinar listados de tres líneas telefónicas de interés (A, B y C).

Este grafo puede intercalarse con otros registros de llamadas entrantes y salientes, de la misma línea telefónica de interés, u otras. Si se incorporan los registros de lla-

<sup>6</sup> Si no hay una plantilla adecuada, se puede crear un nuevo “modelo de importación” que interprete adecuadamente los datos contenidos en el archivo CSV.

madras de varias líneas, se logrará un grafo similar al de la Figura 7, en donde pueden observarse tres “islas” de llamados, que corresponden a tres líneas telefónicas de interés. Se ha marcado en celeste, dentro de este gráfico, a aquellos dos números que son contactos en común con las líneas A, B y C, en naranja los contactos en común entre B y C, y en violeta se han indicado los llamados que se registraron directamente entre B y C.

Eventualmente un grafo de este tipo podría ampliarse realizando un cruzamiento de personas y números de teléfono. De esta manera podría hacerse evidente alguna situación en donde, por ejemplo, un sospechoso utiliza múltiples líneas para tratar de ocultar su identidad (ya sea a las autoridades o a sus contactos).

## 5 Conclusiones y Trabajo Futuro

Desde el 2016 el software INVESTIGA se encuentra disponible para los instructores del Ministerio Público de la Provincia de Buenos Aires a modo de prueba (en una versión *beta* del sistema), y se ha utilizado bajo esta modalidad en causas judiciales con una experiencia positiva de parte de los usuarios finales. Durante el año 2017 se prevé su instalación y capacitación progresiva en los distintos departamentos judiciales de la provincia.

En base a la experiencia de uso, tanto en etapa *beta* como de testeos, se ha logrado corregir *bugs*, así como la detección de algunas potencialidades a mejorar y nuevas necesidades en el software desarrollado. De este *feedback* pueden destacarse los siguientes puntos:

- Ante una carga masiva de datos, la interfaz y la librería de dibujado de grafos pueden verse saturadas, resultando así en un rendimiento inferior al esperado por parte de la aplicación. De esta forma, se han tomado medidas para mejorar la *performance*, y se incorporó un “*Modo Lite*”<sup>7</sup> para ayudar en estas situaciones. En la actualidad también se están explorando nuevas alternativas que van a permitir trabajar de manera más fluida y cómoda sobre grafos mucho más grandes que en la versión actual.
- Dado que el convenio que rige al InFo-Lab pone en disponibilidad sus desarrollos a todas las provincias del país por medio del Consejo Federal de Procuradores, varias provincias han expresado su interés en tener implantaciones propias de INVESTIGA. Con este propósito en mente, se ha establecido contacto con los Ministerios Públicos de CABA, Chaco, Chubut, Entre Ríos, La Pampa, San Luis, Santiago del Estero, Río Negro y Tucumán, a quienes se les ha mostrado el funcionamiento del software, teniendo un acceso a la versión beta.

---

<sup>7</sup> Es un modo donde el dibujado de los grafos se realiza de manera simplificada para facilitar el trabajo con grandes cantidades de nodos y vínculos.

- Hay pruebas en curso para integrar una base de conocimiento junto con el motor de relaciones de INVESTIGA. Esto posibilitará realizar vinculaciones y contrastar la información que se maneja dentro del sistema contra resultados de investigaciones previas de manera rápida. Además, permitirá reutilizar la información, evitando solicitudes duplicadas a los organismos que las proveen, reduciendo esfuerzos y recursos por parte del Ministerio Público.
- Si bien el sistema ya cuenta con algunas plantillas de importación rápida, a medida que su uso sea más extendido se incorporarán a la “biblioteca estándar” aquellas que sean creadas por los usuarios y resulten de mayor utilidad.
- Se está trabajando en una nueva modalidad de trabajo dentro del software que pueda ayudar con las estrategias de investigación y litigación.
- Se encuentra en progreso otro proyecto complementario que agrega a INVESTIGA capacidades de *Big Data*, de manera que se puedan realizar tareas de filtrado y búsqueda de la información previa a la generación del grafo desde un ambiente integrado al sistema. Esto permitirá potenciar el trabajo con la aplicación sin tener que abandonar su interfaz.
- En el marco de un proyecto final de graduación de la carrera Ingeniería en Informática, dos alumnos realizaron una herramienta complementaria, que permite visualizar y explorar los grafos de manera interactiva, incorporando además contenido multimedia de audio y video, como, por ejemplo, escuchas telefónica, fotografías de cualquier índole, o imágenes de cámaras de seguridad viales. Esta herramienta, llamada “Visor Web INVESTIGA”, permite realizar presentaciones más visuales y dinámicas, por ejemplo, para utilizar a la hora de intervenir en un juicio oral o en peticiones a realizarse ante los jueces.

