R E V I S T A

INGENIERIA DE SISTEMAS

Volumen XXIII Septiembre 2009

Publicada por el
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

R E V I S T A
INGENIERIA DE SISTEMAS

ISSN 0716 - 1174

Financiado parcialmente por el Instituto Milenio "Sistemas Complejos de Ingeniería", como reconocimiento a la difusión de las materias abordadas y de sus participantes.

Instrucciones a los autores:
Los autores deben enviar 2 copias del manuscrito que desean someter a referato a: Comité Editorial, Revista Ingeniería de Sistemas, Av. República 701, Santiago, Chile. Los manuscritos deben estar impresos en hojas tamaño carta, a doble espacio, deben incluir un resumen de no más de 150 palabras, y su extensión no debe exceder las 25 páginas. Detalles en www.dii.uchile.cl/∼ris

Correo electrónico: ris@dii.uchile.cl
Web URL: www.dii.uchile.cl/∼ris

Representante legal: Máximo Bosch
Dirección: República 701, Santiago, Chile.
Diagramación: Ma. Fernanda Bravo
Impresión: Ka2 Diseño e Impresión
Mail: contacto@ka2.cl

Carta Editorial Volumen XXIII

Nos es muy grato editar el número XXIII de la Revista de Ingeniería de Sistemas (RIS) dedicado a temas de frontera en Investigación de Operaciones, Gestión y Tecnología. Queremos agradecer al Instituto Milenio "Sistemas Complejos de Ingeniería"por su colaboración para hacer posible esta publicación.

Este número contiene artículos de académicos y estudiantes de nuestro Departamento de Ingeniería Industrial (algunos de ellos incluso son consecuencia de trabajos finales de grado, tesis de magister o tesis de doctorado), y de investigadores del Instituto Milenio. En esta oportunidad también contamos con un artículo de la Universidad de Concepción y otro de la Universidad de Buenos Aires, en Argentina.

La revista muestra trabajos sobre la aplicación de técnicas modernas de gestión de operaciones, programación matemática, tecnologías de información, redes neuronales y teoría de juegos.

Nuestro objetivo a través de esta publicación es contribuir a la generación y difusión de las tecnologías modernas de gestión y administración. La revista pretende destacar la importancia de generar conocimiento propio en tecnología y administración para nuestras problemáticas, junto con adaptar las tecnologías foráneas a las realidades de nuestro país y de otros de la región.

Estamos seguros de que los artículos publicados en esta oportunidad muestran formas de trabajo innovadoras que serán de gran utilidad e inspiración para todos los lectores, ya sean académicos o profesionales, por lo que esperamos que esta iniciativa tenga la recepción que creemos se merece.

Atentamente,

Guillermo Durán
Editor

Web Mining: Análisis sobre la Privacidad del Tratamiento de Datos Originados en la Web

Juan D. Velásquez*
Lorena Donoso**

Resumen

Web mining es el concepto que agrupa a todas las técnicas, métodos y algoritmos utilizados para extraer información y conocimiento desde los datos originados en la Web (web data). Parte de estas técnicas apuntan a analizar el comportamiento de los usuarios, con miras a mejorar continuamente la estructura y contenido de los sitios que son visitados. Detrás de tan altruista idea, es decir, ayudar al usuario a que se sienta lo mejor atendido posible por el sitio web, subyacen una serie de metodologías para el procesamiento de datos, cuya operación es al menos cuestionable, desde el punto de vista de la privacidad de los usuarios de un sitio web determinado. Entonces surge la pregunta ¿hasta donde el deseo por mejorar continuamente lo que se ofrece a través de un sitio web puede vulnerar la privacidad de quien lo visita?. El uso de herramientas de procesamiento de datos poderosas como las que contempla el web mining, puede atentar directamente contra la privacidad de los actos de los usuarios de un sitio web. En este trabajo se analizará el problema del procesamiento de los web data, para concluir con cuáles serían las técnicas que vulneran la privacidad de los usuarios que visitan un sitio y cuales no. A la luz de esta información, se podrá orientar mejor a los profesionales de las TICs para aunar esfuerzos en la mejora de la estructura y contenidos de la Web, pero sin vulnerar la privacidad de las personas.

Palabras Clave: *Web Mining*, *Privacidad*, *Regulación*.

*Departamento de Ingeniería Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile.

**Centro de Derecho Informático, Facultad de Derecho, Universidad de Chile, Santiago, Chile.

1. Introducción

Desde los orígenes de la Web, la creación de un sitio no ha sido un proceso fácil. Muchas veces se requiere de un equipo multidisciplinario de profesionales avocados a una sola misión "asegurar que el contenido y la estructura del sitio le son atractivos al usuario". Lo anterior es la clave del éxito para obtener una adecuada participación en el mercado electrónico, mantener la vigencia del sitio y sobre todo, lograr la tan ansiada y difícil fidelización del cliente digital [9].

La personalización implica que de alguna forma se puede obtener información respecto de los deseos y necesidades de las personas, para luego preparar la oferta correcta en el momento correcto [5]. Entonces, tal vez la solución para entender mejor al cliente digital sería someterlo a varias encuestas de opinión vía e-mail o al llenando formularios electrónicos. Sin embargo, la práctica ha demostrado que los usuarios no gustan de llenar formularios, contestar e-mails con preguntas, etc., a menos que se trate de algún amigo o familiar que quiera ayudar en el análisis, lo cuál no sería un caso real.

Cualquier análisis serio que se pretenda hacer respecto del comportamiento de navegación y preferencias que tiene un usuario en la Web, requiere del uso de datos reales, originados por usuarios reales. La pregunta entonces es: "¿de dónde saldrán estos datos?" . La respuesta es simple: de la misma Web. Ahora el cómo extraer estos datos y procesarlos para obtener un nuevo conocimiento acerca de los usuarios, es el gran desafío detrás de la personalización de la Web [2].

¿Hasta dónde este afán por analizar al usuario en la Web no se transforma en una persecución?. Con los web data adecuados, se puede hacer un completo seguimiento a todas las actividades de los usuarios en la Web, es decir, invadir directamente su privacidad, sin que estos se den cuenta de que están siendo vigilados por un "gran hermano" cibernético [4]. Evidentemente, la tecnología tiene dos caras, una de ellas muy siniestra y que la historia ha demostrado que si no se le regula adecuadamente, se pueden caer en excesos que atentan contra los derechos fundamentales de las personas [14].

En este trabajo se analiza hasta donde las herramientas usadas en web mining vulneran la privacidad de los usuarios, desde un punto de vista de la regulación nacional e internacional. En tal sentido, el capítulo 2 muestra la estructura que poseen los datos que se original en la web. El capítulo 3 muestra cómo estos datos son limpiados y preprocesados antes de que se le apliquen técnicas de web mining, las cuales son detalladas en el capítulo 4. Posteriormente, el capítulo 5 presenta un marco legal genérico respecto de la

regulación nacional e internacional que podría ser utilizada en el contexto del análisis de la privacidad que pueden vulnerar las técnicas de web mining, para luego establecer, en el capítulo 6, una revisión especifica de la regulación existente en torno a la personalización de la Web. Con todos estos antecedentes, el capítulo 7 centra su análisis en aquellas técnicas que estarían vulnerando en mayor medida la privacidad de los usuarios. Finalmente, el capítulo 8 muestra las principales conclusiones del presente artículo.

2. Naturaleza de los datos originados en la Web

La Figura 1 muestra en forma simple el funcionamiento de la Web. El servidor web o web server (1) es un aplicación que está en ejecución continua, atendiendo requerimientos (4) de objetos web, es decir, el conjunto de archivos que conforman el web site (3) y enviándoselos (2) a la aplicación que hace la solicitud, generalmente un web browser (6). En general estos archivos son imágenes, sonidos, películas páginas web que conforman la información visible del sitio. Las páginas están escritas en Hyper Text Markup Language (HTML), que en síntesis es un conjunto de instrucciones, también conocidas como "tags" (5), acerca de cómo desplegar objetos en el browser o dirigirse a otra página web (hyperlinks). Estas instrucciones son interpretadas por el browser, el cual muestra los objetos en la pantalla del usuario [1].

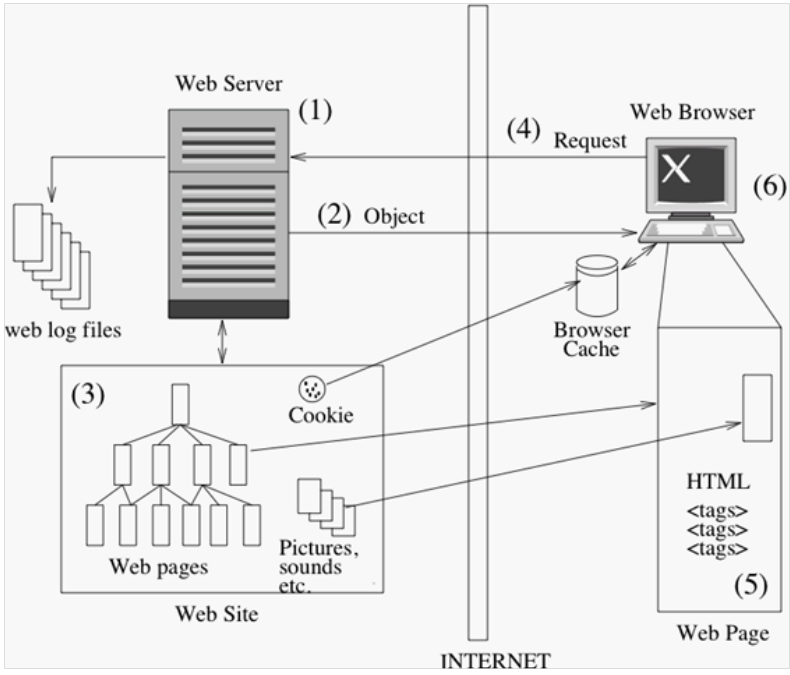


Figura 1: Modelo básico de operación de la Web

Cada uno de los tags presentes en una página, son interpretados por el browser. Algunos de estos tags hacen referencia a otros objetos en el web site, lo que genera una nueva petición en el browser y la posterior respuesta del server. En consecuencia, cuando el usuario digita la página que desea ver, el browser, por interpretación secuencial de cada uno de los tags, se encarga de hacer los requerimientos necesarios que permiten "bajar" el contenido de la página al computador del usuario.

La interacción anterior, ha quedado registrada en archivos conocidos como "web log files" [2], con lo cual es posible saber aproximadamente qué objetos fueron requeridos por un usuario, reconstruir su sesión y en la práctica realizar un verdadero seguimiento a sus actividades de navegación, analizando los contenidos visitados, el tiempo que se ha invertido en ello, qué información no atrae su interés, etc. La Figura 2 muestra un ejemplo del contenido y estructura de un archivo de web log, comenzando por la dirección IP del visitante del sitio, los parámetros ID y Authority (A), que son una forma de autenticar al usuario, siempre y cuando se especifique esa opción en el sitio web; la fecha y hora de conexión (Time), el método de obtención de la página, estatus de la petición, datos transferidos, de qué página procede el usuario (Referer) y finamente el tipo de software utilizado para navegar (Firefox, Mozilla, Explorer, etc.).

#	IP	Id	A	Time	Method/URL/Protocol	Statu	Byte	Referer	Agent
1	164.77.129.50	-	-	12/Apr/2003:23:47:44	GET /img/tab.gif HTTP/1.1	200	89	http://www.thebank.cl	MSIE 6.0; Windows 98
2	200.28.206.200	-	20	12/Apr/2003:23:48:31	GET transa/info.htm HTTP/1.1	200	144	/infoeco/info.html	MSIE 4.01; Windows 95
3	200.86.248.170	-	-	12/Apr/2003:23:48:37	GET /img/gen.gif HTTP/1.1	304	0	/ofert/wines/	MSIE 6.0; Windows 98
4	66.249.65.97	-	-	12/Apr/2003:23:48:41	GET /index.htm HTTP/1.1	200	88	-	Googlebot/2.1; google.com/bot.html
5	216.241.8.179	-	31	12/Apr/2003:23:50:03	GET /tx/infoeco/card.htm HTTP/1.1	200	210	/tx/infoeco/prom/	MSIE 6.0; Windows NT 5.1
6	164.77.129.50	-	-	12/Apr/2003:23:48:34	GET /tx/infoeco/ HTTP/1.1	200	186	/tx/infoeco/card.htm	MSIE 6.0; Windows 98
7	200.28.206.200	-	20	12/Apr/2003:23:51:13	GET transa/account.htm HTTP/1.1	200	180	/transa/info.htm	MSIE 4.01; Windows 95
8	216.241.8.179	-	31	12/Apr/2003:23:51:23	GET /tx/infoeco/ind.htm HTTP/1.1	200	300	/tx/infoeco/card.htm	MSIE 6.0; Windows NT 5.1
9	200.86.248.170	-	-	12/Apr/2003:23:51:41	GET /prom/wine.html HTTP/1.1	404	0	/ofert/wines/	MSIE 6.0; Windows 98
10	164.77.129.50	-	44	12/Apr/2003:23:52:04	GET /tx/infoeco/ind.htm HTTP/1.1	200	186	/tx/infoeco/	MSIE 6.0; Windows 98

Figura 2: Parte de un web log file

Como se puede apreciar, cada registro da cuenta de los movimientos de un usuario en un sitio web. En consecuencia, y en forma casi anónima, los datos generados en el sitio web son tal vez la mayor encuesta que podría tener una empresa por sobre sus eventuales clientes, analizando sus preferencias de información, las cuales están directamente relacionadas con las características de los productos y servicios ofrecidos.

El proceso anterior, no está exento de desafíos, siendo el primero de ellos la preparación de los datos para un proceso de extracción de información. En efecto, los web data, como se les conoce, consideran todos los tipos de datos existentes, lo cual dificulta su procesamiento. Adicionalmente, no siempre contienen datos relevantes e incluso algunos de ellos son más bien ruido, por lo cual se requiere de su pre-procesamiento y limpieza antes de que la información salga a la luz. El procesamiento considera la reconstrucción de la sesión del usuario, la limpieza del contenido de las páginas web, para la identificación

de elementos relevantes (textos clave, imágenes, sonidos, etc.,) y en general la transformación de los web data en vectores de características que modelen el comportamiento del usuario en un sitio web particular.

3. Limpieza y preprocesamiento de los web data

Los datos originados en la Web o web data, corresponden esencialmente a tres fuentes [2]:

- Contenido: Son los objetos que aparecen dentro de una página web, por ejemplo las imágenes, los textos libres, sonidos, etc.
- Estructura: Se refiere a la estructura de hipervínculos presentes en una página.
- Uso: Son los registros de web logs, que contienen toda la interacción entre los usuarios y el sitio web.

Algunos autores argumentan que también se debieran considerar los datos de "perfil de usuario" como parte de los web data [2] y [3]. Se trata de datos personales como el nombre, edad, sexo, etc., los cuales en rigor no deben ser tratados en un proyecto web mining, correspondiendo su procesamiento al uso de otras herramientas (data mining), por lo que no se les considerará en el presente trabajo.

Los web data deben ser pre procesados antes de entrar en un proceso de web mining, es decir, son transformados en "vectores de características" que almacenan la información intrínseca que hay dentro de ellos [3] y [10].

Aunque todos los web data son importantes, especial atención reciben los web logs, ya que ahí se encuentra almacenada la interacción usuario sitio web, sus preferencias de contenido y en síntesis su comportamiento en el sitio. Por esta razón, y concediendo de que es posible que en los otros web data se pueda albergar información que identifique a los usuarios, nos concentraremos esencialmente en los web logs, como fuente de mayor controversia al momento de analizar el comportamiento de los usuarios.

La primera etapa, entonces, corresponde a la reconstrucción de la sesión del usuario a partir de los datos existentes en los registros de web log. Este proceso se denomina sesionización y puede resumirse en los siguientes pasos [2]:

1. Limpiar los registros de los web logs, dejando sólo aquellos concernientes a peticiones de páginas web, eliminando los que indican peticiones de objetos contenidos en éstas

2. Identificar registros de peticiones hechas por web crawlers. Para ello, existen listas oficiales y no oficiales de crawlers en la Web. Pueden ser identificados a través del campo Agent o, en su defecto, por la IP Address.

3. Agrupar los registros por IP Address y por Agent. Cabe señalar que se está asumiendo que no hay más datos acerca de los usuarios que visitan el sitio.

4. Ordenar los registros de menor a mayor Timestamp, de modo que los registros aparezcan cronológicamente.

5. Identificar las sesiones de navegación, para lo cual existen dos alternativas:

   a) Utilizar un criterio estadístico, es decir, asumir que por lo general las sesiones reales de los usuarios no duran más allá de 30 minutos.

   b) Asumir que ninguna sesión tiene páginas visitadas más de una vez.

6. Por último, reconstruir la sesión usando la estructura de hipervínculos del sitio y completando aquellas páginas que no fueron registradas debido a que se utilizó la memoria caché del web browser o la caché corporativa de un servidor proxy.

Existen dos estrategias a seguir para realizar el proceso de sesionización [2],[3] y [9]:

1. **Estrategia Proactiva**: Consiste en utilizar herramientas invasivas para identificar a los usuarios, generalmente se trata de cookies[1] o spybots[2], las cuales permiten identificar al usuario con un identificador único, de modo que es posible conocer su frecuencia de visitas y cómo ha variado su comportamiento en el tiempo. La efectividad de estas técnicas no es muy clara, sobre todo por que existen programas especializados en borrar spybots y la mayoría de los navegadores puede eliminar las cookies y no permitir su funcionamiento.

2. **Estrategia Reactiva**: Consiste en la utilización de los web logs como fuente única de datos para la reconstrucción de sesiones, evitando atentar contra la privacidad de los usuarios. Si bien es la estrategia que provee una riqueza potencial de información menor, pues no se puede identificar al usuario y su frecuencia de visitas, es la que puede ser seguida en cualquier sitio.

Cabe señalar que ciertos sitios han cambiado su estructura con el propósito de identificar a sus visitantes [2],[9] y [12]. Una primera estrategia consiste en implementar un sistema username/password, que promueva el registro de los

[1]Fragmento de información que se almacena en el disco duro del visitante de una página web a través de su web browser, a petición del web server de la página.

[2]Programas que entregan directamente la secuencia de visitas realizadas por el usuario en toda la Web.

usuarios a cambio de nuevos servicios. Sin embargo, sólo es posible reconstruir perfectamente las sesiones de los registrados, quedando los no registrados en el anonimato. Otra estrategia consiste en utilizar páginas dinámicas en el sitio. Con ellas, cada solicitud de abrir una página genera un identificador único para el usuario, sin embargo, ello obliga a reconstruir el sitio y trae complejidades para identificar qué está realmente viendo el visitante, dadas las direcciones URL dinámicamente generadas [3].

4. Web mining

El concepto web mining, agrupa a todas las técnicas, algoritmos y metodologías utilizadas para extraer información y conocimiento desde los web data. En este sentido, se podría decir que es la aplicación de teoría del data mining al caso particular que revisten los web data [7].

Web Mining ha permitido estudiar la creciente cantidad de datos disponibles, donde la estadística clásica y la revisión manual ya resultan ineficientes [5]. La importancia de tener datos limpios y consolidados radica en que la calidad y utilidad de los patrones encontrados por estas herramientas dependerán directamente de los datos que sean utilizados. Por este hecho es que muchas veces los procesos de web mining dan lugar a información errónea, pues a diferencia de las herramientas estadísticas, existen herramientas a disposición de cualquier usuario con poco conocimiento de este campo.

Dentro de las técnicas de Data Mining más utilizadas se puede mencionar [7]:

1. *Reglas de asociación*: Consiste en encontrar correlaciones entre conjuntos de datos bajo una probabilidad de ocurrencia (confianza) para una proporción del total de datos (soporte). Por ejemplo: pan queso (soporte = 5 %, confianza = 42 %) quiere decir que 42 % de las personas que compraron pan también compraron queso y que esta combinación se dio en 5 % de las transacciones. Una extensión a ésta es la asociación multidimensional, dónde se asocia más de un atributo a otro. Ejemplo: pan, mantequilla y queso.

2. *Clasificación*: Consiste en clasificar una serie de registros en alguna categoría previamente definida. Para ello, se suele usar un proceso de aprendizaje, en el cual el algoritmo es aplicado sobre datos ya clasificados, de modo que pueda establecer bajo qué valores de los otros atributos del registro, se trata de una categoría u otra. Una vez realizado el aprendizaje, se procede a efectuar la clasificación sobre registros que no fueron utilizados como input en el aprendizaje.

3. *Clustering*: Consiste en agrupar objetos que tienen características similares, a diferencia de la técnica anterior en el clustering no se conoce la categorización a priori y se espera encontrarla a partir de los datos. Para ello se utiliza una medida de similitud entre registros, que permite separarlos de acuerdo a sus diferencias en esta medida. Existen principalmente 3 técnicas de clustering:

   a) *Particionado*: bajo esta técnica se definen a priori los n clusters en los que se clasificarán los registros.

   b) *Jerárquico*: esta técnica construye los clusters mediante una descomposición jerárquica que puede ser de dos tipos: Aglomerativa, en el que se parte con un cluster por cada registro y estos empiezan a agruparse de acuerdo a la medida de similitud utilizada hasta una condición terminal previamente definida y Divisiva, en el que se parte con un sólo cluster que incluye todos los registros y este empieza a separarse de acuerdo a la medida de similitud utilizada hasta una condición terminal previamente definida.

   c) *Basado en densidad*: Esta técnica toma prestada la definición de densidad de la Física. Consiste en definir una densidad umbral, que no es más que una cardinalidad predefinida para cada cluster, y un radio, que no es más que una distancia predefinida, de modo que los clusters se van formando por registros a una distancia del centroide del cluster menor al radio, los centroides son redefinidos en cada iteración y el algoritmo para cuando todas las cardinalidades de los clusters encontrados son menores a la densidad umbral previamente definida.

5. Marco legal para el análisis de la privacidad en la web

La legislación internacional ha recogido una de las preocupaciones más relevantes de los constitucionalistas de la nueva sociedad, cual es el grado de afectación a la libertad y a la dignidad humana por el nivel de control personal que se puede alcanzar gracias al almacenamiento y tratamiento de datos personales. Cuando ello alcanza al control de los datos que hoy en día transitan a través de Internet la preocupación se hace más aguda, en tanto que a través de esta red podría llegar a conocerse prácticamente toda la información de la persona, incluso su ubicación y estado de salud o situación emocional, todo ello en tiempo real.

5.1. Antecedentes

Desde la perspectiva jurídica, en Chile y la mayoría de los países es susceptible de ser calificado como "dato personal?? cualquier información relativa a personas naturales identificadas o susceptibles de ser identificadas, lo que dependerá del avance de la técnica y de la potencialidad identificativa del dato. En este contexto, los elementos que conforman el concepto de dato personal son los siguientes:

- **Toda información**. Se trata de un concepto amplio, que abarca tanto, imagen, sonido, muestras biológicas, incluso al conjunto de caracteres grafológicos que proporcionen antecedentes que coadyuven al conocimiento de un sujeto.

- **Relativos a una persona Natural**. Si bien hay legislaciones (como la de Argentina) que admiten la protección de datos personales respecto de las personas jurídicas, la generalidad de los países sólo admite la protección de datos personales de personas físicas. Ello principalmente porque hacen derivar la garantía desde los derechos fundamentales emanados de la dignidad humana.

- **Identificada o identificable**. En el ámbito europeo, se ha dicho será identificable "Toda persona cuya identidad pueda determinarse directa o indirectamente, en particular mediante un número de identificación o uno o varios elementos específicos, característicos de su identidad física, fisiológica, psíquica, económica, cultural o social[3]". De su parte, en la legislación española se ha dicho al respecto que estaremos frente a una persona identificable cuando estemos en presencia de "cualquier elemento que permita determinar directa o indirectamente la identidad física, fisiológica, psíquica, económica, cultural o social de la persona física afectada[4]" . En consecuencia el concepto identificable debe adaptarse en cada momento a la realidad tecnológica y social, pues hoy en día existen mecanismos hábiles para la identificación del sujeto, a partir de diversas circunstancias, los que son accesibles al común de las personas y, por tanto, no imponen un esfuerzo exorbitante al responsable del tratamiento o, en su caso, al encargado del tratamiento de datos personales.

Los datos personales podrán ser públicos o sensibles[5], siendo estos últimos aquellos que gozan generalmente de un mayor nivel de protección. Si bien no podemos generalizar respecto de qué datos ocupan una u otra categoría, pues ello dependerá de las condiciones y usos sociales de cada país, lo que queda

[3] 2 letra a) Directiva 95/46 CE, del Parlamento y del Consejo, de 1995

[4] Así se ha reconocido desde el R.D. 1332/94, Artículo 1, inciso 5.

[5] En la nomenclatura europea, estos son "datos especialmente protegidos". Véase al respecto la Directiva 95/46 CE, del Parlamento y del Consejo, de 1995.

claro es que el factor que determina la sensibilidad es el riesgo de que el uso del dato acarree perjuicios a su titular, básicamente por las posibilidades de que terceras personas adopten decisiones arbitrarias a su respecto. De ahí que por regla general y a vía de ejemplo los datos sobre salud, los relativos a las ideologías políticas, al credo religioso y/o la vida sexual sean considerados como datos sensibles.

Otro aspecto conceptual que es importante considerar es que la legislación entiende que el **tratamiento de datos personales** comprende cualquier operación **manual o automatizada**, que se realice sobre los datos personales, desde la recogida y hasta la cancelación definitiva de los datos personales.

Finalmente dentro del contexto es importante destacar que la legislación entiende que el dueño del dato personal es siempre su titular, esto es, la persona a quien se refiere el dato, lo cual no se altera por el hecho de que el titular del banco o base de datos en la cual consta ese dato, le haya agregado valor o lo haya integrado con otros datos personales a través de reglas de negocio específicas y/o procesos tecnológicos de tratamiento de datos.

5.2. Los datos personales y la Web

El tratamiento de datos personales en/o a través de la web conlleva problemáticas de derecho internacional, en el sentido que podrán producirse situaciones en las que los servidores se encuentren en un país diferente a aquel en que se encuentra o pertenece el sujeto fuente de datos. Es más, la situación puede ser más compleja en tanto que el titular del banco de datos puede asimismo estar en un país diverso al de localización del servidor. Esto es más evidente conforme avanzan los sistemas y procesos que adscriben al concepto de Cloud computing.

Para los efectos de responder a la interrogante sobre cuál es el estatuto jurídico de estas actividades de tratamiento de datos personales hemos atendido a lo regulado en la Unión Europea y en Estados Unidos. Ello principalmente porque se han pronunciado expresamente sobre estas materias, mientras en nuestro entorno el debate aún no se ha iniciado.

Al respecto es importante considerar que en la Unión Europea, el Grupo del Artículo 29 del tratado constitutivo de la Unión ha sostenido que si el sitio web o motor de búsquedas o recogidas de datos accede a datos de habitantes de Europa, debe aplicárseles la legislación europea de tratamiento de datos[6].

En Estados Unidos, la Children's Online Privacy Protection Act (COPPA) de 1998 extiende su ámbito de aplicación a los sitios web extranjeros que recogen información personal de niños menores de 13 años establecidos en el

[6]Véase al respecto 5035/01/ES/Final WP 56, Documento de trabajo relativo a la aplicación internacional de la legislación comunitaria sobre protección de datos al tratamiento de los datos personales en Internet por sitios web establecidos fuera de la UE

territorio de los Estados Unidos.

Siguiendo estos criterios, nos queda claro que la tendencia internacional es considerar que la legislación aplicable en esta materia es aquella que corresponde a la del titular de los datos en el momento de la recogida, lo que deberá resolverse de acuerdo a las reglas generales del Derecho internacional.

5.3. Proveedores de medios de conexión y Almacenamiento

Jurídicamente cada uno de los servidores contectados a la red de redes actúa como emisor/receptor de datos (contenidos) y se les denomina nodos de la red. Los servicios de transferencia de información a través de la red suponen la existencia de proveedores de medios de interconexión y de almacenamiento de información. Ambas categorías integran lo que se ha dado en llamar **ISP (Internet Service Providers)** o **PSI (Proveedores de Servicios Internet)**.

En este punto es importante destacar que el marco normativo nacional está contenido en la Resolución Exenta No. 1483, de 22 de octubre del año 1999, que "Fija Procedimiento y Plazo Para Establecer y Aceptar Conexiones Entre ISP". Como podemos apreciar del solo título de la norma ya se desprende que su ámbito de aplicación está restringido a la interconexión entre ISP concebidos en ella como personas naturales o jurídicas que actúan como Proveedores de Acceso a Internet[7], que es aquel que "permite acceder a la información y aplicaciones disponibles en la red Internet"[8].

Estos proveedores proporcionan a los usuarios **capacidad de almacenamiento de datos (Hosting), capacidad de conexión**, esto es aquella que permite hacer transitar la información desde su punto de origen hasta su destino, y que en definitiva son aquellos servicios asociados a las redes de telecomunicaciones.

En consecuencia estos servicios podrán clasificarse como: **Proveedor de Acceso**, que permiten la conexión a la red y por tanto el funcionamiento del usuario como un nodo temporal o continuo; **Proveedores de enlace**: que son aquellos que proporcionan los mecanismos de transmisión y finalmente los **Proveedores de Hosting**, que, como su nombre lo indica dan el servicio de proporcionar un espacio en disco que permite almacenar la información, dejándola a disposición de todo aquel que quiera acceder a ella a través de la Red. Estos últimos en todo caso, se califican en general como suministradores de servicios suplementarios.

Teniendo en consideración las escasas bases jurídicas con que se cuenta, debemos sostener que conforme al principio de **neutralidad tecnológica**, habrá de considerarle un servicio de telecomunicaciones[9] y en consecuencia

[7]Resolución Exenta 1483, 1999, Subsecretaría de Telecomunicaciones, Artículo 1 letra c)
[8]Resolución Exenta 1483, 1999, Subsecretaría de Telecomunicaciones, Artículo 1 letra a)
[9]El Artículo 1 de la ley 19.628 define servicio de telecomunicaciones en los siguientes

seguir el principio de "aplicación de la misma regulación a los servicios de telecomunicaciones con independencia de la tecnología utilizada para la prestación de los mismos"[10].

El marco básico a aplicar será la ley 18.168, general de telecomunicaciones en lo que respecta al servicio propiamente tal y por la ley 19.628, de tratamiento de datos personales en lo que se refiere a lo datos personales que circulan por las redes con ocasión o gracias a este servicio y la normativa complementaria, por ejemplo en lo que se refiere a la obligación de captura y almacenamiento de datos personales para los efectos, por ejemplo, de investigación criminal. Para esta calificación nos basamos en que los medios empleados para realizar esta transmisión son las redes de telecomunicaciones, las que para estos efectos se intercomunican mediante el protocolo TCP/IP. Considerado así, las comunicaciones a través de las cuales opera Internet responden al concepto de Telecomunicación y en consecuencia la instalación, operación y explotación de los servicios de telecomunicaciones ubicados en el territorio nacional, incluidas las aguas y espacios aéreos sometidos a la jurisdicción nacional y, en lo que les sea aplicable, los sistemas e instalaciones que utilicen ondas electromagnéticas con fines distintos a los de telecomunicación[11], quedan bajo la Subsecretaría de Telecomunicaciones como el organismo técnico cuyas finalidades principales son la aplicación y control de la ley y sus reglamentos, la interpretación técnica de las disposiciones legales y reglamentarias que rigen las telecomunicaciones[12], sin perjuicio de las facultades que establece la legislación respecto de otros órganos administrativos o judiciales.

En el contexto que nos interesa, los proveedores del servicio de acceso a Internet son responsables de la custodia de datos personales a los que tienen acceso y que son objeto de tratamiento con ocasión del servicio prestado. Estos serán principalmente aquellos relativos a las navegaciones, comunicaciones efectuadas y en general, los logs generados en el proceso comunicativo. Asimismo, en términos generales los PSI quedan expresamente exonerados de monitorizar la red o el servidor que administran, por entender que dicha labor es imposible ante el volumen de información existente.

Ahora bien, en cuanto al papel de los proveedores de medio en la cadena de

términos: Para los efectos de esta ley, se entenderá por telecomunicaciones toda transmisión, emisión o recepción de signos, señales, escritos, imágenes, sonidos e informaciones de cualquier naturaleza, por línea física, radioelectricidad, medios ópticos u otros sistemas electromagnéticos.

[10]Comentarios de Aniel en Relación con la Comunicación de la Comisión Europea, COM (1999) 539, "Revisión 1999 del Sector de las comunicaciones" en línea en http://europa.eu.int/ISPO/infosoc/telecompolicy/review99/comments/aniel22b.htm [Consulta: 28 Ago. 2004]. En el mismo sentido Directiva 2002/20/CE del Parlamento Europeo y del Consejo, relativa a la autorización de redes y servicios de comunicaciones electrónicas, de 7 de marzo de 2002 (Directiva autorización)

[11]4 Ley 18168, incisos 1 y 2.

[12]Art. 6 incisos 1 y 2 ley 18168

responsabilidad por los contenidos, se ha reconocido que ellos "no controlan de manera directa el contenido disponible en Internet ni qué parte deciden consultar sus clientes"... puede que sea preciso cambiar o clarificar la legislación para ayudar a los suministradores de acceso y los suministradores de servicio de ordenador central, cuya ocupación primordial es prestar servicio al cliente, a abrir un camino que evite, por un lado las acusaciones de censura, y por otro, los riesgos de acciones judiciales"[13].

5.4. Regulación y web mining

La idea básica que subyace detrás del web mining es la extracción de información y conocimiento desde un conjunto de web data. Dependiendo del tipo de web data a minar, el algoritmo de web mining puede estar altamente relacionado con los datos personales del usuario. Lo anterior plantea muchas interrogantes, sobre todo en lo referente a la privacidad del usuario.

Partamos analizando el o los archivos de web log, en especial la dirección IP desde donde accedió el usuario al sitio web. Este parámetro en combinación con otros datos existentes en el registro de web log, ha sido frecuentemente utilizada para identificar la sesión del usuario. Debido a la posibilidad de que se pueda relacionar o identificar a la persona a través de la dirección IP que utiliza para navegar por la Web, es que en la UE se está comenzando a considerar a la IP como un dato personal[14]. En España, la Ley Orgánica 15/1999, en su artículo 3a define al dato personal como "cualquier información concerniente a personas físicas identificadas o identificables".

El TCP/IP versión 4, que es el protocolo con que en la actualidad opera Internet, fue concebido para identificar un computador conectado a la red. Hay que recordar que Internet (Interconection Network) es una "red de redes", así que para identificar un computador, primero se identifica a qué red pertenece. De esta forma, las direcciones IP están compuestas de cuatro números (rango entre 0 y 255 cada uno) con los que se identifica la red y el computador dentro de ésta.

Entonces, por construcción la dirección IP no fue creada para identificar a la persona detrás del computador. Mucho menos ahora que existen sistemas que permiten a varios usuarios acceder a Internet, usando la misma IP y que los ISP entregan direcciones dinámicas, es decir, sólo relacionan una sesión de usuario mientras que está conectado. Incluso más, es posible que durante la sesión, el usuario experimente cambios en la IP asignada. Sin embargo, si se realizan los cruces de datos adecuados, se puede llegar a una aproximación respecto de quien sería la persona que en un determinado momento, estaba

[13]Comunicación de la Comisión Europea al Parlamento Europeo, al Consejo, al Comité Económico y Social y al Comité de las Regiones"de fecha 16 de octubre de 1996, (COM(96)487), Acápite Nos. 1 y 2.

[14]http://www.habeasdata.org/Ipcomodatopersonal

conectada desde un computador, usando una IP específica. Si lo anterior es probatorio en un tribunal, es un tema que escapa a las pretensiones de este estudio, pero hay que dejar en claro de que, desde el punto de vista técnico, no habría una certeza del 100 % de que una persona accedió a un sitio desde una IP determinada.

El escenario anterior debería cambiar una vez que la nueva generación del protocolo TCP/IP, la versión 6, entre en total funcionamiento en Internet, por cuanto se podrán implementar otros recursos de identificación de la persona, por ejemplo transmisión de datos cripto-segurizados y con firma digital.

Asumiendo, entonces, que a través de la dirección IP sólo se puede identificar la sesión y no a la persona que hay detrás, los algoritmos de web mining se orientan a extraer información desde los web data para analizar comportamientos de usuarios en determinados momentos del día, es decir, un mismo usuario se puede comportar diferente en momentos diferentes, con lo cual se argumenta que no se estaría analizando a la persona, sino más bien a grupos de personas para extrapolar comportamientos colectivos [7].

Como en todo aquello donde el ser humano no encuentra consenso, aparecen posturas divididas. Por una parte, los especialistas que aplican algoritmos de web mining para lograr un mejor entendimiento de las preferencias de los usuarios de un sitio web, asumen que todo el procesamiento de los web data es inocuo para el usuario. Por otro lado, si el usuario se entera de que todos sus movimientos en el sitio están siendo monitoreados, tal vez decide no visitar el sitio, porque siente que a través del uso de estas herramientas se produce una intromisión directa en su privacidad.

A partir del uso de algoritmos de web mining, se pueden extraer patrones respecto del comportamiento de grupos de usuarios en la Web. En este sentido el valor del "individualismo" podría verse afectado. Este concepto se relaciona con el de privacidad por cuanto muchos sistemas que usan los patrones extraídos a través del web mining, tienden a clasificar a los usuarios y a tomar decisiones en base a cuan parecido es su comportamiento respecto de un grupo. Por ejemplo, este usuario se comporta como aquellos que pertenecen al grupo de los amantes del rock, entonces las páginas a mostrarle en su navegación son sólo las referentes a ese tipo de música. Lo anterior claramente coarta toda posibilidad al usuario de que pueda tomar decisiones respecto de lo que en realidad quiere ver [13].

Si se analizan los web data utilizados para la extracción de patrones de navegación y preferencias de lo usuarios, estos se pueden agrupar en [6]:

- *Datos explícitos.* Son provistos por los usuarios en forma directa, por ejemplo, su nombre, edad, nacionalidad, etc.
- *Datos implícitos.* Se infieren a partir del comportamiento del usuario en un sitio web, por ejemplo, qué paginas visitará, historia de compras, etc.

¿Qué parte de los web data es público y privado? La respuesta depende casi exclusivamente del país donde se haga la pregunta. Algunos países industrializados, han abordado la privacidad de los datos creando regulaciones específicas, por ejemplo en EEUU se crea la Self-Regulatory Principles for Online Preference Marketing by Network Advisers. NetworkAdvertising Initiative del año 2000, donde la legislación cubre muy pocos tipos de datos, pero se espera que esto cambie rápidamente.

Cabe se?a?ar que el dueño o mantenedor de un sitio, es amo y señor de los registros de web logs que se generen producto de la navegación de los usuarios. Perfectamente podría comercializar estos datos, pero sería muy extraño, pues estaría abriendo al mundo la mayor ventaja competitiva que puede tener una empresa en el mundo digital "conocer el comportamiento de sus clientes virtuales" [8], [9] y [11].

Visto lo anterior, la sola visita de un usuario a un sitio expone la privacidad de su navegación al dueño de los web logs. Lo mismo sucede cuando entramos en una tienda con cámaras de vigilancia. El fin altruista puede ser proteger al cliente ante los robos, pero igual este pierde intimidad al ser filmado, toda vez que su comportamiento de compra también puede ser estudiado para luego formular reglas de fidelización, promociones, etc. Desde el derecho la respuesta es clara, las filmaciones son operaciones de tratamiento de datos personales y por tanto deberán respetar los estándares legislativos correspondientes. Por tanto esas filmaciones sólo podrán ser utilizadas en el marco de la finalidad para las que fueron recogidas.

Otro punto muy importante a dejar en claro, es que los registros de web log no pueden identificar a una persona, pero si a un usuario web, es decir, un ente que posee una dirección IP desde donde se conecta, fecha y hora de visita, las páginas que visitó etc. En este sentido, en forma directa no se estaría trabajando con datos personales, por lo que la regulación que existe al respecto, podría ser insuficiente para los web data.

La utilización de mecanismos de identificación, tales como las conocidas cookies, podría establecer una relación directa entre el ser humano y el usuario web. Sin embargo, es posible que un tercero use el computador de una persona y sin desearlo la suplante en el sitio web que visita, ya que estaría usando la misma cookie que su antecesor.

Otro caso de vinculación usuario web/persona se produce en los ISP[15]. En efecto, cuando contratamos el servicio Internet, datos personales respecto de nosotros quedan consignados en un contrato. Luego para una determinada sesión, el ISP sabe al menos a través de que conexión el cliente está navegando por la Web. Sin embargo, dado que una conexión puede ser compartida, es decir, varias personas saliendo por un mismo lugar, nuevamente no es posible vincular una determinada sesión a un usuario.

