
CARACTERIZANDO LA FIJACIÓN OCULAR DEL USUARIO WEB EN LOS CONTENIDOS DE UNA PÁGINA: UNA APROXIMACIÓN BASADA EN TEORÍA DE GRAFOS

PABLO LOYOLA H. *
GUSTAVO MARTÍNEZ A. *
JUAN D. VELÁSQUEZ *

Resumen

En este artículo se presenta un marco de análisis, basado en la teoría de grafos, para estudiar la evolución temporal de la fijación ocular de los usuarios en un sitio web, con el objetivo de obtener patrones que permitan parametrizar su comportamiento de navegación y preferencias, con miras a la mejora continua de tanto la estructura y contenido de las páginas que visita. En ese sentido, se propone modelar los objetos web, tales como imágenes y párrafos, como nodos, siendo los arcos las transiciones visuales que realiza el usuario cuando mueve la fijación ocular desde un objeto a otro. A los grafos resultantes se les aplican métricas para establecer una noción de importancia entre su nodos. Esta metodología fue probada con datos extraídos desde 23 personas a quienes se les solicitó realizar una actividad de visualización de un sitio web, la cual fue parametrizada usando un dispositivo de rastreo del movimiento ocular (*eye tracker*). Los resultados sugieren que un análisis basado en teoría de grafos puede capturar de forma confiable la dinámica del comportamiento del usuario y la identificación de objetos importantes dentro de un sitio web.

Palabras Clave: Comportamiento usuario web, Teoría de Grafos, Patrones de comportamiento visual, Objetos web.

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

1. Introducción

La necesidad de entender la evolución del uso de la Web, ha llevado a incorporar nuevas fuentes de datos a las ya existentes en los análisis estándares. Como la Web ya no es un ecosistema estático, donde los usuarios sólo podían desarrollar tareas simples, la idea de considerar su uso como un fenómeno evolutivo y dependiente del tiempo nace como una alternativa viable para desentrañar los factores subyacentes que influyen en la toma de decisiones [15].

Originalmente las aproximaciones típicas han utilizado el concepto de las sesiones de usuario, las que son extraídas desde el archivo de registros de *weblogs* almacenados en el servidor web institucional. Por años esta propuesta ha sido la base de la investigación en *web mining*. A pesar de que los resultados han mostrado muchos casos de éxito, existen dos limitaciones principales que amenazan la confiabilidad y mejoras futuras [11, 14]: dificultad en la reconstrucción de la sesión y falta de información respecto de las preferencias de los usuarios.

Para reconstruir la sesión de un usuario en un sitio existen dos grandes conjuntos de técnicas: Proactivas y Reactivas [16]. Las primeras utilizan herramientas de rastreo en el lado del cliente, es decir, permiten conocer en forma directa qué página está siendo visitada por un usuario en particular. Lo anterior vulnera la privacidad del usuario y en general está prohibido en la mayoría de los países [20]. Por otro lado, propuestas reactivas utilizan *weblogs* y desarrollan un análisis posterior para intentar reconstruir la sesión del usuario, basándose en metodologías heurísticas que generalmente consideran tiempo, la topología del sitio web o métricas basadas en ontología para dividir eficientemente los *weblogs* en sesiones. Si bien estos métodos no vulneran la privacidad de los usuarios, tienen una efectividad más baja debido a factores como:

1. Las características de los navegadores web, que han contribuido a enriquecer la experiencia del visitante, dando al usuario opciones y herramientas como visualización de múltiples páginas, botón de regreso (back bottom) y navegación colaborativa, representando todo un desafío en términos de identificar sesiones puras para el análisis [17].
2. Anonimizadores. Se trata de aplicaciones que por su funcionamiento esconden la dirección IP original desde donde se realiza la petición de una página. Por ejemplo un firewall que posee una IP pública y por donde

pasa cualquier solicitud de la red interna de una institución. En este caso se realiza un cambio de la IP privada a IP pública, es decir, cualquier petición de una página que se origine desde un computador de la red interna saldrá con al misma IP pública. Lo mismo ocurre con los sistemas caché implementados con Proxy Server y con la asignación de IP dinámicas, por parte de los Internet Service Providers.

3. Almacenamiento de páginas en caché. Con el fin de asegurar una navegación más rápida para el usuario, los navegadores incorporan un área caché en el disco duro del computador cliente, en la cual se almacenan páginas y objetos previamente visitados, de forma que si el usuario vuelve a consultar la página ya visitada, no se realiza una petición al servidor web, pues primero se consulta al caché. La petición anterior no queda registrada en los registros de weblog, por lo que la reconstrucción de la sesión se dificulta.
4. Mobile: El explosivo incremento en el uso de nuevas plataformas como smartphones y tablets ha forzado a desarrollar nuevas soluciones adaptivas dentro del proceso de desarrollo web. Este dinamismo muchas veces dificulta obtener un registro respecto de qué contenido o qué página está visitando el usuario

Dado lo anterior, usar sólo como fuente de datos los registros de *weblogs*, no necesariamente reflejaría el comportamiento del usuario web [22], debido principalmente a que sólo muestra mayormente qué páginas se visitaron, y por cuánto tiempo, sin dar cuenta el motivo por el cual se tomaron esas decisiones al hacer clic, o qué objeto dentro de la página afectó la navegación del usuario (tanto de duración en la página como de decisión).

Para poder comprender de mejor manera el comportamiento del usuario es necesario analizar qué componentes de cada página detonan las decisiones, o llaman su atención. Es claro que dichas componentes deben estar correlacionados directamente con el flujo de clics posterior [15].

Las sesiones web proveen una estimación de los intereses desde la base de visitas entre páginas, pero no entregan ninguna percepción acerca de las acciones y preferencias de contenido del usuario web [18]. Dado el dinamismo y complejidad de las actuales aplicaciones web, se considera que este problema es altamente relevante y debiera ser tomado en consideración. Por ejemplo, muchos sitios son aplicaciones web de una sola página, que han utilizado el poder los frameworks para javascript para mejorar la experiencia del usuario. En este caso la pregunta es ¿cómo analizar el comportamiento del usuario cuando no hay una transición explícita entre páginas? Dentro de este análisis se incluyen las variables nombradas anteriormente, junto con los tiempos

que toma el usuario para decidir cambiar de página, revisar un contenido en particular, y en general la forma en la que navega en el sitio.