El software INVESTIGA ha probado ser una herramienta de gran utilidad. Sin embargo, su mayor fortaleza será cuando esté implantado en toda la provincia y los operadores judiciales puedan aprovecharla al máximo. El *feedback* de los operadores sobre su uso y necesidades, permitirán retroalimentar el desarrollo y tener un proceso de mejora continua sobre el mismo.

La experiencia en el desarrollo de este tipo de software resulta alentadora para otros proyectos que el InFo-Lab mantiene actualmente, tales como un entorno de análisis de los datos contenidos en dispositivos móviles, así como uno destinado al procesamiento de imágenes, ambos enfocados en las necesidades de los operadores judiciales. La labor interdisciplinaria entre distintas ciencias, en este caso reflejadas en los aportes que puede realizar la ciencia informática a la criminalística y al derecho, coadyuva a una mejor resolución de los conflictos judiciales y en definitiva a una mejor calidad del servicio de justicia.

## Agradecimientos

Se agradece a la Universidad FASTA, al Ministerio Público Fiscal de la Provincia de Buenos Aires y a la Municipalidad de General Pueyrredon por brindar un espacio único de trabajo como es el Laboratorio de Investigación y Desarrollo de Tecnología en Informática Forense. En particular, a Julián Waimann por su desempeño como arquitecto de *software* de la aplicación, y a Santiago Trigo como ingeniero a cargo del control de calidad del sistema.

## Referencias

1. Martin Hilbert, Priscila López, “*The World’s Technological Capacity to Store, Communicate and Compute Information*”, Science 332, 60 (2011). Disponible en <http://www.uvm.edu/pdodds/files/papers/others/2011/hilbert2011a.pdf>.
2. Bret Swanson, George Gilder, “*Estimating the Exaflood - The Impact of Video and Rich Media on the Internet*”. Discovery Institute, Seattle, Washington (2008).
3. Roger Bohn, James Short, “*How Much Information? 2009 Report on American Consumers*”. Global Information Industry Center, University of California, San Diego (2009). Copia disponible en [https://web.archive.org/web/20130422121105/http://hmi.ucsd.edu/pdf/HMI\\_2009\\_ConsumerReport\\_Dec9\\_2009.pdf](https://web.archive.org/web/20130422121105/http://hmi.ucsd.edu/pdf/HMI_2009_ConsumerReport_Dec9_2009.pdf).
4. Pajek (software), información disponible en: <http://mrvar.fdv.uni-lj.si/pajek/>.
5. Gephy (software), información disponible en: <https://gephi.org/>.
6. Cytoscape (software), información disponible en: <http://www.cytoscape.org/>.
7. IBM i2 (software), información disponible en: <http://www-03.ibm.com/software/products/es/analysts-notebook>.
8. Maltego (software), información disponible en: <https://www.paterva.com/>.
9. Roberto Giordano Lerena, Ana H. Di Iorio, Ariel Podestá, Bruno Constanzo, “*InFo-Lab, una laboratorio mixto de investigación y desarrollo de tecnología en Informática Forense*”. III CADI (2016).
10. Leonhard Euler, “*Solutio problematis ad geometriam situs pertinentis*”. (1736). Scan del original disponible en <http://eulerarchive.maa.org/docs/originals/E053.pdf>.
11. Mark Newman, “*The structure and function of complex networks*”. SIAM Review 45 (2003). Disponible en <http://www-personal.umich.edu/~mejn/courses/2004/cscs535/review.pdf>.
12. Ana Di Iorio, Sabrina B. Lamperti, “*InFo-Lab: cooperando en el desarrollo de tecnología nacional en Informática Forense que coadyuve a la investigación criminal*”, en “*Lavado de activos, narcotráfico y crimen organizado. Aportes para la prevención y control de una problemática en desarrollo.*” (2015). Editorial UAI. Red de Editoriales de Universidades Privadas. CRUP. ISBN 978-987-4023-01-8.}