[15]Internet Service Provider

La tendencia mundial en mejores prácticas para el tratamiento de los web data, especifica que se debe [6]:

- Informar al usuario que está entrando en un sistema informático el cual por construcción almacenará datos respecto de su navegación en el sitio y que dichos datos pueden ser usados para hacer estudios posteriores.
- Obtener el consentimiento explícito del usuario para realizar una operación de personalización del sitio web que visita. Por ejemplo "¿desea usted que le enviemos sugerencias de navegación?".
- Proveer una explicación sobre las políticas de seguridad que se aplican para mantener los web data que se generen en el sitio.

Estas prácticas, son un marco mínimo de requerimientos para asegurar una adecuada privacidad del usuario en el tratamiento de los web data.

En Chile, la ley 19.628 sobre datos personales, consagra como tales a "los relativos a cualquier información concerniente a personas naturales, identificadas o identificables". En su sentido amplio, los web data no estarían contemplados como dato personal, salvo los referentes a las direcciones IP que podrían ser utilizadas, en combinación con otros datos para identificar a la persona detrás de la sesión del usuario. Entonces, el tratamiento de los web data podría estar regulado por la citada ley, siendo el responsable del banco de datos, el administrador o dueño del sitio web que el usuario visita.

6. Privacidad en la personalización de la Web

La personalización de la Web es la rama de la investigación en "Web Intelligence" dedicada a ayudar al usuario a que pueda encontrar lo que busca en un sitio web [5] y [9]. Para esto, se han desarrollado sistemas informáticos que ayudan a los usuarios a través de sugerencias de navegación, contenidos, etc. y más aun, entregan información valiosa a los dueños y administradores de sitios para que realicen cambios en su estructura y contenido, siempre con la idea de mejorar la experiencia del usuario, haciéndolo "sentir" como si fuese el visitante más importante del sitio, con una atención personalizada. Para lograr lo anterior, se han desarrollado múltiples esfuerzos tendientes a extraer información desde los web data que se generan con cada visita del usuario a un sitio, siendo los trabajos en web mining, los que han concentrado la mayor atención de empresas e investigadores en los últimos años.

Primero que todo, hay que dejar en claro el fin último que persigue el uso del web mining: aprender del comportamiento de los usuarios en la Web, para mejorar la estructura y contenido de un determinado sitio, personalizando la atención del usuario [11] y [13].

Como se puede apreciar, el fin es bastante altruista, siempre orientado a satisfacer al usuario y en el fondo a ayudarlo a encontrar lo que busca. Ahora bien, el exceso de "ayuda" no sólo puede molestar al usuario, sino que además, para ayudarlo mejor, se requiere de más y más datos, conocer sus preferencias y en buenas cuentas, intrometerse en su privacidad.

Existe evidencia empírica que los sitios que incorporan sistemas de personalización de sus contenidos, logran establecer una relación de lealtad con sus visitantes [6]. Sin embargo, el precio a pagar es permitir que el sistema se inmiscuya en aspectos relacionados con las actividades del usuario en el sitio, sus hábitos anteriores de navegación o de pares parecidos, etc. En algunos casos, el usuario puede llegar a experimentar una verdadera sensación de invasión su privacidad, lo que se traduce en otra razón más por la cual un usuario no visita un sitio web que personaliza la información que muestra a sus visitantes, es decir, el "remedio fue peor que la enfermedad", por lo que el desarrollo de este tipo de sistemas se está tomando con cautela, más allá de las implicancias legales que puede traer el vulnerar la privacidad de los actos de los visitantes de un sitio.

Entonces ¿hasta que punto la personalización de la Web es invasiva de la privacidad de los usuarios? [3].[4] y [6]. La percepción dependerá mucho de las características culturales de cada país o más aun, comunidad de individuos. La solución a la cual más se ha recurrido, es realizar encuestas de opinión a los usuarios de los sitios, pero que van más de acorde a las bondades que trae la personalización, sin explicar en detalle el cómo se logra.

La creación de sistemas para personalizar la navegación en la Web, limita el libre albedrío, por cuanto asume que el usuario no es lo suficientemente avezado como para encontrar información por si solo y necesita ayuda, que al final se transforma en una imposición sublime sobre qué debe ver. Ahora bien, la gran queja de los usuarios es que visitan un sitio y nunca encuentran lo que buscan, pese a que muchas veces el contenido si estaba. La "culpa" es compartida, por cuanto si el usuario no encuentra nada, tal vez se deba a su poca experiencia en la Web, y también es muy posible que el sitio esté mal estructurado y en realidad oculte información en vez de mostrarla [8].

Entonces, ¿dónde está el punto de balance entre vulnerar privacidad y ayudar al usuario?. Tal vez la solución sea muy simple, y todo pase por preguntarle al usuario si necesita apoyo y explicarle que para ayudarlo se requiere involucrarse un poco más en su vida privada.

Lamentablemente lo anterior en la Web es complicado, ya que muchas preguntas cansan al usuario y es ineficaz.

7. Análisis de la operación de las técnicas de minería de datos

Las técnicas de web mining analizadas, utilizan como entrada de datos los web data preprocesados y en forma de vectores de características. Como ya se ha comentado antes, de todos los posibles web data, son los registros de log los que más información aportan para realizar un análisis del comportamiento de los usuarios en un sitio web [10].

Previo al uso de estos registros, se requiere aplicar un proceso de reconstrucción de la sesión de los usuarios: la sesionización. Al respecto, la Tabla 1 muestra un resumen de las técnicas más utilizadas para sesionizar registros de log.

Desde un punto de vista de la privacidad de los web data, todo apunta a que el análisis del comportamiento del usuario debe hacerse utilizando estrategias de reconstrucción de la sesión que no liguen directamente a un ser humano con el usuario web [2] y [3]. En tal sentido, las técnicas más comúnmente aceptadas son las dos primeras de la Tabla 1. Sin embargo, la extracción de patrones de navegación y preferencia de los usuarios, siempre puede ser utilizada como una forma indirecta de extrapolar el comportamiento de un visitante en un sitio web, que a través de la personalización de sus contenidos, puede atentar contra el libre albedrío del usuario, toda vez que la información que verá no será toda la que puede ver, de eso la lógica informática del sitio se va a encargar, tal como "el gran hermano" que vela por lo bueno y lo malo que se le permite ver a las personas.

Luego, asumiendo que sólo se trabajará con datos que identifican sesiones pero no personas, se construyen los vectores de características. Los más frecuentemente usados, contienen información sobre la página visitada, el tiempo que el usuario gasta por página y sesión, más alguna referencia al objeto que se está visitando [12] y [13].

Las técnicas de web mining más frecuentemente usadas, apuntan a la identificación de grupos de usuarios con preferencias de navegación y contenidos similares (uso de clustering), las cuales no permiten identificar en forma directa a la persona detrás de la sesión.

El paso siguiente es analizar cómo usar los patrones que se pueden extraer desde los grupos de usuarios con características a fines. Desde un punto netamente informático, este "cómo" se transforma en reglas "if-then-else", que junto con los patrones, configuran el conocimiento extraído desde los web data [11].

La personalización de la Web se logra a partir del uso del conocimiento

Método	Descripción	Vulneración de la Privacidad	Ventajas	Desventajas
IP + Agente	Asume que cada par IP/Agente único es una sesión	Baja	Siempre disponible. No se requiere tecnología adicional	No garantiza unicidad. Las Ips pueden cambiar
Identificadores de sesión	Uso de páginas web dinámicas para asociar ID a cada hipervínculo	Baja a Media	Siempre disponible independiente de la IP	No puede capturar visitas repetidas. Sobre carga por generación de páginas dinámicas
Registro	Uso explícito de los logs del sitio web	Media	Se pueden seguir las acciones de la persona detrás del browser	Muchos usuarios no se registran.
Cookie	Almacena una ID en el computador del cliente	Medio a Alto	Se puede hacer seguimiento de varias visitas de un mismo browser	El usuario las puede deshabilitar
Robots espías	Código que se carga en el browser y envía toda la navegación del usuario a un destinatario	Alta	Se pueden seguir todos los movimientos de un usuario en el sitio	Existen software que permiten limpiar los robot espías del alojados en el browser

Tabla 1: Técnicas de reconstrucción de la sesión de un usuario.

extraído, el cual permite aplicar técnicas de clasificación de los usuarios por similaridades con los grupos identificados [10] y [11]. En síntesis, cuando un usuario visita el sitio, inmediatamente se procede a analizar su navegación para luego identificar a que grupo de usuarios pertenece. Luego, aplicando las reglas "if-then-else" se procede a crear la recomendación de navegación y preferencia que se enviará al usuario, siempre manteniendo la "sugerencia" por omisión que es el estado de "no sugerencia" es decir, sino hay nada bueno que recomendar, entonces no se perturba al usuario con información anexa.

Es en la preparación de la recomendación donde más se puede vulnerar la

privacidad del usuario, ya que se requiere de un seguimiento de sus acciones en el sitio, para poder clasificarlo en el grupo adecuado y preparar la recomendación de navegación que más se ajuste a lo que el sistema cree que el usuario anda buscando en el sitio.

En concreto, la etapa de preparación de los web data para ser usados en un proceso de web mining, puede ser realizada sin identificar a la persona detrás del browser, a través de técnicas no invasivas [2] y [3]. De hecho, son las más comúnmente aceptadas en investigación y que generan menos controversia en el tratamiento de los web data.

Las técnicas usadas en web mining para analizar el comportamiento del usuario en la Web, trabajan con miles de sesiones, sin importar quién es la persona que generó una determinada sesión. Aquí se aplica el principio estadístico de que el comportamiento de una persona es aleatorio, por lo tanto no sirve para conjeturar nada. Sin embargo, el comportamiento colectivo siempre marca una tendencia, por lo que se puede extrapolar y usar como un estimador probabilístico aceptado.

Respecto de las técnicas de web mining utilizadas para personalizar la Web, involucran que de alguna forma se pueda hacer un análisis del usuario que en ese momento visita el sitio, incluso se podría hasta llegar a la identificación de la persona detrás del usuario (buenos días señora Lorena, ayer compró un libro de web mining electrónico, pero aun no lee nada, ¿algún comentario?), con lo cual se estaría vulnerando abiertamente la privacidad del visitante.

Finalmente, la preparación de la acción de personalización claramente limita el libre albedrío del usuario que visita el sitio, por cuanto implica limitar su exposición a contenidos que "tal vez no le son de interés". En la práctica, esta limitación no ha sido mal recibida por los usuarios, lo cual no quita que igual sea una invasión en la privacidad del visitante del sitio. Sin embargo, en el ciberespacio, ¿existe el libre albedrío?, claramente somos dueños de ir donde queramos, pero en la mayoría de los casos lo hacemos influenciados por una recomendación de un motor de búsqueda, así que al menor podemos decir que el libre albedrío estaría limitado a lo que "el gran hermano" tecnológico quiera mostrarnos.

8. Conclusiones

La identificación de una persona a partir de los web data que se recolectan en un sitio web, no es factible en su totalidad. Utilizando el actual protocolo de comunicaciones de Internet: IP versión 4. A lo más se puede identificar la sesión de un usuario, es decir, un ente que en un determinado momento está navegando en un sitio web. Cuando esté en operación el próximo protocolo de Internet: IP versión 6, será posible establecer mecanismos de autenticación de

los usuarios, por ejemplo firma digital avanzada, y asociar una determinada dirección IP a una persona. En ese momento, tal vez sería conveniente analizar si es factible que este guarismo sea tratado como un dato personal.

El uso de las técnicas de web mining para la extracción de información y conocimiento desde los web data, puede vulnerar la privacidad de los usuarios que visitan un sitio web. Ahora bien, existen formas de minimizar esta vulneración hasta lo estrictamente necesario y con el consentimiento del usuario para ayudarlo en su búsqueda de información en un sitio, comenzando por cómo se pre procesan los web data y finalizando en la forma en que se entregan las recomendaciones de navegación y preferencias de contenido.

La personalización de la Web es el área de investigación en Web Intelligence que por lejos ha acaparado más seguidores en investigación y en el comercio. Se le considera el siguiente paso en la creación de sitios web, por lo que el desarrollo tecnológico a su alrededor ha sido exponencial. Como suele suceder cuando aparece una nueva tecnología, el marco regulatorio no existe y su creación demora, comparativamente hablando, "siglos" en estar listo, y cuando esto ocurre, ya está obsoleto. Sin embargo, existen ciertos principios universales que se podrían cautelar y que son independientes del cambio tecnológico. Uno de ellos es la privacidad de la navegación del usuario en un sitio web. Al menos debería quedar claro que si el usuario acepta que su visita sea personalizada, entonces existirá un seguimiento de sus acciones y que los datos que genere su sesión serán usados para establecer líneas de acción en los cambios que experimentará el sitio web.

Teniendo presente qué técnicas, metodologías y algoritmos vulneran la privacidad de los usuarios, es posible establecer las modificaciones necesarias para hacer del análisis del comportamiento del usuario en la Web, una actividad no invasiva, de mínimo impacto en lo referente a la privacidad de los usuarios que visitan un determinado sitio web.

Agradecimientos: Este trabajo fue parcialmente financiado por el Instituto Milenio Sistemas Complejos de Ingeniería.

Referencias

[1] http://www.dii.uchile.cl/

[2] R. Cooley, B. Mobasher, and J. Srivastava. Data preparation for mining world wide web browsing patterns. *Journal of Knowledge and Information Systems*, 1(1):5-32, 1999.

[3] M. Eirinaki and M. Vazirgannis. Web mining for web personalization. *ACM Transactions on Internet Technology*, 3(1):1-27, 2003.

[4] M. Kilfoil, A. Ghorbani, W. Xing, Z. Lei, J. Lu, J. Zhang, and X. Xu. Toward an adaptive web: The state of the art and science. In Procs. *Annual Conference on Communication Networks & Services Research*, pages 119-130, Moncton, Canada, May 2003.

[5] W. Kim. Personalization: Definition, status, and challenges ahead. *Journal of Object Technology*, 1(1):29-40, 2002.

[6] A. Kobsa. Tailoring privacy to usersñeeds. *In In Procs. of the* 8^{th} *International Conference in User Modeling*, pages 303-313, 2001.

[7] Z. Markov and D. T. Larose. Data Mining the Web: Uncovering Patterns in Web Content, Structure and Usage. *John Wiley & Sons*, 2007.

[8] M. Perkowitz and O. Etzioni. Towards adaptive web sites: Conceptual framework and case study. *Artificial Intelligence*, 118(1-2):245-275, April 2000.

[9] J.D. Velásquez and V. Palade. Adaptive Web Site. *IOS Press*, Netherlands. 2008.

[10] J.D. Velásquez and V. Palade. Building a Knowledge Base for Implementing a Web-Based Computerized Recommendation System. *International Journal of Artificial Intelligence Tools*, 2007.

[11] J.D. Velásquez and V. Palade. A Knowledge Base for the maintenance of knowledge extracted from web data. *Knowledge-Based Systems*, 20(3):238-248.

[12] S.A. Ríos, J.D. Velásquez, H. Yasuda and T. Aoki, Web site off-line structure reconfiguration: A web user browsing analysis. *Lecture Notes in Artificial Intelligence*, 4252(1):371-378, 2006.

[13] J.D. Velásquez, R. Weber, H. Yasuda and T. Aoki. Acquisition and maintenance of knowledge for web site online navigation suggestions. *IEICE Transactions on Information and Systems*, E88-D(5):993-1003, 2005.

[14] A. Vedder. KDD: The Challenge to Individualism. *Ethics and Information Technology*, 1: 275-281, 1999.

Programación Matemática para el Uso Eficiente de Mallas de Cultivo en una Empresa Salmonera

Francisco Cisternas*
Guillermo Durán**
Cristian Polgatiz*
Andrés Weintraub*

Resumen

La industria salmonera en Chile representa uno de los sectores más importantes de exportaciones de alimentos del país. Con una producción de más de 600 mil toneladas y ventas por más de 2.200 millones de dólares en 2008, es uno de los polos de desarrollos más importantes del sur de Chile. Dentro de las actividades más relevantes en el proceso de producción de esta industria, está el cultivo de salmones en el mar. Esto requiere de centros de cultivos especializados cuya mantención es realizada en tierra por talleres artesanales. El presente trabajo presenta la creación de una herramienta basada en modelos de programación entera mixta para optimizar el uso de recursos, mejorar la planificación y realizar evaluaciones económicas que facilitan el análisis y la toma de decisiones con respecto a la mantención, reparación y cambio de mallas de cultivo. El prototipo fue implementado en una de las principales empresas de la industria. Los resultados obtenidos logran reducir los costos de mantención de las mallas alrededor de un 15 % y generan también diversos beneficios cualitativos.

Palabras Clave: Cultivo de Salmones, Mallas de Cultivo, Planificación, Programación Lineal Entera.

*Departamento de Ingeniería Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile.

**Departamento de Ingeniería Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile; Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina y CONICET, Argentina.

1. Introducción

La introducción de especies acuícolas exóticas se produjo en Chile entre 1850 y 1920, pero los primeros salmones llegaron a nuestro país a partir de 1921, gracias a la labor del Instituto de Fomento Pesquero (IFOP).

De todos modos, fue recién a principios de los años 80 que se comenzó a cultivar el salmón en Chile. Hacia 1985, existían en nuestro país 36 centros de cultivo operando y la producción total llegaba a más de 1200 toneladas. Un año más tarde comenzó el auge de la industria salmonicultora y la producción ya superó las 2100 toneladas anuales. Ese mismo año, y como muestra de una consolidación definitiva de la industria salmonicultora, nació la Asociación de Productores de Salmón y Trucha de Chile A.G, hoy SalmonChile, la Asociación Gremial que aglutina a las empresas del sector en el país.

En 1990, la salmonicultura comenzó a desarrollar reproducción en Chile y se obtuvieron las primeras Ovas nacionales de Salmón Coho, uno de los 3 tipos que hoy se cultivan (los otros dos son la Trucha y el Salmón del Atlántico). Este hito se recuerda como el primer adelanto científico chileno en el área y el punto de partida para el despegue definitivo de la industria. Desde ese momento se realizaron las mejoras más importantes en los alimentos para salmones. El aumento de los volúmenes permitió la profesionalización de la industria, incorporando los alimentos secos con crecientes contenidos de lípidos, y un balance más eficiente entre éstos y las proteínas.

No obstante los avances de la industria chilena y de los mercados, en 1998 la industria vivió uno de sus momentos más complicados debido a la crisis asiática, que hizo caer los precios en Japón, y una sobreproducción a nivel mundial. Sin embargo, y gracias a las medidas necesarias para enfrentar la situación y la correcta manera de abordar los desafíos por parte de los diversos productores, la industria pudo sobrellevar el problema y seguir aumentando su producción.

Hasta 2008 la industria salmonicultora era el cuarto sector exportador del país, generaba más de 45.000 empleos directos e indirectos, y era el segundo productor de salmones en el mundo superado sólo por Noruega [1]. La crisis generada en la industria el a no pasado por la aparición del virus ISA en la mayor parte de los centros de cultivo de las empresas en agua-mar ha sido un fuerte golpe para su desarrollo y aún no hay datos de cuánto ha afectado a estas estadísticas.

La empresa donde se implementó la herramienta desarrollada en este trabajo es Salmones Multiexport, parte del holding empresarial Multiexport Foods. Salmones Multiexport es una de las principales empresas chilenas y mundiales

[1] Fuente: SalmonChile A.G - www.salmonchile.cl

del sector. Su inicio de actividades se remonta a 1989, se dedica a la producción, procesamiento y comercialización de salmones y truchas cultivados en Chile. Sus operaciones productivas se extienden desde la IX hasta la XI región de Chile, donde cuenta con centros de pisciculturas, smoltificación, engorda y plantas de procesamiento que se encuentran entre las más eficientes del mundo. Salmones Multiexport esta integrada verticalmente, operando desde la reproducción hasta la distribución al cliente.

El cultivo del salmón se divide en tres etapas:

1. **Reproducción:** cada empresa cuenta con centros especializados de reproducción y genética, donde se producen las ovas (1 a 2 meses)

2. **Crianza en agua dulce:** en esta etapa los peces son criados en pisciculturas, hasta alcanzar los pesos requeridos para su traspaso a centros en agua dulce, donde se smoltifican [2](10 a 14 meses)

3. **Engorda en agua de mar:** es la etapa más larga, en ella los peces son trasladados a los centros de cultivo ubicados en el mar, donde concluirán su desarrollo (16 a 20 meses)

En la etapa de engorda en agua de mar, los salmones son criados en balsas que contienen jaulas flotantes, las cuales poseen mallas de cultivo donde están contenidos los salmones durante todo el proceso (usaremos los términos mallas y redes como sinónimos a lo largo de este trabajo).

El objetivo del presente trabajo es optimizar el uso de recursos, mejorar la planificación y realizar evaluaciones económicas que faciliten el análisis y la toma de decisiones con respecto a la manutención, reparación y cambio de las mallas de cultivo de salmones.

El desarrollo de modelos de optimización ha estado presente dentro de la evolución del cultivo de salmones en grandes volúmenes, lo que ha sido fuente de investigación de varios autores, quienes buscando mejoras en el crecimiento y la productividad ([4], [9], [6], [1], [10], [2]), o bien mejoras en la planificación de sus cultivos ([7], [3], [8], [5]), muestran los efectos positivos, los beneficios y la necesidad de la utilización de modelos de optimización dentro de la industria del salmón.

El trabajo está organizado de la siguiente manera. En la Sección 2 se describe el problema a encarar. En la Sección 3 se detalla el modelo matemático desarrollado. La Sección 4 muestra los resultados y el impacto del trabajo, mientras que en la última Sección se reseñan las principales conclusiones.

[2]Smoltificación: Adaptación natural de los peces para condiciones de mayor salinidad

2. Descripción del problema

2.1. Composición de las mallas de cultivo

La mallas de cultivo están confeccionadas con diversos tipos de **nylon** o **poliéster** (ver figura 1), lo que permite un mejor manejo y traslado por su alta flexibilidad.

Figura 1: Tipos de tejidos de mallas de cultivo

Las dimensiones más usadas son 30 metros de largo por 30 metros de ancho (usualmente llamado 30x30), o 20 metros de largo por 20 metros de ancho (usualmente llamada 20x20), con profundidades que son de 10 metros o de 17 metros (ver figura 2).

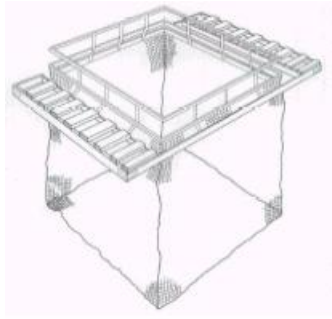


Figura 2: Malla de Cultivo

2.2. Mantención de las mallas de cultivo

Dado que la composición de las redes de cultivo es de nylon o poliéster con el paso del tiempo se deteriora y/o se ensucia con fouling[3], y por lo tanto las mallas deben ser reparadas o cambiadas para realizarles mantención.

[3]Fouling: se refiere al material orgánico que se adhiere en la malla debida a su exposición al agua de mar.

La mantención de las mallas consta de 3 etapas principales, las cuales se realizan en **Talleres de Redes**, empresas dedicadas exclusivamente a esta tarea. Estas etapas son:

Lavado Una vez que las mallas son llevadas a los talleres están deben ser lavadas para poder extraer todo el fouling adherido a las mallas. Este proceso se realiza por medio de lavadoras de tambor, llamadas hidrolavadoras o por mangueras de alta presión. Posteriormente las redes deben ser desinfectadas para cumplir con las normativas de bioseguridad establecidas.

Reparado Una vez realizado la etapa anterior las mallas son revisados por cuadrillas de 3 o más personas, quienes van revisando tramo a tramo las mallas para ver cualquier desperfecto que estas tengan e ir reparando de ser necesario. Una vez que la mallas es revisada completamente esta es derivada a los centros de acopio de mallas para ser retiradas por la empresa dueña de las mallas.

Pintado Una vez lavadas y reparadas las mallas pueden ser pintadas con pintura anti-fouling. El pintado tiene un costo adicional, pero permite que la malla dure más tiempo en el agua.

2.3. Factores de deterioro en las mallas de cultivo

Durante la permanencia de las mallas de cultivo en el mar, los principales factores de deterioro o causales de mantención son:

Crecimiento de algas El crecimiento de algas o desarrollo de fouling en las mallas de cultivo disminuye la resistencia y aumenta el peso de la malla, generando dificultades en su manejo y aumentando el riesgo de ruptura. Asimismo, causa también un impacto negativo en el desarrollo de los peces, pues al estar la malla "sucia" debido al exceso de algas en la misma, la circulación del agua baja creando una disminución en los niveles de oxígeno en el agua, generando así aletargamiento de los peces (disminución de ingesta de alimento) y aumentando la mortalidad. El crecimiento de algas tiene directa relación con el aumento de luz solar, por lo que este efecto se ve incrementado en verano.

Ataque de lobos marinos Uno de los predadores naturales que los salmones tienen en esta zona son los lobos marinos. Estos mamíferos son capaces de romper las mallas de cultivo en busca de los peces, generando grandes pérdidas a las empresas. Para proteger a los centros de los ataques de los lobos marinos, las jaulas de cultivo tienen además por encima una red más resistente, conocida como **malla lobera**.

Condiciones ambientales En general las mallas de cultivo no sufren grandes deterioros con el clima, pero en condiciones de extremo mal tiempo o de mucho sol, las mallas aumentan su desgaste.

2.4. Pérdidas económicas asociadas a los factores de deterioro

Los factores de deterioro pueden ocasionar los siguientes problemas si no son tratados a tiempo:

Ruptura de mallas Una ruptura de malla implica una fuga masiva de salmones, con la consecuente pérdida de la inversión realizada. Si se considera que en cada jaula hay aproximadamente 90.000 peces, y si éstos tienen un peso promedio de 3 Kg. por salmón, a precio del salmón[4] del año 2008, las pérdidas por la ruptura de una malla superarían el medio millón de dólares.

Problemas de crecimiento El crecimiento de algas en la malla de cultivo disminuye los niveles de oxígeno en su interior, esto causa que los peces permanezcan un mayor tiempo en engorda, debido a un desarrollo más lento [2]. Todo esto retarda la recuperación de capital, aumenta los costos por kilogramo de salmón generado y disminuye la disponibilidad del centro para otras cosechas.

2.5. Medidas de prevención

Las medidas de prevención son las siguientes:

Revisión periódica de las mallas Consta en revisiones realizadas por buzos con el fin de ver el estado de las mallas y el nivel de crecimiento de algas.

Mantenimiento periódico de las mallas La forma de realizar las mantenciones periódicas involucra distintas etapas y actores. En primer lugar se solicita a una embarcación destinada exclusivamente para el movimiento y manejo de redes, la cual llamaremos **Barco de Redes**, que vaya a un centro de cultivo y cambie una malla, que está "sucia" o rota en el centro de cultivo, y la reemplace por otra que esté limpia y en óptimas condiciones. Una vez que el Barco de Redes realiza la faena de cambio, se traslada al puerto de carga y la malla se lleva por medio de camiones a talleres de mantenimiento y reparado de redes. Estos talleres limpian y reparan las mallas de cultivo, dejándolas listas para su reutilización. Es en esta etapa donde la empresa dueña de la

[4]Precio de referencia según Fondo Monetario Internacional, estadísticas de producción acuícola.

malla de cultivo decide si pinta o no pinta la malla con pintura que disminuye la adhesión de las algas.

2.6. Situación Actual

Como fue mencionado, las empresas poseen mecanismos de prevención ante los problemas que pudiesen estar asociados al correcto funcionamiento de las mallas de cultivo. Pero estas medidas tienen altos costos para la empresa. En el Cuadro 1 se detallan algunos costos involucrados en el mantenimiento de las mallas de cultivo, a modo de establecer un orden de magnitud.

Costos Directos de Mantencion	**Valor**
Contratar equipo de buceo para inspeccionar las jaulas y verificar la condición de las mallas.	US$ 20.000 mensuales por equipo
Transporte de mallas por los Barcos de Redes.	US$ 30.000 mensuales por barco
Mantención y Reparación de las mallas.	US$ 0,26 por m^2
Pintado de las mallas con pintura antifouling	US$ 0,65 por Kg
Costo de malla nueva (Precio ref.: Malla de cultivo de 30x30)	US$ 4.400

Cuadro 1: Principales costos directos de mantenimiento de mallas de cultivo

Actualmente las decisiones que la empresa debe tomar son: cuándo instalar una malla; cuándo retirar una malla; cuándo cambiar una malla instalada, por una limpia o por otra más grande; pintar o no pintar una malla; cuántas mallas se deben comprar y cuándo comprarlas; qué centros debe visitar un barco por semana; cuándo realizar cambio de mallas loberas.

Estas decisiones, considerando además la disponibilidad de los barco de redes y las capacidades de los talleres de redes, se basan en juicios expertos sin herramientas específicas que los apoyen.

Estas decisiones se ven fuertemente afectadas por condiciones externas, como el nivel de luz solar, lo que aumenta el crecimiento de fouling en las redes; o condiciones climáticas, que impiden que los barcos realicen sus tareas. Además, la falta de programación adecuada de la llegada de las mallas al taller genera inconvenientes en los tiempos de respuesta.

Si consideramos que la cantidad de mallas que la empresa tiene ha ido en aumento (930 mallas de cultivo de 1"y de 2"aproximadamente en la actualidad), y que existe una gran variabilidad de las necesidades asociada a los recursos involucrados, lograr hacer una planificación mensual ya resulta bastante difícil. Realizar proyecciones semestrales o anuales de manera manual lleva varios días de trabajo, generando así una merma en las labores diarias del experto, así como también una incapacidad de determinar de buena

manera soluciones eficientes. Considerar todas las combinaciones posibles sin ningún soporte tecnológico (solamente planillas Excel) es prácticamente imposible, por lo que resulta necesario el uso de herramientas de programación matemática a modo de apoyar al experto en la toma de decisiones.

3. Modelo de Programación Entera Mixta

Se representó el problema mediante un modelo de Programación Entera Mixta (MIP). A continuación describimos algunos puntos conceptuales del modelo y su formulación matemática.

3.1. Descripción conceptual del modelo matemático

Antes de dar la formulación del modelo, es necesario mencionar que el mismo toma conjuntos de *centros*[5] que la empresa tiene, los cuales se denominan *áreas*[6]. Esto nos permite resolver el problema de cada una de estas áreas por separado, pues son independientes en la operación que involucra tanto la mantención como el traslado de las redes de los centros a los talleres (hay un taller distinto asignado a cada área).

El fin de esta sección es describir conceptualmente qué es lo que hace el modelo, cuáles son sus limitantes y los resultados que entrega.

1. Inputs del Modelo

   A continuación se especificarán los inputs base que se le debe entregar al modelo:

   a) Horizonte de tiempo del modelo, expresado en *semanas*.

   b) Tipos de red utilizadas en el área y sus características básicas:

   - Dimensión de la red, expresado en m^2.
   - Peso de la red, expresado en Kg.

   c) La cantidad de jaulas totales del área.

   d) La cantidad de centros del área.

   e) Recursos disponible del área (cantidad de barcos de redes) y su rendimiento de faenas o actividades por día que pueden realizar.

   f) La permanencia máxima que una red puede tener en el agua, definido en *semanas*.

[5] Un **centro de cultivo** es un conjunto de **módulos** en los cuales están contenidos las jaulas. Cada módulo contiene 14 o 28 jaulas y en cada centro hay de uno a dos módulos.

[6] Estas áreas corresponden a las áreas geográficas que la misma empresa ha definido como agrupación para sus centros.

g) Stock inicial de redes de cada tipo en el área.

h) El tiempo de respuesta de los talleres[7], expresado en *semanas*.

i) Costo de mantención de redes en los talleres, expresado en $[\frac{US\$}{m^2}]$.

j) Costo de impregnar pintura antifouling en las redes en los talleres, expresado en $[\frac{US\$}{Kg}]$.

k) Costo de la pintura antifouling[8], expresado en $[\frac{US\$}{Lts}]$.

l) Costo de traslado[9], expresado en $[US\$]$

2. Con la información entregada de la utilización de centros (Plan de Producción, Carta Gantt) se define cuándo y qué tipo de redes serán requeridas de cada uno de los centros. Es decir la demanda por redes (parámetro del modelo) está determinada a partir de esta información y el modelo está obligado a satisfacer los requerimientos de redes.

3. Se generan restricciones relacionadas con los rendimientos que tienen los barcos de redes y los periodos destinados para la instalación de redes loberas. El modelo buscará de que modo cumple la demanda por redes adecuándose a los recursos disponibles.

4. Tomando estas restricciones y consideraciones, el modelo **minimiza** los costos de mantención, de transporte y de compra de redes de cultivo.

5. Además de entregarnos la información necesaria para poder determinar el plan de mantención para las redes, el modelo indica también la cantidad de viajes realizados hacia cada uno de los centros, la cantidad de redes utilizadas en cada periodo de tiempo y la cantidad de redes necesarias para dar cumplimiento a todo el periodo de evaluación incluyendo la cantidad que deben ser compradas (i.e. Stock crítico de redes).

6. Los parámetros necesariamente variarán según el área que deseemos evaluar (distintos rendimientos, cantidad de barcos, número de centros, distintos tiempos de respuesta, etc), lo cual es posible de realizar sin alterar la estructura del modelo.

3.2. Formulación del modelo matemático

1. Índices

   - t = Período, t $\epsilon\{1, .., T\}$, donde T es el número total de períodos en el horizonte de planificación, medidos en meses.

[7] Tiempo que tarda una red en ser traslada y reparada en el taller.

[8] Pintura anti-incrustantes marinos [Algas, moluscos, etc.].

[9] El *costo de traslado* se determina a través de la distancia (Centro - Puerto de Carga) en $[mn]$, el consumo de petróleo de los barcos $[\frac{Lts}{mn}]$ y el precio del petróleo.

- k = Tipo de malla, k $\epsilon\{1,..,K\}$, donde K es el número total de tipos de mallas utilizados.
- j = Jaula perteneciente a los centros abiertos, j $\epsilon\{1,..,J\}$, donde J es el número total de jaulas en todos los centros.
- c = Cento abierto, c $\epsilon\{1,..,C\}$, donde C es el número total de centros.
- d = Número de periodos que permanecerá la malla en el agua cuando la coloque, d $\epsilon\{1,..,D\}$, donde D es el máximo de periodos que la malla puede permanecer en el agua.
- p = Establece el estado de la pintura de la malla.
- b = Número de barcos, b $\epsilon\{1,..,B\}$, donde B es el número total barcos.

2. Parámetros

- DDA_{kjt} = Demanda de mallas tipo k de la jaula j en el periodo t.
- $CPAV_{kdp}$ = Costo incurrido por mantener en el agua una malla de tipo k, por d periodos con estado de pintura p y en periodo de Verano.
- $CPAI_{kdp}$ = Costo incurrido por mantener en el agua una malla de tipo k, por d periodos con estado de pintura p y en periodo de Invierno.
- CPA_{kdpt} = Costo incurrido por mantener en el agua una malla de tipo k, por d periodos con estado de pintura p y empezando en el periodo t. (este parámetro se calcula en función de los 2 anteriores)
- Semana Inicial = Indica el número de semana (correlativa del año) en que se inicia el horizonte de planificación.
- TR = Tiempo que tarda una malla en ser reparada, desde que es retirada de su jaula.
- $CTrans_c$ = Costos de transporte al centro c, este costo se incurre al colocar o sacar una o más jaulas en el centro c.
- $CComp_k$ = Costo de comprar una malla de tipo k.
- J_c = Índices de jaulas que pertenecen al centro c.
- $Hist_{kjp}$ = Cantidad de periodos que una malla tipo k, ha permanecido en la jaula j, con estado de pintura p, hasta hoy.
- $MInv_k$ = Cantidad de mallas tipo k que tengo en inventario.
- $Aux_{kjp} = \begin{cases} 1 & \text{si la jaula j, tiene una malla k, con estado de pintura p} \\ & \text{en el periodo 0.} \\ 0 & \sim \end{cases}$

- $Barcos_{bc}$ = Matriz de asignación de barcos por centro
- $Rend_{ct}$ = Rendimiento del barco b en periodo t
- $DiasLab$ = Cantidad de días laborales en una semana

3. Variables

- $X_{kjtdp} = \begin{cases} 1 & \text{si coloco la malla tipo k, en la jaula j en el periodo t,} \\ & \text{por d periodos, con estado de pintura p} \\ 0 & \sim \end{cases}$
- $W_{bct} = \begin{cases} 1 & \text{si utilizo el barco b en el centro c en el periodo t} \\ 0 & \sim \end{cases}$
- U_k = Cantidad total de mallas tipo k que tengo que disponer para el horizonte de planeación.
- MU_{kt} = Cantidad de mallas tipo k requeridas en el periodo t.
- $MComp_k$ = Cantidad de mallas tipo k que debo adicionar a mi inventario para satisfacer las necesidades del horizonte de planeación.

4. Función Objetivo

$$\text{mín} \sum_{kjtdp} CPA_{kdpt} \cdot X_{kjtdp} + \sum_{k} CComp_k \cdot MComp_k + \sum_{bct} CTrans_c \cdot W_{bct} \cdot 2 + \sum_{k,j,d,p} X_{kj0dp} \cdot CPA_{k,Hist_{kjp}+d,p,0} \tag{1}$$

5. Restricciones

a) Naturaleza de las variables.

$$X_{kjtdp} \quad Binaria \qquad W_{bct}, U_k, MU_{kt}, MComp_k \text{ positivas} \tag{2}$$

b) Satisfacción de la demanda.

$$\sum_{p,\theta\epsilon[0..T]:\theta\leq t, d>t-\theta} X_{kj\theta dp} \geq DDA_{kjt} \qquad \forall k, j, t \tag{3}$$

c) Mallas tipo k que están siendo utilizadas este periodo.

$$\sum_{j,p,\theta\epsilon[0..T]:\theta\leq t, d>t-\theta-TR} X_{kj\theta dp} = MU_{kt} \qquad \forall k, t \tag{4}$$

d) Máximo número de mallas tipo k que están siendo utilizados.

$$MU_{kt} \leq U_k \qquad \forall k, t \tag{5}$$

e) Indica si hay utilización de transporte desde o hacia el centro c.

$$\sum_{k,j\epsilon J_c,d,p} X_{kjtdp} + \sum_{k,j\epsilon J_c,\theta<t,\delta=t-\theta,p} X_{kj\theta\delta p} \leq W_{bct} \cdot Rend_{ct} \quad \forall c,t \quad (6)$$

f) Uso máximo de un barco en un centro en un periodo

$$W_{bct} \leq 2 \cdot Barco_{bc} \cdot DiasLab \quad \forall b,c,t \quad (7)$$

g) Calcular el número de mallas que se deben comprar.

$$MComp_k \geq U_k - MInv_k \qquad \forall k \quad (8)$$

Cabe observar que el modelo maneja las semanas destinadas a cambios de mallas loberas asignando rendimiento 0 a los barcos de talleres en esas semanas (lo que implica que durante esas semanas no hay ningún cambio de mallas de cultivo).

4. Resultados e impacto

Se comparó el impacto del modelo contra lo realizado por la empresa, tomando como situación base una de las áreas de cultivo.

La aplicación se ejecutó en OPL 1.71, y se utilizó el solver de programación lineal entera CPLEX 11.2, en un procesador Intel®Core™DUO T7200 de 2.00 GHz y 1 GB de RAM. Los tiempos de corrida de la aplicación varían de acuerdo al tamaño de las áreas entre 15 y 30 segundos.

4.1. Situación Base

4.1.1. Antecedentes

Para la realización de este análisis se tomó como área de evaluación el área **Dalcahue**[10]. Los valores de los parámetros son:

- En el área operan 3 barcos de redes.
- Cada barco puede realizar 7 actividades [11] en un día de trabajo.
- El área posee 6 centros de cultivo.
- El horizonte de evaluación es de un **semestre**[12].

[10] Ubicación Geográfica: Isla Grande de Chiloé, X Región de los Lagos

[11] Las actividades que un barco realiza son **instalación** y **retiro** de mallas. Un **recambio** de mallas involucra las dos actividades mencionadas anteriormente

[12] Equivalente a 24 semanas

- Durante 6 semanas los barcos no realizarán cambios de mallas sino que estarán asignados a realizar labores de cambio de mallas loberas[13].
- Existen 2 tipos de mallas de cultivo, mallas de 1" y mallas de 2".
- El tiempo de respuesta de los talleres[14] es de **1 semana**.
- Stock Inicial de Mallas

Mallas 1"	Mallas 2"
35	98

Cuadro 2: Stock Inicial

4.1.2. Resultados

A continuación se presentarán los resultados obtenidos para la situación base:

- Situación Base

  La situación base puede ser considerada de dos formas:

  - **Gasto Presupuestado**: Es aquel gasto que considerando los parámetros descritos en 4.1.1, se debieron haber efectuado en Mantención de Mallas en un periodo de tiempo de 24 semanas, según el plan de mantención que la empresa estimó.
  - **Gasto Real**: Es aquel gasto que efectivamente la empresa incurrió en Mantención de Mallas en el mismo periodo de 24 semanas. Los parámetros descritos en 4.1.1 se mantuvieron iguales.