Una alternativa viable que ha sido considerada, es el análisis del movimiento o fijación ocular del usuario en los contenidos de una página web. Al respecto, se han propuesto variados modelos para estudiar cómo el usuario explora una interfaz web y busca información que le es de interés [12]. De esta forma, la búsqueda de los objetos web clave, del punto de vista del usuario, así como su distribución en una página web, han contribuido a enriquecer el entendimiento del comportamiento del usuario en un sitio web y también han provisto nuevas líneas de investigación, específicamente en el área del desarrollo de nuevas aplicaciones basadas en la Web.

A pesar de que esta línea de investigación ha mostrado importantes resultados [7] y ha contribuido a incrementar el conocimiento acerca del comportamiento del usuario en una interfaz visual, en particular una página web, aun no considera las variaciones temporales propias de la navegación web, es decir, centra su análisis en fotos instantáneas de una página, obviando el cambio de contenido en el tiempo, cuando el usuario navega el sitio.

El análisis del movimiento ocular de un usuario en una página web, muestra que revisa los objetos web siguiendo una especie de patrón o camino [19] a través de los diferentes elementos presentes, siendo a veces lineal y otras cíclicos, además de presentarse el fenómeno de fijación ocular en ciertos objetos. Este camino presenta la trayectoria de fijación ocular que siguió el usuario, y es modelable a través de un grafo con las relaciones de atención entre los elementos visitados. De ahí mismo viene implícito el uso directo de la centralidad de grafos, en el sentido de que entrega qué objetos son los más relevantes dentro de un conjunto interrelacionado[10]. Por lo tanto la hipótesis de la investigación del presente artículo es que las métricas de centralidad utilizadas en teoría de grafos, podrían dar a conocer qué elementos son importantes dentro de los patrones de observación de un sitio web. El concepto de importancia es relativo, por lo que las diferentes métricas de centralidad, como *eigenvector* o *closeness centrality*, pueden dar resultados opuestos, por lo que es necesario analizarlas y ver cuáles se ajustan en mejor grado a los datos obtenidos desde los usuarios a analizar.

En este trabajo el principal objetivo es encontrar una medida de centralidad que se ajuste con los tiempos de fijación ocular en los objetos, y como objetivos secundarios, se propone apoyar el entendimiento del proceso de ubicación de la atención del usuario [10]. Estos intereses se refieren particularmente a aquellos componentes del sitio web que son los que guían la mirada y que finalmente son inspeccionados, incluyendo los textos que son leídos, las imágenes revisadas y los objetos evitados durante la inspección ocular. Al res-

pecto, los experimentos realizados consideraron la participación de 23 personas a quienes se les analizó la respuesta visual frente a estímulos presentes en una serie de páginas web (tiempos de fijaciones en objetos, movimientos oculares y dilatación pupilar) los cuales fueron registradas utilizando tecnologías del lado del servidor y dispositivos de seguimiento visual.

Este artículo está compuesto por una introducción que explica los componentes generales de la propuesta de investigación. La sección marco conceptual busca interiorizar al lector en conceptos y términos que fueron utilizados durante la investigación, así como el estado del arte de los dispositivos de rastreo ocular. Luego en la sección trabajo relacionado se muestran algunas investigaciones dentro del área de análisis de comportamiento del usuario y *eye tracker* que fueron citados en esta investigación. La sección metodología explica cómo fueron realizados los experimentos, la toma de datos y los dispositivos utilizados para ello. La sección análisis presenta los resultados obtenidos de la sección anterior y por último en conclusiones y trabajo futuro se entrega una discusión acerca de los resultados del experimento y de la investigación, junto con los posibles temas que se podrían abordar en trabajos posteriores.

2. Marco Conceptual

En los últimos años, el movimiento ocular en torno a estímulos visuales ha sido estudiado como una forma de analizar qué elementos atraen la atención de los usuarios. En el ámbito de la Web, lo que se busca es dar una respuesta aproximada a una vieja pregunta ¿qué elementos de un sitio atraen y retienen a un usuario web?. En la presente sección, se analizarán las distinciones básicas presentes en el análisis del movimiento ocular utilizando dispositivos de eye tracking, que luego serán necesarios para adentrarse en la propuesta de investigación que persigue este artículo.

2.1. Movimientos oculares

El ojo puede realizar diferentes tipos de movimiento para fijar, recibir y seguir estímulos visuales, y de esta forma el sujeto puede crear una imagen mental del escenario [12].

Estos movimientos están clasificados en tres categorías principales:

- **Movimientos de seguimiento:** Son aquellos realizados por el ojo para seguir el rastro de un objeto en movimiento.

- **Movimientos de convergencia:** Son realizados por ambos ojos con el objetivo de mantener la visión binocular.
- **Movimientos sacádicos:** Ocurren cuando el sujeto escanea una escena tratando de encontrar las partes más interesantes para así poder enfocarse en ellas. La velocidad de este movimiento no puede ser controlada voluntariamente, por lo que los ojos transitan tan rápidamente como pueden [1].

Hay otro tipo de mecanismo ocular no considerado movimiento llamado fijación la cual ocurre cuando el sujeto mantiene el foco en un objeto para así poder obtener una imagen en alta resolución de los detalles importantes. Generalmente cuando un sujeto recibe un estímulo visual, empieza a rastrear la escena, haciendo fijaciones en los detalles y realizando movimientos sacádicos entre los distintos puntos de fijación [12].

2.2. Atención visual y movimientos oculares

Si la atención visual fuera considerada en término de “qué” y “dónde”, es decir, “dónde” están los objetos destacados en una escena (observadas a través de la visión parafoveal) y “qué” son esos objetos destacados (a través de la visión foveal), sería normal considerar que los movimientos oculares trabajan de modo que apoyen la hipótesis de atención dual. Esto significa que la visión trabaja en un proceso cíclico compuesto de los siguientes pasos:

1. Dada una escena como estímulo, el encuadre completo es visto a través de la visión periférica y, por lo mismo, en baja resolución mayormente. En este paso los componentes más interesantes de la escena “destacan” o “aparecen” en el campo visual y atraen la atención hacia esa ubicación en particular para así poder realizar una inspección más detallada.
2. La visión foveal es desacoplada de la atención y los ojos se mueven hacia la región donde fue atraída la atención.
3. Después de que los ojos son apuntados hacia el área de interés, la fóvea es dirigida hacia la región de interés y la atención es centrada en ese punto para así poder percibirlo en alta resolución.