Gasto Presupuestado Mantención de Redes [US$]	Gasto Real Mantención de Redes [US$]
541.460	**580.520**

Cuadro 3: Situación base

La diferencia entre el gasto real versus el presupuestado se debe a la poca información con que se realizan las estimaciones. Las estimaciones

[13]Se pueden dar escenarios en que el uso de las embarcaciones sea bajo, es decir menores que 2 actividades por semana los cuales también podrían ser considerados para las labores de cambio de mallas loberas

[14]Incluye tiempo de traslado

de capacidad de talleres y los movimientos de los barcos no estaban consideradas dentro del plan de mantención.

Las compras de nuevas mallas se realiza en función de los requerimientos de todas las áreas, pero la asignación de las mallas no quedaba registrada pues la distribución era en función de las contingencias que la empresa tiene en todas sus áreas. Presupuestariamente se estimaron aproximadamente 111 mallas de 2"(equivalentes a [US$]478.500) para las 2 áreas de cultivo de la X Región de los Lagos, las cuales presentan tamaños similares lo que hace suponer que el gasto incurrido por el área en concepto de mallas nuevas compradas es de [US$]239.250 por área.

Por lo tanto, para comparar nuestra situación base con las futuras propuestas de mejora que nos entrega el modelo, consideraremos los **gastos reales** del periodo,las mallas utilizadas y los parámetros descritos a continuación, definiendo de esta manera la **Situación Base**:

Gasto Mantención [US$]	Gasto Transporte [US$]	Gasto en Compra mallas 1"[US$]	Gasto en Compra mallas 2"[US$]
580.520	**85.072**	**0**	**239.250**

Cuadro 4: Valores de Gastos de la Situación base

Nº barcos operando	Rendimiento barco	Total mallas 1" utilizadas	Nº faenas con mallas 1"pintadas	Total mallas 2" utilizadas	Nº faenas con mallas 2"pintadas
3	**7**	**35**	**0**	**121**	**32**

Cuadro 5: Parámetros de la Situación base

- Situación Base Optimizada

  Considerando los parámetros definidos en 4.1.1 y utilizando los resultados obtenidos tras aplicar el MIP desarrollado definimos la **Situación Base Optimizada**. Los resultados se describen a continuación::

Gasto Mantención [US$]	Gasto Transporte [US$]	Gasto en Compra mallas 1"[US$]	Gasto en Compra mallas 2"[US$]
528.207	**12.214**	**0**	**222.639**

Cuadro 6: Valores de Gastos de la Situación Optimizada

- Comparación Situación Base v/s Situación Base Optimizada

  Considerando los cuadros 4 y 6 podemos construir el siguiente resumen:

Nº barcos operando	Rendimiento barco	Total mallas 1" utilizadas	Nº faenas con mallas 1"pintadas	Total mallas 2" utilizadas	Nº faenas con mallas 2"pintadas
3	**7**	**35**	**0**	**145**	**0**

Cuadro 7: Parámetros de la Situación Optimizada

Gasto total Sit. Base [US$]	Gasto total Sit. Optimizada	Diferencia [US$]
904.842	**763.060**	**141.782**

Cuadro 8: Diferencia en gastos

Como era de esperarse, el modelo mejoró la situación base disminuyendo los costos asociados a la mantención de las mallas de cultivo en un 16 %. Además por construcción del modelo, ahora se entrega un plan de mantenimiento más acabado, ya que actualmente los planes de mantenimiento son aproximados (proyectados en su gran mayoría a estimaciones mensuales de requerimientos de mallas de cultivo por centro).

4.2. Resultados del Análisis de Sensibilidad

Con el fin de lograr identificar el impacto que tiene en los resultados del problema original determinadas variaciones en los parámetros, analizamos los distintos resultados ante variaciones del rendimiento de los barcos, del número de los mismos que se utilizan, del tiempo de respuesta de los talleres o de la permanencia de una malla en el agua.

4.2.1. Barcos de redes y rendimiento

A continuación se muestra un cuadro comparativo que muestra el costo calculado por el modelo, las cantidades de mallas requeridas y la cantidad de mallas que deben ser "pintadas"para distintos rendimientos de los barcos (expresados en actividades por barco) y distinto número de barcos utilizados[15]:

Del Cuadro 9 podemos realizar los siguientes comentarios:

- Aunque en general todos los resultados tienen similar cantidad de mallas utilizadas de ambos tipos, las estrategias de traslado de mallas son distintas y eso se ve reflejado en la columna de Costo del Cuadro 9, donde se aprecia la diferencia de los costos por concepto de transporte y de

[15]Cuando el número de barcos de redes es igual a 2, se descontará al valor de la función objetivo [US$] 40.000, por el ahorro del arriendo de la embarcación

Nº barcos operando	Rend. Barco redes	Costo Func. Obj. [US$]	Total mallas 1" utilizadas	Nº faenas con mallas 1"pintadas	Total mallas 2" utilizadas	Nº faenas con mallas 2"pintadas
3	10	762.295	35	0	145	0
3	9	762.591	35	0	145	0
3	8	762.557	35	0	145	0
3	**7**	**763.060**	**35**	**0**	**145**	**0**
3	6	760.142	36	0	144	0
3	5	765.209	35	0	145	0
3	4	762.191	35	0	143	1
2	10	722.295	35	0	145	0
2	9	722.591	35	0	145	0
2	8	722.557	35	0	145	0
2	7	722.512	35	0	145	0
2	6	723.241	35	0	145	0
2	5	723.744	35	0	144	0
2	4	726.744	35	0	145	0

Cuadro 9: Resultados análisis de sensibilidad para barcos de redes y su rendimiento

utilización de mallas. Un buen control de los rendimientos de los barcos de redes permitirá que existan ahorros significativos por este rubro en el mediano plazo.

- El costo de tener 3 barcos en el área no genera el ahorro suficiente que justifique la incorporación de una embarcación más. En periodos de tiempo que existan necesidades de embarcación con el fin de brindar apoyo a otro tipo de actividades, puede justificarse el uso de 3 embarcaciones, pero si no es tal el escenario, es recomendable dejar solamente 2 embarcaciones en el área.

4.2.2. Tiempo de Respuesta de Talleres

Con el fin de entender el impacto que tiene dentro de nuestra solución los tiempos en que las mallas **no** se encuentran en las jaulas por motivos de reparado y/o traslado, se analizó como varía la solución base frente a cambios en los tiempos de respuesta:

Del Cuadro 10 podemos concluir lo siguiente:

- A medida que aumentan los tiempos de respuesta mayor es el requerimiento de mallas y de mallas pintadas para satisfacer las necesidades del área.

Tiempo de Respuesta de Taller	Costo Func. Obj. [US$]	Total mallas 1" utilizadas	Nº faenas con mallas 1"pintadas	Total mallas 2" utilizadas	Nº faenas con mallas 2"pintadas
1	763.060	35	0	145	0
2	910.136	36	14	168	6
3	1.006.532	36	25	168	54
4	1.192.391	36	28	224	0

Cuadro 10: Resultados análisis de sensibilidad para el tiempo de respuesta de talleres

- Dentro de este parámetro existen 2 variables que **deben** ser objeto de constante control. El primero es el cumplimiento de los tiempos de reparación de los talleres y el segundo los tiempos de traslados de las mallas de cultivo desde y hacia los centros de cultivo (poniendo también reparo en los tiempos de muertos del proceso), ya que cualquier retraso en alguno de estas variables hará que la estrategia, y por ende los costos asociados, cambien.

- Los resultados obtenidos hacen pensar que es necesario crear un estándar de tiempo de permanencia de una malla en un taller. Esto nos permitiría lograr una planificación más certera pero requeriría de mayor control sobre los talleres.

4.2.3. Variación de la Permanencia de una Malla en una Jaula

Con el fin de entender como el modelo varía su solución a partir de distintos rangos de permanencia de una malla en el agua, se mostrarán a continuación los resultados obtenidos. Estos resultados se obtienen a partir de modificar el parámetro CPA del modelo. La forma de entender esta variación es esencialmente modificando el tiempo promedio que una malla se conserva en el agua (o sea, aumentando o disminuyendo el número de semanas que una malla se mantiene en promedio en el agua).

- Con los resultados presentados en el Cuadro 11 podemos ver el impacto que tiene la estrategia a seguir según sea el criterio definido para el tiempo de permanencia de una malla en una jaula. Criterios conservadores, es decir más cautos al momento de decidir cuándo realizar un cambio de la mallas, generan mayores costos sobre aquellos criterios en que ven de modo más optimista la permanencia de las mallas en el agua.

Con la información reunida y tras el análisis realizado queda demostrado que la variable que mayor impacto causa marginalmente es la **Variación de**

Variación en el Rango de Permanencia	Costo Func. Obj. [US$]	Total mallas 1" utilizadas	Nº faenas con mallas 1"pintadas	Total mallas 2" utilizadas	Nº faenas con mallas 2"pintadas
+2 semanas	457.823	35	0	133	0
+1 semana	590.104	35	0	140	0
...	**765.060**	**35**	**0**	**145**	**0**
-1 semana	1.024.851	35	0	150	36
-2 semanas	1.481.737	35	28	168	140

Cuadro 11: Resultados análisis de sensibilidad para la variación de la permanencia de una malla en una jaula

la permanencia de las mallas en una jaula. Esta es una variable que las personas encargadas de la toma de decisiones deben determinar según sea el comportamiento que ellos perciban del área geográfica en que se encuentran. El juicio de expertos en este caso es un buena aproximación a la realidad.

Los tiempos de respuestas de talleres deben ser controlados por la empresa productora para que la estrategia definida no varíe y no se deban tomar decisiones apresuradas (tanto en pintado como en compra de mallas).

Finalmente, los barcos de redes deben mantener altos rendimientos para no caer en gastos adicionales. Una manera de hacer que los administradores de barcos de redes enfoquen sus esfuerzos en mejorar sus rendimientos, es el de tener contratos con tarifas variables (por ejemplo en función de las actividades realizadas) pudiendo así mantener un alto estándar en servicio y menores costos finales para la empresa.

5. Conclusiones

Cómo se puede ver en este trabajo, la solución propuesta tiene múltiples ventajas con respecto a la situación base.

- Usar un MIP mejora la solución actual del problema y ayuda a una mejor planificación. La utilización de este modelo permite realizar una planificación a largo plazo (un horizonte de seis meses) mientras que antes se resolvía prácticamente semana a semana.

- Permite elaborar escenarios con variables que no son controlables, cómo variaciones climáticas, o condiciones laborales extraordinarias en el arriendo de barcos o reparación de las mallas. Con esto se pueden realizar simulaciones y evaluar configuraciones a un mucho menor costo y en un menor tiempo.

- Ayuda a responder una de las preguntas largamente sin respuesta de la industria: "¿Es rentable utilizar antifouling?". La respuesta depende del periodo de permanencia de la red en el agua y la época del año en que se instalará la malla (debido a la intensidad del sol). Esto sorprendió a la contraparte en la empresa, pero se ajustó a la intuición que se tenía al respecto y respaldó la toma de decisiones al respecto.

- El uso de eficiente de los barcos es muy importante debido a su alto costo de arriendo y mano de obra. La herramienta generada permite evaluar el número de barcos y el riesgo que puede involucrar tener una cantidad muy ajustada a la demanda.

- Se puede involucrar a los talleres de reparación de mallas dentro de la cadena de suministro, informando con mayor anticipación el flujo de mallas, lo que les permite ajustar sus recursos, para disminuir costos y tiempos de reparación, y sobre todo tener menos incertidumbre del tiempo que tardarán en reparar las mallas. Como se vio en el análisis de sensibilidad de este parámetro, el efecto sobre los costos que tiene cada semana adicional que los talleres tardan en reparar las mallas es muy alto.

- Los tiempos de planificación disminuyen notablemente, por lo que se pueden evaluar mayores escenarios y a más largo plazo. Además, los expertos encargados del tema tienen mayor tiempo disponible para realizar otras actividades.

Todos estos beneficios detectados se resumen en una herramienta para la toma de decisiones con respecto al mantenimiento de las mallas de cultivo. La potencialidad de ahorro es importante si se utiliza con una mirada en el largo plazo.

La herramienta fue implementada de manera piloto en Salmones Multiexport, pero la extensión de esta solución a otras empresas de la industria, como empresas productoras, transportistas de redes y reparadoras de redes es evidente y representaría un beneficio importante para la industria.

Adicionalmente a los beneficios mostrados de manera explícita en este trabajo, se deben considerar otras mejoras que incorpora la solución propuesta que son difíciles de cuantificar, como lo son los beneficios asociados al mayor crecimiento de los peces debido a menores condiciones de stress y mayor limpieza de las mallas; mejores relaciones con empresas contratistas (talleres de redes, barcos de redes, etc); menor uso de personal en la creación de estos escenarios y menor tiempo de resolución de los mismos, lo que posibilita evaluar múltiples escenarios de manera rápida y eficiente, ya que actualmente la creación de cada escenario tarda aproximadamente 2 días y con la herramienta los

tiempos de resolución son menores a 30 segundos (y se consigue una planificación de mejor calidad y con mayor nivel de detalle, dado que la herramienta planifica a nivel de jaula y la elaboración manual, a nivel de centro).

Un aspecto interesante de esta herramienta es que permite al usuario forzar decisiones. Lo que permite por un lado incorporar aspectos que no fueron modelados explícitamente, como es el caso del cambio de mallas loberas. También puede ser utilizado para que el usuario experto considere en la solución imprevistos o cosas muy difíciles de modelar, cómo una huelga, o acciones que ya fueron tomadas, para que se adapte de mejor forma a la realidad. Esta flexibilidad permite que la herramienta sea mucho más aplicable, puesto que la realidad es mucho más compleja de lo que se puede modelar y de esta forma se puede adaptar con mucho mayor facilidad a las condiciones observadas.

Otro aspecto muy importante en la implementación de este trabajo, es que se desarrolló una aplicación muy flexible, que genera automáticamente gráficos de uso de mallas, costos en que se incurre, una matriz con la planificación de la mantención de las mallas, permite ingresar la historia de planificaciones previas, permite al usuario fijar algunos valores de variables de forma sencilla, usa colores para indicar pintado de mallas, entre otras cosas. Todo lo cual representa parte fundamental de la solución, puesto que sin esta interfaz amigable, a pesar de sus beneficios evidentes, sería difícil que se utilice cotidianamente.

Durante el 2008 la industria salmonera en Chile sufrió una profunda crisis impulsada por la aparición del virus ISA que detiene fuertemente el crecimiento de los salmones e incluso puede aumentar la mortalidad. Este flagelo, sumado a la crisis económica a nivel mundial, hizo que la industria tuviera que disminuir drásticamente sus dotaciones y reducir la cantidad de centros de cultivos y personal que no estuviera dedicado al proceso productivo. Es por esto que la empresa piloto congeló la implementación de la herramienta y cambió algunos procesos operativos del proceso de mantención. Por esto la herramienta a la fecha no se encuentra en uso, aunque se espera que se retome su utilización una vez que la industria se normalice.

Agradecimientos: A Corfo, Salmones Multiexport S.A. y el Instituto Científico Milenio "Sistemas Complejos de Ingeniería", por el apoyo financiero para la realización de este proyecto. A Diego Delle Donne por el trabajo realizado en la programación de la aplicación; a Javier Marenco, quien brindó su apoyo en diversas etapas del trabajo; y a Rodrigo Niklitschek y Fredi Espinoza, de Salmones Multiexport, por toda su colaboración para la concreción de este proyecto. El segundo autor es parcialmente financiado por Fondecyt 1080286 y el cuarto autor es parcialmente financiado por Fondecyt 1085188.

Referencias

[1] W. Brown and A. Hussen. A Production Economic Analysis of the Little White Salmon and Willard National Fish Hatcheries, Report 428, Oregon State University, 1974.

[2] V. Crampton, A. Bergheim, M. Gausen, A. Næss and P.M. Hølland. Effects of low oxygen on fish performance. Oxygen levels in commercial salmon farming conditions. *EWOS Perspective*, pp. 8-12, UK Edition Nº 2, 2003.

[3] O. Forsberg. Optimal harvesting of farmed Atlantic salmon at two cohort management strategies and different harvest operation restrictions, *Aquaculture Economics and Management*, Volume 3, Number 2, pp. 143 – 158, 1999.

[4] O. Forsberg and A. Guttormsenb. Modeling optimal dietary pigmentation strategies in farmed Atlantic salmon: Application of mixed-integer nonlinear mathematical programming techniques, Aquaculture 261(1), pp. 118-124, 2006.

[5] M. Gustavson. Maximizing Profits for a Commercial Salmon Rearing Facility Using Linear Programming, Master's thesis, Naval Posgraduate School, Monterrey, California, 1972.

[6] R. Hilborn and C. Walters (editors). Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty, Chapman & Hall, London, 1992.

[7] P. Jensson and E. Gunn. Optimization of production planning in fish farming, Technical Report, University of Iceland, 2001.

[8] P. Leung and R. Yu Optimal Partial Harvesting Schedule for Aquaculture Operations, *Marine Resource Economics*, Volume 21, pp. 301 – 315, 2006.

[9] R. Pomeroy, B.Bravo-Ureta, D. Solís and R. Johnston. Bioeconomic modelling and salmon aquaculture: an overview of the literature, *Int. J. Environment and Pollution*, Volume 33, Number 4, pp 485 - 500, 2008.

[10] M. Ryea and I. Maob. Nonadditive genetic effects and inbreeding depression for body weight in Atlantic salmon, Livest. Prod. Sci. 57, pp. 15–22, 1998.

Modelo Aplicado de Teoría de Juegos para el Estudio del Crimen en la Vía Pública

Jose L. Lobato*
Richard Weber*
Nicolás Figueroa*

Resumen

En este trabajo se plantea un modelo de teoría de juegos que emula la interacción entre criminales y policías en la vía pública. El modelo se basa en teoría de juegos en una red de transporte, en donde las rutas de la red simulan a las opciones de delitos disponibles y los flujos a los criminales, los cuales pueden actuar organizada o desorganizadamente. El juego supone efectos de congestión en las rutas tanto por un aumento de flujo (criminales) o por la asignación de recursos policiales. Adicionalmente, se plantea una metodología genérica de aplicación del modelo. Esta consta de cinco pasos la cual finaliza con una sugerencia de asignación de recursos policiales bajo criterios de optimalidad. Los pasos metodológicos proveen de flexibilidad de modelamiento, lo que permite que la aplicación del modelo pueda ser altamente ajustable a diferentes problemáticas. La metodología fue probada utilizando datos reales de denuncias de delitos de la Región Metropolitana. Además, se realiza un análisis de robustez de la metodología planteando diferentes escenarios de modelación. Los resultados se analizaron en base a los parámetros de calibración del modelo y a los resultados de las asignaciones óptimas de recursos policiales. En todos los escenarios, se muestran mejoras teóricas significativas en la asignación de recursos policiales, lo cual sugiere que es posible incorporar modelos de esta naturaleza en el rol de disuasión del crimen que tengan resultados potencialmente efectivos.

Palabras Clave: Crimen, Teoría de Juegos, Selfish Routing

*Departamento Ingeniería Industrial, Universidad de Chile

1. Introducción

El estudio del crimen visto como un fenómeno social es llamado en las ciencias sociales *criminología*. En general, sus estudios involucran diversas disciplinas como sociología, matemáticas y economía entre otras y por lo tanto, consideran tanto elementos cualitativos como cuantitativos.

En el marco de la criminología cuantitativa, las primeras investigaciones comenzaron desde el ámbito de la estadística. Estos estudios se enfocaron en encontrar patrones de comportamiento de los criminales. En la actualidad, los modelos cuantitativos se han expandido hacia múltiples áreas y han aprovechando los avances teóricos y tecnológicos para el análisis del fenómeno. Destaca el uso de minería de datos y simulación para el descubrimiento complejo de patrones de comportamiento criminal[1] y el uso de teoría de juegos para modelar la interacción de agentes involucrados en el crimen. Sin embargo, estas disciplina mencionadas aportan sólo desde sus propias ópticas de resultados y pocas veces se integran para complementarse.

En el presente artículo, se propone un enfoque innovador del estudio del crimen, visto como un fenómeno de interacción entre individuos criminales y policiales. En particular, se plantea un modelo de teoría de juegos que simula dicha interacción y una metodología de carácter genérica para llevar a la aplicación el modelo. Luego, este trabajo se lleva a la práctica utilizando datos reales de denuncias de delitos.

Gracias al enfoque de modelación, el estudio aporta no solo a un mejor entendimiento del fenómeno del crimen, sino que también provee de sugerencias de accionar policial bajo criterios de optimalidad. Del mismo modo, la metodología de aplicación del modelo permite un alto grado de flexibilidad en las decisiones, lo que permite poder ser adaptado a múltiples escenarios de trabajo.

El artículo se organiza de la siguiente forma: el capítulo 2 presenta una breve revisión bibliográfica de importantes estudios cuantitativos del crimen. El capítulo 3 presenta el modelo teórico de éste trabajo y su metodología de aplicación. El capítulo 4 muestra los resultados obtenidos al utilizar la metodología de aplicación con datos reales. Finalmente, las conclusiones y trabajos futuros son presentados en el capítulo 5.

[1] Ver por ejemplo http://paleo.sscnet.ucla.edu/ucmasc.htm

2. Estudios Cuantitativos del Crimen

2.1. La Economía y el Crimen

Uno de los teóricos más influyentes en la criminología contemporánea es Gary Becker, quien en su ensayo *"Crime and Punishment: An Economic Approach"* [1], rompe con el paradigma que sostenía que el acto criminal era una acción cometida por personas socialmente oprimidas o mentalmente enfermas.

En sus estudios, se asume que los criminales son agentes racionales y que poseen una función de utilidad que desean maximizar. Los factores que influyen en la acción de delinquir son entre otras cosas, la probabilidad de ser atrapado, el castigo potencial que recibirían y las otras opciones de actividades que tienen disponibles.

Según Becker, la principal contribución de su ensayo es la demostración de que las políticas óptimas para combatir el comportamiento ilegal son parte de una decisión eficiente de asignación de recursos. En su estudio utiliza la teoría económica clásica, en donde existen múltiples modelos para la asignación eficiente de recursos, pero agregando a tales modelos, aspectos particulares del fenómeno del crimen (*e.g.* el castigo) los cuales son elementos no monetarios que afectan los costos que enfrenta la sociedad.

2.2. Teoría de Juegos y Criminalidad

La teoría de juegos posee un razonamiento analítico altamente aplicable al fenómeno criminal puesto que en todo delito coexisten al menos dos agentes con intereses contrapuestos (*e.g.* el criminal y la víctima). El desarrollo profundo de estos modelos comienza en los años 90 y destacan los juegos de inspección, estudios de relaciones sociales y de comportamiento terrorista [9].

Por ejemplo, el juego de inspección más básico proviene del modelo de George Tsebelis [16], quién plantea que existen infractores en potencia que podrían ser disuadidos mediante la aplicación de una multa. La medida busca aumentar los costos para aquellos quienes infringen la ley, pero al mismo tiempo se generan costos asociados al hecho de aplicar la multa. Paradójicamente, los resultados obtenidos de la interacción (equilibrio de Nash) muestran que la aplicación de una multa no tienen ningún efecto sobre el comportamiento del potencial infractor.

La gran mayoría de los modelos de teoría de juegos no llegan a una aplicación real, puesto que son complejos de manipular y resolver. A pesar de ello, algunos modelos han logrado ser aplicados con éxito en situaciones reales. Un trabajo reciente se llevó a cabo en el aeropuerto de Los Ángeles (LAX) [12],

en donde mediante un sistema computacional en base a un modelo de Stackelberg, se determina eficientemente cuáles puntos de monitoreo del aeropuerto deben ser abiertos cada día y hora de manera de que los criminales no detecten el patrón de comportamiento de seguridad.

2.3. Minería de Datos para el Estudio del Crimen

Múltiples técnicas de minería de datos han sido utilizadas en criminología con este fin, incluyéndose tanto modelos supervisados como no supervisados. Los principales han sido modelos de predicción, clustering y clasificación.

La predicción del crimen se ha usado como cualquier problemática en donde se tiene datos en una serie de tiempo. La técnica más utilizada para esto son las redes neuronales [4] en reemplazo de técnicas de series de tiempo tradicionales. Las series de datos que usualmente se manejan en criminología son a nivel de denuncias de delitos. La técnica se basa en la hipótesis de la existencia de patrones de comportamiento de los criminales, como por ejemplo el aumento de robos en la vía pública el día de pago de salarios.

La utilización de clustering en el crimen se ha usado de muchas maneras. La más utilizada es la identificación de *hot-spots* (imagen 1 [3]), la cual se realiza mediante algoritmos de análisis de densidad de puntos para formar los clusters [4]. Otra técnica de clustering que se ha empleado es el análisis del crimen es el algoritmo *k-medias*. En particular, se ha utilizado para identificar distintos tipos de delitos que a simple vista parecen ser los mismos, pero que en realidad pueden ser diferenciados si se agrupan apropiadamente [11].

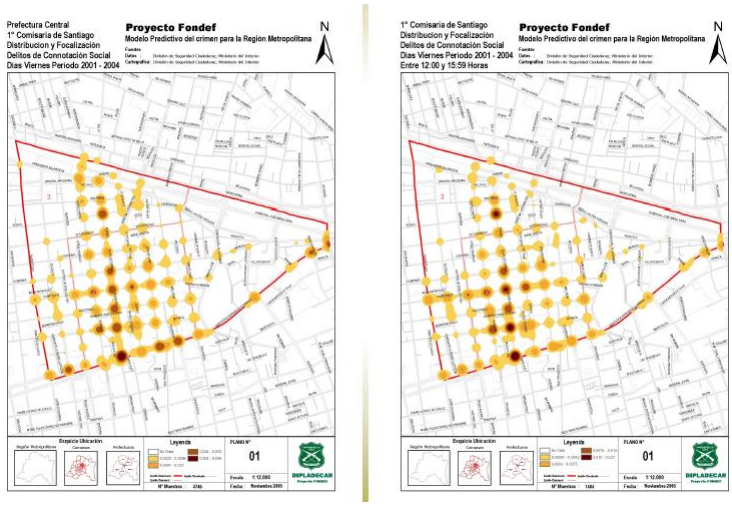


Figura 1: *Hot-spots* en el Centro de Santiago.

Las técnicas de clasificación en general se han utilizado para los crímenes "de cuello blanco", en donde la expresión se utiliza para los delitos cometidos por personas de nivel socioeconómico alto, en el cuadro de sus actividades

profesionales y con el objetivo de llegar a una ganancia más importante. Ejemplos de estos delitos son el blanqueo de dinero, falsificación de dinero y estafas en general. Su objetivo es identificar de la manera más certera posible cuáles montos de dineros en transacciones son ilícitas.

2.4. La Criminalidad y la Computación

La computación cada vez se vuelve más indispensable en el trabajo contra la criminalidad. Su aporte ha sido en el almacenamiento y procesamiento de información entre otras cosas.

Además, la computación ha permitido también el desarrollo de herramientas de simulación, lo que consiste en la emulación de situaciones reales, con el objetivo de observar virtualmente fenómenos y evitar experimentos en la vida real. En la criminología actual, esta técnica se ha convertido en una herramienta crucial, ya que permite crear laboratorios en donde se simulan situaciones de crimen y no se compromete la integridad de las personas y evita cuestiones éticas o morales.

Una de las técnicas de simulación que más se está usando en esta materia son los *agent based models* (modelos basados en agentes) [8]. Estos modelos combinan elementos de teoría de juegos, ecuaciones diferenciales y sistemas complejos. La base de estos modelos es la construcción de agentes definidos por una serie de características, que se relacionan con otros individuos y con el medio. Ejemplos de estos sistemas se han empleado para el estudio dinámico de *hot-spots* bajo distintos escenarios [17] y evoluciones de sistemas reveladores de patrones de crimen [6].

3. Modelo de Interacción entre Criminales y la Policía

3.1. Definición del Juego

El modelo que interacción entre criminales y la policía se contextualiza bajo los concepto de teoría de juegos. A continuación se definen los supuestos y elementos que enmarcan este trabajo:

1. Agentes: Criminales y la policía. Los criminales se definen como una masa continua de agentes y la policía como un agente que controla un conjunto continuo de recursos.

2. Estrategias: Las estrategias de los criminales se definen como las opciones de delitos disponibles. Estas pueden ser por ejemplo, elecciones de delinquir según áreas geográficas, tipos de delitos, momentos del día o combinaciones de

éstas. La estrategia de la policía es distribuir sus recursos sobre las opciones de delito de manera de dificultar a los criminales a que los efectúen.

3. Utilidades de los agentes: Los criminales obtienen un beneficio por delinquir en alguna de las opciones. La policía obtiene un beneficio por prevenir el crimen.

4. Supuestos del Juego: Congestión, actuar criminal y secuencialidad de las decisiones.

Congestión: Las opciones de crimen se tornan menos atractivas a medida que más criminales las escogen. Es decir, se produce un efecto de congestión entre los criminales. Los recursos policiales también provocan un efecto de congestión sobre las opciones de delito. En otras palabras, las estrategias se congestionan ya sea porque más criminales las escogen o porque hay recursos policiales asignados.

Actuar Criminal: Se debe definir *a priori* si los criminales actúan como agentes organizados o desorganizados, es decir, si maximizan el bienestar común o cada uno de ellos maximiza su propio bienestar[2].

Secuencialidad de las decisiones: Los agentes no deciden sus estrategias simultáneamente. Primero lo hace la policía enunciando su distribución de recursos sobre las opciones de delito y luego los criminales actúan óptimamente de acuerdo a aquella distribución.

3.2. El Modelo Teórico

En base a los elementos planteados anteriormente, el modelo que se plantea es en base a un grafo. En detalle, el modelo está compuesto por un nodo fuente, un nodo demanda y E arcos que los conectan, con funciones de costos asociadas c_e con $e \in E$ (ver figura 2). A continuación, se presenta la relación entre los elementos del crimen definidos anteriormente y el modelo de grafos.

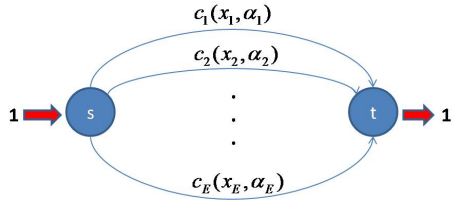


Figura 2: Grafo de elecciones criminales.

1. Los arcos representan las estrategias de los criminales (opciones de crimen).

2. Los flujos factibles a través de la red representan la cantidad de criminales que escogen cada una de las estrategias.

[2]Entiéndase bienestar por la utilidad que reciben los criminales por delinquir

3. La tasa de tráfico es unitaria, lo que se interpreta como que la cantidad de criminales que atraviesan el grafo es 1 (cada criminal representa una parte infinitesimal de flujo).

4. Las utilidades de los criminales son vistas como los costos asociados a cada arco. Es decir, los criminales en vez de maximizar su utilidad, minimizan su costo de viaje a través de la red.

5. La distribución de flujo resultante debido a los efectos de congestión, se supone que es una situación de equilibrio de Wardrop[3] o un flujo a costo mínimo, dependiendo si se considera el comportamiento de los criminales no organizado u organizado respectivamente.

6. Los recursos policiales modifican las funciones de costos aumentando su valor, lo cual afecta en la congestión de cada arco. Se considera que se cuenta con una unidad de recursos policiales que se distribuyen en las estrategias.

Lo cual, matemáticamente se formaliza con una instancia (G, d, c) que se especifica como sigue:

- **Grafo:** $G = (V, E)$ es un grafo dirigido donde $V = \{s, t\}$ representa a los vértices *fuente* y *demanda* (s y t respectivamente).
- **Arcos:** Todos los arcos $e \in E = \{1, ..., E\}$ están conectados desde s a t y se denotan como $(s, t)_e \quad \forall e \in E$.
- **Tasa de tráfico:** $d = 1$.
- **Función de costos:** $c_e \quad \forall e \in E$, funciones continuas, no negativas y no decrecientes.
- G posee un sólo *commodity*, el cual es representado por los flujos $\overrightarrow{x} = [x_1, \ldots, x_E]^t$.

Y la secuencia de interacción entre entre los criminales y la policía, en el contexto de la formulación como grafo, se explica en los siguientes pasos:

1. **Estrategia Policial**

La policía distribuye su unidad de recursos sobre las opciones de crimen de manera de modificar el valor de las funciones de costos. Se denota $\overrightarrow{\alpha} = [\alpha_1, \ldots, \alpha_E]^t$ a los recursos policiales y $\sum_{e \in E} \alpha_e = 1$, donde α_e es la cantidad de recursos asignados a la estrategia e.

Se le llama $\overrightarrow{\alpha}^*$ a la asignación de recursos escogida *ex-ante*. Entonces, las funciones de costo quedan definidas por $c_e(x_e, \alpha_e^*) \quad \forall e \in E$.

Además, para cumplir con el efecto de congestión provocado por la policía, se debe satisfacer que $\frac{\partial c(x,\alpha)}{\partial \alpha} \geq 0$, pues a mayor cantidad de recursos policiales en un arco, mayor es el costo de atravesarlo.

[3]Un flujo es un *Equilibrio de Wardrop* (flujo de Nash) si ningún agente puede mejorar su costo total cambiando su ruta.

2. **Estrategia Criminal**

Criminales desorganizados: Dada la distribución de recursos policiales, los criminales escogen individualmente la manera de minimizar su costo del viaje a través del grafo. Esto conlleva a un flujo en equilibrio de Wardrop para (G, d, c), el cual se alcanza cuando se iguala el costos de todos los arcos con flujo positivo. Esto se formaliza como sigue:

Equilibrio de Wardrop en Modelo de Crimen: Sea la instancia $(G, d, c)^{crimen}$ con funciones de costos $c_e(x_e, \alpha_e^*) \quad \forall e \in E$ donde $\sum_{e \in E} \alpha_e^* = 1$.

El flujo $\overrightarrow{x}^*$ es un equilibrio de Wardrop si se cumplen los siguientes puntos:

- $c_e(x_e^*, \alpha_e^*) = C \quad \forall e \in E$ tal que $x_e^* > 0$.
- $c_e(x_e^*, \alpha_e^*) \geq C \quad \forall e \in E$ tal que $x_e^* = 0$.

Para algún $C > 0$.

Criminales organizados: El flujo resultante se obtiene de la minimización del costo grafo. Es decir, $\overrightarrow{x}^*$ es el flujo a costo mínimo si y sólo si:

$$\begin{aligned} \sum_{e \in E} x_e^* \cdot c_e(x_e^*, \alpha_e^*) &= \min_{\overrightarrow{x}} \sum_{e \in E} x_e \cdot c_e(x_e, \alpha_e^*) \\ s.a. \quad \sum_{e \in E} x_e &= 1 \\ x_e &\geq 0 \quad \forall e \in E \end{aligned} \tag{1}$$

3.2.1. Estrategia Policial Óptima

Hasta ahora, el modelo planteado no supone que la policía actúa de manera óptima en la distribución de sus recursos. En esta sección se define la formulación matemática que representa la estrategia óptima de la policía tanto en situaciones de criminales organizados o desorganizados.

Para obtener este propósito, se quiere maximizar el costo social de los criminales. Esto se obtiene asignando los recursos de manera de maximizar el costo del grafo a sabiendas de la actuación criminal posterior.

Si los criminales actuaran de forma desorganizada, la estrategia óptima de la policía sería maximizar el costo del grafo sujeto a que los criminales alcanzarán posteriormente el equilibrio de Wardrop, es decir:

$$
\begin{aligned}
\max_{\vec{\alpha},\vec{x},\vec{y},C} \quad & \sum_{e\in E} x_e \cdot c_e(x_e,\alpha_e) \\
s.a. \qquad c_e(x_e,\alpha_e) &\geq C \quad \forall e \in E \\
c_e(x_e,\alpha_e) - (1-y_e)\cdot M &\leq C \quad \forall e \in E \\
x_e &\leq y_e \quad \forall e \in E \\
\sum_{e\in E} x_e &= 1 \\
\sum_{e\in E} \alpha_e &= 1 \\
C &> 0 \\
x_e &\geq 0 \quad \forall e \in E \\
\alpha_e &\geq 0 \quad \forall e \in E \\
y_e &\in \{0,1\} \quad \forall e \in E
\end{aligned}
\tag{2}
$$

Donde $M >> 0$ y las variables binarias $y_e \quad \forall e \in E$ obligan a que los costos se iguales sólo sobre los arcos ocupados.

Cabe destacar que la formulación anterior puede ser altamente compleja de resolver ya que el problema tiene funciones de costos no lineales y presenta variables del tipo continuas y binarias. No obstante, ésta puede ser simplificada si se agregan condiciones sobre las funciones de costos.

En particular, si $c_e(0,\alpha_e) = c_1(0,\alpha_1)$ y $\frac{\partial c_e(x_e,\alpha_e)}{\partial x_e} > 0 \quad \forall e \in E$ (todas las funciones de costos tienen el mismo punto de origen y son estrictamente crecientes), todos los arcos en equilibrio tienen un valor positivo y en consecuencia, la formulación 2 se simplifica al siguiente problema:

$$
\begin{aligned}
\max_{\vec{\alpha},\vec{x},\vec{y},C} \quad & \sum_{e\in E} x_e \cdot c_e(x_e,\alpha_e) \\
s.a. \qquad c_e(x_e,\alpha_e) &= C \quad \forall e \in E \\
\sum_{e\in E} x_e &= 1 \\
\sum_{e\in E} \alpha_e &= 1 \\
C &> 0 \\
x_e &\geq 0 \quad \forall e \in E \\
\alpha_e &\geq 0 \quad \forall e \in E
\end{aligned}
\tag{3}
$$

Lo que es equivalente a que el resultado de la formulación 2 se cumpla que $y_e^* = 1 \quad \forall e \in E$. Es decir, en el óptimo todos los arcos del grafo son utilizados $(x_e^* > 0 \quad \forall e \in E)$[4].

[4]Demostración en [7]

Finalmente, si los criminales actuaran organizadamente, la estrategia óptima de la policía estaría dada por la maximización del costo mínimo del grafo:

$$\begin{aligned} \max_{\vec{\alpha}} \min_{\vec{x}} \quad & \sum_{e \in E} x_e \cdot c_e(x_e, \alpha_e) \\ s.a. \quad & \sum_{e \in E} x_e = 1 \\ & \sum_{e \in E} \alpha_e = 1 \\ & x_e \geq 0 \quad \forall e \in E \\ & \alpha_e \geq 0 \quad \forall e \in E \end{aligned} \tag{4}$$

Las tres formulaciones anteriores entregan los valores óptimos $\vec{\alpha}^*$ de recursos policiales y los flujos de criminales $\vec{x}^*$ como mejor respuesta a dicha distribución de recursos.

3.3. Metodología de Aplicación

Anteriormente se plantea el modelo teórico que explica la interacción entre los criminales y la policía usando un enfoque de redes. En esta sección, se presenta una metodología genérica para aplicar tal modelo a situaciones reales.

A continuación, se describe en detalle la metodología de aplicación del modelo, ésta consta de cinco pasos y son enumerados en lo que sigue:

1. Selección de datos.
2. Construcción de estrategias.
3. Selección de funciones de costos.
4. Calibración del modelo.
5. Cálculo de la estrategia óptima de recursos policiales.

1. Selección de datos

Datos de información criminal: El primer paso de la metodología, es establecer una selección de datos relacionados con el comportamiento criminal que se adecúen de la mejor manera posible a los supuestos del modelo teórico. Para ello, se deben considerar elementos del tipo espacial, temporal y del tipo de crímenes seleccionados, de manera de mantener una homogeneidad en los datos y minimizar sesgos de comportamiento de los criminales.

Datos de información policial: Los datos de recursos policiales deben ser relacionados con la información de delitos y que proporcionen información respecto a los esfuerzos que se toman para la disuasión del crimen. Esto quiere decir que ésta información debe poder ser asociada a las mismas zonas

geográficas e intervalos temporales que se seleccionaron para la información criminal.