Este tipo de atención visual representa un enfoque del tipo “*bottom-up*”. Esto significa que las áreas interesantes de un estímulo visual, o en este caso una escena visual, atraen suficientemente la atención para prevenir mirar el resto de la escena. Por otra parte, un enfoque “*top-down*” es dirigido por otros tipos de factores cognitivos tales como conocimientos, expectativas, objetivos,

etc. Bajo este modelo, los sujetos están más predispuestos a mirar alrededor usando su visión periférica. Por ejemplo una persona que maneja regularmente notará las estaciones de bencina mucho más que alguien que no maneja tan seguido [9].

2.3. Técnicas de eye tracking

Los sistemas de rastreo ocular son dispositivos que permiten, a través de diferentes técnicas, medir los movimientos oculares de una persona [7]. De este modo es posible saber hacia dónde el sujeto apunta su mirada durante todo el intervalo de tiempo de uso del dispositivo, y usando estos datos, se obtiene la secuencia de movimientos oculares, junto con los tiempos en que se observaron los objetos dentro del sitio.

2.3.1. Foto-oculografía o video-oculografía

Esta categoría de *eye tracking* agrupa una amplia variedad de técnicas para medir los movimientos oculares, que incluyen la medición de diferentes características del ojo durante los movimientos de rotación y traslación. Algunas de ellas son la forma de la pupila, la posición del limbo o el reflejo de la córnea bajo la exposición a la luz (generalmente luz infra roja).

Estas técnicas trabajan de forma diferente, pero están agrupadas porque generalmente no entregan medidas sobre el punto de atención. Además sus resultados pueden ser obtenidos automáticamente, pero en algunos casos pueden incluir hacer inspecciones visuales de los movimientos registrados. Este tipo de análisis puede ser extremadamente tedioso y propenso a error, y está limitado a la tasa de muestreo del dispositivo utilizado.

2.3.2. Sistemas basados en video pupilar / Reflexión Corneal

La principal innovación de esta técnica es que provee una medida del punto de mirada. La medición de esta característica requiere tener la cabeza del sujeto fija de modo que la posición de los ojos y el punto de mirada coincidan. De otro modo sería necesario medir otras características de los ojos para poder separar los movimientos de la cabeza de los oculares.

Dos características útiles para medir son el reflejo corneal y en el centro de la pupila. Dispositivos basados en video que utilizan cámaras y hardware para procesamiento de imágenes relativamente baratos, son utilizados para procesar el punto de mirada en tiempo real. Los dispositivos y su configuración pueden ser montados en una mesa o en la cabeza. Ambas técnicas son similares al usar los mismos sistemas, y la única diferencia es el tamaño de los dispositivos.

Generalmente los sistemas montados en la cabeza son más precisos, porque se mueven en conjunción, pero a la vez son mucho más invasivos.

La reflexión corneal de la fuente de luz (generalmente luz infra-roja) es medida de manera relativa a la ubicación del centro de la pupila, donde un disco brillante (la pupila), fácilmente identificable, y uno pequeño (la reflexión corneal) pueden ser detectados. Usando un sistema de calibración de 9 puntos y algunos cálculos trigonométricos básicos es posible estimar la posición de la córnea y de la pupila, por lo tanto el punto de mirada del sujeto.

La arquitectura típica de un sistema de rastreo ocular basado en video se puede observar en la figura 1

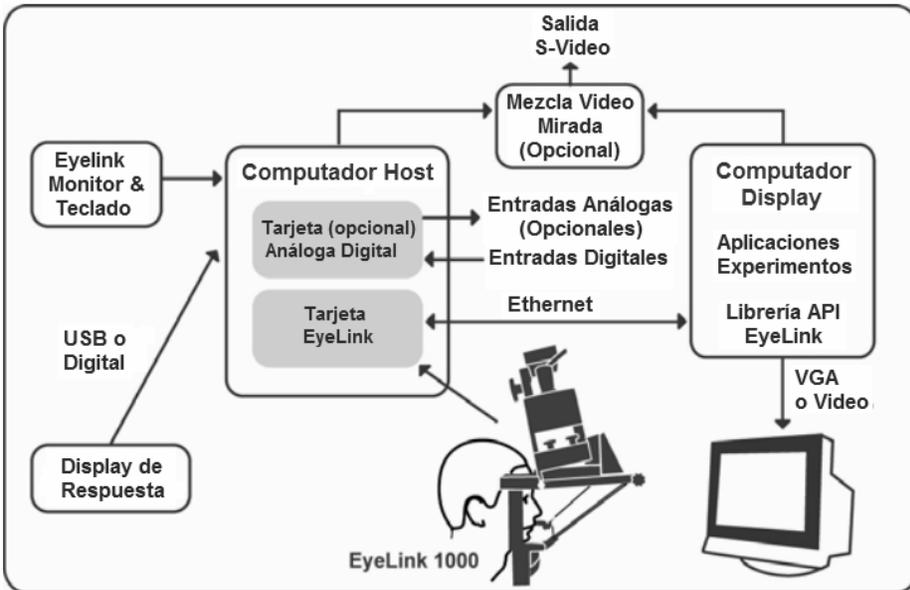


Figura 1: Arquitectura de un sistema de rastreo ocular usando video-oculografía

3. Trabajo relacionado

La literatura provee muchos ejemplos de cómo los sistemas de *eye tracking* han sido aplicados para enriquecer el análisis del usuario web. En las secciones subsiguientes, se mostrará un resumen de los más importantes casos, basados en el nivel de generalización y de novedad entregado.

3.1. Identificación de elementos web destacados

La principal motivación viene de la necesidad por entender cómo las personas centran la atención visual en páginas web, tomando en consideración su importancia tanto para desarrolladores web como para publicistas y anunciantes.

Un estudio del año 2009 realizó un análisis basado en rastreo visual en el que a 20 usuarios les fueron mostradas 361 páginas mientras desarrollaban búsqueda de información y tareas de inspección [4], donde el principal supuesto fue que los datos asociados a la fijación ocular podrían representar un centro (proxy) de atención. Desde esto, un marco de análisis fue desarrollado para, primeramente, generar una herramienta que permitiera caracterizar los elementos DOM, de las siglas en inglés de Document Object Model, para representar e interactuar con objetos dentro de archivos HTML, XHTML Y XML, independientemente de la plataforma o del lenguaje, por ejemplo, imágenes, párrafos, tablas, títulos, etc., con los datos asociados al movimiento ocular. La segunda parte incluye el uso de técnicas de *machine learning* para predecir los elementos importantes en una página web.