2. Construcción de estrategias

Una estrategia de delito puede ser vista como una combinación de múltiples variables que escoge un criminal al momento de delinquir (*e.g.* día de la semana, hora, lugar y tipo de delito). Entonces, luego de realizar una adecuada selección de datos, debe decidirse cómo se utilizarán estos datos para la construcción de las estrategias de manera de incluir al mismo tiempo la multiplicidad de variables.

Para ello, en esta metodología se propone utilizar métodos de clasificación mediante clustering. Así, se asume que existen patrones "escondidos" en los datos de modo que puedan ser clasificados en base a similitudes subyacentes. De esta manera, cada cluster y su magnitud representarán respectivamente, una estrategia de delito y la intensidad (proporción de criminales) con la que fue escogida.

3. Selección de funciones de costos

Luego de que se tienen seleccionadas las estrategias, se escogen las funciones de costos para los arcos de manera de modelar el efecto de congestión producido por los criminales y por los recursos policiales. Para esto, se imponen tres condiciones que deben satisfacer las funciones de costos:

1. Función no decreciente: $\frac{\partial c(x,\alpha)}{\partial x} \geq 0$. A mayor número de criminales, mayor es el costo que paga cada uno de ellos.

2. Función convexa: $\frac{\partial^2 c(x,\alpha)}{\partial x^2} \geq 0$. Congestión por efecto de mayor cantidad de criminales.

3. Efecto policial : $\frac{\partial c(x,\alpha)}{\partial \alpha} \geq 0$. Incremento del costo al aumentar los recursos policiales.

4. Calibración del modelo

La calibración del modelo consiste en determinar los parámetros de las funciones de costos que mejor se ajusten a los datos. Para este objetivo, se debe utilizar toda la información disponible de los pasos previos: los datos de información criminal y policial, las estrategias definidas y las funciones de costos.

Existen múltiples métodos de calibración de parámetros, sin embargo para ésta problemática, cualquiera sea el escogido, éste debe considerar la unidad de tiempo en la cual se asume la existencia de un equilibrio entre criminales y la policía. Este supuesto debe ser cuidadosamente escogido en base al estudio de los datos y al conocimiento experto.

5. Cálculo de la estrategia óptima de recursos policiales

Una vez calibrado el modelo, se quiere calcular la estrategia óptima de recursos policiales, esto es, la mejor asignación de recursos con el fin de encarecer la acción criminal. Específicamente basándose en técnicas de optimización sobre las ecuaciones 2, 3 y 4.

4. Aplicación

La metodología es aplicada utilizando los datos de denuncias de delito en el centro de Santiago (cuadrantes 1, 2 y 3), durante el período junio 2006 a mayo 2007 (51 semanas), ésta área cubre aproximadamente 3 km^2.

Respecto a la información policial, para este trabajo no se cuenta con información respecto a los recursos que empleó la policía durante el período y lugar en cuestión, por lo que se emplea una simulación semanal de datos basados en una suavización exponencial. El modelo se presenta a continuación.

$$\tilde{P}_e^t = \eta\left(\frac{P_e^{t-1} + P_e^{t-2} + P_e^{t-3}}{3}\right) + (1-\eta)P_e^{t-4} \Rightarrow \alpha_e^t = \frac{\tilde{P}_e^t}{\sum_{e\in E}\tilde{P}_e^t} \quad \forall e \in E. \quad (5)$$

Donde $\tilde{P}_e^t$ representa el nivel de crimen (cantidad de denuncias) estimado para la estrategia e en la semana t[5]. Esto se basa en la cantidad de denuncias observadas P_e de los cuatro períodos previos y η es el ponderador de la serie. Luego, la asignación de recursos policiales α_e^t se determina normalizando los valores de $\tilde{P}_e^t$.

Construcción de Estrategias

Para la construcción de estrategias, se emplea el método *k-medias* con distancias euclidianas. Para ello, las variables a utilizar son clasificadas según su tipo para luego transformarlas para minimizar los efectos de magnitud. La clasificación de las variables es la siguiente:

Variable espacial: Cuadrantes 1, 2 o 3.

Variables temporales: Día de la semana y rango horario (6 rangos horarios).

Variable circunstancial: Tipos de delito (hurto, robo con fuerza, robo con violencia).

Funciones de costos

Las funciones de costos que se proponen para asociar a los arcos son las funciones de congestión BPR y CCF [15], las cuales se modifican para tomar

[5]La simulación de datos se realizó posterior a la construcción de las estrategias.

en cuenta el efecto policial[6]:

$$f^{BPR}(x,\alpha) = 1 + x^{(\beta-\alpha)} \tag{6}$$

$$f^{CCF}(x,\alpha) = 2 + \sqrt{(\beta-\alpha)^2 \cdot (1-x)^2 + \tilde{\gamma}^2} - (\beta-\alpha)\cdot(1-x) - \tilde{\gamma} \tag{7}$$

Donde $\tilde{\gamma} = \frac{2(\beta-\alpha)-1}{2(\beta-\alpha)-2}$ y $\beta \geq 1$.

Por lo tanto, los parámetros a ajustar en la calibración del modelo son los valores de β_e de estas funciones.

Calibración del modelo

Se plantea un método que itera sobre una grilla de valores de los parámetros de las funciones de costo, de manera de seleccionar todas las combinaciones posibles que ofrece la grilla. En cada iteración, el algoritmo selecciona un conjunto de parámetros β y la información semanal de recursos policiales α simulados, para así determinar los parámetros de las funciones de costos en cada arco.

Luego, en base a esos costos parametrizados, se computa el flujo de Nash correspondiente (o de costo mínimo) y se compara con el flujo semanal observado[7]. Luego se registra con alguna medida de error la diferencia entre el valor observado y la del equilibrio computado. Finalmente, entre todas las ejecuciones se selecciona aquel conjunto de parámetros de la grilla que minimiza aquel error.

Cálculo de la estrategia óptima de recursos policiales

Para el modelo teórico en que los criminales no son organizados (problemas 2 y 3), se realizan los modelos de optimización con el solver de MS Excel 2007 desde diferentes puntos iniciales. Para el modelo teórico en que los criminales están coludidos (problema 4), la maximización se obtiene iterando sobre grillas de valores de los recursos policiales y luego computando el flujo a costo mínimo (de manera análoga al algoritmo de calibración de parámetros) para luego seleccionar los parámetros α que maximicen el costo mínimo del grafo.

4.1. Análisis de Robustez

Como se observa en los pasos metodológicos, existe una gama de decisiones que deben tomarse para calibrar el modelo teórico. Por ello, se plantea un caso base que considera un conjunto decisiones fijas y luego, se analizan otros

[6]Esto se hizo considerando un β real ajustado, que incorpora la asignación de recursos policiales α. Es decir, se plantea $\beta_{real} = \beta - \alpha$.

[7]El cual se obtiene según la intensidad semanal (cantidad de delitos normalizada) de cada cluster.

resultados sensibilizando algunas de estas decisiones. Arbitrariamente, el caso base se define como sigue:

Estrategias de crimen: Se obtienen utilizando *k-medias* y con un número definido de clases *a priori*.

Funciones de Costos: Se considera la función de costos *BPR* modificada.

Heurística de recursos policiales: Se utiliza el modelo de suavización exponencial planteado.

Comportamiento criminal: Se consideran criminales no organizados.

Y para el análisis de robustez se varían sobre el caso base como lo muestra la tabla 1.

Escenario	**Clustering**	**Func. Costos**	**Rend. Policial** $(\beta - \alpha^{\delta})$	**Org. Criminales**
Caso Base (C-base)	Cluster base	BPR	$\delta = 1$	Desorganizados
Caso C-estrategias	Cambio en N° clases	BPR	$\delta = 1$	Desorganizados
Caso C-costos	Cluster base	CCF	$\delta = 1$	Desorganizados
Caso C-rendimientos	Cluster base	BPR	$\delta = 0.95$	Desorganizados
Caso C-mafia	Cluster base	BPR	$\delta = 1$	Organizados

Tabla 1: Notación escenarios investigados.

4.2. Resultados

Resultados Clustering

Las estrategias criminales se determinaron utilizando el algoritmo *k-medias*, en donde se determinó que se utilizarán siete clases (C7), dada la magnitud e interpretación de éstas. La tabla 2 muestra una síntesis de los resultados obtenidos, en donde a cada clase se le bautiza con nombre que intenta explicar su comportamiento. La notaciones del campo *Delito* corresponden a H, V, F a los delitos hurto, robo con violencia y robo con fuerza respectivamente.

Para el escenario C-estrategias, se utiliza el resultado del algoritmo de *k-medias* de ocho clases [8].

4.2.1. Resultados Caso Base

Calibración de Parámetros

El resultado de la calibración de parámetros es el vector de valores del parámetro β de la función BPR. La tabla 5 muestra el valor promedio de las preferencias de los criminales en las 51 semanas y el valor de β obtenido y se observa que los parámetros β capturan el efecto de las preferencias de los criminales, en donde los mayores valores de β representan a las preferencias

[8] El detalle del resultado del clustering puede ser visto en [7]

Cluster	Magnitud	Cuadrante	Delito	Momento del día	Día de la Semana	Bautizo
1	8.8 %	1-2-3	F-V	Mañana	L-Ma-S-D	Robo temprano
2	14.8 %	3	H-V	Tarde	L-Ma-Mi-J-V-S-D	Robo en la Tarde en cuandrante 3
3	26.3 %	2	H-V	Tarde	L-Ma-Mi-J	Robo hora de almuerzo o vuelta del trabajo
4	19.9 %	2	H-V	Noche	V	Robo en la tarde-noche en cuadrante 2
5	8.6 %	1-2-3	H	Día	L-Ma-Mi-J-V	Hurto en horario de trabajo
6	8.3 %	3	V	Noche	S-D	Robo en fin de semana nocturno
7	13.3 %	1	V	Tarde	S-D	Robo en paseo de fin de semana

Tabla 2: Tabla resultados clustering.

más escogidas (clases 3 y 4). Además, se muestra que entre las clases de menor magnitud (1, 5 y 6), estos parámetros presentan un valor distinto a pesar de la cercanía en magnitud.

Escenario/Cluster	1	2	3	4	5	6	7
Promedio C7	8.8 %	14.8 %	26.3 %	19.9 %	8.6 %	8.3 %	13.3 %
β C-base	1.83	2.37	3.36	2.87	1.80	1.79	2.24

Tabla 3: Resultados β Caso Base.

Optimización de Recursos Policiales

La optimización de recursos policiales se obtiene resolviendo el problema 3, puesto que la función BPR modificada cumple con las condiciones requeridas. El resultado se despliega en la tabla 4 y compara entre el escenario actual y el optimizado los valores de la distribución de los recursos policiales (α) y de las preferencias criminales (x).

Escenario/Cluster	1	2	3	4	5	6	7
α prom., actual C7	8.9 %	14.7 %	26.0 %	20.3 %	8.6 %	8.2 %	13.3 %
x prom., actual C7	9.0 %	14.5 %	26.3 %	19.8 %	8.7 %	8.3 %	13.4 %
α prom., C-base	1.3 %	56.4 %	0.0 %	0.0 %	0.0 %	0.0 %	42.4 %
x prom., C-base	9.8 %	9.7 %	28.5 %	23.0 %	9.6 %	9.5 %	9.8 %

Tabla 4: Resultados de α óptimo en Caso Base.

En base a lo anterior, se observa que el caso base muestra las asignaciones óptimas de recursos policiales se distribuyen sólo en las estrategias intermedias

(2 y 7) y en muy pequeña cantidad en la 1, las magnitudes son 56.4 %, 46.4 % y 1.3 % respectivamente. Esta distribución provoca una reacción teórica de los criminales que disminuye los ataques en las estrategias intermedias y aumenta los ataques en todas las demás. Como consecuencia, los ataques en todas las estrategias, a excepción de las mayores (3 y 4), se equiparan en magnitud. (9.6 %±0.1 %).

Finalmente, para medir la efectividad de la optimización, se compara las diferencia relativa del costo del grafo entre la situación actual, la maximización y la minimización, de manera de obtener la posición porcentual en que se encuentra la situación actual de la óptima. Los resultados obtenidos fueron respectivamente 0.01453, 0.01477 y 0.01441, lo que indica que la optimización aumenta en un 65 % el costo variable del grafo respecto a la situación actual, cumpliendo ampliamente el objetivo del modelo en términos de encarecer y dificultar el actuar criminal.

4.2.2. Resultados de Robustez

Los resultados de los parámetros β de los escenarios de robustez se despliegan en la tabla 5. Los análisis se explican comparativamente respecto al caso base.

Escenario/Cluster	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Promedio C7	8.8 %	14.8 %	26.3 %	19.9 %	8.6 %	8.3 %	13.3 %	-
β C-base	1.83	2.37	3.36	2.87	1.80	1.79	2.24	-
β C-rendimientos	1.88	2.43	3.45	2.95	1.85	1.84	2.30	-
β C-mafia	1.64	2.23	3.42	2.80	1.60	1.62	2.08	-
β C-costo	1.48	2.29	3.87	3.10	1.43	1.43	2.10	-
Promedio C8	7.6 %	13.7 %	17.2 %	17.1 %	14.1 %	11.7 %	6.0 %	12.7 %
β C-estrategias	1.84	2.44	2.73	2.71	2.43	2.27	1.65	2.32

Tabla 5: Resultados β escenarios C7.

En el escenario C-mafia, la brecha entre los parámetros asociados a las estrategias de crimen menos numerosas y las más numerosas se amplifica: las diferencias se dan entre los valores 1.62 y 3.42, mientras que en el caso base los βs toman los valores 1.79 y 3.36 respectivamente. Además, este escenario presenta una leve diferencia entre los parámetros β del cluster 5 y 6, en donde su relación de orden es inversa a las demás. Este fenómeno se debe a la sensibilidad que presenta el algoritmo propuesto, en donde el criterio de minimización de errores absolutos provoca aquel efecto.

El escenario C-costos presenta el mismo efecto de amplificación de C-mafia, pero aun más pronunciado: los β mínimo y máximo son 1.43 y 3.87 respectivamente. Adicionalmente, los valores β mínimos tienen el mismo valor (1.43), lo que se interpreta como que las pequeñas fluctuaciones de valores no son captadas con parámetros de dos dígitos significativos.

En el escenario C-rendimientos, los β se amplifican de manera no lineal en todas las clases, lo que es consistente con el efecto inducido con el exponente γ en la expresión α^{γ}.

El escenario C-estrategias, es equivalente al caso base en términos de las funciones de costo, rendimientos policiales y organización criminal. Los comportamientos en magnitud de los parámetros en este caso son muy similares a las del caso base.

Respecto a los resultados de la optimización, se presentan los resultados de las estrategias óptimas comparando los valores promedios del escenario actual. Los resultados son presentados en la tabla 6.

Escenario/Cluster	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
α prom., C-rend.	8.3 %	56.1 %	0.0 %	3.7 %	3.6 %	3.4 %	24.9 %	-
x prom., C-rend.	8.9 %	9.7 %	28.6 %	22.6 %	9.2 %	9.1 %	11.9 %	-
α prom., C-costos	0.0 %	0.0 %	0.0 %	0.0 %	43.0 %	0.6 %	56.4 %	-
x prom., C-costos	9.5 %	16.0 %	27.2 %	22.1 %	6.1 %	9.1 %	10.0 %	-
α prom., C-mafia	0.0 %	19.3 %	0.0 %	76.3 %	0.0 %	0.0 %	4.3 %	-
x prom., C-mafia	9.8 %	14.4 %	28.2 %	14.4 %	9.3 %	9.6 %	14.4 %	-
α prom., C-estrat.	0.0 %	32.5 %	0.0 %	0.0 %	31.5 %	15.5 %	0.0 %	20.5
x prom., C-estrat.	8.6 %	11.8 %	19.1 %	18.8 %	11.8 %	11.8 %	6.4 %	11.8 %

Tabla 6: Resultados de α óptimo análisis de robustez C7 y C-estrategias.

El análisis de los resultados es presentado a continuación de la misma manera como fue presentado el caso base.

C-rendimientos: Se menciona que el efecto del exponente γ sobre los recursos policiales α genera una amplificación no lineal de los parámetros β. Esto provoca que los recursos policiales sean menos eficientes a medida que aumentan en una estrategia. En consecuencia, los recursos que antes se repartían en las opciones de crimen de magnitud intermedia, ahora se reparten en todas las estrategias exceptuando en la de mayor tamaño (clase 3). Además, se observa que el efecto de ineficiencia provoca que la reacción criminal no sea equiparada como en el caso base, presentando una mayor varianza de magnitud de las estrategias menores e intermedias (9.8 %±1.3 %).

C-costos: Las estrategias óptimas policiales bajo este escenario son distintas a las anteriores. En este caso, se priorizan las estrategias 5 y 7, las cuales son de categoría intermedia y baja. La diferencia también se da en que la clase 1, que es la que en magnitud se ubica entre las clases 5 y 7, no tiene asignado recursos policiales. Esto sugiere la posibilidad de haber obtenido un óptimo local que probablemente se encuentra cercano al global.

C-mafia: En el escenario que se modela a los criminales organizados, la asignación óptima de recursos policiales prioriza también a las estrategias de tamaño intermedio (clases 4, 2 y 7). La diferencia está en que este caso, la estrategia más efectiva para asignar recursos policiales es la 4, que es la se-

gunda en tamaño. Esto significa que bajo criminales organizados, los recursos policiales se vuelven más eficientes designándolos a las clases intermedias de mayor tamaño (clases 4 y 2), que en contraste de los dos escenarios anteriores, se priorizan las clases de intermedias pero de menor tamaño (clases 2, 7 y 1).

C-estrategias: Este escenario presenta una diferencia de configuración de cluster respecto al caso base. El aporte de este escenario está en observar como cambia la asignación óptima de recursos policiales, cuando sólo se cambia el tamaño y el número de estrategias criminales. El resultado, es equivalente al obtenido en el caso base: se prioriza la asignación de recursos en las estrategias de magnitud intermedia. Sin embargo, al haber más estrategias en dicha categoría, la distribución óptima resultante queda mejor distribuida: 32.5 %, 31.5 %, 20.5 % y 15.5 % en las estrategias 2, 5, 8 y 6 respectivamente.

Finalmente, de manera análoga al caso base se realiza la medición cuantitativa de las efectividades de las asignaciones óptimas de recursos policiales. La tabla 7 muestra las posiciones relativas del caso actual frente al mínimo y máximo estimado.

Escenario	C-base	C-rendimientos	C-mafia	C-costos	C-estrategias
Dif. Actual-Máximo	65.0 %	16.3 %	24.1 %	34.8 %	7.5 %

Tabla 7: Brecha del costo del grafo entre caso original y máximo en cada escenario.

De estos resultados, se observa que en general los resultados de la maximización presentan una gran brecha en el costo del grafo respecto a la situación actual, exceptuando el escenario C-estrategias. Se destacó ya que el caso base presenta una brecha de un 65.0 % entre en costo del grafo actual y el obtenido de la maximización, pero se observa que la variante C-rendimientos presenta una brecha de costos claramente menor (16.3 %). Los escenarios C-costos y C-mafia presentan brechas respectivas de 34.8 % y 24.1 % de mejora respecto a la situación actual. Luego, todos estos casos cumplen con el objetivo de dificultar o encarecer la reacción criminal.

Para concluir, es importante señalar que los resultados de la optimización de recursos policiales obedecen sólo a los criterios planteados e ignoran todo efecto externo al modelo. Además, los resultados pueden no reflejar necesariamente una decisión real y deben entenderse como un resultado teórico que debiese complementarse con un juicio experto.

5. Conclusiones y Trabajos Futuros

En este trabajo se estudia el fenómeno del crimen en la vía pública como una interacción competitiva entre criminales y la policía. Para esto, se presenta un modelo matemático basado en teoría de juegos en grafos, el cual es calibrado

según datos reales de crímenes.

El estudio es realizado sobre cinco escenarios distintos que representan diversos supuestos sobre el modelo teórico. De esta manera se obtiene una comprensión amplia del fenómeno del crimen y se evalúa la robustez del modelo planteado y de su metodología de aplicación. Como resultado, es determinan estrategias óptimas del actuar policial considerando el comportamiento criminal.

El modelo consiste en un grafo que representa a los criminales por un flujo unitario que debe viajar desde un nodo fuente a un nodo demanda. Las diferentes rutas del grafo representan a las opciones de crimen. Las utilidades son vistas como los costos de los caminos incurridos por los criminales al momento de viajar por el grafo. Los recursos policiales son representados como parámetros que alteran las funciones de costos *a priori* y se consideran también unitarios.

Los supuesto del modelo se basan en que las opciones de crimen se saturan, ya sea porque más criminales las escogen o porque hay mayor cantidad de recursos policiales. Este hecho se representa con el flujo de Nash o con el de costo mínimo, dependiendo si se asumen criminales desorganizados u organizados respectivamente.

Para la metodología de aplicación del modelo teórico usando datos reales, se proponen cinco pasos: selección de datos, construcción de las estrategias, selección de las funciones de costo, calibración del modelo y cálculo de la estrategia óptima de recursos policiales. Ésta se testea con datos de denuncias de delitos de la Primera Comisaría de Santiago durante el período de un año.

La construcción de las estrategias se lleva a cabo utilizando el algoritmo *k-medias*. Se definieron para el caso base siete clases que representan siete opciones de delitos potenciales que escogen los criminales. Las funciones de costos de los empleadas se inspiraron en las clásicas utilizadas en ingeniería de transporte. Estas debieron ser modificadas de manera de incluir el efecto de los recursos policiales y al mismo tiempo satisfacer las condiciones que provocan el efecto de congestión. Finalmente las funciones de costos son calibradas sobre el caso baso y cuatro escenarios propuestos, que capturan variantes de la metodología de aplicación.

Los resultados de la asignación óptima de recursos policiales entrega resultados novedosos. Para todos los escenarios, se obtuvo que las asignaciones óptimas estaban en las opciones de crímenes de magnitud intermedia, dejando sin recursos tanto a las estrategias mayores como a las de menor intensidad. Estos resultados son analizados caso a caso en cada escenario y es posible diferenciar los elementos diferenciadores en cada uno de ellos.

De este trabajo se han desprendido potenciales temáticas a investigar tanto en ámbitos teóricos como aplicados. Dentro de las líneas teóricas, destacan principalmente los elementos relacionados con el modelamiento en redes, los

modelos de optimización y los algoritmos de calibración. De las líneas aplicadas, los desafíos aparecen en la ampliación del modelo y en la extrapolación a otros fenómenos de interacción de crimen.

Agradecimientos: Un especial agradecimiento al grupo CEAMOS y al Ministerio del Interior por el financiamiento parcial y en la transmisión de experiencia para este estudio.

Referencias

[1] Becker, G. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 78: 169-217. 1968.

[2] Beirne, Piers. Adolphe Quetelet and the Origins of Positivist Criminology. *American Journal of Sociology* 92(5): pp. 1140-1169. 1987.

[3] Benavente, J. Análisis Espacial de la Criminalidad Basado en Georeferenciación de Denuncias. *Presentación en Taller de Inteligencia de Negocios.* Santiago de Chile. Junio 2008.

[4] Corcoran, J. J., Wilson, I. D., Ware, J. A. Predicting the Geo-temporal Variations of Crime and Disorder. *International Journal of Forecasting*, 19: 623-634. 2003.

[5] Guerry, A. Essai sur la Statistique Morale de la France. 1833.

[6] Liu, L., Eck, J. Artificial Crime Analysis Systems: Using Computer Simulations and Geographic Information Systems. Publicado por Idea Group Inc (IGI). Enero 2008.

[7] Lobato, J. L. Modelo Aplicado de Teoría de Juegos para el Estudio del Crimen en la Vía Pública. Tesis para optar al grado de Magíster en Gestión de Operaciones. Universidad de Chile. Junio 2009.

[8] Macal, C. M., North, M. J. Tutorial on Agent-Based Model Modeling and Simulation. Proceedings of the 2005 Winter Simulation Conference.

[9] Molina, E. Aplicaciones de la Teoría de Juegos al Análisis de Comportamientos Criminales. Escuela de verano de la Universidad Complutense *"Matemáticas para la Seguridad Contra la Criminalidad"*. El Escorial, España. Julio 2008.

[10] Nuño, J. C. Aportación de las Matemáticas a la Lucha Contra la Criminalidad. Escuela de verano de la Universidad Complutense *"Matemáticas para la Seguridad Contra la Criminalidad"*. El Escorial, España. Julio 2008.

[11] Perversi, I. Aplicación de Minería de Datos para la Exploración y Detección de Patrones Delictivos en Argentina. Tesis de Grado en Ingeniería Industrial, Instituto Tecnológico de Buenos Aires. 2007.

[12] Pita, J., Jain, M., Marecki, J., Ordóñez, F., Portway, C., Tambe, M., Western, C., Paruchuri, P., Kraus, S. Deployed ARMOR Protection: The Application of a Game Theoretic Model for Security at the Los Angeles International Airport. International Conference on Autonomous Agents. *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems: industrial track*. Pág. 125-132. 2008.

[13] Roughgarden, T. Selfish Routing and the Price of Anarchy. *Department of Computer Science, Standford University*. 7 de Enero, 2007.

[14] Schmeidler, D. Equilibrium Points of Nonatomic Games. *Journal of Statistical Physics*. 7(4):259-300. 1973.

[15] Spiess, H. Conical Volume-Delay Functions. *Transportation Science*, Vol 24. No. 2. 1990.

[16] Tsebelis, G. Penalty Has No Impact on Crime: A Game Theoretic Analysis. *Rationality and Society 2*, 255-86. Julio 1990.

[17] UC MaSC Project. Matematical and Simulation Modeling of Crime. [Página Web] http://paleo.sscnet.ucla.edu/ucmasc.htm.

[18] Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. En Proceedings of the Institute of Civil Engineers, Pt. II, vol. 1, páginas 325-378. 1952.

Optimización de la Recolección de Residuos en la Zona Sur de la Ciudad de Buenos Aires

Flavia Bonomo[*]
Guillermo Durán[**]
Federico Larumbe[***]
Javier Marenco[****]

Resumen

En este trabajo presentamos una propuesta para optimizar el recorrido de los camiones encargados de la recolección de los contenedores de residuos en la zona sur de la ciudad de Buenos Aires (Argentina), utilizando técnicas de programación matemática. Se describe el proceso de depuración de los datos disponibles y los pasos realizados para aplicar herramientas de programación lineal entera para este problema. Se obtienen propuestas que minimizan la distancia total recorrida por cada camión y muestran una disminución de los trayectos de entre un 10 % y un 40 % con respecto a los recorridos actuales. Además, estos nuevos trayectos disminuyen sustancialmente el desgaste final de los camiones.

Palabras Clave: Programación Entera, Recolección de Residuos, Ruteo de Vehículos.

[*]Departamento de Computación, FCEyN, Universidad de Buenos Aires y CONICET, Argentina
[**]Departamento de Ingeniería Industrial, FCFM, Universidad de Chile; Departamento de Matemática, FCEyN, Universidad de Buenos Aires y CONICET, Argentina
[***]Departamento de Computación, FCEyN, Universidad de Buenos Aires, Argentina
[****]Instituto de Ciencias, Universidad Nacional de General Sarmiento, Argentina

1. Introducción

La recolección de residuos puede dividirse en 3 grandes tipos: domiciliaria, comercial e industrial [5]. La recolección domiciliaria consiste en atender fundamentalmente casas particulares. Un mismo camión puede atender en una misma ruta desde 100 hasta 1000 casas cada día. La frecuencia del servicio varía según las ciudades, aunque en general las rutas suelen repetirse una vez todos los días. La recolección comercial consiste en atender shoppings, restaurants o edificios de oficinas. Cada ruta comercial puede atender entre 60 y 400 clientes, con 2 o 3 visitas al depósito en cada día. Cada cliente puede tener una ventana de tiempo prefijada para ser visitado. La recolección industrial consiste en atender fábricas, obras y edificios en construcción. La diferencia principal con la recolección comercial es que los contenedores a ser vaciados suelen ser 4 o 5 veces más grandes, y en muchos casos sólo un contenedor puede ser atendido por vez, lo que convierte al problema en uno de ruteo de vehículos muy diferente al de la recolección comercial. Una buena recopilación sobre artículos de ruteo de camiones para recolección de residuos aparece en [7].

Los objetivos en este tipo de problemas suelen ser diversos: minimizar el número de camiones, minimizar la distancia recorrida, conseguir rutas compactas o balancear la carga entre los distintos camiones.

Existen en la literatura varios trabajos que muestran la aplicación de sistemas basados en técnicas de programación matemática para la recolección de residuos en diferentes ciudades del mundo. Algunos ejemplos son el de Chicago, en los Estados Unidos [4], el de una importante ciudad en Taiwan [2] y el de Lisboa, en Portugal [10]. La mayor parte de los artículos que muestran la solución de problemas reales de recolección de residuos consiste en la implementación de heurísticas, debido a la dificultad de los problemas.

En la Ciudad de Buenos Aires se encuentra en implementación un nuevo sistema para la recolección de residuos domiciliarios, que consiste en ubicar contenedores distribuidos por la ciudad, para que los vecinos depositen en estos contenedores las bolsas de residuos. Cada contenedor tiene 1000 litros de capacidad. La intención de este proyecto de la Ciudad es reemplazar la deposición de residuos en cestos individuales por el uso de estos contenedores generales, contribuyendo así a la higiene general y a la eficiencia en la recolección.

A los efectos de la recolección de residuos, la ciudad se divide en 6 zonas (figura 1). En cada zona mostrada en la figura, se ven pintadas con un color oscuro los lugares donde el nuevo sistema se encuentra actualmente implementado. Las zonas 1 a 4 y la zona 6 tienen asignada una empresa cada una, cuya responsabilidad es gestionar la recolección de residuos en esa área. La zona 5

está asignada al Ente de Higiene Urbana (EHU), un organismo del Gobierno de la Ciudad de Buenos Aires con las mismas responsabilidades.

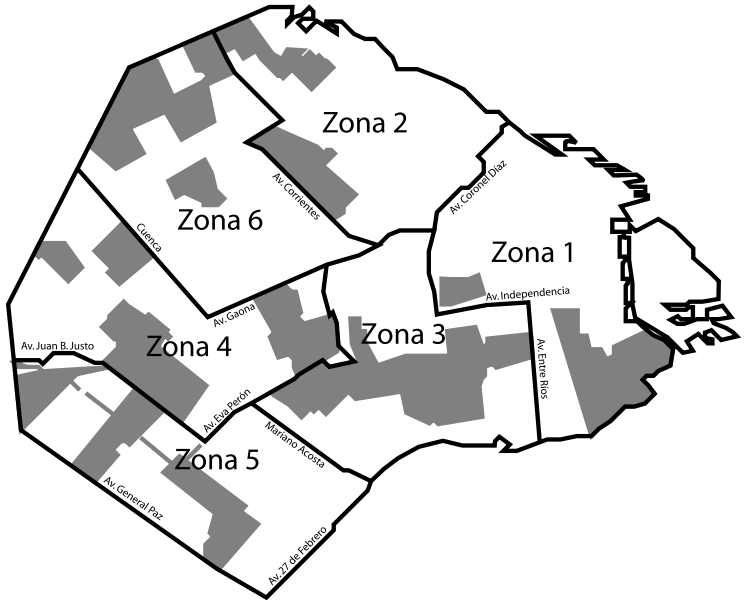


Figura 1: Zonas para la recolección de residuos en Buenos Aires.

La ubicación de los contenedores está prefijada de antemano, y los contenedores que recolecta el EHU están agrupados en 4 sub-zonas. Se tienen 4 camiones de recolección de contenedores. Cada camión tiene una sub-zona asignada, y hace dos recorridos iguales, uno por la mañana y otro por la tarde. El recorrido se inicia en el EHU, se dirige al primer contenedor, recolecta todos sus contenedores asignados, descarga en el depósito los residuos y vuelve al EHU. Aproximadamente, el camión tarda 3 minutos en recolectar un contenedor y limpiar el entorno donde se encuentra. A la mañana, los camiones salen a las 7 hrs. del EHU y vuelven aproximadamente a las 15 hrs. A la tarde, salen a las 18 hrs. Las sub-zonas tienen 47, 133, 134 y 161 contenedores, respectivamente.

En un camión entran 15000 kg. de residuos. A la mañana los camiones recolectan entre 10000 kg. y 15000 kg., y a la tarde cargan entre 2500 kg. y 5000 kg. Hay un camión de limpieza que pasa cada 10 días detrás del camión de recolección y limpia los contenedores vacíos. En el área de descarga, se compactan los residuos y un camión con mayor capacidad los lleva al lugar final de deposición, ubicado en el conurbano de la ciudad.

El propósito del presente trabajo es mejorar la ruta de cada camión con respecto a las rutas hoy existentes, de modo que cada camión parta del EHU, recolecte todos los contenedores de su sub-zona, vaya al depósito y vuelva al

EHU. Los objetivos consisten en minimizar la distancia recorrida y disminuir el desgaste de los camiones, en ese orden. Cuando el camión se encuentra cargado, existe un desgaste considerable de varias de sus partes, con lo cual es conveniente que no recorra una gran distancia con demasiado peso.

Cabe destacar que los empleados cobran por jornada laboral de medio tiempo, por lo que la optimización no influye en el costo laboral, aunque permite realizar la misma tarea recorriendo menos distancia, provocando un menor desgaste de los camiones y permitiendo un mejor tráfico en la ciudad.

Con los datos disponibles actualmente, no es posible estimar apropiadamente los tiempos de recorrido en las calles y avenidas que deben recorrer los camiones. Por este motivo, no se plantea en este trabajo la optimización de los tiempos de recorrido, y en cambio se trabaja con la distancia total como objetivo principal.

Para cuantificar el desgaste del camión a lo largo del recorrido, utilizamos el *trabajo* realizado por el camión. El trabajo en un tramo se calcula como el producto de la distancia recorrida por la fuerza realizada en esa dirección. La unidad básica de trabajo en el Sistema Internacional es el Newton por metro, que se denomina Joule (J). En cada tramo del itinerario entre dos contenedores consecutivos, el trabajo que realiza el camión es el producto de la distancia por el peso de la carga del camión. El trabajo total realizado es la suma de los trabajos de cada tramo. Hasta lo mejor de nuestro conocimiento, no hay trabajos en la literatura de recolección de residuos donde se plantee la disminución del trabajo realizado por los camiones.

En este artículo describimos la aplicación de herramientas computacionales para la resolución efectiva de este problema por medio de técnicas de programación lineal entera, para los cuatro camiones del EHU en las sub-zonas actualmente cubiertas por contenedores. Se describen las herramientas utilizadas y el trabajo computacional de depuración e interpretación de los datos. Como se verá en la sección 5, se obtuvieron recorridos muy eficientes desde el punto de vista de los objetivos planteados, dado que se consiguió disminuir la distancia y el trabajo realizado por los camiones entre un 10 % y un 45 %, con respecto a los recorridos actuales.

Este trabajo está organizado de la siguiente forma. En la sección 2 definimos una representación natural del mapa de la ciudad por medio de grafos. En la sección 3 mostramos la forma de reducir nuestro problema al Problema del Vendedor Viajero. En la sección 4 se describe brevemente la implementación del programa desarrollado para realizar tareas de procesamiento del mapa, cálculo de itinerarios óptimos y visualización de los resultados. En la sección 5 se presentan los resultados obtenidos, y el trabajo se cierra en la sección 6 con las conclusiones obtenidas y los próximos pasos.

2. Modelo del mapa

En esta sección se describen los datos disponibles para la realización del estudio, y se define en detalle cómo representar el mapa por medio de un grafo, con la intención de que este modelo permita calcular recorridos en vehículo por la ciudad. Esta representación es la base de todo el trabajo de implementación dado que, además del cálculo de recorridos, permite validar la información procesada y depurar los datos incorrectos.

La Unidad de Sistemas de Información Geográfica (USIG) del Gobierno de la Ciudad de Buenos Aires proveyó el mapa de la ciudad en formato Shape (SHP), uno de los formatos standard para la representación de mapas e información geográfica. El archivo contiene una base de datos con cada cuadra de la ciudad. Para cada cuadra se tiene la ubicación de sus esquinas, el nombre de la calle, la altura inicial, la altura final y el sentido. Por otro lado, se conoce la posición de los semáforos en la ciudad.

La información sobre los semáforos contiene un punto geográfico por cada uno de ellos, que indica aproximadamente dónde se encuentra el mismo. Si hay dos calles que se intersecan en el punto p y hay un semáforo en la intersección, entonces en la base de datos de semáforos hay un punto cercano a p. En nuestro modelo, necesitamos saber si en la intersección de dos calles cualesquiera hay un semáforo. Para eso, buscamos si hay un punto en dicha base de datos cuya distancia al punto de intersección de las dos calles sea menor que un cierto valor D. Realizamos varias pruebas con distintos valores de este parámetro y encontramos que $D = 3{,}5$ metros detecta correctamente los semáforos.

2.1. Construcción del grafo y cálculo de recorridos

Describimos a continuación una representación natural del mapa de la ciudad por medio de un grafo, de manera tal que esta representación permita calcular distancias recorridas en vehículo. Dado que necesitamos tener en cuenta el sentido de circulación en las cuadras, trabajamos con un grafo dirigido. Los nodos se definen como las posiciones significativas donde puede ubicarse un vehículo, dados por los extremos de cada cuadra y las posiciones de los contenedores.

Si un vehículo está en una intersección de cuadras, se marca como dos posiciones diferentes cuando está en una u otra cuadra de la intersección. Por otro lado, en las cuadras doble-mano, la posición del vehículo cambia si está en una mano de la cuadra o en la contramano. Luego, la posición del vehículo se define como la cuadra, la mano y las coordenadas geográficas. Existe un arco entre dos nodos que representan dos posiciones si un vehículo puede pasar

directamente de una a otra. Por esta razón, un camino en el grafo representa un recorrido válido de un vehículo en la ciudad. El peso de un arco es la distancia en metros entre las posiciones que representan sus nodos extremos.

En una intersección de dos calles doble-mano, tenemos entonces 8 posiciones: mano y contramano de cada una de las 4 cuadras. Si no hay giros prohibidos, estos puntos se conectan de forma que se pueda girar para un lado y para el otro. Sin embargo, cuando hay un semáforo en la intersección, el vehículo no puede girar a la izquierda. En la figura 2 se ve este caso, mostrando los giros que un vehículo que va por la Av. Rivadavia puede realizar hacia Boyacá. Notar que los 8 puntos que aparecen en la figura tienen las mismas coordenadas geográficas, lo que varía es la cuadra y/o la mano.

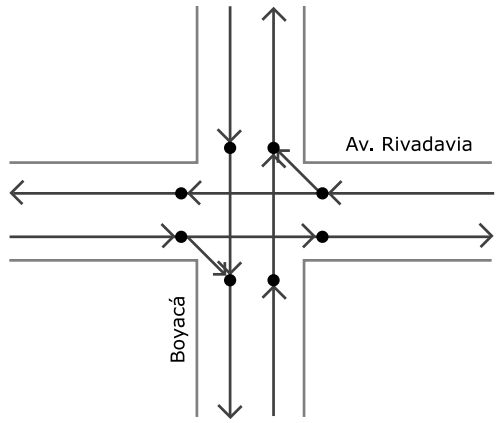


Figura 2: Arcos de giros de Av. Rivadavia a Boyacá con semáforo.

Formalizamos a continuación la definición del grafo $G = (V, A)$ que modela los recorridos en vehículo. Sea $C = \{(s_1, p_1), \ldots, (s_n, p_n)\}$ el conjunto de cuadras de la ciudad, donde $s_i \in S = \{$*creciente, decreciente, doble-mano*$\}$ y $p_i = \left(q_{i_1}, \ldots, q_{i_{t_i}}\right)$ es el conjunto de puntos geográficos de la cuadra i, para $i = 1, \ldots, n$. Los puntos q_{i_1} y $q_{i_{t_i}}$ corresponden a las esquinas, y los puntos $q_{i_2}, \ldots, q_{i_{t_i-1}}$ corresponden a los contenedores ubicados en la cuadra. Los puntos de cada cuadra están ordenados en el sentido creciente de la altura de la cuadra. Si el sentido es creciente, el vehículo puede transitar de q_{i_1} a q_{i_2}, de q_{i_2} a q_{i_3} y así hasta $q_{i_{t_i}}$. Si el sentido es decreciente, el vehículo puede transitar de $q_{i_{t_i}}$ a $q_{i_{t_i-1}}$, de $q_{i_{t_i-1}}$ a $q_{i_{t_i-2}}$ y así hasta q_{i_1}. Si el sentido es doble-mano, el vehículo puede transitar en ambos sentidos.

Sea $Q_i = \left\{q_{i_1}, \ldots, q_{i_{t_i}}\right\}$ el conjunto de puntos de la cuadra c_i (es decir, el vector p_i considerado como un conjunto), y definimos M_i como el conjunto de manos por las que se puede circular en la cuadra:

$$M_i = \begin{cases} \{creciente\} & \text{si } s_i = creciente \\ \{decreciente\} & \text{si } s_i = decreciente \\ \{creciente,\ decreciente\} & \text{si } s_i = doble\text{-}mano \end{cases}$$

Para cada cuadra i, el conjunto de nodos del grafo que origina la cuadra i es $V_i = \{c_i\} \times Q_i \times M_i$, dados por las distintas posiciones que puede ocupar un camión en la cuadra en función de sus puntos geográficos y las manos de la cuadra. El conjunto de nodos de G está dado por $V = \bigcup_{i=1}^{n} V_i$.

Decimos que dos nodos (c_i, q_{i_k}, m) y (c_j, q_{j_l}, m') son *consecutivos* si y sólo si $i = j$, $m = m'$, $m = creciente \Rightarrow l = k + 1$ y $m = decreciente \Rightarrow l = k - 1$. Decimos que un nodo $(c_i, q_{i_k}, m) \in V$ es un *extremo de salida* de la cuadra $c_i = (s_i, p_i)$ con $p_i = \left(q_{i_1}, \ldots, q_{i_{t_i}}\right)$ si y sólo si $m = creciente \Rightarrow k = t_i$ y $m = decreciente \Rightarrow k = 1$. Decimos que un nodo $(c_i, q_{i_k}, m) \in V$ es un *punto de entrada* a la cuadra c_i si es una esquina y no es un extremo de salida de c_i.

Dados dos nodos $v_1 = (c_i, q_{i_k}, m) \in V$ y $v_2 = (c_j, q_{j_l}, m') \in V$, decimos que hay un *giro* de v_1 a v_2 si $q_{i_k} = q_{j_l}$, v_1 es extremo de salida de c_i y v_2 es punto de entrada a c_j. Para definir el ángulo de este giro, consideramos $v_1' = (c_i, q_{i_x}, m) \in V$ y $v_2' = (c_j, q_{j_y}, m') \in V$ tales que v_1' y v_1 son consecutivos y v_2 y v_2' son consecutivos. El *ángulo de giro* es el ángulo entre la semirrecta con origen q_{i_k} y sentido de q_{i_x} a q_{i_k} a y la semirrecta con origen en q_{i_k} y que contiene q_{j_y}. Consideramos que este ángulo está comprendido en el intervalo $(-\pi, \pi]$, con lo cual los ángulos positivos corresponden a giros la izquierda y los ángulos negativos corresponden a giros hacia la derecha.

Para restringir los arcos que corresponden a giros que no puede realizar un vehículo, definimos S como el conjunto de los puntos de salida de las cuadras con semáforos. Dados dos nodos $v_1 = (c_i, q_{i_k}, m) \in V$ y $v_2 = (c_j, q_{j_l}, m') \in V$, decimos que hay un *giro prohibido* de v_1 a v_2 si hay un giro de v_1 a v_2, si $q_{i_k} \in S$ y el ángulo de giro es mayor que $\pi/4$. Dados dos nodos $v_1 \in V$ y $v_2 \in V$, decimos que hay un *giro permitido* de v_1 a v_2 si hay un giro de v_1 a v_2 y no hay un giro prohibido de v_1 a v_2.

Definimos ahora los arcos del grafo considerando si un nodo es consecutivo del otro en la misma cuadra o si se puede girar de un nodo al otro.

$$A = \{(v_1, v_2) \in V \times V : \begin{array}{l} v_1 \text{ y } v_2 \text{ son consecutivos, o bien} \\ \text{hay un giro permitido de } v_1 \text{ a } v_2\} \end{array}$$

Como los arcos del grafo corresponden a todos los movimientos permitidos que puede realizar un vehículo, si utilizamos un algoritmo de camino mínimo sobre este grafo, tenemos como resultado un recorrido mínimo que puede realizar un vehículo para ir de una posición a otra en la ciudad.

2.2. Ubicación de los contenedores

Como parte de los datos, contamos con cuatro listas de direcciones de contenedores; cada una corresponde a una sub-zona y las direcciones están ordenadas de la forma en que se recorren actualmente. En la figura 3, pueden verse las cuatro sub-zonas de contenedores actuales.

Estos datos nos permiten calcular la longitud de los itinerarios actuales para comparar con los resultados obtenidos en este trabajo. Aún si no sabemos si el camino que realizan entre cada par consecutivo de contenedores es el mínimo posible, podemos asumirlo y así calcular una cota inferior de la distancia y el trabajo de los itinerarios actuales. Análogamente, podemos asumir que los caminos entre el EHU y el primer contenedor, entre el último contenedor y el depósito, y entre el depósito y el EHU son también mínimos.

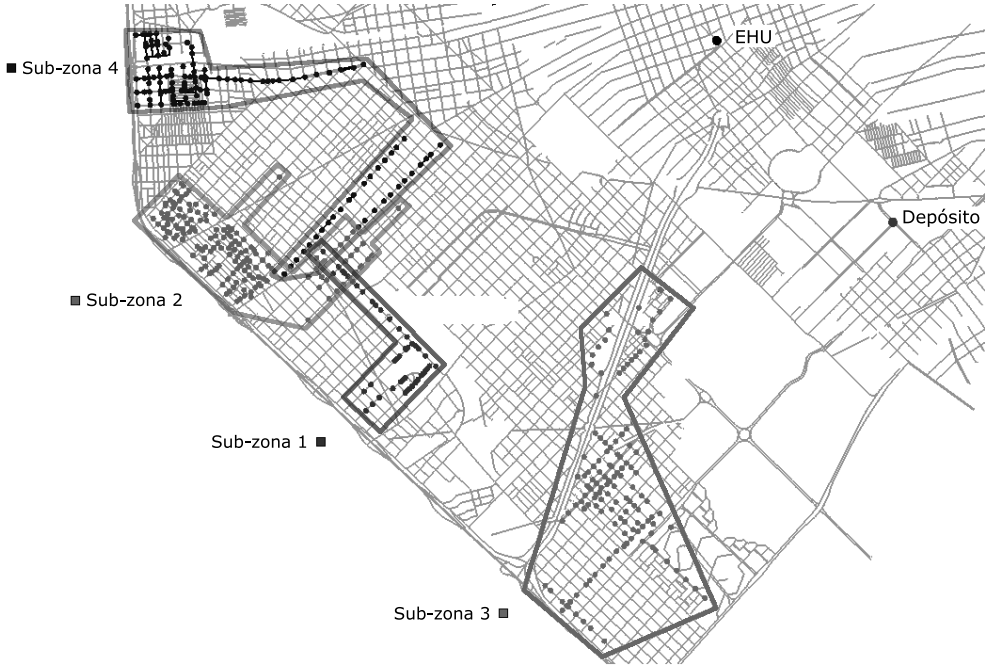


Figura 3: Sub-zonas de contenedores actuales.

Como las listas de contenedores no estaban pensadas para ingresar a un sistema informático, las direcciones no estaban normalizadas. En algunos casos, una dirección se definía por alguna referencia del mapa, por ejemplo nombrando instituciones sin dar su ubicación. Estos casos se corrigieron manualmente en la base de datos. Por otra parte, se detectaron múltiples denominaciones y variaciones de los nombres de las calles, lo cual dificultó su sistematización para el procesamiento automático. Para solucionar este problema, se reemplazó manualmente cada nombre con su correspondiente en la base de datos.

El resultado de esta traducción son listas de direcciones de los contenedores dadas por nombre de la calle y altura o por intersección de calles. Una vez completado este proceso, es necesario traducir cada dirección a una posición del mapa *(punto, cuadra, mano)*, con el objetivo de generar dentro del grafo que representa la ciudad los nodos correspondientes a los contenedores.

Para una dirección por altura, se busca la cuadra de la calle cuyo intervalo [*altura inicial*, *altura final*] incluye la dirección en cuestión. En las cuadras doble-mano, se le asigna la mano que va en sentido creciente de altura o la otra mano según si la altura es impar o par, respectivamente.

Para una dirección denotada por intersección de calles, se buscan las cuadras de ambas calles que se intersecan en un punto. En el caso más simple, hay cuatro cuadras que se intersecan en un punto: dos de una calle y dos de la otra. Es necesario elegir una de las cuatro cuadras para definir la posición del contenedor. El criterio que usó el EHU para confeccionar las listas es que la dirección nombra primero la calle sobre la cual está el contenedor y el mismo se encuentra antes de la intersección. Para resolver las ambigüedades que se presentan cuando la primera calle es doble-mano, entre los dos puntos posibles (mano y contramano) se escoge el de menor coordenada x y, si ambas coordenadas x son iguales, se escoge el de menor coordenada y. Este criterio fue definido arbitrariamente.

Como toda base de datos que contiene información de la realidad, el mapa tiene fallas. Se verificaron muchas de las calles que están en las zonas donde hay contenedores para revisar los sentidos. Se revisaron especialmente las que se utilizan en los recorridos óptimos que encontramos. Para eso se comparó con otros mapas digitales y con fotos aéreas, que permitieron validar los sentidos de las calles. Este proceso manual permitió depurar los datos, con el objetivo de contar con información lo más precisa posible para el proceso de optimización.

3. Estrategia de resolución

Dado que el primer objetivo es minimizar la distancia recorrida por cada camión, la opción más natural es transformar el problema de diseñar el recorrido de cada camión en una instancia adecuada del Problema del Vendedor Viajero (TSP). El TSP consiste en encontrar un circuito Hamiltoniano de longitud mínima en un grafo, y es un problema NP-hard. El segundo objetivo es disminuir el trabajo total realizado por el camión (como aproximación del desgaste total). Como las instancias generadas resultan tener una gran cantidad de óptimos alternativos para el TSP, entonces se resuelve muchas veces en forma óptima el TSP utilizando una estrategia no determinística, y se selecciona la solución que realiza el menor trabajo total. Para que este esquema tenga éxito, es necesario contar con una implementación que permita resolver el TSP en

forma óptima y no determinística. En nuestra implementación utilizamos el software *Concorde* [1].

Se describe en esta sección el proceso de generación del grafo y resolución del TSP. En la Sección 3.1 se describe la construcción de una instancia del problema de camino Hamiltoniano mínimo entre dos nodos particulares, que modela nuestro problema. En la Sección 3.3 se describe el uso del software *Concorde*, que resuelve el TSP sobre grafos no dirigidos. Por este motivo, es necesario transformar el problema en un TSP simétrico, proceso que se describe en la Sección 3.4 y la Sección 3.5.

3.1. Construcción de la instancia

Para modelar el problema adecuadamente, construimos un digrafo completo con los contenedores como nodos. El peso de un arco del nodo A al nodo B se define como la distancia del recorrido mínimo en vehículo del contenedor A al contenedor B (Ver Sección 3.2). Agregamos a este grafo el EHU y el Depósito. Agregamos un arco desde el EHU hacia cada contenedor, y desde cada contenedor al depósito. Los pesos de estos arcos también quedan definidos por la distancia del recorrido mínimo de un elemento al otro. Llamamos a este grafo $G_1 = (V_1, A_1)$ y denotamos por $w_1 : A_1 \to \mathbb{R}$ la función de distancia que a cada arco le asocia la distancia del recorrido correspondiente. Puede verse un ejemplo en la figura 4.

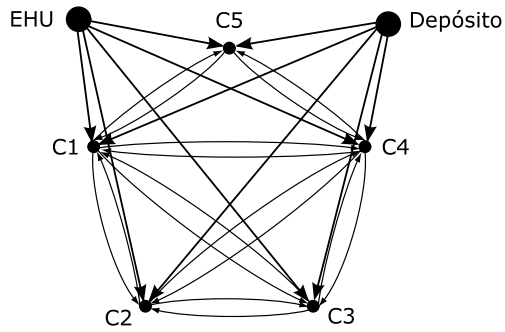


Figura 4: G_1. Grafo con el EHU, cinco contenedores y el depósito.

Nuestro objetivo es buscar un camino Hamiltoniano mínimo en este grafo, que parta del nodo que representa al EHU, pase por todos los nodos que representan contenedores, y termine en el nodo que representa al depósito.

3.2. Algoritmo de camino mínimo

Una vez definido el grafo asociado al mapa de la ciudad aplicamos la variante A^* del algoritmo de Dijkstra [3, 6] para obtener el camino mínimo en

vehículo entre cada par de nodos en el digrafo completo que tiene a los contenedores, al EHU y al depósito como nodos. El problema que surge al aplicar el algoritmo de Dijkstra sin dicha variante es que durante su ejecución el conjunto de nodos pendientes de analizar se expande en una forma similar a un rombo con el punto de origen como centro.

Este comportamiento hace que sea menos eficiente cuanto mayor sea la distancia entre los nodos a analizar. Por ejemplo, para calcular un recorrido mínimo entre dos puntos que dé como resultado 10 km., calculará los caminos mínimos a todos los puntos que estén a menos de 10 km. del origen. En el grafo de la ciudad, este cálculo demora más de una hora realizado con una implementación en SmallTalk en una computadora Intel Dual Core de 1.60 GHz.

Este problema se soluciona usando la variante A*, ya que dicha heurística cambia la forma de elegir el siguiente nodo a analizar. En lugar de tener en cuenta sólo la distancia al origen, también tiene en cuenta la distancia al destino. Para que funcione correctamente, utiliza la propiedad de que la distancia del recorrido es siempre mayor o igual que la distancia euclidiana entre ambos puntos. Luego, en un punto intermedio del recorrido, al cual ya le calculó el recorrido mínimo, se deduce que la distancia del recorrido mínimo del origen al destino es mayor o igual que la suma de la distancia del recorrido del origen al intermedio, más la distancia euclidiana.

Cuando se hace el mismo cálculo, en la misma computadora, de un recorrido de más de 10 km. utilizando la variante A* se reduce el tiempo de ejecución drásticamente a 280 milisegundos, lo cual verifica que la elección de la misma es satisfactoria para este trabajo.

3.3. Utilización de Concorde

Concorde [1] es un programa realizado por David Applegate, Robert Bixby, Vašek Chvátal y William Cook en el Instituto de Tecnología de Georgia (Georgia Institute of Technology), que permite resolver instancias del TSP en forma exacta mediante programación lineal entera. Para esto se le adjunta un paquete para resolver problemas de programación lineal, por ejemplo *QSOpt* (de libre uso y desarrollado por los mismos autores). *Concorde* puede resolver instancias de hasta 1000 elementos en forma eficiente. Por ejemplo, una instancia de nuestro problema con 98 nodos es resuelta en 200 milisegundos en una computadora Intel Dual Core de 1.60 GHz.

Concorde también permite resolver el TSP mediante la heurística Chained Lin-Kernighan [9]. Esta heurística es muy eficiente y proporciona resultados óptimos o muy cercanos al valor óptimo para instancias pequeñas. Por ejemplo, para la instancia de 98 nodos que consideramos en este trabajo, el circuito obtenido resulta ser el óptimo y el tiempo de ejecución es similar al tiempo de

resolución del algoritmo exacto.

Por otro lado, como *Concorde* utiliza aleatoriedad para obtener una solución, si lo ejecutamos varias veces, se obtienen distintos óptimos alternativos.

3.4. Transformación de camino Hamiltoniano a circuito Hamiltoniano

En la Sección 3.1 modelamos nuestro problema como un problema de camino Hamiltoniano mínimo dirigido. Sin embargo *Concorde* resuelve el TSP, es decir el problema de circuito Hamiltoniano mínimo no dirigido y recibe como entrada una matriz de distancias de un grafo completo. Por lo tanto, debemos traducir nuestro modelo de camino Hamiltoniano mínimo dirigido a un circuito Hamiltoniano mínimo no dirigido. Lo haremos en dos pasos: primero pasaremos de un camino dirigido a un circuito dirigido y luego de éste a un circuito no dirigido.

Calcularemos el camino Hamiltoniano mínimo en función del circuito Hamiltoniano. Necesitamos que en este circuito luego del depósito se ubique el EHU. Es decir, a partir del nodo del EHU, se pase por todos los contenedores, luego por el depósito y nuevamente por el EHU. Además, queremos que el grafo sea completo. Para esto, construimos el grafo completo $G_2 = (V_2, A_2)$ a partir de G_1 tomando $V_2 = V_1$ y $A_2 = V_2 \times V_2$. Definimos una función de peso $w_2 : A_2 \to \mathbb{R} \cup \{+\infty\}$ de la siguiente forma:

$$w_2((v_1, v_2)) = \begin{cases} w_1((v_1, v_2)) & \text{si } (v_1, v_2) \in A_1 \\ 0 & \text{si } v_1\text{=EHU y } v_2\text{=Depósito, o viceversa} \\ +\infty & \text{si } (v_1, v_2) \notin A_1 \end{cases}$$

De esta forma, si no hay un arco entre dos nodos de G_1, agregamos un arco en G_2 con peso infinito y así G_2 es completo. De esta forma, todo camino Hamiltoniano de G_1 que comience en el EHU y termine en el depósito genera un circuito Hamiltoniano en G_2 con distancia finita, y viceversa.

3.5. Transformación de grafo dirigido a grafo no dirigido

En esta sección reduciremos el problema de encontrar un circuito Hamiltoniano mínimo de un grafo dirigido, al problema de encontrar un circuito Hamiltoniano mínimo de un grafo no dirigido [8]. Para esto definimos el grafo $G_3 = (V_3, A_3)$ con un nodo ficticio y un nodo real por cada nodo de G_2. Más formalmente, supongamos que $V_2 = \{v_1, \ldots, v_p\}$ y definamos $F = \{f_1, \ldots, f_p\}$, de modo tal que el nodo f_i es el nodo ficticio asociado a v_i. Definimos $V_3 = V \cup F$ y $A_3 = V_3 \times V_3$. Sea $M \in \mathbb{R}$ un número suficientemente grande, y definimos la función de pesos asociados a las aristas $w_3 : A_3 \to \mathbb{R} \cup \{+\infty\}$ del siguiente modo:

$$w_3(x,y) = \begin{cases} +\infty & \text{si } x \in V \wedge y \in V, \\ +\infty & \text{si } x \in F \wedge y \in F, \\ -M & \text{si } \exists i / x = v_i \wedge y = f_i, \\ w_2(v_i, v_j) & \text{si } \exists i / x = v_i \wedge y = f_j \wedge i \neq j \end{cases}$$

En la figura 5 se presenta un ejemplo donde se define un grafo no dirigido G_3 a partir de un grafo dirigido G_2, con $M = 10000$. En G_3 se tienen 3 nodos reales y 3 nodos ficticios. Las aristas entre nodos reales tienen valor $+\infty$, lo que implica que los caminos mínimos no utilizarán esas aristas. Lo mismo sucede para las aristas entre nodos ficticios. Un camino mínimo en este grafo siempre alternará nodos ficticios con reales. Además, como las aristas entre un nodo real y su correspondiente nodo ficticio tienen el valor $-M$, los caminos mínimos siempre utilizarán esas aristas. Como el camino mínimo no contendrá los arcos con peso infinito y sí contendrá los que tienen peso negativo, a continuación de un nodo real siempre aparecerá su nodo ficticio en un camino mínimo.

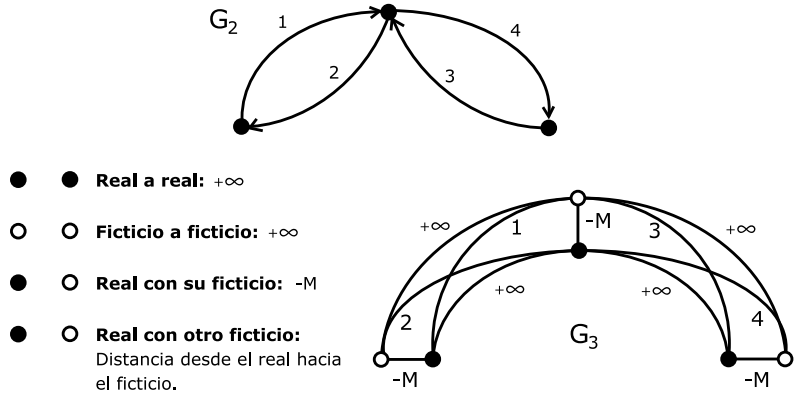
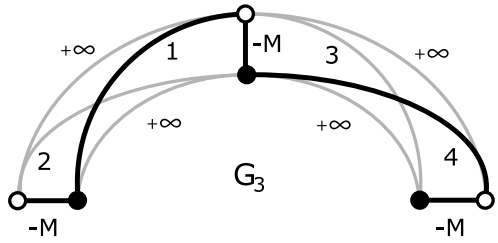


Figura 5: Modelo de grafo dirigido con un grafo no dirigido.

El peso de las aristas entre un nodo real v_i y uno ficticio f_j con $i \neq j$ es el peso del arco que va de v_i a v_j en G_2. Esto implica que el conjunto de aristas del nodo real v_i en G_3 representa los arcos de salida de v_i en G_2. Recíprocamente el conjunto de aristas del nodo ficticio f_j en G_3 representa los arcos de entrada de v_j en G_2. Entonces, un camino $(f_{i_1}, v_{i_1}, f_{i_2}, v_{i_2}, \ldots, f_{i_k}, v_{i_k})$ en G_3 se interpreta como un camino $(v_{i_1}, v_{i_2}, \ldots, v_{i_k})$ en G_2. Se puede ver un ejemplo en la figura 6.

Por medio de los procedimientos descritos en esta sección, pasamos entonces de un grafo dirigido G_1 en el que queremos calcular un camino Hamiltoniano mínimo, a un grafo dirigido completo G_2 tal que un circuito Hamiltoniano mínimo de G_2 nos permite calcular el camino Hamiltoniano mínimo



Longitud con arcos ficticios: $-3M + 5$
Longitud sin arcos ficticios: 5

Figura 6: Camino en un grafo no dirigido.

de G_1. Luego pasamos a G_3, un grafo no dirigido completo. Si calculamos el circuito Hamiltoniano mínimo de G_3 podemos obtener fácilmente el circuito Hamiltoniano mínimo de G_2. De esta forma reducimos nuestro problema al Problema del Vendedor Viajero en un grafo completo no dirigido y podemos utilizar *Concorde* para obtener el resultado eficientemente.

4. Descripción de la implementación

Se implementó un programa en el lenguaje de programación *SmallTalk* que realiza tareas de procesamiento del mapa, cálculo de itinerarios óptimos y visualización de los resultados. En este proyecto utilizamos la implementación *VisualWorks NonCommercial, 7.4.1* disponible sobre los sistemas operativos *Microsoft Windows XP/Vista* y *GNU-Linux*.

El programa consta de un conjunto de objetos, de clases y de métodos que modelan el mapa de la ciudad, el grafo de los contenedores, los algoritmos necesarios para calcular los caminos mínimos y los itinerarios mínimos. También incluye interfaces con una base de datos del mapa en *PostgreSQL*, con el programa *Concorde* y con el paquete de visualización gráfica *Takenoko*.

El sistema realiza las siguientes tareas para procesar la información, calcular los itinerarios mínimos y visualizarlos:

1. Lee la base de datos del mapa de la ciudad.
2. Almacena las direcciones de los contenedores de cada itinerario.
3. Para cada una de las 4 sub-zonas realiza los siguientes pasos:
   a) Calcula el recorrido mínimo en vehículo entre cada par de elementos: ente, contenedores y depósito.
   b) Calcula la distancia y el trabajo del itinerario actual.

c) Construye el grafo G_1 con las distancias entre los elementos.

d) Construye el grafo G_2 a partir de G_1 como se define en la Sección 3.4.

e) Construye el grafo completo G_3 a partir de G_2 como se define en la Sección 3.5.

f) Ejecuta *Concorde* con la matriz de distancias de G_3.

g) Interpreta el resultado para construir un itinerario de distancia mínima.

4. Además, para cada itinerario se permite:

a) Calcular su distancia.

b) Calcular su trabajo.

c) Generar una lista con el orden de los contenedores.

d) Generar una lista de las cuadras por las que pasa el camión.

e) Visualizar una imagen del mapa con el recorrido marcado.

f) Visualizar una animación del recorrido del camión sobre el mapa.

5. Resultados y discusión

En esta sección reportamos los resultados obtenidos para las cuatro instancias de nuestro problema, correspondientes a las cuatro sub-zonas, que contienen 47, 133, 138 y 161 contenedores, respectivamente. A su vez, comparamos las distancias del itinerario que utiliza actualmente el EHU con la distancia de un itinerario de distancia mínima. También comparamos el trabajo ejercido en ambos itinerarios.

En las 4 instancias se obtienen a través de Concorde los recorridos mínimos. Utilizando la aleatoriedad del software se realizan múltiples corridas de cada caso, para luego quedarnos con el camino mínimo que da el menor trabajo.

En la siguiente tabla se muestran los resultados calculados para cada uno de los itinerarios. Se reportan la distancia y el trabajo para el itinerario que actualmente realiza el EHU y para el itinerario de distancia mínima (los casos reportados son siempre los que dan el menor trabajo posible dentro de los múltiples recorridos mínimos obtenidos). Se ven importantes mejoras tanto en la distancia, como en el trabajo.

Vemos que hay una diferencia muy grande entre la mejora de la sub-zona 1 y la mejora de la sub-zona 4 que radica en la complejidad de las mismas. La primera tiene menos que un tercio de la cantidad de contenedores de la segunda y sus contenedores están sobre sólo 7 calles distintas. Estas características

		Recorrido actual		Nuestra propuesta		Porcentaje de mejora	
	Contenedores	Distancia	Trabajo	Distancia	Trabajo	Distancia	Trabajo
1	47	27.010	$6,08 \times 10^8$	24.126	$5,50 \times 10^8$	10,68 %	9,52 %
2	133	68.102	$4,77 \times 10^9$	41.752	$2,98 \times 10^9$	38,69 %	38,14 %
3	134	52.270	$4,28 \times 10^9$	39.762	$2,86 \times 10^9$	23,93 %	33,23 %
4	161	61.692	$5,44 \times 10^9$	40.841	$3,11 \times 10^9$	33,80 %	42,80 %

Tabla 1: Tabla de resultados

hacen que un itinerario cercano al mínimo sea bastante intuitivo. En cambio, en la sub-zona 4 hay un área con mucha densidad de contenedores y el método intuitivo no es practicable. Por esta razón, un itinerario de distancia mínima provoca una gran mejora (33,80 %).

Con respecto a los tiempos de resolución, la mayor parte la insume el cálculo de recorridos mínimos entre cada par de elementos. En la sub-zona 4, demora 40 minutos la primera vez y sólo 7 segundos cuando los recorridos ya están calculados. Si agregamos o movemos un contenedor del recorrido, sólo hará falta calcular 162 recorridos, lo que insumiría alrededor de 24 segundos según cuál sea la posición nueva del contenedor. El tiempo que demora el cálculo de los caminos mínimos no sólo depende de la cantidad de caminos, sino también de la longitud de los mismos. Cuanto más dispersos están los contenedores, habrá caminos más largos y, por lo tanto, este cálculo demorará más.

Los resultados nos permiten concluir que el problema fue abordado en forma satisfactoria ya que la distancia y el trabajo totales se redujeron en forma muy importante. En el caso de contar con información de las velocidades de las calles, se esperaría una mejora similar en la duración de los itinerarios.

6. Conclusiones y próximos pasos

En esta aplicación vemos niveles de mejora significativos en los recorridos de los camiones. La distancia de los itinerarios se reduce hasta un 39 % y el trabajo, aunque no es la variable que optimiza el modelo, también se redujo hasta un 43 % por tener a la distancia como uno de sus factores.

Una buena parte del trabajo realizado consistió en modelar el grafo e implementar el algoritmo de camino mínimo, considerando todos los detalles para producir recorridos en vehículo válidos en el mapa de la ciudad. La interfaz gráfica para producir imágenes y animaciones es un elemento relevante de la implementación, dado que permite visualizar la información de diversas formas, característica fundamental al tener gran cantidad de información. Tanto la interfaz gráfica como la capa de datos fueron abstraídas en objetos pro-

pios del sistema, por lo cual las partes que cambiarían en caso de elegir otras tecnologías están bien localizadas.

Una posible continuación del proyecto sería desarrollar un software que permita el cálculo de itinerarios en tiempo real. Al ocurrir eventos como cortes de calles, congestiones o manifestaciones, se podrían recalcular los itinerarios en el momento. La eficiencia de los algoritmos utilizados permite esta posibilidad en forma directa.

También podría agregarse un módulo de localización por GPS que permita ingresar al programa la posición de los camiones y, mediante dispositivos móviles, indicar a los conductores los próximos contenedores a recolectar. Estos agregados hacen que el sistema sea escalable a la hora de agregar nuevos contenedores y vehículos.

Para disminuir aún más el desgaste de los vehículos, podría considerarse el problema de optimizar el trabajo total realizado a lo largo del itinerario. Para esto, se pueden diseñar heurísticas que busquen itinerarios alternativos partiendo de la solución óptima del TSP.

El proyecto de recolección de residuos por medio de contenedores prevé agregar contenedores anualmente para llegar a un contenedor por cuadra en toda la ciudad. Por esta razón, será necesario agregar nuevas zonas con sus respectivos itinerarios. El problema de definir la zona asignada a cada camión es un nuevo problema que resulta muy interesante. La solución a este problema puede utilizar la suma de las distancias de los itinerarios mínimos como forma de validar la zonificación.

Si al mapa de la ciudad le agregamos la información de las velocidades de circulación promedio en las cuadras, podemos saber cuánto demora el camión en transitar una cuadra. De esta forma, podemos utilizar el programa para calcular un itinerario de tiempo mínimo. Para esto, sería necesario modificar la implementación del algoritmo A^* para que utilice otra función de cota inferior, como por ejemplo la distancia euclídea multiplicada por la velocidad máxima de circulación.

Como conclusión general, se puede afirmar que las herramientas de investigación de operaciones propuestas deberían ser de suma utilidad si se aplican para implementar un sistema que organice la recolección de residuos de la Ciudad.

Agradecimientos: A Bill Cook, por sus múltiples sugerencias acerca del uso de Concorde. A Daniel Iglesias, Jorge Takashashi y Julián Dunayevich, por proveernos de toda la información necesaria de la Ciudad de Buenos Aires. A Florencia Fernández Slezak, por su colaboración en la realización de este trabajo. Parcialmente financiado por los proyectos UBACyT X069 y X606, ANPCyT PICT-2007-0518 y PICT-2007-0533 (Argentina), FONDECyT 1080286 e Instituto de Ciencias Milenio "Sistemas Complejos de Ingeniería" (Chile).

Referencias

[1] Applegate D. L., Bixby R. E., Chvatal V., and Cook W.J. Who's Who in Networks: The Key Player. *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, Princeton, New Jersey, 2006.

[2] Chang N., Lu H., and Wei L.. GIS technology for vehicle routing and scheduling in solid waste collection systems. *Journal of Environmental Engineering*. Vol 123: 901-933. 1997.

[3] Dijkstra E. W. A note on two problems in connexion with graphs. *Numerische Mathematik*. Vol 1(1): 269-271. 1959.

[4] Eisenstein D. and Iyer A. Garbage collection in Chicago: a dynamic scheduling model. *Numerische Mathematik*. Vol 43: 922-933. 1997.

[5] Golden B., Assad A., and Wasil E.. Routing vehicles in the real world: applications in the solid waste, beverage, food, dairy, and newspaper industries. *In: Toth P, Vigo D, editors. The vehicle routing problem. Philadelphia, PA: SIAM*. 245-286. 2002.

[6] Hart P. E., Nilsson N. J., and Raphael B. A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics, IEEE Transactions on*. Vol 4(2): 100-107. 1968.

[7] Kim B., Kim S., and Sahoo S. Waste collection vehicle routing problem with time windows. *Computers and Operations Research*. 3624-3642. 2006.

[8] Kumar R. and Li H. On Asymmetric Tsp: Transformation to Symmetric Tsp and Performance Bound. Manuscript. 1996.

[9] Martin O., Otto S. W., and Felten E. W. Large-step Markov Chains for the Traveling Salesman Problemws. *Complex Systems*. Vol 5: 299-326. 1996.

[10] Mourao M. and Almeida M. Almeida, Lower-bounding and heuristic methods for a refuse collection vehicle routing problem. *European Journal of Operational Research*. Vol 121: 420-434. 2000.

Identificación de Patrones en Series de Tiempo Usando Redes Neuronales en Datos de una Empresa Petroquímica

Leonardo Buschiazzo*
Lorena Pradenas**

Resumen

En este artículo se presenta un estudio para la identificación de patrones en series de tiempo de datos de proceso basado en redes neuronales, con el objetivo de apoyar a la gestión del control automático en una industria de procesos continua. Para implementar la identificación de patrones, se realizó un sistema computacional que contiene un algoritmo de redes neuronales para la clasificación de segmentos de una serie de tiempo en casos conocidos predefinidos. Así, con el sistema computacional se evaluó la temperatura de salida del producto de un horno automatizado de una empresa petroquímica de la octava región, considerando un mes de datos en crudo y filtrados. Al clasificar los patrones de la salida del horno, se obtuvo un rendimiento general de un 83,33 % para los datos crudos y un 82,96 % para los datos filtrados, para patrones conocidos y desconocidos. En el caso particular de los datos conocidos, la aplicación tuvo un rendimiento de clasificación similar para entradas filtradas y no filtradas, lo que es un buen indicador acerca de la tolerancia de la red a conjuntos de datos sin filtrar. A partir de los resultados y de la experiencia del estudio, se concluyó que es válido el uso de la herramienta para clasificación de patrones, siendo un apoyo para la evaluación del comportamiento del control automático de un proceso, además, esta herramienta es aplicable como apoyo en algunos tipos de cartas de control ocupadas en gestión de calidad. En cuanto a trabajos futuros, se puede mencionar que la herramienta es perfectible en cuanto a un mayor tipo de patrones de entrenamiento, arquitecturas de redes neuronales, mejores datos de entrenamiento y la implementación de un esquema que permita evaluar en tiempo real el proceso en estudio para propósito de alarmas de proceso.

Palabras Clave: Redes Neuronales Artificiales, Patrones, Procesos Continuos, Series de tiempo.

*Magíster en Ingeniería Industrial, Universidad de Concepción.
**Departamento Ingeniería Industrial, Universidad de Concepción.

1. Introducción

En las industrias de procesos continuos y semi continuos es común el uso de tecnologías de automatización para el control de la producción. Existen empresas donde se han implementado sistemas de control distribuido (conocidos como DCS por sus siglas en inglés) o sistemas de Control supervisor y adquisición de datos (conocidos como SCADA, por sus siglas en inglés) para manejar los principales procesos productivos. Esto implica mejoras en la calidad del proceso, ahorros en costos de producción, manejo adecuado de los tiempos de producción y, por ende, de logística, entre otras ventajas con respecto a una producción realizada en forma manual por personal capacitado. En otras palabras, implementar automatización aumenta cuantitativamente la competitividad de una empresa. Sin embargo esto es en el escenario ideal donde el control automático funciona sin problemas y maneja el proceso de producción tal que su salida es homogénea y acotada alrededor de un valor objetivo deseado. En el escenario real, el control automático del proceso no es infalible y es propenso a perturbaciones, debido a causas al azar o causas especiales (Gutierrez [4]), que provocan un comportamiento anormal en la salida del proceso controlado.

Esta situación lleva a plantearse varias preguntas: ¿Quién controla al controlador, en cuanto a que realice su labor apropiadamente?, ¿Cómo detectar cambios en el proceso que no sean gruesos pero que implican una anormalidad? y ¿Cómo realizar la detección en forma oportuna? o visto desde otro punto de vista, cómo realizar una gestión operativa de estos sistemas de forma cuantitativa y objetiva.

Entonces, a modo de respuesta de las interrogantes planteadas, se propuso que un camino para apoyar a la gestión operativa de un sistema de control automático es por medio de la revisión el patrón de salida en el tiempo de la variable del proceso que se controla. Un comportamiento normal de la salida del proceso tiene un patrón de salida característico, luego, otros patrones de salida implican la existencia de alguna anormalidad en la capacidad del sistema de control para manejar el proceso.

Para realizar la detección, se planteó el desarrollo de una herramienta computacional que tuviera como núcleo un modelo de redes de neuronas artificiales. Modelo que tienen una amplia experiencia en la clasificación de elementos de un mismo conjunto (Matich [10], Viñuela [22], Hann [6], por mencionar algunos autores) sin la necesidad de entender como se originan éstos, sino que aprende por medio de ejemplo de elementos ya existentes y que se cuentan con variadas experiencias en el trabajo con Patrones de datos (Fu [3], Guh [5], Olmedo [15] y Singh [18], entre otros) y en la aplicación en el ámbito de

sistemas industriales (Muñoz [13], Silva [19], Ruiz [16], entre muchos otros).

Se trabajó con datos de un proceso productivo real de una empresa de la octava región. Al ser datos de terreno, éstos contienen datos poco confiables por que se definió evaluar la herramienta computacional con los datos del proceso en forma directa y con los mismos datos pero con un pre procesamiento para descartar valores poco confiables (picos en la señal principalmente).

En la siguiente sección se presenta una introducción al problema en estudio, describiendo un proceso real donde se evaluó la solución propuesta. Luego, se describe el diseño de la solución propuesta, se detalla el sistema computacional desarrollado específicamente para implementar el modelo de solución. A continuación se muestran los principales resultados obtenidos al evaluar datos del proceso real de estudio con la herramienta computacional. Finalmente, se dan a conocer las conclusiones obtenidas en base a los resultados y se presentan una serie de propuestas para trabajos futuros.

2. Problema en Estudio

El problema de investigación es la detección de patrones en una serie de tiempo para una variable controlada de un proceso real de producción industrial por medio de un sistema computacional. En particular, la variable controlada es la salida de un proceso productivo en una planta continua. Básicamente, se tiene un patrón de señal "normal" y patrones que se alejan de una distribución normal, lo que indica la presencia de alguna perturbación que no ha podido ser compensada por el ente controlador.

Así, un patrón normal de la variable controlada en el tiempo sería como el ilustrado en la figura 1, donde la señal varía alrededor de un valor objetivo.

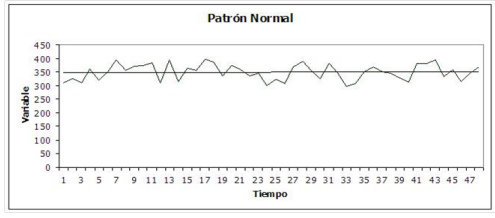


Figura 1: Patrón normal.

Por otro lado, comportamientos anormales en la salida serían los ilustrados en las figuras 2 y 3. En la figura 2 se tiene un salto (escalón) de un valor objetivo a otro y en la figura 3 un patrón creciente, que se va alejando del valor objetivo a través de una recta de pendiente positiva.

Los escalones pueden ser originados debido a un cambio operacional del proceso lo cual es una causa normal, por otro lado, si no se tenía programado,

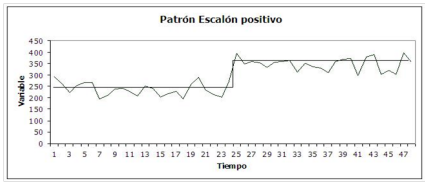


Figura 2: Patrón escalón positivo.

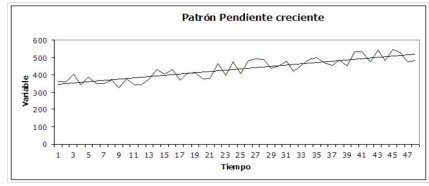


Figura 3: Patrón de pendiente creciente.

es una anormalidad de la operación del control.

2.1. Proceso en Estudio

El estudio se realizó en un proceso real, de una empresa de la región del Bío-Bío, correspondiendo a un horno de calentamiento de materia prima, donde se controla la temperatura de salida de la materia de forma automática por medio de un sistema mecánico-computacional dedicado (DCS). La figura 4 muestra un esquema del proceso, donde se aprecia la entrada y salida de materia prima, líneas verdes continuas desde el tren de crudo hasta el equipo E-1; varias línea de alimentación del horno (aire, fuel gas, fuel oil) y las líneas punteadas que representan la red de control sobre la temperatura de salida del horno.

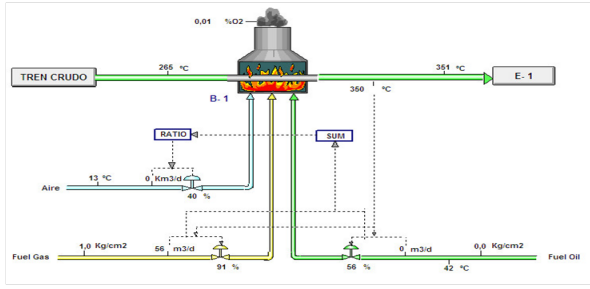


Figura 4: Horno de precalentamiento de materia prima.