En este estudio el concepto de *fixation impact*[4] es utilizado para identificar los elementos que están bajo la mirada del usuario en un determinado momento. Los resultados empíricos muestran que la visión humana está caracterizada por una estrecha ventana de alta agudeza visual, junto con un área estándar alrededor del centro de la mirada. Así, cuando se visualiza un elemento, también implica que los objetos a su alrededor son considerados. Por lo tanto, dado un punto de fijación, un área DOM es seleccionada para así poder considerar todos los elementos en ella. Un puntaje basado en la distancia entre ellos es dado a cada objeto elemento basado en su cobertura, asumiendo una distribución gaussiana, la cual fue asumida debido a que la visión humana no tiene un alcance lineal, sino que está caracterizada por un punto de alta agudeza llamado fovea. Si bien la fijación está concentrada principalmente en este punto, la visión humana es capaz de capturar de manera parcial elementos que se encuentran alrededor de la fovea. Luego, en el caso de la navegación Web, cuando un usuario enfoca su atención en un punto en particular, una imagen, implícitamente está capturando información sobre los elementos que la rodean. Esta información adicional puede influir en la siguiente fijación que el usuario lleve a cabo, por lo tanto es importante considerarla.

Si bien existen varias propuestas para conseguir una buena estimación del *fixation impact*, la mayoría usa el concepto de gradiente de captura de información, es decir, asigna más peso al objeto que está más cerca del punto de fijación. En ese sentido, se opta por utilizar una distribución de pesos

que sigue una Gaussiana debido que es fácil de interpretar, es simétrica y ha demostrado buenos resultados en estudios previos [4]. El *fixation impact* es procesado usando la distancia y también incluyendo la dimensión del tiempo, que es la duración de la fijación.

La información obtenida en el paso anterior es usada para estimar los elementos importantes. Después de realizada una selección de las diez características que entregaron mayor ganancia de información, se aplica una regresión lineal para identificar las medidas que más influyeron en el puntaje del *fixation impact*. Los resultados mostraron que las características de posición obtuvieron los pesos más altos.

Otra línea de investigación ha sido desarrollada por Velásquez et al., donde el principal objetivo es identificar los objetos más relevantes en un sitio web usando el concepto de *website keyobjects* [21]. Un objeto web o *website object* es considerado como un elemento de una página web el cual puede ser caracterizado a través de un metadato. De esta forma, un párrafo, imagen o incluso un video puede ser un objeto web el cual está definido a través de datos sobre su contenido. Luego la definición de *website keyobject* deriva desde la de *website object* de la siguiente manera: *Uno o más website objects que atraen la atención del usuario y que caracterizan el contenido de una página o un sitio*. Esta definición establece qué objetos web obtienen mayor atención y son más interesantes desde el punto de vista del usuario y, por lo tanto, su identificación podría ayudar a mejorar la presentación, usabilidad y contenido de un sitio web. La identificación de los *keyobjects* involucra primeramente el análisis de los *weblogs* y una medida del tiempo utilizado. Para poder validar los descubrimientos fueron realizadas encuestas, las que no proveen un nivel alto de confianza para los resultados. Los autores abordaron este problema, y en [19] incorporaron metodologías de seguimiento visual para reemplazar el uso de encuestas. Con esto fueron capaces de validar el acercamiento teniendo una medida objetiva de la atención del usuario.

3.2. Implicit Feedback en la recuperación de información

En [5], Buscher et al. exploraron la aplicación de técnicas de *eye tracking* en el análisis del comportamiento del usuario en la página principal de un motor de búsqueda. Específicamente se trata obtener información respecto a los textos en los documentos que efectivamente son leídos por los usuarios. A lo anterior se le conoce como *implicit feedback* y puede ser usado para personalizar la búsqueda de documentos en un motor.

Entonces se recolectaron datos respecto de la fijación ocular para concluir qué partes de los documentos recuperados por el motor de búsqueda fueron

leídas y cuáles eran relevantes para el usuario. Adicionalmente el concepto de *attentive documents* es explicado para introducir una categoría especial de documentos que mantienen un registro de cómo y cuándo son vistos por los usuarios. Basándose en esa información adicional se pueden generar personalizaciones de los documentos leídos. Este concepto está basado en estudios previos [13, 23] donde la idea de intensidad de la fijación ocular está relacionada con la duración en el tiempo invertido.

El uso de *implicit feedback* es interesante ya que no sobrecarga al usuario y no interfiere con sus actividades. Este tipo de retroalimentación es capturado a través del análisis de la interacción del usuario con los sistemas en estudio y luego analizados para desarrollar determinadas acciones de asistencia (certain assistive action).

3.3. Identificación de la influencia de la publicidad en las páginas de resultados

La mayoría de los motores de búsqueda incorporan elementos de publicidad conocidos como ADS (advertisement o aviso publicitario) para entregar información comercial relacionada con las palabras clave ingresadas por los usuarios. Los ADS corresponden a publicidad pagada por los patrocinadores del motor de búsqueda y contienen hipervínculos a sus sitios web institucionales. No está claro si esta técnica realmente genera algún tipo de decisión de compra por parte del usuario, pero al menos se puede medir el comportamiento de clic, que queda almacenado tanto en el motor como en el sitio institucional visitado. En [6] se presenta un análisis de la respuesta ocular de los usuarios frente a los ADS de un motor de búsqueda, examinando qué variables influyen la fijación ocular del usuario en un nivel más alto.

Los descubrimientos claves del estudio están relacionados con la influencia de las tareas en las cuales el usuario está envuelto mientras navega, la calidad de los ADS y la secuencia en que aparecen a lo largo de los resultados de búsqueda. De ese modo encontraron evidencia cuantitativa de sesgo, por parte de la atención del usuario hacia los resultados cuando estos están ubicados en la parte superior de las listas. La influencia de la calidad de los ADS fue medida en término de la distancia semántica con la consulta de búsqueda. Basado en los resultados experimentales, los usuarios prestan menos atención a los resultados orgánicos¹ cuando la calidad de los avisos es buena.

¹Hacen referencia a aquellos resultados de una búsqueda que se basan directamente en el algoritmo imparcial utilizado por el buscador. Se diferencian de los resultados patrocinados en el sentido que estos últimos dependen de la cantidad de dinero invertido en ellos.