Se pretende mantener la temperatura de salida de la materia prima en un valor cercano al valor definido por el departamento de operaciones de la empresa como el óptimo. El sistema de control manipula una serie de entradas a los quemadores del horno, tal que se aumente, disminuya o mantenga la temperatura de salida.

Al ser un proceso real, en la adquisición de datos es inherente tener datos poco confiables, debido al ruido eléctrico que afecta a los instrumentos de terreno. Dado esto, se trabajó con dos grupos de datos. Un grupo corresponde a los datos obtenidos de terreno y otro corresponde a los mismos datos de terreno pero con un pre-procesamiento para prevenir datos no confiables por medio de un filtro de suavizamiento exponencial.

3. Solución

3.1. Red de neuronas artificial

El problema a resolver es diseñar un sistema que tuviera como entradas el muestreo discreto de una señal de proceso en el tiempo, realizado en ventanas fijas de tiempo, como se muestra en la figura 5, y como salida una clasificación de la forma de la señal ingresada. Se diseñó un sistema computacional con un algoritmo de redes neuronales como núcleo, con la capacidad de reconocer patrones en las señales discretas ingresadas.

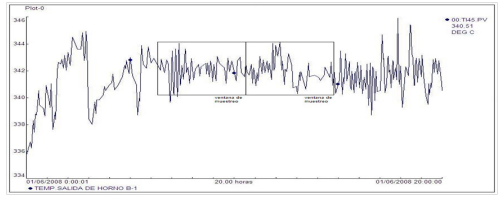


Figura 5: Serie de tiempo de variable de proceso.

El modelo de red neuronal seleccionado es una arquitectura de tipo Perceptron Multi Capa, el cual es un modelo ampliamente difundido y probado en lo relativo a la clasificación de entradas en un red ([22], [17], [5], [6], entre otros).

En la figura 6 se ilustra un diagrama de entradas y salidas, con el diseño del proceso de clasificación de patrones con la red neuronal como núcleo de procesamiento. En esta el muestreo para un periodo de tiempo fijo (ventana de tiempo) de la tendencia en el tiempo de la variable de proceso es ingresado como un vector de valores a la red; siendo procesado por ésta, entregando como salida de la operación un vector de elementos binarios que corresponde al código del patrón que mejor se ajusta a la forma de la señal de entrada.

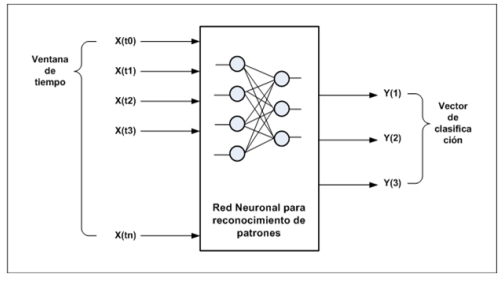


Figura 6: Red Neuronal Artificial como solución para el problema.

La red tiene como entrada un vector que representa el muestreo discreto en el tiempo de una señal de una variable de proceso controlada. El número de elementos del vector de entrada es fijo y se inicia desde el primer valor en la ventana de tiempo analizada hasta el último valor de la ventana de tiempo, es decir, el vector es la segmentación de la serie de tiempo dentro de un marco de tiempo fijo. Los valores en sí, corresponden a variables numéricas de tipo real. Dado esto, y sumado el hecho de que la red sólo puede tener entradas reales entre valores comprendidos en el siguiente intervalo $[-1, 1]$, se consideró una transformación lineal de los valores del vector de entrada de modo de acotar el rango permitido para la red. Esta transformación se implementa en el sistema computacional.

La salida de la red es un vector que indica a que patrón se ajusta mejor la entrada, según los patrones que la red ha aprendido previamente en su etapa de entrenamiento. La salida está compuesta por un vector de tres elementos, donde los elementos son variables de tipo binario. Se tiene un conjunto de combinaciones del grupo de elementos para definir un tipo de patrón en particular. Así, se definieron los patrones que aparecen en la tabla 1. Con estos patrones se entrenó la red de modo que fuera capaz de reconocer cuando alguno de ellos se encuentra presente en una tendencia de una variable.

Nombre del Patrón	**Clase**		
	Y(1)	**Y(2)**	**Y(3)**
Normal	0	0	1
Pendiente Creciente	0	1	0
Pendiente Decreciente	0	1	1
Escalón positivo	1	0	0
Escalón negativo	1	1	0

Tabla 1: Patrones definidos para el estudio.

3.1.1. Datos de entrenamiento y validación

Los datos para entrenar la red y para su validación, se generaron a partir de un conjunto de fórmulas propuestas en [17] y que ocupan información estadística de valor verdadero y de la desviación estándar de la variable en estudio. Las fórmulas utilizadas son las (1) a (5).

(1) Patrón Normal $f(x) = \eta + r(x) \cdot \sigma \quad \ldots(1)$

(2) Patrón de Pendiente Creciente $f(x) = \eta + r(x) \cdot \sigma + g \cdot x \quad \ldots(2)$

(3) Patrón de Pendiente Decreciente $f(x) = \eta + r(x) \cdot \sigma - g \cdot x \quad \ldots(3)$

(4) Patrón de Escalón positivo $f(x) = \eta + r(x) \cdot \sigma + b \cdot s \quad \ldots(4)$

(5) Patrón de Escalón negativo $f(x) = \eta + r(x) \cdot \sigma - b \cdot s \quad \ldots(5)$

Donde:

η: es el valor verdadero de la variable de proceso para un año de datos.

σ: es la desviación estándar de la variable de proceso para un año de datos.

g: es el gradiente de la pendiente en ecuación (2) y (3).

b: es el pivote para el escalón, sus valores posibles son 0 y 1. En el caso de la ecuación (4), toma el valor 0 hasta el valor x donde se desea el escalón, luego toma el valor 1, el caso inverso se da para la ecuación (5).

s: es el valor de la magnitud del escalón de las ecuaciones (4) y (5).

Con las ecuaciones (1) a (5) se generó un conjunto de patrones de entrenamiento y validación para la red. En total se crearon 200 vectores de 48 elementos para cada tipo de patrón, 150 para entrenamiento y 50 para validación del entrenamiento. Las figuras 7, 8, 9, 10 y 11 muestran gráficos de ejemplo para cada tipo de patrón mencionado.

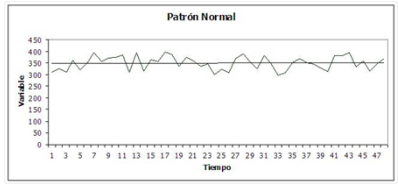


Figura 7: Patrón normal.

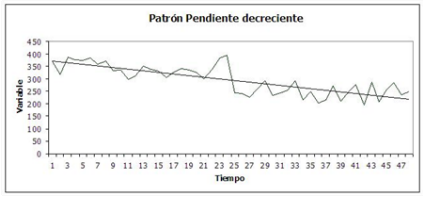


Figura 8: Patrón de pendiente decreciente.

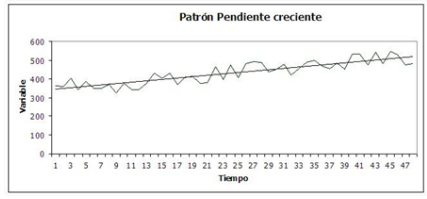


Figura 9: Patrón de pendiente creciente.

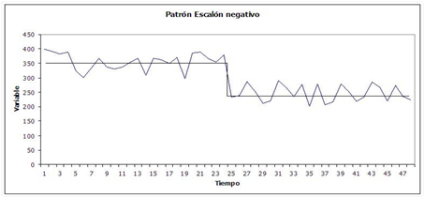


Figura 10: Patrón de escalón negativo.

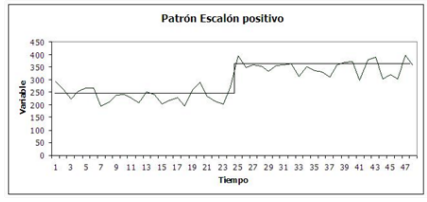


Figura 11: Patrón de escalón positivo.

3.1.2. Filtrado de Datos

Debido a que los datos del proceso real son obtenidos desde sensores de terreno, están expuestos a distorsiones que implican valores poco confiables en éstos, [21] y [14]. Así, para filtrar los datos, se utilizó un algoritmo de suavizamiento exponencial de primer orden, según antecedentes en [21]. Así, se suavizan los picos de datos poco confiables y la señal en general, para efectos del ruido inherente. A continuación se entrega la ecuación que define el suavizamiento exponencial, [23].

$$A_t = \alpha \cdot x_t + (1 - \alpha) A_{t-1} \qquad ...(6)$$

Donde:

A_t:	valor pronosticado para la señal en el instante t. Variable dependiente.
A_{t-1} :	valor pronosticado para la señal en el instante anterior a t, $(t-1)$.
α:	constante de suavizamiento, con valores dentro del rango $(0, 1)$.
x_t:	valor muestreado de la señal en el instante t.

3.1.3. Pseudo algoritmo de entrenamiento de la Red Neuronal

El entrenamiento es el proceso por el cual se fijan los pesos de la red, es un proceso iterativo en busca de los mejores valores para estos pesos tal que se obtenga el menor error de entrenamiento y de validación de la red.

El pseudo algoritmo usado para el entrenamiento de la red se muestra a continuación:

```
Inicio;
Se inicializan los pesos y umbrales en la red en forma aleatoria;
Repetir{iniciar contador de repeticiones en n=1
        Repetir{ iniciar contador de repeticiones en i=1
                Calcular Y(i)= ANN(X(i));
                Calcular E(i) = Diferencia entre S(i) y Y(i);
                Aplicar regla delta para modificar los pesos y umbrales;
                }hasta i=n;
        Evaluar Et;
        Repetir{ iniciar contador de repeticiones en j=1
                Calcular Z(i)= ANN(W(i));
                Calcular Ev(i){
                        Si Z(i) = V(i) entonces Ev(i) = 0
                        Sino Ev(i) = 1}
                }hasta j=m;
        Evaluar Evt como porcentaje de aciertos del total (m);
        } Hasta Et < Cota de Error o n = máximo de repeticiones permitido;
Fin;
```

Donde:

- El par $(X(i), S(i))$ corresponde a un vector de parámetros de entrenamiento (X) con su respectiva salida (S).

- $ANN()$ es la función de transferencia de la red neuronal, es decir, la propagación del vector X por la red, dando como resultado el vector Y.
- E es el error para la evaluación de un patrón de entrenamiento en particular y es el error cuadrático medio entre la salida conocida del patrón y la propuesta por la red.
- E_t es el error total de un ciclo completo de entrenamiento.
- $W(i)$ es el i-esimo vector de entrada con un patrón de validación.
- $Z(i)$ es el i-esimo resultado entregado por la red para un patrón de validación.
- E_v es un vector traspuesto que almacena dos valores, 0 ó 1. donde cero implica que la validación no fue exitosa y 1 que si lo fue. Esto es debido a que se trata de clasificar el patrón dentro de un conjunto de formas conocidas por la red, ergo, las opciones son 100 % o 0 %.
- E_{vt} es el porcentaje de validaciones acertadas con respecto al número total de elementos del conjunto de validación.

3.1.4. Diagrama resumen de la solución implementada

A modo de consolidar la solución, en la figura 12 se muestra un diagrama que resume la solución implementada. Así, la serie de tiempo es segmentada en forma secuencial, luego, un segmento de la serie de tiempo es ingresado como un vector discreto a un algoritmo de clasificación en patrones, el cual es la red de neuronas artificial, para ser clasificado en alguna de las clases conocidas de formas de onda de serie de tiempo. La clasificación es entregada como una codificación en un vector binario.

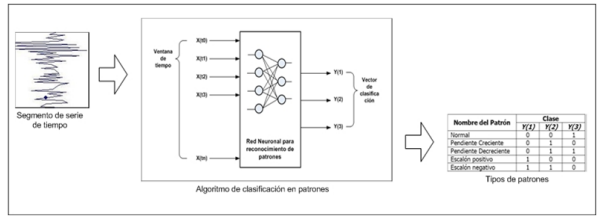


Figura 12: Diagrama conceptual de la solución.

3.2. Sistema computacional

Como ya se mencionó, para poder llevar a cabo el estudio en la práctica, se debió construir un sistema computacional que permitiera recopilar la información de procesos, formatearla y evaluarla según el algoritmo de clasificación basado en redes neuronales. Así, se diseñó un sistema computacional

conformado por tres entidades que interactúan entre ellas, donde se tiene un repositorio de datos de proceso en tiempo real e históricos, denominado PI System; planillas Excel con lógica imbuida para la recolección de los datos desde el repositorio y para formatear los vectores de entrada que son clasificados por un programa, de desarrollo especial para el estudio, que contiene el algoritmo de la red neuronal para clasificar la entrada en alguno de los patrones previamente definidos.

En cuanto a PI system, este es un sistema de almacenamiento de datos en tiempo real, orientado principalmente para manejar una gran cantidad de datos de variables de procesos productivos (hasta 1 millón de variables) y consiste en un servidor con una base de datos en tiempo real e histórico que se conecta a la mayoría de los sistemas de control automático del mercado, de esta forma concentra la información de una o varias plantas productivas en un único repositorio de datos. Para este trabajo se ocupó el repositorio para consultar información por la variable de temperatura de salida de un horno bajo control, su identificación es "00:ti45.pv" que es un tag de control.

Excel se utilizó para la gestión de los datos a ser ingresados al programa clasificador y para la visualización del resultado de la clasificación y como graficador. De tal modo se desarrollaron dos planillas con lógica imbuida (código Visual Basic para Aplicaciones) en este programa; una para extraer los datos de proceso desde el sistema PI y preparar archivos de entrada para el programa que los procesa y clasifica en algún patrón conocido con un algoritmo de redes neuronales y la otra planilla se usó para crear los patrones de entrenamiento y validación por medio del uso de las formulas descritas en el punto 3.1.1. El uso de Excel fue de carácter mandatorio, puesto que para poder accesar a los datos era necesario tener instalado en este programa un complemento dedicado a la conectividad con PI. Además, la facilidad que entrega Excel para estructurar los datos en matrices y graficar series de éstos es un valor agregado a tomar en cuenta.

El programa que clasifica los patrones de entradas en alguno del conjunto de los conocidos por él es una aplicación de consola sin interfaz gráfica para el usuario que lee archivos de datos de entrada tanto como para la evaluación de patrones como para el entrenamiento de la red. Los resultados los entrega como archivos compatibles con Excel, de modo de usar este último programa para revisar los resultados y graficar. Este programa corresponde a un desarrollo específico para este estudio, siendo construido bajo el lenguaje de programación Visual C# y denominado "AnnCore". El lenguaje de programación fue seleccionado luego de revisar la literatura relacionada, destacándose los autores Chesnokov [1], Madhusudanan [8] y Kirillov [7], y encontrando que este lenguaje está totalmente vigente y conteniendo los desarrollos más reciente. Dado que es un lenguaje relativamente nuevo, se debió aprender su uso, principalmente por medio de documentación del proveedor, [11] y [12].

3.2.1. Arquitectura

La arquitectura de red (mostrada en la figura 13), donde se apoya el sistema computacional, es de dos capas y de tipo cliente-servidor en cuanto al flujo de datos. Así se tiene una red de control del proceso productivo, que corresponde a la primera capa, donde se origina la información y es enviada al repositorio de datos históricos y en línea de producción (PI System) por medio de una interfaz entre esta red de control y la red computacional donde se encuentran los demás equipos. En esta red computacional, segunda capa, se encuentra el servidor PI y el equipo donde reside la aplicación Excel y el programa clasificador de redes neuronales.

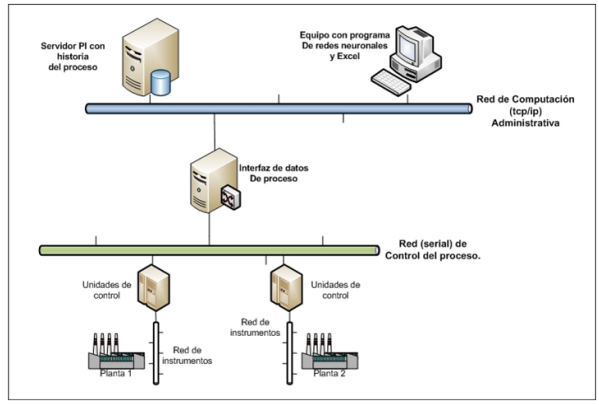


Figura 13: Arquitectura física del sistema computacional.

Para efectos de la aplicación de redes neuronales, tienen injerencia directa el servidor PI y el equipo tipo PC que contiene Excel y el programa de redes neuronales. Así, Excel saca datos desde el servidor PI por medio de librerías de comunicación, sobre TCP/IP, propietarias de PI. Luego, el procesamiento de los datos se realiza en forma local en el equipo PC.

3.2.2. Diagrama de Procesos Principales

La figura 14 ilustra los procesos principales en el sistema y su interacción entre estos y con los resultados. Se tiene el repositorio de datos de producción PI desde donde Excel rescata esta información y la prepara para se ingresada al programa "AnnCore" para ser procesada.

Excel colecta la información desde el repositorio PI por medio de un complemento especial, luego, prepara la información para ser ingresada en el programa de clasificación. Una vez que este programa realiza la identificación de patrones, el resultado se visualiza en Excel.

El programa denominado "AnnCore" es un desarrollo específico para este estudio y es el que contiene el algoritmo matemático de la red de neuronas para la clasificación en patrones la serie de tiempo ingresada en forma discreta.

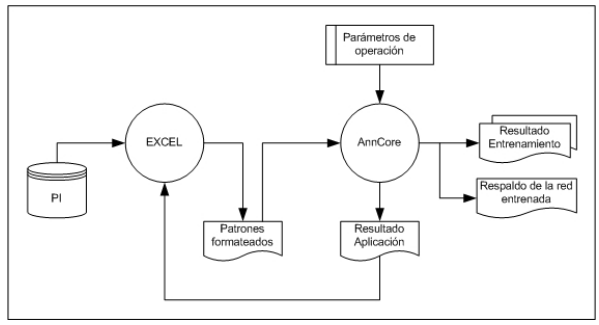


Figura 14: Diagrama de procesos principales.

Tiene dos modos de funcionamiento; uno de entrenamiento para sintonizar la red con la mejor combinación de parámetros con respecto al error de validación y entrenamiento y otro de ejecución que es para la clasificación de patrones de entrada.

En el modo de entrenamiento se ingresan patrones conocidos con sus respectivas salidas, también conocidas, para que el programa corra un algoritmo de entrenamiento de modo que la red aprenda a identificar los diferentes patrones por medio de ejemplos conocidos. Al finalizar el proceso, entrega como resultados archivos con el error de entrenamiento y validación y un archivo binario con el respaldo de la red ya entrenada, para ser usada en el modo de ejecución posteriormente.

En el modo de ejecución, el programa evalúa una serie de tiempo ingresada como un vector discreto y la clasifica según los patrones ya conocidos por éste o, en caso que no coincida con ninguno, en desconocido. El resultado es entregado como un archivo con formato compatible con Excel para su visualización.

4. Resultados

4.1. Resultados del Entrenamiento de la Red

Para que la red neuronal adquiera la inteligencia necesaria para clasificar patrones en series de tiempo, se debió "educar" por medio de un entrenamiento repetitivo. Entonces, se realizaron una serie de pruebas, alterando los parámetros constitutivos y de entrenamiento de la red, buscando la combinación de pesos en la red que entregara el mayor porcentaje de aciertos en la validación de la red y el menor error de entrenamiento, siendo el primer punto el preponderante a la hora de la selección de la mejor configuración. El porcentaje de acierto de validación corresponde a la razón entre el número de patrones identificados correctamente y el número total de patrones a validar, en términos porcentuales, donde lo que se deseaba era obtener un porcentaje cercano

al 100 %. En el caso del error de entrenamiento, expresado como cuadrático medio, éste corresponde al error generado entre la salida dada por la red para un vector de entrada y el vector de salida conocido para aquella entrada.

En cuanto al número de pruebas, 46 en total, se contrastó con experiencias de autores (Sagiroglu [17]; Guh [5], Hann [6]; entre otros) en diversos documentos técnicos y libros encontrándose que el número de pruebas estaba dentro de lo usual. Para cada prueba se modificó la cantidad de capas ocultas de la red (1 ó 2), el número de neuronas ocultas de cada capa (desde 48 hasta 96), la razón de aprendizaje de la red (desde 0,05 a 0,4, para un rango de 0 a 1), el momentum de la red (desde 0 a 0,6) y el número de ciclos de entrenamiento (de 1000 a 20000).

En particular, para el caso de los valores de la razón de aprendizaje, se siguió las recomendaciones de Elman, [2] con respecto a utilizar valores bajos de este parámetro para prevenir falta de generalización de la red ya entrenada.

La configuración de red más simple corresponde a la con menor número de capas y neuronas ocultas, la más completa a la con mayor número de capas y neuronas. Entre los extremos anteriores se sitúan las 44 pruebas restantes.

Así, del conjunto de pruebas realizados se seleccionó la prueba 12 como el de mejor desempeño, dado que fue la que mejor porcentaje de aciertos de validación obtuvo (82,96 %). Sin embargo, no fue la del menor error: 0,00011476 versus 1,349E-05 de la prueba 14, pero donde el error era menor, se perdía generalización lo que se reflejó en un menor porcentaje de validación. Así, la arquitectura de red elegida es la de una red Perceptron multicapa de 1 capa oculta y 72 neuronas en dicha capa.

El comportamiento del error de entrenamiento y el porcentaje de validación se ilustran en las figuras 15 y 16 respectivamente. Ambas se entregan en escala logarítmica.

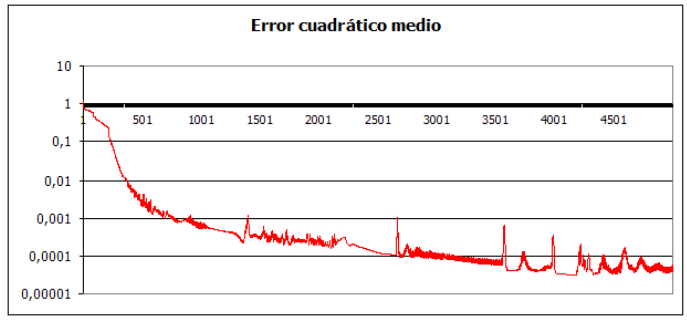


Figura 15: Gráfico de error de entrenamiento.

De la figura 15 se puede apreciar la convergencia del error hasta alcanzar un valor estable alrededor de $0,00011476$, se observa que la convergencia es pausada y no brusca, esto se debe a una baja razón de aprendizaje, lo que garantizó la estabilización del error al final del bucle.

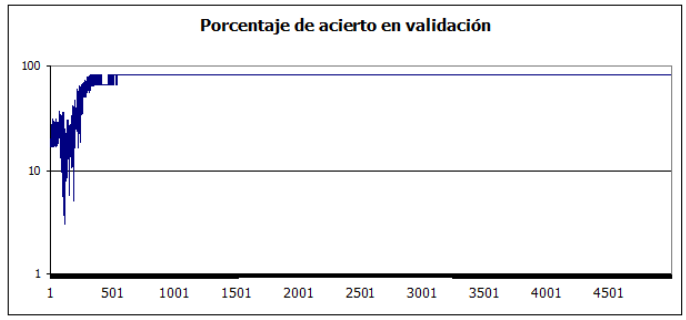


Figura 16: Gráfico del porcentaje de acierto en validación.

A diferencia del error de entrenamiento, el porcentaje de validación converge antes, aproximadamente en el ciclo 700, tal y como se observa en la figura 15. Sin embargo, el entrenamiento debe esperar que el error convergiera adecuadamente alrededor de un valor.

4.2. Resultados del Desempeño del Sistema Computacional

El estudio se llevo a cabo ocupando un computador personal con una CPU Intel Pentium dual-core de 1,73 GHz, memoria de trabajo RAM 1Gb y 60 Gb de disco duro, configuración que corresponde a un equipo de escritorio de una gama media.

En el caso del entrenamiento de la red, el tiempo de CPU utilizado para las pruebas varió dentro del rango de 56 segundos hasta 1 hora con 19 minutos y 56 segundos, según la complejidad de la topología de la red, correspondiendo el menor tiempo a una red simple y el mayor tiempo a la red más compleja. La carga de la CPU estuvo bajo el 70 % para todas las pruebas y, de igual forma, la memoria de trabajo no paso el umbral de los 25 MB.

En el escenario de la ejecución de la red ya entrenada para clasificar una entrada dada, los recursos fueron notoriamente menores, siendo el tiempo de CPU menor a un minuto, la carga de CPU menor a un 30 % y la memoria de trabajo menor a 5 MB en todas las ejecuciones.

Por lo tanto, un computador de escritorio de gama media entrega la plataforma de hardware necesaria para una correcta ejecución del sistema computacional de la red neuronal.

4.3. Resultados del Problema en Estudio

4.3.1. Caso Datos en Bruto

En general, la red identificó adecuadamente un 83,33 % del total de entradas. El desglose se ilustra en la figura 17, donde se muestra el número de patrones identificados según el tipo ingresado. Siendo el mejor desempeño para

el patrón normal (91,23 %. de efectividad) y el peor en el caso de un patrón de "Escalón positivo" (0 %).

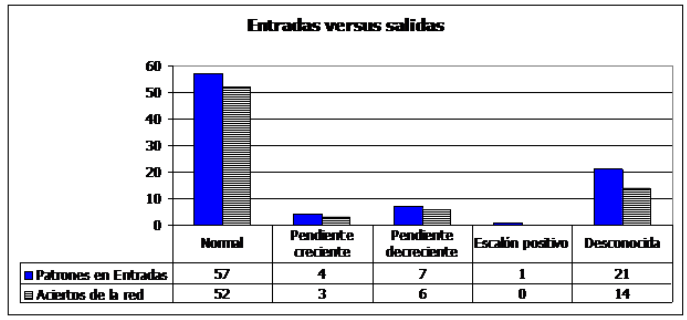


Figura 17: Desglose de patrones junto con clasificaciones realizadas por la red.

En los diferentes tipos de patrón conocidos, menos el escalón, se obtuvo un 66 % de acierto como mínimo.

En particular, se tiene que un 23,33 % del total de entradas corresponde a patrones no definidos ni conocidos. Se observa que las formas de las curvas son compuestas entre varios patrones a lo que debe sumarse una alta dispersión de los datos en algunos casos y, en general, ruido en los datos.

Dentro de los patrones conocidos existe el de "Escalón negativo" pero que no se hizo presente en el conjunto de datos analizados y la red tampoco lo identificó erróneamente.

En la siguiente matriz de cuatro cuadrantes (tabla 2) se ilustra los porcentajes en que la red identificó correctamente los patrones de entrada cuando eran conocidos e incorrectamente los mismos. Además, por otro lado entrega los porcentajes en que la red no identificó correctamente los patrones de entrada conocidos y los desconocidos.

	Identificación correcta de la red	Identificación incorrecta de la red
Entrada de Patrón conocido	**88,40 %**	**11,6 %**
Entrada de Patrón desconocido	**66,66 %**	**33,3 %**

Tabla 2: Eficiencia de la red en la clasificación de patrones.

De la tabla 2 se aprecia que la red obtiene una eficiencia de un 88,40 % para el caso de patrones conocidos, lo cual es un valor bueno. La eficiencia bajó ostensiblemente para los patrones desconocidos, en el fondo, no es que la red los identifique sino que no computa una solución de un patrón conocido en esta situación.

4.3.2. Caso Datos Filtrados

En general, la red identificó adecuadamente un 70 % del total de entradas. El desglose se ilustra en la figura 18, donde se muestra el número de patrones identificados según el tipo ingresado. Siendo el mejor desempeño para el patrón de "Pendiente creciente" (100 %. de efectividad) y el peor en el caso de un patrón de "Escalón positivo" (0 %).

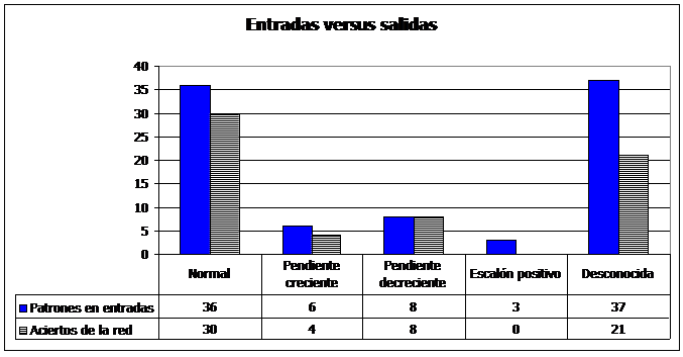


Figura 18: Desglose de patrones junto con clasificaciones hechas por la red.

De la figura 18 se puede apreciar que para los patrones conocidos, exceptuando el cado del "Escalón positivo", la red tuvo un buen rendimiento en la clasificación de éstos.

En particular, se tiene que un 41,22 % del total de entradas corresponde a patrones no definidos ni conocidos. Se ve que las formas de las curvas son compuestas entre varios patrones a lo que debe sumarse una alta dispersión de los datos en algunos casos.

Es de mencionar que dentro de los patrones conocidos existe el de "Escalón negativo" pero que no se hizo presente en el conjunto de datos analizados y la red tampoco lo identificó erróneamente.

En la siguiente matriz de cuatro cuadrantes (tabla 3) se ilustra los porcentajes en que la red identificó correctamente los patrones de entrada cuando eran conocidos e incorrectamente los mismos. Además, por otro lado entrega los porcentajes en que la red no identificó correctamente los patrones de entrada conocidos y los desconocidos.

	Identificación correcta de la red	Identificación incorrecta de la red
Entrada de Patrón conocido	**83,33 %**	**16,67 %**
Entrada de Patrón desconocido	**55,26 %**	**44,74 %**

Tabla 3: Porcentajes de acierto en clasificación.

De la tabla 3 se puede apreciar que la red obtiene una eficiencia de un 83,33 % para el caso de patrones conocidos, lo cual es un valor muy apropiado bueno. La eficiencia baja considerablemente para los patrones desconocidos; en el fondo, no es que la red los identifique sino que no computa una solución de un patrón conocido en esta situación.

5. Conclusiones y trabajo a futuro

5.1. Conclusiones

Con respecto a la clasificación de la salida del horno por medio de la herramienta computacional, se encontró que el porcentaje de eficiencia fue de un 83,33 %, que corresponde a un porcentaje aceptable y, además, es consistente con lo obtenido en la fase de entrenamiento (82, 96 %) de la red neuronal contenida en el programa. Dado que los datos eran de un proceso real, lo que implica una componente de ruido inherente en los valores, se volvieron a procesar los datos pero con un filtrado previó, obteniendo un rendimiento similar para patrones conocidos, lo que indica que la red puede de trabajar con datos en crudo.

Además, al analizar los resultados, se encontraron falencias en el proceso de entrenamiento, siendo éstas: definición no adecuada del patrón "escalón" puesto que el cambio de valor objetivo es gradual en la realidad y no directo como en los datos de entrenamiento, falta de otros patrones típicos como cíclico por ejemplo y flexibilizar el esquema de ventana de muestreo fija a uno dinámico, pues en algunas ventanas fijas se mezclan el inicio y el fin de un patrón conocido, lo cual inducía a un error de clasificación por parte de la red.

En cuanto al campo de aplicación de esta metodología, es aplicable para evaluar el funcionamiento del control automático de cualquier proceso continuo o semi continuo. Además, es aplicable también en el ámbito del control de calidad para buscar patrones en cartas de control, por ejemplo, en las cartas I individual se puede aplicar directamente.

Acerca del uso de la metodología, se puede comentar que una vez entrenada la red, la utilización de la herramienta de clasificación es directa sin la necesidad de manejar conceptos de redes neuronales.

En lo relativo a las dificultades o requerimientos que se presentaron, se puede decir que el principal costo para la aplicación de la red es la necesidad de una gran cantidad de datos conocidos de entradas y salidas para poder entrenarla adecuadamente, por otro lado, en cuanto a las competencias requeridas para la interpretación de resultados se requiere un conocimiento de la operación del proceso bajo evaluación para poder explicar las causas a los patrones "anormales" que puedan encontrarse.

5.2. Trabajos futuros

En cuanto a la topología de la red, dado que sólo se trabajó con una arquitectura Perceptron multicapa, se propone probar con arquitecturas de red de base radial y de neuronas recurrentes para contrastar el porcentaje de efectividad y los tiempos de entrenamiento de cada caso.

Con respecto al conjunto de patrones, se propone incrementar el número de elementos, con otros tipos de patrones de origen teórico y/o práctico, para cubrir un amplio rango de patrones en señales de terreno.

Cambiar el sistema computacional a un esquema de una ventana de muestreo de tamaño fijo y estática en el tiempo por uno de una ventana dinámica en el tiempo, que se realice el muestre de la señal de salida del proceso en tiempo real, de modo de abolir la mezcla de patrones conocidos debidos a una ventana fija de tiempo.

Se puede reorientar fácilmente el sistema computacional desarrollado para ser utilizado en la simulación de procesos industriales, lo que permite evaluar los procesos en cuanto a rendimiento y en predicción de las salidas de estos para entradas de interés.

Agradecimientos: a Mónica Meneses por su apoyo y más.

Referencias

[1] Chesnokov, Y. Backpropagation Artificial Neural Network in C++. *http://www.codeproject.com.*, 2008.

[2] Elman, J. Learning and development in neural networks: the importance of starting small. *University of California.*, 1992.

[3] Fu, T., Chung, F., Ng, V., Luk, R. Pattern Discovery from Stock Time Series Using Self-Organizing Maps. *Department of Computing Hong Kong, Polytechnic University*, 2001.

[4] Gutiérrez, H., Salazar, R. Control estadístico de calidad y seis sigma. *Mc Graw Hill*, 2005.

[5] Guh, R., Shiue, Y. Effective Pattern Recognition of Control Charts using a Dynamically Trained Learning Vector Quantization Network. *Journal of the Chinese Institute of Industrial Engineers*, Vol. 25, Nro. 1, 2008.

[6] Hann, F., Kostanic., I. Principles of neurocomputing for science & engineering. *Mc Graw Hill*, 2000.

[7] Kirillov, A. Aforge.Net Framework. *http://code.google.com/p/aforge/*, 2008.

[8] Madhusudanan, O. Brainnet 1 - A Neural Network Proyect - With Illustration And Code - Learn Neural Network Programming step By Step And develop a simple Handwriting Detection System. *http://www.codeproject.com*, 2006.

[9] Madhusudanan, O. NXML - Introducing an XML Based Language To Perform Neural Network Processing, Image Analysis, Pattern Detection, Etc. *http://www.codeproject.com*, 2006.

[10] Matich, D. Redes Neuronales: Conceptos básicos y aplicaciones. 2001.

[11] Microsoft, MSDN Library. 2003.

[12] Microsoft, Introduction to Visual Basic 2005. 2005.

[13] Muñoz, A. Aplicación de Técnicas de Redes Neuronales Artificiales al Diagnóstico de Procesos Industriales. *Universidad Pontificia comillas Madrid*, 1996.

[14] Ogata, K. Ingeniería de Control Moderna. *Prendice Hall*, Segunda Edición, 1993.

[15] Olmedo, E., Valderas, J. Utilización de de Redes Neuronales en la Caracterización, Modelización y Predicción de Series Temporales Económicas en un Entorno Complejo. 2004.

[16] Ruiz, C. Sistemas inteligentes en la Industria de Procesos: Experiencia, Actualidad y Perspectivas. *Argentine Symposium on Artificial Intelligence*, 2001.

[17] Sagiroglu, S., Besdok, E., Erler, M. Control Chart Pattern Recognition Using Articial Neural Networks. *Turk J Elec Engin,* Vol.8, No.2., 2001.

[18] Singh, S., Fieldsend, J. Pattern Matching and Neural Networks based Hybrid Forecasting System. *PANN Research,* Department of Computer Science, University of Exeter, UK, 2001.

[19] Silva, R., Cruz, A., Hokka, C., Giordano, R. A Hybrid Feedforward Neural Network Model for the Cephalosporin C Production Process. *PANN Research,* Department of Computer Science, University of Exeter, UK, 2001.

[20] Soto, L. Tutorial Visual C# 2005. *http://www.programacionfacil.com*, 2007.

[21] Tham, M. Dealing with measurement noise. *http://lorien.ncl.ac.uk/ming/filter/filter.htm*, 1998.

[22] Viñuela, P., Galván, I. Redes de Neuronas Artificiales, Un enfoque Práctico. *Prentice Hall*, 2004.

[23] Winston, W. Investigación de operaciones: Aplicaciones y algoritmos. *Thomson*, 2004.

Un método de optimización lineal entera para el análisis de sesiones de usuarios web.

Pablo E. Román*
Juan D. Velásquez*
Robert F. Dell**

Resumen

"Web usage mining" es una nueva área de investigación que ha producido importantes avances en la industria del e-Business, mediante la entrega de patrones de comportamiento de compra y sugerencias de navegación que mejoran la experiencia del usuario web en el sitio. Una de las principales fuentes de datos usadas en web mining, son las sesiones (secuencias de páginas) de los usuarios web que deben ser reconstruidas a partir de los archivos de Log. El problema con los archivos de Logs es que incluyen una componente de ruido al no identificar explícitamente a los usuarios que generan los registros. Con este trabajo, se desarrolla una aplicación basada en modelos de optimización como el como el problema de "maximum cardinality matching" y programación entera, que comparamos con una heurística comúnmente usada. Se analizan variaciones de los modelos de optimización presentados para explorar la verosimilitud de sesiones específicas y características de las sesiones. Se obtiene como resultado sesiones de mejor calidad que las obtenidas con los métodos tradicionales, además de una metodología de análisis de ellas.

Palabras Clave: Web Usage Mining, Web User Session, Maximum Cardinality Matching, Network Flow Model, Integer Programming, Web Logs.

*Departamento Ingeniería Industrial, Universidad de Chile, Santiago, Chile
**Operations Research Department, Naval Postgraduate School, Monterey, California, USA

1. Introducción

Los archivos de Log de un servidor web contienen registros de las operaciones que realizan los usuarios al navegar por un sitio web, convirtiéndose en una potencial gran fuente de datos acerca de sus preferencias [23]. Un Log [2] es un gran archivo de texto donde cada línea (registro) contiene los siguientes campos: Tiempo de acceso al objeto web (Ej. página web), la dirección IP del usuario, el agente que es la identificación del navegador usado, y el objeto web. También contiene evidencia de las actividades de de los usuarios web y se le puede considerar como una gran encuesta sobre sus preferencias en relación a la información que aparece en el sitio web. Lo anterior ha motivado gran parte de la investigación que se realiza en web mining, y define un nuevo campo de investigación denominado Web Usage Mining [23].

Un archivo de Log por si mismo no necesariamente refleja las secuencias de páginas que acceden los usuarios web i.e., se registra cada acceso pero sin un único identificador que represente al cliente. Esto se debe a que muchos usuarios distintos pueden compartir la misma dirección IP y tipo de navegador (agente), generando la necesidad de reconstruir las sesiones de usuario usando los datos disponibles. En la actualidad se utilizan métodos heurísticos para reconstruir las sesiones desde los archivos de Logs, las que se basan principalmente en limitar la duración de las sesiones [3], [6] y [20]. Este trabajo se centra en proponer modelos de optimización para recuperar las sesiones y estudiar sus propiedades.

El presente artículo se organiza de la siguiente manera: La sección 2 resume el estado del arte en relación a la sesionización. La sección 3 presenta nuestro modelo de optimización. La sección 4 muestra variaciones del modelo de optimización para explorar la verosimilitud y propiedades específicas de las sesiones. La sección 5 describe los datos experimentales usados. La sección 6 presenta los resultados. La sección 7 concluye el trabajo y sugiere futuras investigaciones.

2. Estado del Arte

Las estrategias de sesionización, pueden ser clasificadas en reactivas y proactivas [20]. La sesionización proactiva captura todas las actividades realizadas por los usuarios durante su visita al sitio web, sin embargo, son invasivas y en general con poco resguardo a la privacidad de los usuarios. El uso de estas estrategias se encuentra regulado por ley en algunos países [20] de forma de proteger la privacidad de las personas [15]. Un ejemplo corresponde al uso

de cookie[1] [4] que registran las actividades del cliente de las cuales se puede extraer la sesión exacta del usuario. Otra técnica usada es la re-escritura de URL[2] [7], donde se incluye información en el URL que se envía al servidor que reconstruye la sesión. La forma mas invasiva de obtener sesiones es a través de los llamados Spyware, que son programas que registran cualquier actividad del usuario (Teclado, Mouse, etc.). Sin embargo son actualmente considerados como una actividad criminal en la mayoría de los países [16].

Las estrategias de sesionización reactivas tienen un alto nivel de resguardo a la privacidad de los usuarios ya que sólo usan los registros de Log y no manejan explícitamente los datos personales de los usuarios [20]. Sin embargo, los archivos de Log son una forma aproximada de obtener las sesiones por muchas razones. Los usuarios pueden tener el mismo IP debido a que los ISP comparten un limitado número de direcciones entre sus clientes. Los usuarios web pueden también hacer uso de los botones back y forward que en la mayoría de las veces no producen registros en los Logs del servidor. Otro factor que introduce ruido en los datos son los servidores Proxy[3] [9] que mantienen en cache un cierto numero de páginas frecuentemente visitadas para optimizar las velocidades de acceso, por lo cual nunca son registrados en los archivos de los del sitio web.

Los métodos que se manejan en la actualidad para reconstruir sesiones desde los archivos de Logs están basados en heurísticas que consideran un límite de tiempo para la duración de las sesiones (30 minutos) [20]. Otras heurísticas se basan en la estructura semántica del sitio y las sesiones se construyen de forma de seguir una semántica común [13].

Se han realizados estudios empíricos en relación al comportamiento estadístico de las sesiones. La función de probabilidad de distribución del largo n (número de saltos entre páginas) tiene un buen ajuste con un ley de potencia $(n^{\alpha}/\sum_k k^{\alpha})$ [10][22]. La distribución parece ajustarse a una variedad de sitios web, aunque con cambio en el parámetro α. Nosotros usamos esta propiedad que parece universal de las sesiones como una medida de calidad de éstas [18].

Existe una gran variedad de literatura para el minado de las sesiones una vez que estas han sido identificadas. Técnicas como análisis estadístico, reglas de asociación, clustering, clasificación, patrones secuenciales y modelamiento de dependencias han sido usados para descubrir patrones de comportamiento de usuarios web [12][14][21].

[1] Archivos que se almacenan en el computador del cliente que almacenan datos

[2] Dirección web de la página, e.g. http://www.dii.uchile.cl

[3] Servidor que almacena copias de páginas mas acezadas por los usuarios de una red, de forma distribuirlas en forma más rápida.

3. Modelos de optimización para la sesionización

Se presentan dos modelos de optimización para la sesionización, los cuales agrupan registros de Logs de un mismo IP y agente, así como también consideran la estructura de links del sitio web. A diferencia de la heurística, que construye las sesiones una por una, los algoritmos de optimización propuestos construyen en forma simultánea. Cada sesión así construida es una lista de registros de Logs, donde cada registro es usado una sola vez en una única sesión. En la misma sesión, un registro r_1 puede ser un predecesor inmediato de r_2 si: los dos registros poseen la misma IP y agente, un link existe desde la página asociada al registro r_1 hasta la página del registro r_2, y el registro r_2 se encuentra en una ventana de tiempo permitida según el registro r_1.

3.1. Bipartite Cardinality Matching

El primer modelo de optimización que presentamos está basado en el conocido problema "Bipartite Cardinality Matching" (BCM) (e.g. [1]), el cual consiste en encontrar en un grafo no dirigido el subconjunto de máxima cardinalidad que tenga la propiedad de "matching" (no hay 2 vértices que compartan la misma arista). Existen varios algoritmos especializados para resolver el problema BCM en un tiempo de computación $O(\sqrt{n}m)$ donde "n" es el número de vértices y "m" es el número de arcos (e.g. [1]). En nuestra red, cada registro es representado por dos nodos, unos que representan el predecesor inmediato y otros a los sucesores inmediatos. La figura 1 muestra un ejemplo con 6 registros.

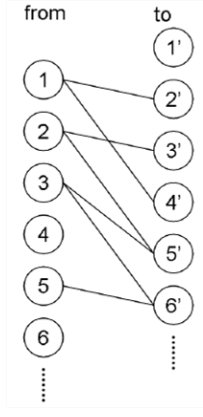


Figura 1: BCM: Cada registro es representado por dos nodos. Un arco existe si un nodo puede ser predecesor (from) de un vértice que puede ser su sucesor, según las restricciones del problema.

En cada lado, se ordenan los nodos (registros) en orden creciente de tiempo de acceso en los Logs del servidor web. Un arco existe de un nodo r_1 (from) a un nodo r_2 (to) si el registro correspondiente a r_1 puede ser un inmediato predecesor de r_2. Para el caso de la figura 1 asumimos que existen siete arcos.

Dada una solución, se construyen las sesiones de acuerdo al "matching" encontrado. Un vértice que no es sucesor de otro vértice, es el primer registro de la sesión. El resto de la secuencia de registros se reconstruye identificando los pares de vértices que corresponden a un mismo registro y siguiendo consecutivamente los sucesores habilitados por un arco. La figura 2 provee una solución factible de acuerdo a la figura 1. Los vértices 4 y 6 son los últimos en una sesión (no tienen sucesores), los vértices 1'y 4'van primero en sus respectivas sesiones ya que no tienen antecesores, las sesiones resultantes son entonces 1-2-3-5-6 y 4.

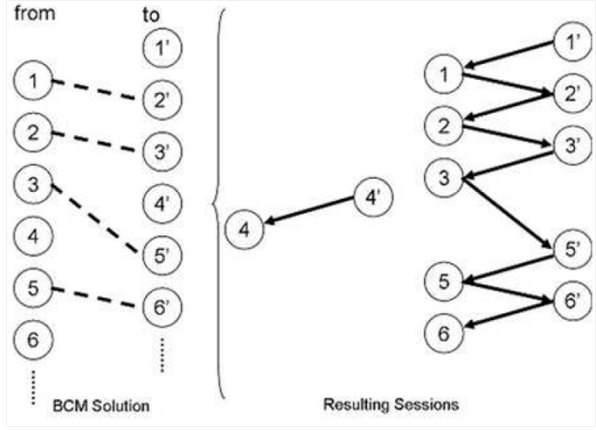


Figura 2: BCM tiene cuatro arcos. Se construyen dos sesiones desde el matching: 1-2-3-5-6 y una sesión con un sólo registro 4.

3.2. Un modelo de programación entera para la sesionización

Una solución óptima para el modelo BCM provee un límite inferior en el número de sesiones para un archivo de Log dado. Esta solución óptima tiene la propiedad que pocas sesiones tendrán largos pequeños (1 o 2). El limite superior en el número de sesiones para un archivo de Logs dado, es el numero de registros donde cada registro es una sesión por si mismo. De forma de construir otras sesiones con un límite superior en su cantidad, formulamos un modelo de programación entera.

El modelo de sesionización desarrollado, utiliza en su formulación programación entera (SIP) usa una variable binaria X_{ros} que tiene valor "1" si el registro de *Logr* es asignado en la posición o durante la sesión s y "0" en cualquier otro caso. Cada índice r identifica a un único registro, cada índice

s identifica a una única sesión de usuario, y el índice o es la posición de un registro en una sesión.

Presentamos a continuación la formulación del problema en formato NPS [5].

- **Índices**

  o Orden de un registro de Log durante una sesión (e.g. *o=1,2,...,20*). La cardinalidad de este conjunto define el máximo tamaño de una sesión.

  p,p' Numera a las páginas web.

  r,r' Numera a los registros de Logs.

  s Numera a las sesiones de usuarios.

- **Conjuntos de Índices**

  $r' \in bpage_r$ Es el conjunto de registros que pueden estar inmediatamente antes del registro r en la misma sesión. Basado en:

  - Páginas que tienen un link a la página del registro r.
  - La dirección IP que comparten r y r'.
  - El agente que comparten r y r'.
  - El tiempo del registro r y r', r debe ocurrir antes de r' dentro de una ventana de tiempo.

  $r \in first$ Es el conjunto de registros que puede ocurrir en primer lugar en una sesión.

- **Datos**

  Estos son usados para generar los conjuntos anteriores.

  $time_r$ El tiempo del registro r.

  ip_r La dirección IP del registro r.

  $agent_r$ El agente del registro r.

  $page_r$ La página del registro r.

  $\underline{mtp}$, mtp El tiempo mínimo y máximo entre páginas (segundos).

- **Variables Binarias**

  X_{ros}: "1" si el registro de Log r es asignado en el lugar o en la sesión s y "0" en otro caso.

- **Formulación**

Maximizar $\sum_{ros} C_{ro} X_{ros}$
s.a.

$$\sum_{os} X_{ros} = 1 \quad \forall r \tag{1}$$

$$\sum_{r} X_{ros} \leq 1 \quad \forall o, s \tag{2}$$

$$X_{r,o+1,s} \leq \sum_{r' \in bpage_r} X_{r'os} \quad \forall r, o, s \tag{3}$$

$$X_{ros} \in \{0,1\}, \quad \forall r, o, s. \quad X_{ros} = 0, \quad \forall r \in first, o > 1, s$$

La función objetivo expresa cuánto se considera a las sesiones de mayor largo donde$\sum_{r,o' \leq o} C_{ro'}$ es el beneficio por una sesión de largo o'. Por ejemplo, si fijamos $C_{ro'} = 1, \quad \forall r, o = 3$ y $C_{ro'} = 0, \quad \forall r, o \neq 3$ se tendrá una función objetivo que maximiza el número de sesiones de largo tres. La sección 5.2 muestra una variedad de resultados asociados a diferentes elecciones de parámetros C_{ro}.

El conjunto de restricciones (1) aseguran que cada registro sea usado una vez. Las restricciones (2) aseguran a cada sesión al menos un registro asignado a cada posición o. Las restricciones (3) aseguran un orden apropiado de registros en la misma sesión.

4. Otras variantes del modelo

Durante la investigación, se desarrollaron otras variantes del modelo para explorar la verosimilitud de sesiones específicas y características de las sesiones. Específicamente, se logra establecer una aproximación para el máximo número de copias de una respectiva sesión, el máximo número de sesiones de un mismo largo y el máximo número de sesiones con una determinada página web en un cierto lugar.

4.1. El máximo numero de copias de una sesión

Para encontrar el máximo número posible de copias de una sesión dada en un registro de Log, tenemos que analizar dos casos, según si se considera repetición de páginas:

1. Cuando cada página en la sesión es visitada solo una vez, esto puede ser modelado como un problema de máximo flujo (e.g. [1]). El problema

de máximo flujo busca una solución que envía el máximo flujo desde un vértice artificial llamado fuente hasta un vértice sumidero. Se construye el grafo con un vértice por cada registro de Log que corresponden a las páginas de la sesión escogida, se además incluye la fuente y el sumidero. Los arcos se construyen de la siguiente forma: los que parten de la fuente apuntan a registros que pueden ser registros iniciales de una sesión, se incluyen arcos desde cada registro hasta el sumidero y entre dos registros que puede ser inmediatos predecesor y sucesor. Los arcos relacionados con la fuente y sumidero tienen infinita capacidad, el resto tiene capacidad máxima de uno.

2. En el caso que las páginas en la sesión se repiten, se introducen restricciones adicionales para el problema de flujo máximo. En este caso, la red es similar al anterior pero se debe registrar el orden en que se acceden las páginas debido a la repetición. Entonces para páginas repetidas se repiten registros que correspondan a la página repetida manteniendo el orden y se restringe que el máximo flujo total que pueda salir de estos registros repetidos sea menor o igual a uno.

4.2. Maximizando el número de sesiones de un mismo largo

Usando el modelo SIP de la sección 3.2, se puede encontrar el máximo número de sesiones de un largo l ajustando la función objetivo a $C_{ro} = 0, \quad \forall r, o \neq l$y $C_{ro} = 1, \quad \forall r, o = l$. Con estos coeficientes, sólo las sesiones de largo mayor o igual que l tienen valor uno. Una solución optima puede incluir sesiones mayores que el que este largo, pero pueden ser separadas en dos sesiones donde una de ellas es de largo l.

4.3. Maximizando el número de sesiones con una página fija en cierta posición

Usando el modelo SIP, se puede encontrar el máximo número de sesiones que visitan una página fija en la posición o. Para ello, se fijan los coeficientes $C_{ro} = 1$ cuando el registro r corresponde a la página fija y/o a la posición escogida, en caso contrario se anula.

5. Los datos experimentales

Durante la presente investigación, se trabajó con un sitio web universitario (http://www.dii.uchile.cl) que mantiene el Departamento de Ingeniería Industrial de la Universidad de Chile. El cual esta compuesto por los sitios web corporativos, de proyectos, programas de diplomados y postgrado, páginas

personales, y webmail entre otras. Corresponde a un sitio web con alta diversidad y un tráfico de Internet suficientemente alto para satisfacer los objetivos de este estudio, aun sin considerar el sistema de webmail cuyo registro de Log no contiene sesiones muy variadas.

5.1. Estadísticas preliminares

Se obtuvieron 3.756.006 registros durante el mes de abril de 2008. Para observar las transiciones entre páginas, se deben filtrar los accesos a objetos multimedia y otros que no corresponden a páginas web. Se deben eliminar registros de spider[4] o robot, herramientas automáticas de monitoreo de seguridad[5], ataques de virus[6] que no son propiamente considerados como pertenecientes a sesiones realizadas por agentes humanos. Se obtuvo un total de 102.303 registros de páginas estáticas HTML con un total de 172 páginas, de estos 9.044 registros corresponden a accesos a la raíz del sitio web.

Se obtuvo que tan sólo unas pocas direcciones IP las que recogen la gran mayoría de los accesos a registros. Sobre un 98 % de todas las direcciones tiene menos de 50 registros por mes. De la misma forma, se constató que pocas direcciones IP tienen el acceso más diverso a páginas. Se almacenó la estructura de link del sitio web usando un web crawler[7] [17], obteniéndose 172 páginas con 1.228 link entre ellas considerando solo páginas registradas en los Logs.

5.2. Pre-procesamiento de datos

Como mencionamos existe una gran cantidad de registros triviales de analizar en el sentido que corresponden a sesiones de largo, ya sea por la baja cantidad de accesos o por la poca variedad de páginas accedidas. En este estudio nos enfocamos en el subconjunto de registros que muestren la mayor diversidad de patrones de acceso. Para ello, se consideran registros para nuestra sesionización mediante filtrado por numero IP. Para cada registro IP se propone una medida de diversidad de páginas visitadas basada en la entropía de distribución de páginas $S = \sum_p f_p Log_N (1/f_p)$, donde f_p es la frecuencia de accesos a la página p y N es el número total de páginas. La entropía S adquiere su valor máximo (igual a 1) cuando la distribución de acceso a páginas

[4]Programas automáticos que recorren la web visitando sitios, típicamente usados por los buscadores (e.g. Google) para mantener su base de datos de la web actualizada.

[5]Programas automáticos usados por los administradores de sistemas para chequear el estatus del sitio web.

[6]Existen virus que se propagan visitando sitios web e intentando ingresar y modificar las páginas para infectarlas, de forma que algún otro usuario que la visite se infecte.

[7]Un crawler es un programa automático que recorre las páginas web de un sitio almacenando su estructura y contenido.

es uniformemente distribuida, es decir existe diversidad de acceso. En cambio, cuando su valor es cercano a cero, solo unas pocas páginas son accedidas frecuentemente. En este caso agrupamos los registros por numero IP que tuviesen alta entropía ($S > 0,5$) y un alto numero de registros ($Log(N) > 3,8$) asegurando diversidad de comportamiento.

6. Resultados obtenidos

Se presenta los resultados de la sesionización y usamos una medida de la calidad de los resultados.

6.1. Medición de la calidad de las sesiones

Sin tener acceso a comparar con las verdaderas sesiones de usuarios, no es posible conocer en forma exacta cuán realista es el método desarrollado en el presente articulo, se propone comparar la distribución de largos de sesiones con la distribución empírica [10][22] observada en forma natural. Para ello se usa regresión lineal en el logaritmo del largo y logaritmo del número de sesiones. Se entregan entonces como resultado de la regresión, el coeficiente de correlación y el error estándar como medidas de la calidad de las sesiones.

6.2. Procesamiento

Es fácil construir instancias del modelo SIP propuesto que no puedan ser resultas por computador alguno debido al gran número de variables involucradas. Por ejemplo un servidor web (como el estudiado) con 100.000 registros, con un máximo de 5.000 sesiones y un largo máximo de sesiones de 20, genera 1010 variables binarias y un número aun mayor de restricciones. Un problema de esa envergadura requeriría cerca de 1Tb de memoria de acceso directo solo para almacenar las variables, las cuales deberían además modificarse en cada iteración. Afortunadamente se puede subdividir el problema separando el archivo de Log en unidades mas pequeñas que se agrupen por numero IP y agente; y fijando un limite máximo de tiempo entre registros (15 min.). Con esto se obtuvieron 403 unidades de registros fijando además que cada unidad disponga un mínimo de 50 registros para evitar por otro lado un número mayor de unidades. Todos los procesamientos fueron realizados en un PC de 1,6Ghz con 2Gb RAM. Se resolvieron las instancias usando el software GAMS [8] en los 403 programas lineales utilizando CPLEX [11] con la versión 10.1.0. El sistema generador de instancias fue realizado en PHP y los datos se almacenaban en una base de datos MySQL 5.0.27 [24].

6.3. Resultados del algoritmo BCM

Se resolvieron las 403 unidades usando CPLEX, donde la instancia más grande consistía en 1.500 variables y 200 restricciones. El tiempo total de resolución para las 403 unidades fue de menos de 5 minutos. Usando un algoritmo BCM especializado, se puede reducir mayormente el tiempo y además puede ser paralelizada resolviendo las 403 unidades concurrentemente. La solución entregó 12,366 sesiones. La sesión más larga tiene 41 registros, después vienen sesiones menores o iguales a 14 registros. Considerando todas las sesiones, obtuvimos un coeficiente de correlación de $R^2 = 0,88$ y un error estándar de 1,23. Considerando largos de sesiones hasta los 14 registros obtenemos un coeficiente de correlación $R^2 = 0,98$ y un error estándar de 0,39 (ver figura 3). Claramente la única sesión de 41 registros debe ser considerada especialmente ya que se encuentra en el rango fraccionario de la ley de distribución de potencias y el no considerarla para la regresión es una buena aproximación.

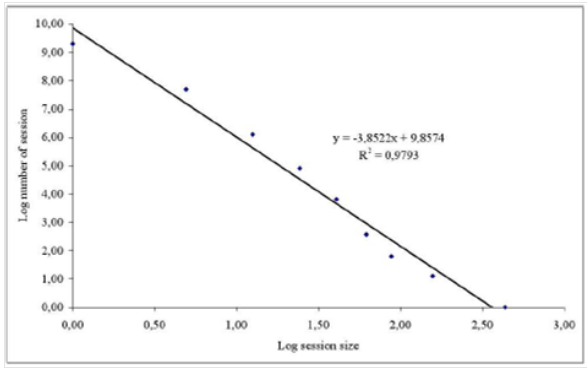


Figura 3: Regresión lineal de los resultados BCM.

6.4. Resultados SIP

Este modelo depende de la elección de los coeficientes lineales en la función objetivo. Por ejemplo si se escoge $C_{ro} = 0, \quad \forall r, o \neq 1$y $C_{ro} = 1, \quad \forall r$; entonces se busca maximizar el número de sesiones entregando como solución trivial sólo sesiones con un solo registro. Se requiere, por tanto, que los coeficientes C_{ro} sean una función monótona creciente en "o", con el fin de valorizar más a las sesiones de mayor largo en el problema de optimización. Para estos fines, se experimentó con varias funciones con distintas tazas de crecimiento obteniendo los resultados de la tabla 1:

El tercer conjunto mostró el mejor ajuste. Cabe señalar que corresponde a la elección de $C_o = Log(1/P_o)$ la cantidad de información siendo Po la probabilidad de tener una sesión de largo o. Como hemos mencionado, esta probabilidad tiene como aproximación una Gausiana inversa, cuyo valor se

	C_{ro}	R^2	StdErr	Total Sesions
1	$1/\sqrt{o}$	0.88	1.10	12.502
2	$Log(0)$	0.93	0.66	12.403
3	$3/2Log(o) + (o-3)^2/12o$	0.94	0.59	12.403
4	o	0.93	0.63	12.409
5	o^2	0.92	0.72	12.410

Tabla 1: Conjuntos de coeficientes para la función objetivo

ajusto en este caso. Consideramos como argumento de plausibilidad parea esta elección que la función objetivo se aproximaba a la entropía de la distribución de largos de sesiones, en cuyo óptimo se aproxima a la distribución deseada. Finalmente los tiempos totales de procesamiento fueron de 3 a 5 horas.

6.5. Comparación con la heurística comúnmente usada

Se evaluaron los resultados comparando con la heurística basada en el tiempo máximo de sesiones. Claramente la heurística consume mucho menos recursos computacionales resolviéndose en menos de 10 segundos para todos los registros y sin preprocesar. Pero obtiene el peor ajuste a la distribución estadística de largos de sesiones ($R^2 = 0,92$ y $std = 0,64$) teniendo aproximadamente el doble del error estándar obtenido con el método BCM (ver figura 4).

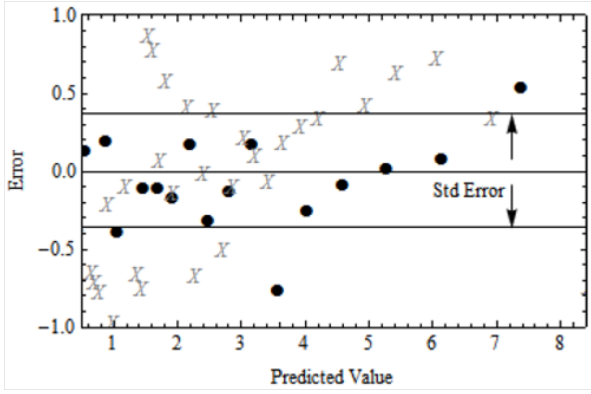


Figura 4: Error de ajuste a la ley de potencia del largo de sesiones.

Si se comparan los resultados logrados por SIP (mejor caso) y BCM en el número de sesiones logradas, tendríamos un GAP del 0,3 % entre ambos (SIP: cota superior, BCM: cota inferior) lo que indica la calidad de las sesiones obtenidas por este nuevo método.

6.6. El máximo número de copias de una sesión

Con el algoritmo propuesto, se rankea el máximo número de secuencias iguales de páginas obtenidas por todas las sesiones obtenidas con los métodos anteriores. Así, el procesamiento de este ranking de verosimilitud tardo menos de 9 horas, cuyos resultados son consignados en la tabla 2.

Session	BCM	SIP	max
1	41	41	186
2	5	3	43
3	4	5	39
4	7	4	34
5	4	4	34
6	1	0	22
7	1	0	22
8	7	0	19
9	1	0	16
10	1	0	16

Tabla 2: Las 10 sesiones más verosímiles y su comparación con su ocurrencia según el algoritmo.

6.7. El máximo número de sesiones según el tamaño

Como se indicó anteriormente, se ajusta la función objetivo en el modelo SIP para obtener el máximo número de sesiones de un mismo tamaño. Con ello se obtuvieron los resultados mostrados en la tabla 3.

Size	Num. Sessions
2	3000
3	1509
4	755
5	435
6	257
7	181
8	135
9	98
10	82

Tabla 3: Máximo número de sesiones dado un largo fijo.

6.8. Máximo número de sesiones que visitan una página en cierto orden

Se usó este algoritmo para rankear las páginas que aparecen en tercer lugar dado el máximo número de sesiones según este algoritmo. Para ello se proceso

para todas las páginas factibles de aparecer en tercer lugar (fueron 58) y se demoró aproximadamente de 29 horas.

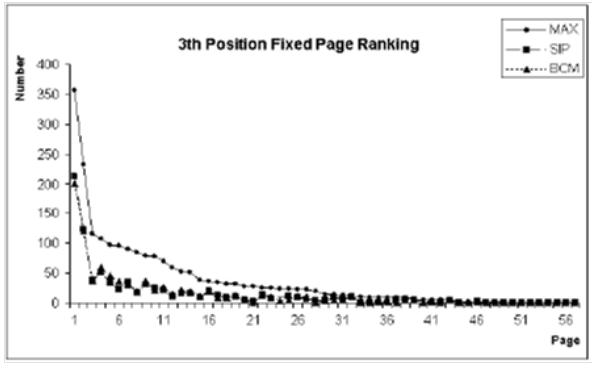


Figura 5: Ranking de páginas más verosímiles de aparecer en tercer lugar.

Este ranking entrega la noción de páginas más verosímiles de encontrar en tercer lugar en una sesión dado la posición examinada. La consistencia se compara con el número de apariciones en los otros algoritmos (figura 5).

7. Conclusiones

Se presenta un nuevo enfoque para la sesionización usando algoritmos de programación entera. Cuando se compara con los métodos heurísticos tradicionales, se reduce en un 50 % el error estándar de la distribución de largo de sesiones con respecto a la distribución esperada. Se dispone, además, de una versión altamente eficiente basada en el problema de "Bipartite Cardinality Matching". Esto nos permite obtener máximos sesiones con máxima probabilidad de ocurrencia dado largo fijo, secuencia de páginas fija y página fija en cierta posición. Los modelos no contemplan la posibilidad de uso de los botones back y forward del navegador, sin embargo dada la flexibilidad de los modelos de optimización en incorporar restricciones es posible lograr recuperar sesiones que incluyan este efecto.

Agradecimientos: Se agradece el apoyo del Instituto Milenio de Sistemas Complejos de Ingeniería y la Beca de Doctorado Nacional Conicyt.

Referencias

[1] Ahuja, T. Magnanti, J. Orlin. Network Flows: Theory, Algorithms, and Applications. *Prentice Hall*, 1993.

[2] Aulds, C. Linux Apache Web Server Administration. *Sybex.* 2002.

[3] Berendt, B., A. Hotho, G. Stumme. Data preparation for mining world wide web browsing patterns. *Journal of Knowlegde and Information Systems*, Vol.1 5-32. 1999.

[4] Berendt, B., B. Mobasher, M. Spiliopoulou, J. Wiltshire. Measuring the accuracy of sessionizers for web usage analysis. *Proc.of the Workshop on Web Mining*, First SIAM Internat.Conf. on Data Mining. 7-14. 2001.

[5] Brown, G., Dell, R. Formulating integer linear programs: A rogues'gallery. *Informs Transactions on Eduacation*, 7. 2007.

[6] Cooley, R., B. Mobasher, J. Srivastava. Formulating integer linear programs: A rogues'gallery. *Towards semantic web mining.* Proc. in First Int. Semantic Web Conference, 264-278. 2002.

[7] Facca, F., P. Lanzi. Recent developments in web usage mining research. *DaWaK*, 140-150. 2003.

[8] GAMS Development Corporation. General algebraic modeling system (gams). *www.gams.com*, Accessed December 2008.

[9] Glassman, S. A caching relay for the world wide web. *Computer Networks and ISDN Systems*, Vol.27 165-173. 1994.

[10] Huberman, B., P. Pirolli, J. Pitkow, R. Lukose. Strong regularities in world wide web surfing. *Science*, Vol.280 95-97. 1998.

[11] ILOG. 2008. Cplex 2008. Strong regularities in world wide web surfing. *www.ilog.com/products/cplex*, Accessed December 2008.

[12] Joshi, A., R. Krishnapuram. On mining web access Logs. *Proc. of the 2000 ACM SIGMOD Workshop on Research Issue in Data Mining and knowledge Discovery*, 63-69. 2000.

[13] Jung, J., Geun-Sik Jo. Semantic outlier analysis for sessionizing web Logs. *ECML/PKDD Conference*, 13-25. 2004.

[14] Kosala, R., H. Blockeel. Web mining research: A survey. *SIGKDD Explorations: Newsletters of the special Interest Group (SIG) on Knowledge Discovery and Data Mining*, 1 1-15. 2000.

[15] Langford, D. Internet Ethics. *MacMillan Press Ltd.* 2000.

[16] Mayer-Schonberger, V. Nutzliches vergessen. *Goodbye Privacy Grundrechte in der digitalen Welt (Ars Electronica)*, 253-265. 2008.

[17] Miller, R., K. Bharat. Sphinx: A framework for creating personal, site-specific web crawlers. *Proceedings of the Seventh International World Wide Web Conference (WWW7)*, 119-130. 1998.

[18] Román, P., R. Dell, J. Velásquez. Web user sesión reconstruction using integer programming. *Proc. of the 2008 Web Inteligence Conference*, Sidney, Australia. 2008.

[19] Román, P., J. Velásquez. Markov chain for modeling web user browsing behavior: Statistical inference. *XIV Latin Ibero-American Congress on Operations Research (CLAIO).* 2008.

[20] Spiliopoulou, M., B. Mobasher, B. Berendt, M. Nakagawa. MA framework for the evaluation of sesión reconstruction heuristics in web-usage analysis. *INFORMS Journal on Computing*, Vol.15 171-190. 2003.

[21] Srivastava, J., R. Cooley, M. Deshpande, P. Tan. Web usage mining: Discovery and applications of usage patterns from web data. *SIGKDD Explorations*, Vol.2 12-23. 2000.

[22] Vazquez, A., J. Oliveira, Z. Dezso, K. Goh, I. Kondor, A. Barabasi. Modeling bursts and heavy tails in human dynamics. *PHYSICAL REVIEW E 73*, 2006.

[23] Velásquez, J., V. Palade. Adaptive Web Sites: A Knowledge Extraction from Web Data Approach. *IOS Press*, Amsterdam, NL. 2008.

[24] Zawodny, J., D. Balling. High Performance MySQL. *O'Reilly*, 2004.

Un modelo de asignación de árbitros para el torneo de fútbol chileno y un enfoque de resolución en base a patrones

Fernando Alarcón*
Guillermo Durán*
Mario Guajardo*

Resumen

La asignación de árbitros en campeonatos deportivos es una tarea que, en instancias reales, generalmente conlleva a un problema de decisión combinatorial de difícil resolución. En la práctica, la asignación suele ser realizada en forma manual por una comisión experta, en base a criterios poco estructurados. Recientemente, diversas técnicas del Sports Scheduling se han abocado a mejorar esta situación, desarrollando modelos y comparando enfoques de resolución.

En este artículo estudiamos el problema de asignación de árbitros para el campeonato de la Primera División del fútbol chileno y lo abordamos mediante un modelo de optimización lineal entera. El modelo logra capturar criterios que otorgan transparencia y objetividad a la asignación, por ejemplo, balanceando la cantidad de partidos dirigidos por cada árbitro y sus distancias de viaje, y tomando en cuenta su categoría para arbitrar partidos especiales. Para su resolución, proponemos un enfoque basado en la construcción de patrones, inspirados en el enfoque de patrones de localía utilizado exitosamente en la programación del fixture de varias ligas deportivas alrededor del mundo. Según nuestro conocimiento, este es el primer artículo que utiliza un enfoque similar para la asignación de árbitros a un torneo.

Implementamos el modelo para instancias reales del problema y desarrollamos una herramienta ágil para el usuario, reportando resultados que mejoran significativamente la asignación tradicional. Además, mediante el enfoque de patrones resolvemos el problema significativamente más rápido que la resolución directa de la formulación original.

Palabras clave: asignación de árbitros, fútbol, patrones, programación entera.

*Departamento de Ingeniería Industrial, Universidad de Chile

1. Introducción

La disciplina que estudia el diseño eficiente de campeonatos deportivos es conocida como *Sports Scheduling*. Uno de los temas de investigación de los cuales se ocupa corresponde a la asignación de los árbitros que dirigirán cada uno de los partidos de un torneo. Usualmente, estas decisiones generan un problema combinatorial que, según las dimensiones y restricciones de cada torneo, puede conllevar a un problema complejo, prácticamente imposible de resolver a mano.

El problema de asignación de árbitros fue reportado por primera vez en la literatura cuando Wright [14] propuso un modelo básico para asignar los árbitros de la liga de cricket de Inglaterra. Varios años más tarde, Dinitz y Stinson [4] discutieron el problema de asignación de árbitros a un torneo programado previamente, utilizando determinados tipos de Room Squares. Trick y Yildiz [12] definieron un problema similar al bien conocido *Traveling Tournament Problem* (Easton et al. [7]), pero minimizando la distancia total a recorrer por los árbitros en vez de equipos y lo llamaron *Traveling Umpire Problem* (TUP). Más tarde, Duarte et al. [5] le dieron otro enfoque al problema y definieron el *Referee Assignment Problem* (RAP) enfocado en la asignación eficiente de árbitros a partidos de una liga deportiva, minimizando la suma sobre todos los árbitros de la diferencia entre los partidos que se quiere que dirija un árbitro y los que finalmente le son asignados.

Por otra parte, Gil y Rojas [8] propusieron un método de asignación que utiliza intervalos de confianza para representar la información considerando el grado de incertidumbre existente (dado que la experiencia y el nivel de cada árbitro es una cuestión subjetiva, los autores de este trabajo resuelven esta situación incorporando la incertidumbre al modelo).

Recientemente, Yavuz et al. [13] presentaron un modelo que busca la asignación justa de árbitros, considerando principalmente la frecuencia con que un árbitro es asignado a dirigir al mismo equipo.

En este artículo presentamos un modelo de optimización lineal entera para el problema de asignación de árbitros, en el marco del campeonato de la Primera División del fútbol chileno y un enfoque para resolverlo. En la Sección 2, describimos el contexto en que se origina el problema y en la Sección 3, desarrollamos un modelo que lo aborda. En la Sección 4, presentamos un enfoque de resolución en base a patrones. Finalmente, exponemos los resultados de nuestra implementación y las conclusiones en las secciones 5 y 6, respectivamente.

2. El Contexto Chileno

El torneo más importante del fútbol profesional chileno es el de Primera División, organizado por la Asociación Nacional de Fútbol Profesional de Chile (ANFP). Una de las tareas a cargo de la ANFP es la asignación de árbitros a los partidos del torneo. Para esto, existe una Comisión Arbitral compuesta por expertos, usualmente ex-árbitros que alguna vez dirigieron en el fútbol profesional chileno y ahora realizan esta labor organizativa. La Comisión realiza esta asignación semana a semana, basada en criterios poco estructurados y completamente a mano, sin herramientas de decisión sofisticadas. Como consecuencia de esto, la asignación resultante presenta varias desventajas. Por ejemplo, algunos árbitros dirigen una cantidad de partidos significativamente mayor a otros a lo largo del torneo, aun en el caso que su experiencia y trayectoria en el referato profesional sean similares. También a menudo sucede que un árbitro dirige una alta cantidad relativa de partidos a un mismo club; otros por el contrario, nunca dirigen a ciertos equipos. Adicionalmente, dado el carácter particular de la geografía chilena, una asigación defectuosa puede conducir a que algunos árbitros viajen una distancia total por torneo considerablemente mayor a la de otros.

Este tipo de factores dan cuenta de la relevancia de realizar una asignación correcta de los árbitros a los partidos del torneo. Aun más, no es extraño observar disconformidad de jugadores, dirigencias e hinchas sobre el desempeño de los árbitros que dirigen sus partidos. Esto a la vez genera presiones adicionales a la tarea de los árbitros. Si bien estos problemas no son netamente atribuibles a la asignación, creemos que si esta tarea es realizada mediante un modelo de optimización con criterios objetivamente definidos, contribuiría a que la asignación sea más justa y transparente para todos los actores involucrados, y eleve el nivel de profesionalismo del torneo profesional chileno.

A este respecto, un antecedente importante de aplicación de programación lineal entera en el fútbol chileno es reportado en [6], donde se desarrolla un modelo de optimización para la confección del fixture del torneo de Primera División. El modelo desarrollado ha logrado un impacto importante desde que fue implementado por primera vez en 2005, y se sigue utilizando hasta la fecha, extendiéndose su aplicación a partir de 2007 a la Segunda División. Mayores detalles sobre la estructura del campeonato chileno pueden ser consultadas en dicho trabajo. En la siguiente sección desarrollamos un modelo para la asignación de árbitros de este torneo.

3. Modelo de Programación Entera

En base a la revisión de literatura, al análisis de asignaciones de árbitros en torneos anteriores y a conversaciones con la dirigencia de la ANFP y su Comisión Arbitral, hemos definido las condiciones que una programación de árbitros apropiada para este campeonato debe cumplir.

El modelo desarrollado toma ideas de diferentes trabajos pero es diferente a otros similares conocidos en la literatura.

El fixture del torneo se asume conocido de antemano. La asignación resultante debe establecer qué árbitro dirigirá cada uno de los partidos y debe estar disponible antes que el torneo comience. En la práctica, a medida que el torneo avanza, esta podría sufrir algunas modificaciones de acuerdo a situaciones coyunturales que, por ejemplo, afecten la disponibilidad de ciertos árbitros.

Cabe mencionar que en el fútbol se necesita un árbitro y dos asistentes. Al inicio de cada temporada, las ligas generalmente definen ternas arbitrales cuya composición no varía durante el torneo. En lo que sigue, hablaremos entonces de "árbitro" en forma singular, entendiendo que la asignación puede también corresponder a un árbitro y a los dos asistentes que completen su terna arbitral.

El problema lo abordamos mediante un modelo de optimización lineal entera, que presentamos a continuación.

- **Conjuntos y Parámetros**

  P: Conjunto de partidos.

  A: Conjunto de árbitros.

  E: Conjunto de equipos.

  F: Conjunto de fechas.

  $SIFIJO(a,p)$: Conjunto de pares (a,p) tal que se fija de antemano que el árbitro a debe dirigir el partido p.

  $NOFIJO(a,p)$: Conjunto de pares (a,p) tal que se fija de antemano que el árbitro a no puede dirigir el partido p.

  $PN(p)$: Subconjunto de partidos que requieren la máxima categoría de árbitro.

  C^a: Categoría del árbitro a.

  N^p: Mínima categoría de árbitro requerida para dirigir el partido p.

  $$fechas^{p,f} = \begin{cases} 1 & \text{si el partido } p \text{ se juega en la fecha } f \\ 0 & \sim \end{cases}$$

$$juega^{p,e} = \begin{cases} 1 & \text{si el equipo } e \text{ juega en el partido } p \\ 0 & \sim \end{cases}$$

$DIST^{a,p}$: distancia (ida y vuelta) entre la localidad considerada como origen para el árbitro a y la localidad de realización del partido p, medida en kilómetros.

T^a: Meta de partidos a ser asignados para dirigir por el árbitro a.

$MINT^a$: Mínima cantidad total de partidos a ser asignados al árbitro a.

$MAXT^a$: Máxima cantidad total de partidos a ser asignados al árbitro a.

$MINP^{a,e}$: Mínima cantidad de incidencias entre el árbitro a y el equipo e.

$MAXP^{a,e}$: Máxima cantidad de incidencias entre el árbitro a y el equipo e.

$DISTMAX$: Máxima diferencia permitida entre las distancias promedio por partido a recorrer por un árbitro y otro.

S^a: Máxima cantidad de fechas consecutivas que pueden pasar sin que el árbitro a sea asignado a dirigir un partido.

- **Variables**

$$x_{a,p} = \begin{cases} 1 & \text{si el árbitro } a \text{ es asignado para dirigir el partido } p \\ 0 & \sim \end{cases}$$

dif_a = Diferencia entre la meta predefinida de partidos a dirigir por el árbitro a y la cantidad de partidos efectivamente asignados.

- **Restricciones y Función Objetivo**

Restricciones básicas. Cada partido necesita un y sólo un árbitro principal para ser dirigido.

$$\sum_{a=1}^{|A|} x_{a,p} = 1 \qquad \forall\, p \in P. \tag{1}$$

Restricciones de fecha por árbitro. Cada árbitro puede ser asignado como máximo una sola vez por fecha.

$$\sum_{p=1}^{|P|} fechas^{p,f} \cdot x_{a,p} \leq 1 \qquad \forall\, a \in A,\ f \in F. \tag{2}$$

Categoría de árbitros y nivel de partidos. En base a la rivalidad histórica o a la expectativa que generan, algunos partidos requieren ser

dirigidos por árbitros de mayor experiencia o categoría. Para esta restricción, se establecen dentro de un mismo rango, números naturales que representan niveles de los partidos y categorías de los árbitros. En este rango, el número más bajo representa mejor nivel o mayor categoría, mientras que el número más alto representa peor nivel o menor categoría. El número de niveles y categorías a utilizar depende de cómo se defina la instancia. Agregamos entonces la siguiente restricción:

$$C^a \cdot x_{a,p} \leq N^p \qquad \forall\ a \in A,\ p \in P. \tag{3}$$

Además, dependiendo del número de partidos clasificados dentro del mejor nivel y de la frecuencia con que estén programados en una determinada cantidad consecutiva de fechas, se requiere que no se repitan los mismos árbitros en dos de este tipo de partidos consecutivos, lo que se logra con la siguiente restricción:

$$x_{a,p} + x_{a,\hat{p}} \leq 1 \qquad \forall\ a \in A,\ p, \hat{p} \in PN(p). \tag{4}$$

($\hat{p}$ es el siguiente partido de mayor nivel, después de p.)