4. Metodología

Fue desarrollado un análisis exploratorio basado en el monitoreo de la actividad visual durante la navegación del usuario. El experimento consistió en pedir a un conjunto de sujetos que exploren un sitio web definido, realizando tareas estándar de obtención de información mientras utilizaban un dispositivo de rastreo ocular que registra y almacena toda la actividad visual y de la mirada. Una línea típica del registro es la siguiente:

```
<time> <xpl> <ypl> <psl> <xpr> <ypr> <psr> <xvl> <yvl> <xvr>
<yvr> <xr> <yr>
```

donde cada componente corresponde a:

- **<time>**: marca de tiempo en milisegundos.
- **<xpl>**: información de posición del eje x para el ojo izquierdo.
- **<ypl>**: información de posición del eje y para el ojo izquierdo.
- **<psl>**: tamaño de la pupila para el ojo izquierdo (área o diámetro).
- **<xpr>**: información de posición del eje x para el ojo derecho.
- **<ypr>**: información de posición del eje y para el ojo derecho.
- **<psr>**: tamaño de la pupila para el ojo derecho (área o diámetro).
- **<xvl>**: velocidad instantánea del ojo izquierdo en el eje x (grados/segundos).
- **<yvl>**: velocidad instantánea del ojo izquierdo en el eje y (grados/segundos).
- **<xvr>**: velocidad instantánea del ojo derecho en el eje x (grados/segundos).
- **<yvr>**: velocidad instantánea del ojo derecho en el eje y (grados/segundos).
- **<xr>**: resolución eje x (posición unidades/grados).
- **<yr>**: resolución eje y (posición unidades/grados).

Esta información fue registrada por el dispositivo cada 2 milisegundos. Los sujetos sólo recibieron la instrucción de revisar cada página web, sin un propósito particular, y de manera secuencial, para así obtener una visión completa de las 29 páginas del sitio. El usuario podía decidir cuánto tiempo destinaba a revisar cada página y luego pasar a la siguiente.

Los pasos específicos son descritos a continuación:

1. El dispositivo de *eye tracking* es calibrado y configurado de acuerdo a las características físicas del sujeto (altura, forma de la cabeza, posición al sentarse) y a sus movimientos oculares, de modo de disminuir el error al momento de hacer las mediciones. Luego la página inicial del sitio es presentada.
2. El sujeto comienza la exploración. Su comportamiento es seguido con el dispositivo de *eye tracking* y adicionalmente con el servidor que almacena la sesión web.
3. Dos procedimientos de recalibración son realizados durante el experimento para verificar la calidad de la información extraída por el dispositivo.

4.1. Implementación

Grupo experimental: Veinte y tres personas fueron elegidas con un promedio de edad de 26,1 años (con una varianza de 2,2 años). El grupo estaba compuesto principalmente por estudiantes universitarios y profesionales de diferentes áreas del conocimiento (ingenieros, biólogos y kinesiólogos). De acuerdo a la propia apreciación de cada sujeto (ningún tipo de test se realizó para medir esto), uno se consideraba experto en navegación web, 18 situaban su conocimiento como promedio y los otros 4 pensaban que eran usuarios básicos.

Sitio web escogido: El análisis fue realizado en el sitio web del programa de MBA ofrecido por el Departamento de Ingeniería Industrial de la Universidad de Chile ². Este sitio ha estado funcionando desde Enero del 2011 y entrega información acerca de los cursos, la metodología, el plan de estudios, los profesores, el perfil de los estudiantes, etc., como también páginas para poder postular al programa.

Al momento del análisis el sitio estaba compuesto por 29 páginas y 359 objetos DOM, que aparecieron un total de 1014 veces, es decir, un objeto puede estar presente en más de una página. Esto sucede bastante en objetos comunes como banners, menús, pies de páginas, cabeceras, etc. El promedio de objetos por página es de 31.9 y su tamaño promedio es de 418.6 pixeles de ancho por 100.1 pixeles de alto.

Dispositivo de seguimiento: Para poder capturar los movimientos del ojo y medir la dilatación pupilar fue utilizado un sistema de *eye tracker* basado en video que ocupa la pupila y la reflexión corneal para realizar registros. Este tipo de dispositivos son considerados herramientas avanzadas para medir el movimiento ocular, debido al alto grado de precisión que se obtiene. En particular se utilizó un *eye tracker* modelo Eyelink 1000, desarrollado por SR

²<http://www.mbauchile.cl>

Research³. Este dispositivo está compuesto de una pantalla principal con una cámara de alta velocidad y un emisor de infrarrojos, conectados a un host que se encarga del procesamiento de los datos. Esto permite al investigador desarrollar los experimentos y obtener la información de una manera simple. También muestra en tiempo real los mismos estímulos que el participante ve. El tiempo total que toma configurar el sistema para realizar un registro está entre los 2 y 5 minutos.

5. Análisis

Después de obtener los datos generados a partir del movimiento ocular de cada sujeto, se procedió a transformar la fijación ocular en una representación de grafo. Cada elemento de una página web, como un párrafo de texto o una imagen, es representado como un nodo, asignándosele un identificador único. Cada nodo tiene asociado un valor numérico que indica el tiempo de fijación en el elemento que representa. Los arcos entre elementos no tienen peso dado que el movimiento sacádico realizado entre fijaciones es de alta velocidad, por lo que el tiempo entre estas transiciones es despreciable. Es importante destacar que el tiempo de fijación incluye todo el tiempo que se miró un objeto, junto con los tiempos donde hubo movimientos oculares dentro del objeto (como en el caso de objetos de texto que necesitan ser leídos).

Las conexiones o arcos en el grafo representan las transiciones en la mirada entregadas por el dispositivo de *eye tracking*, tal como se aprecia en la Fig. 2. En este caso, el usuario fija la atención inicialmente en la foto superior izquierda, luego se mueve hacia el párrafo central, para finalmente mirar la foto inferior central. Esta transición es representada como un arco dirigido entre dos nodos.

Por lo tanto cada transición añade un nuevo nodo al correspondiente grafo, creándose una estructura dirigida. Se podría pensar que en vez de una representación por grafos sólo se estaría obteniendo una lista secuencial de los objetos vistos, pero la evidencia muestra que los usuarios tienden a retornar a los elementos vistos previamente, lo que genera ciclos dirigidos. Toda la información de la generación de grafos es almacenada, para luego inspeccionar el proceso completo.

Cada usuario genera un grafo por página visitada, por lo tanto el número máximo de grafos por usuario es 29. Esto fue diseñado de esta manera dado que los elementos de cada página son diferentes (a pesar de existir elementos

³<http://www.sr-research.com>

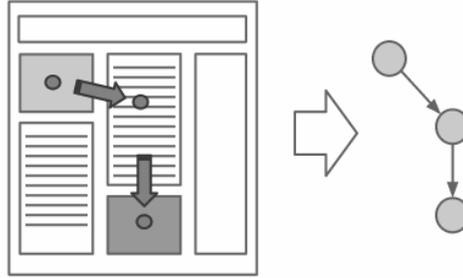


Figura 2: Transformación desde actividad de la mirada hacia una representación de grafos

comunes como los menús, cabeceras, etc.) y por lo mismo no se podrían agregar nodos de forma directa para generar un sólo grafo. Además se busca encontrar la importancia relativa de los elementos dentro de una página, lo que se anula al aunar toda la información, junto con perder el elemento de temporalidad que se busca inicialmente, al no poder diferenciar cuándo aparecen nuevos objetos. Claramente los grafos que genera un usuario en una página varían respecto de otros usuarios, puesto que el proceso de ubicación de la atención es personal y diferente en cada ser humano. Sin embargo, se aprecian ciertas similitudes que a la larga representarán patrones de comportamiento común.

5.1. Reconocimiento de objetos destacados

Es uno de los más importantes tipos de análisis que pueden ser realizados cuando se estudia la atención del usuario. Para lograr tal tarea se propone el uso del concepto *Network Centrality*, muy utilizada en teoría de grafos para expresar la relevancia de un nodo. Este tipo de métrica se originó desde el campo de análisis de redes sociales para poder identificar el grafo de influencia dentro de un grupo de individuos [8].

Una de las más comunes medidas de centralidad es la llamada centralidad del vector propio [3, 2]. Dada la matriz de adyacencia A del grafo G , esta centralidad es definida como el vector propio principal según la ecuación 1

$$\lambda v = Av, \quad (1)$$

donde λ es el valor propio (constante) y v es el vector propio. La interpretación de la ecuación 1 es que un nodo tiene un puntaje de vector propio alto si es adyacente a otros nodos con valores altos. La idea intrínseca de esta métrica es la noción de influencia, es decir, si un nodo influye sólo un nodo, el cual subsecuentemente influye en múltiples nodos, entonces el primer nodo

en esa cadena es altamente influyente. Otra interpretación, que podría estar más relacionada al análisis de control de flujo, es que la centralidad del vector propio entrega un modelo para el riesgo nodal (nodal risk), en el sentido de que el equilibrio a largo plazo del nodo de recibir tráfico va en función del riesgo de sus vecinos.

A pesar que la aplicación de esta métrica parece promisoría en el contexto del flujo de datos, debe notarse que, dada su naturaleza, existen algunas suposiciones a ser tomadas en cuenta. Primero, se asume que el tráfico está habilitado para moverse a través de caminos no restringidos. Segundo, un nodo influye a todos sus vecinos al mismo tiempo.

Otra medida es el grado de centralidad. Dado el grafo $G = (V, E)$, donde V es el conjunto de vértices y E es el conjunto de arcos, el grado de centralidad para el nodo i está definido como $CD = deg(i)$, que es el número de relaciones (ties) que inciden sobre un nodo i . La interpretación es que el grado de centralidad representa el efecto inmediato en una red. Por ejemplo, si cierta proporción de los nodos están infectados, la probabilidad de infección es una función en base del número de nodos que están adyacentes.

Se puede ver que el grado de centralidad y la centralidad del vector propio comparten ciertos aspectos. Aunque mientras el grado de centralidad exagera en el riesgo inmediato/influencia, la centralidad del vector propio está relacionado al riesgo/influencia en el largo plazo.

La proximidad central para un nodo es definida como la suma teórica (graph-theoretic) de las distancias desde todos los otros nodos. La noción de distancia en este caso es definida como el largo del camino más corto entre dos nodos. En el contexto de flujo de información, esta métrica es relevante en el sentido de que puede ser interpretada como el tiempo de llegada estimado de la información a un nodo en particular. Nodos con puntaje de cercanía bajo tienen distancias más cortas del resto, lo que permite recibir la información antes.

Debe notarse que el uso de esta métrica implica una discusión del concepto de accesibilidad, en el sentido de que sólo entrega resultados significativos si el grafo estudiado es uno conectado.

Finalmente se tiene la centralidad intermedia. Esta métrica se define como la proporción de veces que un nodo i necesita un nodo k para poder alcanzar un nodo j a través del camino más corto. Si g_{ij} es el número de caminos geodésicos desde i hasta j y g_{ikj} es el número de estos caminos que pasan a través del nodo k , entonces la centralidad intermedia del nodo k está definida como lo expresa la ecuación 2

$$C_B = \sum_i \sum_j \frac{g_{ikj}}{g_{ij}}, \tag{2}$$

con $i \neq j$ and $i \neq k$.

Las métricas anteriormente expuestas, fueron aplicadas a los grafos resultantes para cada página visitadas por cada usuario. El objetivo principal es analizar si los nodos que tienen un puntaje alto están correlacionados con los que el usuario fija su atención más tiempo.

Para generar la comparación, cada métrica de centralidad fue procesada para cada nodo en el grafo, y los puntajes fueron agregados junto con los grupos de usuarios.

El tiempo empleado en observar cada objeto fue utilizado como medida de la atención del usuario. Como una medida de la atención del usuario fue calculada el tiempo gastado en cada objeto como un sustituto para la fijación.

Por lo tanto, para cada nodo que fue visualizado por cualquier usuario existe un conjunto de cuatro puntajes de centralidad, junto con una medida del tiempo gastado en él.