Partidos fijos. Un árbitro puede ser sancionado e impedírsele dirigir algún partido o puede no estar disponible para dirigir en cierta(s) fecha(s), por ejemplo, debido a una lesión. Para esto agregamos la siguiente restricción:

$$x_{a,p} = 0 \qquad \forall\ (a,p) \in NOFIJO. \tag{5}$$

También la Comisión Arbitral puede decidir que cierto árbitro deba dirigir determinado partido. Incorporamos esta condición como sigue:

$$x_{a,p} = 1 \qquad \forall\ (a,p) \in SIFIJO. \tag{6}$$

Restricciones de equidad sobre los equipos. Consideramos un número mínimo y máximo de incidencias entre los árbitros y los equipos.

$$\sum_{p=1}^{|P|} juega^{p,e} \cdot x_{a,p} \geq MINP^{a,e} \qquad \forall\ a \in A,\ e \in E. \tag{7}$$

$$\sum_{p=1}^{|P|} juega^{p,e} \cdot x_{a,p} \leq MAXP^{a,e} \qquad \forall\ a \in A,\ e \in E. \tag{8}$$

Restricciones de equidad sobre los árbitros. Imponemos un número mínimo y máximo de asignaciones para un mismo árbitro durante todo el campeonato.

$$\sum_{p=1}^{|P|} x_{a,p} \geq MINT^a \qquad \forall\, a \in A. \tag{9}$$

$$\sum_{p=1}^{|P|} x_{a,p} \leq MAXT^a \qquad \forall\, a \in A. \tag{10}$$

Además, el promedio de distancias recorridas por partido por los árbitros durante el campeonato deben ser similares y pueden estar acotadas.[1]

$$\frac{1}{T^a}\sum_{p=1}^{|P|} DIST^{a,p}\cdot x_{a,p} - \frac{1}{T^r}\sum_{p=1}^{|P|} DIST^{r,p}\cdot x_{r,p} \leq DISTMAX \quad \forall\, a, r \in A. \tag{11}$$

También consideramos una cantidad máxima de fechas en que los árbitros pueden permanecer sin dirigir.

$$\sum_{s=0}^{S^a}\sum_{p=1}^{|P|} fechas^{p,f+s}\cdot x_{a,p} \geq 1 \qquad \forall\, a \in A,\ f \leq |F| - S^a. \tag{12}$$

Restricciones lógicas y función objetivo. En general, a los árbitros se les paga por partidos dirigidos. Por esta razón, para cada árbitro se predefine una cantidad de partidos a dirigir como meta, que depende principalmente de su categoría. Tal como presentan Duarte et al. [5] en el RAP, nuestra función objetivo establece que la suma sobre todos los árbitros de la diferencia absoluta entre su meta y el número de partidos efectivamente asignados debe ser minimizada. Para lograr esto, incorporamos las siguientes restricciones:

$$\sum_{p=1}^{|P|} x_{a,p} + dif_a \geq T^a \qquad \forall\, a \in A. \tag{13}$$

$$\sum_{p=1}^{|P|} x_{a,p} - dif_a \leq T^a \qquad \forall\, a \in A. \tag{14}$$

[1] Esta desigualdad considera la distancia promedio por partido a recorrer por árbitro, siempre y cuando cada árbitro dirija un número de partidos igual a su meta, lo que no está asegurado *a priori*. Sin embargo, en la función objetivo intentaremos lograr que esto se cumpla, por lo cual consideramos que es una buena estimación con la que no se pierde la linealidad del problema.

Finalmente, presentamos la función objetivo según dicho criterio:

$$\text{mín } g = \sum_{a=1}^{|A|} dif_a \tag{15}$$

Nos referiremos a este modelo de asignación de árbitros al torneo chileno como MATCH.

4. El enfoque de patrones

En términos computacionales, la resolución de la formulación anterior puede ser difícil, debido a la dimensión y a la estructura combinatorial del problema. Por ejemplo, para un campeonato de 6 equipos y 4 árbitros disponibles para dirigir cada fecha, en que los equipos juegan todos contra todos en dos rondas (por lo que 3 de los 4 árbitros deben dirigir cada fecha), existen más de 63 billones de asignaciones posibles. Para el torneo de Primera División A del fútbol chileno, que actualmente cuenta con 18 equipos, 34 fechas de fase regular por año y alrededor de 15 árbitros, el número de alternativas posibles se hace prácticamente inimaginable.

El carácter combinatorial y el tamaño de los problemas del mundo real, suelen ser las principales dificultades en los problemas de *Sports Scheduling*. En la programación de fixtures, un enfoque que ha sido amplia y exitosamente utilizado en la literatura es el de generar estructuras que definen secuencias de localías para cada equipo, denominadas *patrones* (ver, por ejemplo, [1], [2], [3], [9], [10]). Luego de que éstos han sido construidos, son fijados a los equipos, para posteriormente definir el fixture propiamente tal.

Esta técnica suele disminuir significativamente el tiempo de resolución y conlleva a soluciones que, si bien no necesariamente son óptimas, presentan buen desempeño en la F.O. (posteriormente, procedimientos de búsqueda local pueden ser utilizados para mejorar su calidad, o procedimientos exactos pueden usar esta solución como punto de partida).

Una buena revisión al respecto y un resumen sobre las variadas maneras en que el enfoque de patrones es utilizado en la resolución de problemas de programación de fixtures de torneos deportivos, es entregada por Rasmussen y Trick [11].

Nos remitimos a explicar brevemente el enfoque en este contexto, tanto en la programación de fixtures como en la asignación de árbitros. En el primer caso, existen muchos trabajos en la literatura que abordan esta técnica. En el segundo caso, entendemos que es una novedad de este trabajo.

4.1. El enfoque de patrones en la programación de fixtures

Un patrón puede ser entendido como un arreglo ordenado de caracteres L y V, que denotan "Local" y "Visita", respectivamente. La dimensión del arreglo corresponde al número de fechas del torneo. Un patrón asignado a un equipo dado, indica en su componente n-ésima si dicho equipo juega de local o de visita en la fecha n.

$$P(\textit{equipo 1}) = (L, V, L, V, V)$$

Por ejemplo, el patrón P indica que el *equipo 1* juega de local la primera y la tercera fecha, y de visita las fechas 2, 4 y 5.

En la programación de fixtures, la literatura reporta distintas estrategias para generar estos arreglos. Por ejemplo, pueden ser construidos mediante reglas lógicas o mediante modelos de programación entera. Independiente de como sean construidos, su uso generalmente confluye en que son fijados a los equipos asegurando que las restricciones de secuencias de localías y otras condiciones particulares de cada problema se cumplan, para posteriormente correr el modelo original a partir de esta fijación (lo que permite eliminar una serie de restricciones) y obtener factibilidad en el resto de las restricciones. Las restricciones que se suele asegurar en el primer paso son aquellas que en la práctica otorgan mayor dificultad a la resolución de la formulación original del problema.

En problemas del mundo real, la disminución del tiempo de resolución mediante el uso de patrones suele ser significativa. Aún más, en varios casos es crucial para obtener soluciones factibles, pues conseguirlas corriendo la formulación original se hace impracticable.

Inspirados en esta idea, desarrollamos para el problema de la asignación de árbitros un modelo en base a patrones. Según nuestro conocimiento, a pesar de que ambas problemáticas han sido un tema de estudio de la misma disciplina, la literatura no reporta la resolución de este tipo de problemas mediante un enfoque similar.

4.2. Un enfoque de patrones para la asignación de árbitros

Lógicamente, los conceptos de "local" o "visita" no aplican para el caso de los árbitros. La variable que hemos definido para generar un patrón para un árbitro es en qué zona del país dirige en cada fecha. Para ello, segmentamos el país en tres zonas: *Norte* (N), *Centro* (C) y *Sur* (S). Clasificamos a cada equipo en una de estas zonas, de acuerdo a la ubicación geográfica en que juega de local.

Además, dado que el número de árbitros es mayor al número de partidos jugados en cada fecha, definimos un valor *Libre* (L) que denota cuándo el árbitro no es asignado para dirigir partidos. Luego, definimos un patrón para

un árbitro como un arreglo ordenado de caracteres en $\{N,\ C,\ S,\ L\}$ y dimensión igual al número de fechas del torneo, por ejemplo:

$$Q(\textit{árbitro 1}) = (C, N, S, N, L, C, S, S, N)$$

El patrón Q indica que el *árbitro 1* debe arbitrar en el Norte en las fechas 2, 4 y 9; en el Centro, en las fechas 1 y 6; en el Sur, en las fechas 3, 7 y 8; y queda libre en la quinta fecha.

Enfatizamos que la segmentación de equipos puede ser hecha en base a criterios distintos del geográfico. Dada las particularidades del territorio chileno, hemos clasificado a los equipos de esta manera, pero en otros casos podría ser incluso hecha en forma aleatoria.

Dadas estas definiciones, desarrollamos un nueva metodología de resolución. Primero, formulamos un modelo que genera los patrones para cada árbitro, incorporando algunas de las restricciones del problema (las que *a priori* nos parecen más relevantes para efectos de la definición del conjunto de patrones). Luego, formulamos un segundo modelo que incorpora el resto de las restricciones y asigna definitivamente qué árbitro dirigirá cada partido.

4.2.1. Modelo para la generación de patrones

En este primer modelo, introducimos una familia de variables para construir los patrones y seleccionamos algunas de las restricciones del modelo original, capturándolas completa o parcialmente en la modelación.

Los conjuntos y parámetros que utilizamos del modelo anterior siguen teniendo el mismo significado y definimos otros adicionales.

- **Conjuntos y Parámetros adicionales**

  $K = \{N, C, S, L\}$: Conjunto de caracteres que denotan el cluster geográfico en que debe dirigir un árbitro o si queda libre.

  $PCluster(k, f)$: Número de partidos que se juegan en el cluster k durante la fecha f.

  $$AC_{a,Cat} = \begin{cases} 1 & \text{si el árbitro } a \text{ es de categoría } Cat \\ 0 & \sim \end{cases}$$

  (donde $Cat \in \{Alta, Media, Baja\}$).

  $PFC(k, f, Cat)$: Número de partidos de categoría Cat que se juegan en el cluster k durante la fecha k (donde $Cat \in \{Alta, Media, Baja\}$.)

  PAC: Conjunto de tuplas (a, k_1, k_2, f_1, f_2) tal que el árbitro a es de categoría *Alta*, en el cluster k_1 se juegan partidos de nivel *Alto* en la fecha f_1 y en el cluster k_2 se juegan partidos de nivel *Alto* en la fecha f_2 (siendo f_2 la primera fecha posterior a la fecha f_1 tal que hay partidos de nivel *Alto*).

$NOFIJO\_C$: Conjunto de tuplas (a, k, f) tal que el árbitro a no puede arbitrar en el cluster k en la fecha f.

$SIFIJO\_C$: Conjunto de tuplas (a, k, f) tal que el árbitro a debe arbitrar en el cluster k en la fecha f.

$$KJuega(e, k, f) = \begin{cases} 1 & \text{si el equipo } e \text{ juega en el cluster } k \text{ en la fecha } f \\ 0 & \sim \end{cases}$$

$DClus(a, k)$: Distancia entre el lugar de origen del árbitro a y el promedio de la distancia de los equipos que juegan de local en el cluster k.

- **Variables**

$$z_{a,k,f} = \begin{cases} 1 & \text{si el patrón del árbitro } a \text{ indica que debe dirigir en el} \\ & \text{cluster } k \text{ en la fecha } f \\ 0 & \sim \end{cases}$$

dif_a = Diferencia entre la meta predefinida de partidos a dirigir por el árbitro a y la cantidad de partidos efectivamente asignados.

- **Restricciones y Función Objetivo**

Restricciones básicas sobre los patrones y el fixture. La suma del número de patrones que indican jugar en un cluster k en la fecha f debe ser igual a la suma de partidos que según el fixture se juegan en el cluster k en la fecha f.

$$\sum_a z_{a,k,f} = PCluster(k, f) \qquad \forall\ k \in K,\ f \in F. \tag{16}$$

Esta familia de restricciones es similar a las restricciones (1) del modelo MATCH, ahora en el contexto de las variables z.

Restricciones de fecha por árbitro. En toda fecha, un árbitro debe ser asignado a dirigir en algún cluster geográfico o quedar libre.

$$\sum_{k \in K} z_{a,k,f} = 1 \qquad \forall\ a \in A,\ f \in F. \tag{17}$$

Esta familia de restricciones es análoga a las restricciones (2) del modelo MATCH.

Categoría de partidos y nivel de partidos. Intentando capturar las restricciones (3) del modelo MATCH, imponemos que la suma del número de patrones asignados a los árbitros de categoría *Alta* que arbitran en el cluster k en la fecha f correspondan al número de partidos que demandan esta categoría de árbitros jugados en el cluster k en la fecha f.

$$\sum_{a} AC_{a,Alta} \cdot z_{a,k,f} \geq PFC(k,f,Alta) \qquad \forall\, k \in K,\ f \in F. \tag{18}$$

Del mismo modo, imponemos esta condición para los partidos que demandan árbitros de calidad al menos *Media* (lógicamente, un árbitro de categoría *Alta* también puede arbitrar estos partidos).

$$\sum_{a} (AC_{a,Alta} + AC_{a,Media}) \cdot z_{a,k,f} \geq PFC(k,f,Alta) + PFC(k,f,Media)$$
$$\forall\, k \in K,\ f \in F. \tag{19}$$

Adicionalmente, para asegurar que se cumplan las restricciones (4) del modelo MATCH, imponemos que si un árbitro dirigió un partido de *Alta* categoría en una fecha dada, en la fecha siguiente en que hay partidos de nivel *Alto* no dirija en el cluster en que estos se juegan.

$$z_{a,k_1,f_1} + z_{a,k_2,f_2} \leq 1 \qquad \forall\, (a,k_1,k_2,f_1,f_2) \in PAC. \tag{20}$$

Partidos fijos. Intentamos capturar las restricciones (5) del MATCH, imponiendo que el árbitro no dirija en el cluster correspondiente:

$$z_{a,k,f} = 0 \qquad \forall\, (a,k,f) \in NOFIJO\_C. \tag{21}$$

Análogamente, para las restricciones (6), imponemos que el árbitro dirija en el cluster al que pertenece el equipo que juega de local:

$$z_{a,k,f} = 1 \qquad \forall\, (a,k,f) \in SIFIJO\_C. \tag{22}$$

Restricciones de equidad sobre los equipos. Intentando capturar parcialmente las incidencias entre árbitros y equipos consideradas en las restricciones (7) del modelo MATCH, imponemos una cota mínima al número de veces que cada patrón es asignado al cluster donde juega cada equipo.

$$\sum_{k,f} z_{a,k,f} \cdot KJuega(e,k,f) \geq MINP^{a,e} \qquad \forall\, a \in A,\ e \in E. \tag{23}$$

Restricciones de equidad sobre los árbitros. Tal como impusimos las restricciones (9) y (10) en el modelo MATCH, fijamos cotas para el

número mínimo y máximo de partidos que cada árbitro debe dirigir en el torneo.

$$\sum_{f,k} z_{a,k,f} \geq MINT^a \qquad \forall\, a \in A,\ k \in \{N, C, S\}. \tag{24}$$

$$\sum_{f,k} z_{a,k,f} \leq MAXT^a \qquad \forall\, a \in A,\ k \in \{N, C, S\}. \tag{25}$$

Intentando capturar el balance sobre las distancias promedio viajadas por cada árbitros, incorporamos la siguiente condición, similar a las restricciones (11) del MATCH:

$$\frac{1}{T^a} \sum_{f,k} DClus(a,k) \cdot z_{a,k,f} - \frac{1}{T^r} \sum_{f,k} DClus(r,k) \cdot z_{r,k,f} \leq DISTMAX$$

$$\forall\, a, r \in A. \tag{26}$$

Al igual que en las restricciones (12), consideramos que todo patrón respeta el número de fechas libres consecutivas que puede pasar un árbitro sin dirigir.

$$\sum_{s=0}^{S^a} z_{a,Libre,f+s} \leq S^a \qquad \forall\, a \in A,\ f < |F| - S^a. \tag{27}$$

Restricciones lógicas y función objetivo. Calculamos la variable dif_a análogamente a como en las restricciones (13) y (14) del modelo MATCH, esta vez sumando las variables z en las regiones geográficas en vez de las variables x.

$$\sum_{f} \sum_{k \neq L} z_{a,k,f} + dif_a \geq T^a \qquad \forall\, a \in A. \tag{28}$$

$$\sum_{f} \sum_{k \neq L} z_{a,k,f} - dif_a \leq T^a \qquad \forall\, a \in A. \tag{29}$$

Finalmente, la función objetivo sigue siendo la misma que en el MATCH.

$$\text{mín}\, g = \sum_{a=1}^{|A|} dif_a \tag{30}$$

Nos referiremos a este modelo que genera los patrones como GP_MATCH.

4.2.2. Modelo de asignación en base a patrones

Luego de haber generado los patrones según el GP_MATCH, procedemos a realizar la asignación de árbitros a partidos según el modelo de optimización lineal entera que exponemos a continuación.

Al igual que en el modelo anterior, los parámetros y conjuntos que ya hemos definido siguen teniendo el mismo significado.

- **Conjuntos y Parámetros adicionales**

  Q: Conjunto de patrones.

  $PKF(k, f)$: Conjunto de partidos que se juegan en el cluster k en la fecha f.

  $TKF(k, f)$: Conjunto de patrones que asignan dirigir en el cluster k en la fecha f.

- **Variables**

  $$x_{a,p} = \begin{cases} 1 & \text{si el árbitro } a \text{ es asignado para dirigir el partido } p \\ 0 & \sim \end{cases}$$

  $$y_{a,w} = \begin{cases} 1 & \text{si el patrón } w \text{ es asignado al árbitro } a \\ 0 & \sim \end{cases}$$

  dif_a = Diferencia entre la meta predefinida de partidos a dirigir por el árbitro a y la cantidad de partidos efectivamente asignados.

- **Restricciones y Función Objetivo**

  Restricciones sobre los patrones y su relación lógica con las variables x**.** A cada árbitro se le asigna un patrón:

  $$\sum_{q \in Q} y_{a,q} = 1 \qquad \forall\, a \in A. \tag{31}$$

  En la práctica, fijamos el valor de estas variables y según la generación de patrones que realizamos en el GP_MATCH.

  Lógicamente, imponemos una condición que relacione las variables x e y:

  $$\sum_{p \in PKF(k,f)} x_{a,p} - \sum_{q \in TKF(k,f)} y_{a,q} = 0 \qquad \forall\, a \in A,\ k \in K,\ f \in F. \tag{32}$$

Adicionalmente, en este modelo consideramos en forma explícita las restricciones del MATCH, que no necesariamente hemos asegurado según la generación de patrones realizada en el GP_MATCH: (1), (3), (6), (7), (8), (11).

Finalmente, mantenemos la función objetivo (15) del MATCH. Nos referiremos a este modelo como AP_MATCH.

5. Resultados

Utilizamos como instancia la fase regular del torneo de fútbol chileno de Primera A del año 2007. Esta instancia cuenta con 21 equipos (uno queda libre cada fecha), 16 árbitros y 420 partidos (42 fechas -incluyendo Apertura y Clausura-, en cada una de las cuales se juegan 10 partidos). El MATCH de esta instancia contiene alrededor de 6.700 variables y 10.000 restricciones.

Implementamos los modelos de optimización en GAMS 22.7 y para su resolución utilizamos el software comercial Cplex 11.0, en un computador de procesador Intel Pentium de 1.86GHz y 2GB de RAM.

A continuación discutimos los resultados obtenidos.

5.1. Características de la solución

El problema fue resuelto a optimalidad, utilizando el enfoque de resolución sin patrones, alcanzando un valor igual a cero en la función objetivo. Esto significa que la solución cumple con que todos los árbitros dirigen su meta de partidos. Además, la solución presenta varias otras bondades.

Primero, imprime un equilibrio relativo en la distancia recorrida por los árbitros. Considerando que en un escenario perfectamente balanceado el promedio de partidos dirigidos por árbitro es 26,25 (calculado como $|P|/|A|$), cada árbitro recorrería una distancia promedio por partido de 847,8 km. En tanto, el promedio de partidos dirigidos por cada árbitro a cada equipo sería igual a 2,5. Los parámetros correspondientes (cotas superiores e inferiores de las variables que se quieren balancear) fueron entonces fijados de manera de acercarse lo mayor posible a estas cifras. Al comparar los valores presentes en la asignación real del año 2007 (que la Comisión Arbitral de la ANFP confeccionó en forma tradicional), con los que se consiguieron como resultado del modelo se observan mejoras en todo ellos, según se muestra para la instancia de resultados de la Tabla 1.

Ítems	Asignación 2007	Resultado Modelo
Mínima cantidad de partidos que dirige en total un árbitro	24	26
Máxima cantidad de partidos que dirige en total un árbitro	36	28
Mínima cantidad de partidos que dirige en total un árbitro a un mismo equipo	0	1
Máxima cantidad de partidos que dirige en total un árbitro a un mismo equipo	7	4
Mínima distancia por partido promedio recorrida por un árbitro	308	571
Máxima distancia por partido promedio recorrida por un árbitro	1.192	1.002
Número de veces que un árbitro fue asignado a dirigir dos partidos consecutivos del nivel más alto	1	0
Máxima cantidad de fechas transcurridas en que un árbitro no es asignado para dirigir	4	2

Tabla 1: Comparación de los distintos ítems entre la asignación de árbitros real del 2007 y el resultado del modelo de optimización.

Para las distancias promedio recorridas por árbitros, la asignación del año 2007 presenta una desviación estándar de 268,96, mientras que la asignación resultante del modelo una de 128,02.

De igual manera, para la cantidad total de partidos dirigidos por árbitro, la desviación estándar disminuyó de 2,96 en la asignación real del 2007 a 0,58 en el resultado del modelo. Notar que el árbitro que más dirige en nuestra asignación lo hace en 28 partidos, mientras que el que menos dirige lo hace en 26 oportunidades, mientras que en la asignación del 2007 esos números son 36 y 24, respectivamente.

A su vez, el año 2007 la varianza de la cantidad de asignaciones por árbitro a equipos fue de 2,64, mientras que en la instancia reportada en la Tabla 1, el modelo obtuvo una asignación con 1,32 de varianza, es decir un 50 % menor. Este es un tema particularmente sensible para la prensa y los simpatizantes de los clubes por lo que es bueno equilibrarlo. En la instancia real, los valores mínimo y máximo de la cantidad de veces que dirige un árbitro a un mismo equipo son 0 y 7, respectivamente, mientras que en nuestra propuesta esos valores se llevan a 1 y 4. En otra de las instancias que resolvimos (en la que también se obtuvo F.O. igual a 0), redujimos la brecha entre la mínima y la máxima cantidad de veces que un árbitro le dirige a un mismo equipo al mínimo posible (2 y 3 respectivamente), alcanzando una varianza de 0,25.

5.2. Tiempo de resolución

A los efectos de comparar los tiempos de resolución del modelo original y el modelo con patrones, generamos 4 nuevas instancias del problema modificando algunos de los valores de los parámetros. En todos los casos ambos modelos alcanzaron el valor óptimo. La Tabla 2 muestra los tiempos obtenidos para dichas instancias.

Inst.	T_{MATCH}	T_{GP_MATCH}	T_{AP_MATCH}	$T_{GP\_MATCH} + T_{AP\_MATCH}$	Δ %
1	173	2	127	129	-25,4 %
2	625	4	353	357	-42,9 %
3	280	5	93	98	-65,0 %
4	243	4	100	104	-57,2 %

Tabla 2: Tiempos de resolución (medido en segundos).

La segunda columna muestra el tiempo en que resolvemos el MATCH y la quinta columna (suma de la tercera y la cuarta) muestra el tiempo total en que resolvemos el problema mediante el enfoque de patrones. Por último, en la sexta columna, se calcula la reducción de tiempos. En promedio, redujimos el tiempo de resolución en aproximadamente 48 %.

6. Conclusiones

Este artículo contribuye en dos líneas. Primero, hemos desarrollado un modelo de optimización lineal entera para la asignación de árbitros a los partidos del torneo de la Primera División A del fútbol chileno. El modelo combina algunas de las ideas presentadas en la literatura con conceptos aplicados por la Comisión Arbitral del fútbol chileno.

La metodología que tradicionalmente ha utilizado la ANFP para realizar esta tarea carece de herramientas sofisticadas y de criterios objetivamente estructurados. Semana tras semana, la Comisión Arbitral designa los árbitros que dirigirán cada partido en forma manual. Analizando las asignaciones implementadas en los últimos torneos, constatamos varias desventajas: grandes diferencias relativas entre las distancias promedio de viaje de los árbitros, desbalance en el número de partidos totales que arbitra cada uno, frecuencias indeseables de asignación de un mismo árbitro a un mismo equipo (en algunos casos, una cantidad relativamente alta de partidos; en otros casos, nula). Nuestro modelo mejora significativamente todos estos aspectos, logrando una asignación mucho más equitativa, tanto para los árbitros como para los equipos. Además, facilita la tarea de asignación y la hace más transparente, al fijar criterios de decisión claros.

Junto con dichas mejoras, el modelo también presenta la ventaja de poder generar rápidamente varias alternativas de asignaciones, que cumplen con todas las condiciones (aun más, para las instancias estudiadas, generamos más de una solución que conduce al valor óptimo en la F.O.). En la práctica, esto permitiría ofrecer una variedad de opciones a la ANFP para su elección final de cuál sería la asignación a implementar. Además, el desarrollo realizado de una interfaz amigable para el usuario permite que la herramienta sea manejada sin problemas directamente por la Comisión Arbitral.

Otra ventaja del modelo es que la solución entregada puede ser utilizada tanto para la asignación completa del campeonato, como para una determinada cantidad de fechas a partir del momento de ejecución. La restricción de fijación de variables permite modelarlo en distintos instantes del tiempo, incorporando las asignaciones que efectivamente han sido realizadas y modificando los parámetros que corresponda.

Por otro lado, el modelo presentado permitiría incorporar más de un campeonato en la instancia a resolver. Esto es interesante para ligas de más de una división que compartan árbitros entre ellas. Por ejemplo, en el caso de Chile, se podría resolver el campeonato de Primera A y Primera B en forma conjunta, modificando los parámetros de niveles de partidos y categorías de árbitros para permitir que ciertos árbitros que *a priori* están destinados para

dirigir partidos en Primera A, lo puedan hacer en la Primera B y viceversa.

Un tema interesante de explorar atañe a la localidad que es considerada como origen para los árbitros. Este es un parámetro del modelo y puede ser modificado si es deseable, por ejemplo, que un árbitro sea asignado a dirigir dos partidos muy cercanos en fecha y lejanos de la localidad origen, de manera de aprovechar el largo viaje que tendrá que hacer. Esto podría conseguirse considerando en la función objetivo del modelo las distancias o tiempos totales a recorrer por los árbitros.

La segunda contribución de este artículo es el desarrollo de un enfoque de resolución basado en patrones. Un enfoque similar ha sido amplia y exitosamente utilizado en la programación de fixtures, otro de los problemas que aborda el *Sports Scheduling*. En dicho caso, los patrones definen secuencias de localías y visitas para los equipos que, generadas y fijadas bajo ciertos criterios, facilitan enormemente la resolución de los problemas de tamaño real. Aún más, el enfoque de patrones suele ser crucial para encontrar soluciones factibles en esos problemas. En nuestro caso, construimos patrones que definen para cada árbitro y cada fecha del torneo, la zona geográfica en que dirige o si queda libre. Implementamos el enfoque mediante un modelo de programación lineal entera de dos fases: la primera, que genera los patrones y la segunda, que los usa para encontrar la asignación de árbitros a los partidos del torneo. El tiempo de resolución que logramos usando este enfoque, reduce notablemente el tiempo que toma la resolución de la formulación original del problema.

Según nuestro conocimiento, el presente artículo es el primero en proponer un enfoque de patrones para la resolución de un problema de asignación de árbitros a un torneo deportivo.

Actualmente, la ANFP está evaluando utilizar el modelo y la técnica de resolución presentados en este artículo, para realizar la asignación de árbitros a los torneos del fútbol profesional chileno.

Agradecimientos: A Salvador Imperatore, Harold Mayne-Nicholls y Carlos Morales, de la ANFP, por su colaboración para la concreción de este proyecto. El segundo autor es parcialmente financiado por el proyecto FONDECyT 1080286 y por el Instituto Milenio “Sistemas Complejos de Ingeniería”.

Referencias

[1] Bartsch T., Drexl A., Kröger S. Scheduling the professional soccer leagues of Austria and Germany. Computers and Operations Research (2006) 33 (7), 1907-1937.

[2] Cain Jr W.O. The computer-assisted heuristic approach used to schedule the major league baseball clubs. In: Ladany SP, Machol RE., editors. Optimal strategies in sports. Amsterdam: North-Holland (1977), 33-41.

[3] De Werra D., Jacot-Descombes L., Masson P. A constrained sports scheduling problem. Discrete Applied Mathematics (1990) 26 (1), 41-49.

[4] Dinitz J.H., Stinson D.R. On assigning referees to tournament schedules. Bulletin of the Institute of Combinatorics and its Applications (2005) 44, 22-28.

[5] Duarte A., Ribeiro C., Urrutia S. Referee Assignment in Sport Tournaments. Lecture Notes in Computer Science (2007) 3867, 158-173.

[6] Durán G., Guajardo M., Miranda J., Sauré D., Souyris S., Weintraub A., Wolf R. Scheduling the Chilean Soccer League by Integer Programming. Interfaces (2007) 37 (6), 539-552.

[7] Easton K., Nemhauser G., Trick M. The traveling tournament problem: description and benchmarks. In Proceedings of the 7th. International Conference on Principles and Practice of Constraint Programming, Paphos (2001), 580-584.

[8] Gil La Fuente J., Rojas Mora J. La idónea asignación arbitral con altos niveles de incertidumbre. Empresa global y mercados locales: XXI Congreso Anual AEDEM (2007) 1, 69-80.

[9] Goossens D., Spieksma F. Scheduling the Belgian Soccer League. Interfaces (2009) 39 (2), 109-118.

[10] Nemhauser G.L., Trick M.A. Scheduling a major college basketball conference. Operations Research (1998) 46, 1-8.

[11] Rasmussen R.V., Trick M.A. Round robin scheduling - a survey. European Journal of Operational Research (2008) 188, 617-636.

[12] Trick M.A., Yildiz H. The Traveling Umpire Problem. Informs Annual Conference, Pittsburgh, Noviembre 2006.

[13] Yavuz M., Inan U.H., Figlali A. Fair referee assignments for professional football leagues. Computers & Operations Research (2008) 35 (9), 2937-2951.

[14] Wright M.B. Scheduling English cricket umpires. Journal of the Operational Research Society (1991) 42 (6), 447-452.

Programas de Postgrado y Postítulos Impartidos por el DII

DOCTORADO

DOCTORADO EN SISTEMAS DE INGENIERIA
UNIVERSIDAD DE CHILE

La carencia de capital humano avanzado es una preocupación nacional. Por ello, la Universidad de Chile ha creado el Doctorado en Sistemas de Ingeniería (DSI), el cual forma profesionales de excelencia en el área de sistemas de ingeniería. Su foco es el desarrollo de habilidades de resolución de problemas con técnicas avanzadas y metodologías multidisciplinarias.

Aportando en esta línea, el Instituto Sistemas Complejos de Ingeniería (ISCI) entrega al DSI conocimientos en materias de ingeniería industrial, transporte, matemáticas y eléctrica, además de la experiencia de académicos de primer nivel.

El DSI es impartido por los Departamento de Ingeniería Industrial, Ingeniería Matemática, Ingeniería Eléctrica y la división Transporte de Ingeniería Civil de la Facultad de Ciencias Físicas y Matemáticas (FCFM) de la Universidad de Chile y está acreditado por la Comisión Nacional de Acreditación de Posgrados.

Más información en www.sistemasdeingenieria.cl/doctorado/

MAGÍSTERES

Magíster en Gestión de Operaciones MGO

El Magíster en Gestión de Operaciones (MGO) busca formar profesionales de excelencia en esquemas de gestión, uso de modelos y tecnologías de información, con capacidad de resolver problemas complejos en gestión de operaciones.

El Magíster en Gestión de Operaciones es impartido por el Centro de Gestión de Operaciones (CGO) del Departamento de Ingeniería Industrial de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile, que recoge la vasta experiencia de sus integrantes

en las áreas forestal, minera, de servicios y manufactura, tanto a través de proyectos de investigación como de consultoría.

Los egresados del Magíster en Gestión de Operaciones se desempeñan en cargos de primer nivel en empresas de servicios y manufactura nacionales e internacionales. También trabajan en universidades como académicos e investigadores, y algunos siguen doctorados en prestigiosas universidades como MIT y Columbia.

Requisitos de Admisión

Poseer el grado de licenciado en Ciencias de la Ingeniería o su equivalente.

Calendario de Postulaciones

Semestre Otoño
Periodo de postulaciones: de octubre a 15 de diciembre.
Inicio de clases: marzo de cada año.

Semestre Primavera
Período de postulaciones: de abril a 15 de junio.
Inicio de clases: julio de cada año.

Mayor información: julie@dii.uchile.cl | www.dii.uchile.cl/mgo | Tel: (56 2) 978 4017 | (56 2) 978 4073

Magíster en Economía Aplicada MAGCEA

El Magíster en Economía Aplicada (MAGCEA) busca formar profesionales de gran competencia analítica y una sólida base en economía.

Es impartido por el Centro de Economía Aplicada (CEA) del Departamento de Ingeniería Industrial, de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile. El CEA es líder en investigación en economía en Chile y en el desarrollo de propuestas para las políticas públicas.

Muchos graduados del programa han seguido estudios de doctorado en universidades como Harvard, Stanford, MIT, Princeton y Yale, entre otras.

Nuestros egresados son altamente requeridos en el mercado laboral, se desempeñan en empresas líderes a nivel nacional e internacional, en destacados organismos estatales y en importantes organismos internacionales.

Requisitos de Admisión

Poseer un título profesional, nacional o extranjero, que exija al menos cinco años de estudios, o el grado de licenciado en campos disciplinarios afines a la especialidad.

Calendario de Postulaciones

Semestre Otoño
Período de postulaciones: de octubre a 15 de diciembre.
Inicio de clases: marzo de cada año.

Semestre Primavera
Período de postulaciones: de abril a 15 de junio.
Inicio de clases: julio de cada año.

Mayor información: magcea@dii.uchile.cl | www.magcea-uchile.cl | Tel: (56 2) 978 4084 | (56 2) 978 4073

Magíster en Ingeniería de Negocios con Tecnologías de Información MBE

El Magíster en Ingeniería de Negocios con Tecnologías de Información (MBE: Master in Business Engineering) tiene como objetivo formar a un tipo de profesional especialista, que integre gestión y tecnología en el diseño de los negocios y que tenga, además, las habilidades necesarias para iniciar y facilitar innovaciones en la empresa. Todas éstas, herramientas imprescindibles dentro del mundo emprendedor en el que se inserta.

Los egresados de este Magíster, por lo tanto, estarán en condiciones no sólo de generar, dirigir y ejecutar proyectos innovadores de transformación, sino además de rediseñar empresas tradicionales usando las TI y de desarrollar nuevas empresas basadas en tales tecnologías.

El programa incluye, como parte integral del mismo, el desarrollo de un proyecto de este tipo.

Requisitos de Admisión

Poseer un título profesional con al menos cinco años de estudio, o licenciatura en una disciplina afín a la especialidad. También, buenos antecedentes académicos y/o laborales.

Calendario de Postulaciones

Semestre Otoño
Hasta el 15 de enero de cada año.
Inicio de clases: marzo de cada año.

Semestre Primavera
Hasta el 15 de junio de cada año.
Inicio de clases: julio de cada año.

Mayor información: mbe@dii.uchile.cl | www.mbe-uchile.cl | Tel: (56 2) 978 4835 | (56 2) 978 4935

Magíster en Gestión **para la Globalización**

A partir del reconocimiento del factor humano de alta calificación como elemento diferenciador de los países exitosos, el programa se orienta a la formación de jóvenes profesionales chilenos con antecedentes de excelencia académica para desempeñarse eficazmente en la empresa global. Se persigue con ello contribuir a abordar los desafíos de capital humano y social que enfrenta el país en esta etapa de su desarrollo.

El MGPG, impartido por laUniversidad de Chile a través del Departamento de Ingeniería Industrial, tiene una duración de 19 meses con dedicación full time. 9 de ellos se realizan en el extranjero, incluyendo estudios en prestigiosas universidades de EE.UU., Australia e Inglaterra, y un study tour. Todos los estudiantes reciben la Beca Minera Escondida (operada por BHP Billiton), que cubre los principales gastos académicos y de manutención. El riguroso proce- so de selección busca promoverla excelencia y la diversidad.

Mayor información: www.magisterglobalizacion.cl

Magíster en Gestión y Políticas Públicas (**MGPP**)

El Magíster en Gestión y Políticas Públicas, tiene como propósito la formación avanzada de profesiona les interesados en la formulación y ejecución de políticas públicas.

El MGPP forma líderes y servidores públicos del más alto nivel, capaces de conceptualizar, pensar y discutir sus visiones e ideas sobre el futuro de América Latina.

Se imparte en dos modalidades: Diurna y Ejecutiva.

Características Distintivas

* Excelencia Académica
* Cuerpo docente de primer nivel
* Orientado a profesionales de formación diversa
* Alta tasa de graduación (83%)
* Reconocido entre los mejores en su área en América Latina.
* Acreditado por el CNA
* 15 años formando líderes

Requisitos de Admisión

* Poseer el grado de licenciado o título universitario

Versión en Horario Diurno:
Inicio: junio de cada año
Duración: 19 meses

Versión en Horario Ejecutivo:
Inicio: julio de cada año
Duración: 24 meses

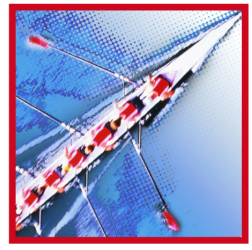
Postulaciones:
Hasta el **15 de octubre** para personas que postulan a becas de instituciones

Hasta el **15 de abril** para personas que cuentan con fondos propios

Mayor información: mgpp@dii.uchile.cl | www.mgpp.cl | Tel.: (562) 978 4067 | Fax 689 4987

MBA | MAGÍSTER EN GESTIÓN Y DIRECCIÓN DE EMPRESAS

El Magíster en Gestión y Dirección de Empresas (MBA) se orienta a formar directivos de nivel internacional, que jueguen un rol protagónico en la dirección o creación empresarial.



A través de la formación en gestión de negocios, habilidades directivas y tecnologías de apoyo a la gestión, este Magíster persigue apoyar el desarrollo de líderes innovadores, que sepan adaptarse a los cambios, detectar oportunidades y conducir en forma motivadora a toda la organización, en el mundo competitivo y global en que se desenvuelve la empresa actual.

El programa se imparte en modalidad de dedicación completa (full time), en conjunto con la Facultad de Economía y Negocios de la Universidad de Chile; y en modalidad de dedicación parcial.

Requisitos de Admisión

Grado de licenciado o título universitario que exija al menos cinco años de estudio; examen de admisión y entrevista personal; dos cartas de recomendación. Modalidad de medio tiempo: experiencia laboral mínima de dos años a jornada completa.

Calendario de Postulaciones

Hasta diciembre de cada año (primer cierre), para ingreso en marzo del año siguiente.
Modalidad de dedicación parcial: hasta junio de cada año para Ingenieros Comerciales e Industriales que puedan convalidar la fase de Nivelación, con ingreso en julio.

Mayor información: mba@dii.uchile.cl | www.mbauchile.cl | Tel: (56 2) 978 4002 | Fax: (56 2) 978 4080

EDUCACIÓN EJECUTIVA

INGENIERÍA INDUSTRIAL
UNIVERSIDAD DE CHILE

Claves de Diferenciación Profesional

DIPLOMAS DE POSTÍTULO 2010

- Gestión de Empresas
- Preparación y Evaluación de Proyectos
- Estrategias y Control de Gestión
- Marketing Decisional
- Gestión de Retail, un Enfoque Táctico
- Inteligencia de Negocios

CURSO DE ESPECIALIZACIÓN

- Finanzas y Administración de Riesgo

PROGRAMAS IN COMPANY

INFORMACIONES:
- Teléfono: 978 4002
- Email: diplomas@dii.uchile.cl
- Web: www.diplomas.uchile.cl