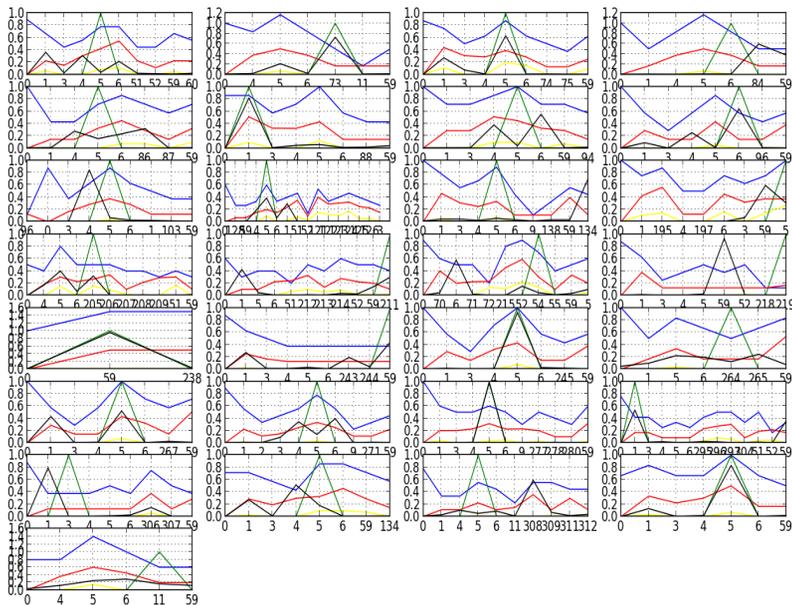


Figura 3: Puntajes de métrica de centralidad y tiempo gastado para cada nodo

La figura 3 muestra los resultados promediados para el reconocimiento de elementos destacados para las 29 páginas que conforman el sitio web estudiado. La representación el gráfico es la siguiente: *degree*, *closeness*, *eigenvector* y *betweenness* (listadas en orden desde la parte superior hasta la inferior de cada gráfico). El tiempo gastado en cada objeto es representado por la línea negra.

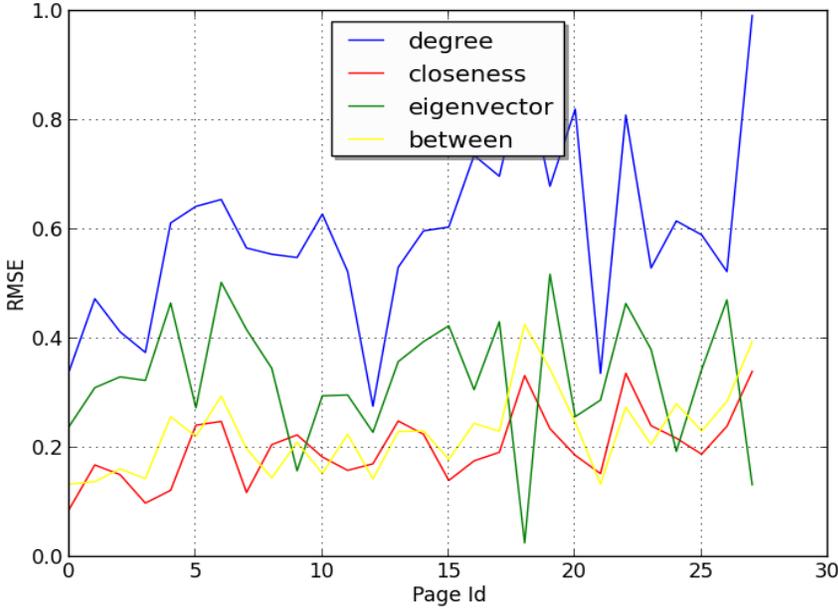


Figura 4: Ejemplo de los valores RMSE para todas las métricas de centralidad para cada página para un sujeto

Fue desarrollado un procesamiento similar para poder encontrar qué métricas de centralidad se comportan de manera más parecida al tiempo gastado por los usuarios en cada nodo. Al ser un estudio exploratorio, fueron elegidas dos maneras de comparar las métricas de centralidad con el tiempo gastado.

La primera manera es usar RMSE⁴ como métrica de cercanía entre la centralidad y el tiempo. La idea es encontrar, de manera agregada, que curva de centralidad tiene el menor RMSE. La Fig. 4 muestra, para un sujeto, los valores de RMSE para todas las métricas de centralidad junto con la sesión web.

Fueron procesados los valores del RMSE promedio para cada métrica de centralidad, y los resultados muestran que la métrica con el puntaje más bajo

⁴Del inglés Root-mean-square error, se refiere a la desviación cuadrática.

es la de centralidad del vector propio, seguido de la proximidad central.

Este resultado parece lógico ya que el concepto de influencia en interfaces gráficas, tales como sitios web, está relacionada a cómo la atención en elementos específicos afecta la atención hacia el resto. Un ejemplo de esto podría ser una imagen colorida que es mostrada en un sitio web. Eventualmente podría guiar la atención del usuario hacia los elementos que la rodean, por lo que los elementos que están adyacentes, recibirán más atención (en términos del tiempo gastado en una fijación), que los que están más alejados. Así un elemento será influyente si al mismo tiempo está conectado con otros que sean también influyentes .

La segunda forma de explorar los resultados es usar una métrica de correlación entre los valores de centralidad y el tiempo gastado en cada elemento. Este análisis es diferente al anterior en el sentido que no tienen como objetivo encontrar el nivel de ajuste entre dos curvas. En este caso se quiso explorar si existe una dependencia entre las variables.

Fue usado el coeficiente de correlación de Pearson, que mide el grado de relación de dos variables siempre y cuando ambas sean cuantitativas, para calcular el puntaje de dependencia, por lo tanto se asumió el escenario de una correlación lineal y que la información sigue una distribución normal.

Los resultados muestran que el valor más alto para la correlación fue obtenido por la centralidad del vector propio ($r=0.41$), a pesar que el nivel de confianza no es estadísticamente significativo. Ninguna otra métrica de centralidad obtuvo un puntaje de correlación aceptable.

6. Conclusiones y trabajo futuro

La métrica presentada en este trabajo es una exploración inicial para poder entender las correlaciones entre la actividad registrada y los resultados obtenidos. Por lo mismo esta métrica es en sí un estimador de la atención que los usuarios invierten en los objetos, pero no podría tomarse como un valor final, ya que en ciertos casos extremos los resultados no se correlacionan con los tiempos de fijación obtenidos en los objetos. Sin embargo, este tipo de modelamiento es novedoso ya que técnicas similares habían utilizados representaciones de grafos y métricas de centralidad en imágenes, pero nunca en análisis de páginas web y sus elementos. Lo anterior deja una puerta abierta para futuras modificaciones o acercamientos similares que permitan obtener mejoras en cuanto a las correlaciones de la métrica y los resultados.

Otro concepto importante que puede someterse a análisis es la validez de la

granularidad de los elementos. En este caso particular se utilizó la definición de Velásquez para *website objects* y *website keyobjects*, la cual aún carece de una estandarización que permita obtener objetos similares en páginas similares. El análisis de un sitio escrito en HTML y que utilice elementos DOM permite una gran variedad de elecciones en cuanto a objetos, dependiendo del nivel en el que estén dentro del árbol o lo que representen, por lo que hay algunas preguntas subyacentes acerca de ¿cómo identificar *web objects* automáticamente?, ¿cómo analizar el HTML y agrupar elementos de la jerarquía DOM? Actualmente se está trabajando en estos métodos, analizando las diferencias al subir o bajar niveles dentro del árbol DOM y otros que incluyen técnicas de *clustering*, pero que escapan al alcance de este trabajo.

Cabe destacar que el sistema que se utilizó mostraba las páginas de manera secuencial, dándole la opción al sujeto de pasar a la siguiente página cuando él lo estimase conveniente, hasta completar el conjunto de 29 páginas. Esto podría afectar la validez del experimento ya que es una simulación de una navegación real dentro de la página, donde un usuario ingresa por voluntad propia y navega a través de clics en los diferentes links existentes dentro de cada página. La restricción impuesta fue necesaria para establecer los límites de navegación del usuario en cuanto a tiempo y a cantidad de páginas que debía visitar. Al “imponer” una tarea al usuario, la navegación se ve sesgada y eso fue lo que se trató de evitar en esta investigación.

En este trabajo se exploró cómo se comporta el usuario web desde una perspectiva de la fijación ocular. Para poder formalizar el análisis fue usado un marco basado en teoría de grafos que permitió la aplicación de métricas para establecer la importancia relativa de los objetos web. Los resultados iniciales muestran que en efecto, las métricas de centralidad pueden ser usadas para identificar los objetos destacados de una página web. Este es un resultado útil ya que siempre es deseable conocer en que objetos el usuario se centra mayormente.

Finalmente los resultados muestran que existe una correlación alta entre la métrica de centralidad del vector propio, y que las otras métricas de centralidad no obtuvieron un valor de correlación aceptable como para poder obtener conclusiones directas desde ellas. En este sentido las contribuciones del presente artículo van por el lado de abrir una nueva ventana de investigación, utilizando técnicas muy conocidas como la teoría de grafos y seguimiento del movimiento ocular usando herramientas de eye tracking.

Agradecimientos: Este trabajo fue financiado por el proyecto FONDEF-CONICYT CA12I10061 y por el Instituto Sistemas Complejos de Ingeniería (ICM: P-05-004-F, CONICYT: FBO16).

Referencias

- [1] George Adler. Psychology: The science of behaviour. *Canadian Psychology*, 43(4):278, 2002.
- [2] Phillip Bonacich. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1):113–120, 1972.
- [3] Stephen P Borgatti. Centrality and network flow. *Social networks*, 27(1):55–71, 2005.
- [4] Georg Buscher, Edward Cutrell, and Meredith Ringel Morris. What do you see when you're surfing?: Using eye tracking to predict salient regions of web pages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 21–30, New York, NY, USA, 2009. ACM.
- [5] Georg Buscher, Andreas Dengel, Ralf Biedert, and Ludger V. Elst. Attentive documents: Eye tracking as implicit feedback for information retrieval and beyond. *ACM Trans. Interact. Intell. Syst.*, 1(2):9:1–9:30, January 2012.
- [6] Georg Buscher, Susan T. Dumais, and Edward Cutrell. The good, the bad, and the random: An eye-tracking study of ad quality in web search. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '10, pages 42–49, New York, NY, USA, 2010. ACM.
- [7] A. Duchowski. *Eye Tracking Methodology: Theory and Practice*. Methods in molecular biology. Springer, 2007.
- [8] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1979.
- [9] Simone Frintrop, Erich Rome, and Henrik I. Christensen. Computational visual attention systems and their cognitive foundations: A survey. *ACM Trans. Appl. Percept.*, 7(1):6:1–6:39, January 2010.
- [10] Pablo Loyola and Juan D Velásquez. Characterizing web user visual gaze patterns: A graph theory inspired approach. In *Brain Informatics and Health*, pages 586–594. Springer International Publishing, 2014.
- [11] H. Brendan et al McMahan. Ad click prediction: A view from the trenches. In *Proceedings of the 19th ACM SIGKDD International Conference on*

- Knowledge Discovery and Data Mining*, KDD '13, pages 1222–1230, New York, NY, USA, 2013. ACM.
- [12] J. Nielsen and K. Pernice. *Eyetracking Web Usability*. Voices That Matter. Pearson Education, 2010.
- [13] Takehiko Ohno. Eyeprint: support of document browsing with eye gaze trace. In *Proceedings of the 6th international conference on Multimodal interfaces*, pages 16–23. ACM, 2004.
- [14] Pablo E. Román, Juan D. Velásquez Robert F. Dell, and Pablo S. Loyola. Identifying user sessions from web server logs with integer programming. *Intelligent Data Analysis*, 18(1):43–61, 2014.
- [15] Pablo E. Román and Juan D. Velásquez. A neurology-inspired model of web usage. *Neurocomputing*, 131(0):300 – 311, 2014.
- [16] Myra Spiliopoulou, Bamshad Mobasher, Bettina Berendt, and Miki Nakagawa. A framework for the evaluation of session reconstruction heuristics in web-usage analysis. *Inform journal on computing*, 15(2):171–190, 2003.
- [17] Yu-Hui Tao, Tzung-Pei Hong, Wen-Yang Lin, and Wen-Yuan Chiu. A practical extension of web usage mining with intentional browsing data toward usage. *Expert Systems with Applications*, 36(2):3937–3945, 2009.
- [18] Juan D Velásquez. Web site keywords: A methodology for improving gradually the web site text content. *Intelligent Data Analysis*, 16(2):327–348, 2012.
- [19] Juan D. Velásquez. Combining eye-tracking technologies with web usage mining for identifying website keyobjects. *Engineering Applications of Artificial Intelligence*, 26:1469 – 1478, 2013.
- [20] Juan D Velásquez. Web mining and privacy concerns: Some important legal issues to be consider before applying any data and information extraction technique in web-based environments. *Expert Systems with Applications*, 40(13):5228–5239, 2013.
- [21] Juan D. Velásquez, Luis E. Dujovne, and Gastón L’Huillier. Extracting significant website key objects: A semantic web mining approach. *Engineering Applications of Artificial Intelligence*, 24(8):1532 – 1541, 2011.
- [22] Sungjoon Steve Won, Jing Jin, and Jason I Hong. Contextual web history: using visual and contextual cues to improve web browser history. In

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 1457–1466. ACM, 2009.

- [23] Songhua Xu, Hao Jiang, and Francis Lau. User-oriented document summarization through vision-based eye-tracking. In *Procs. of the 14th Int. Conf. on Intelligent user interfaces*, pages 7–16. ACM, 2009.