

***MyAdChoices* — Transparencia y Control de la Publicidad en Línea**

**Dr. Javier Parra Arnau¹
Dr. Jagdish Prasad Achara²
Dr. Claude Castelluccia³**

**Agencia Española de Protección de Datos
Premio de Investigación de Protección de Datos Personales Emilio Aced**

¹ Universitat Rovira i Virgili, Tarragona, España

² École Polytechnique Fédérale de Lausanne, Suiza

³ INRIA, Grenoble, Francia

Barcelona, 28 de noviembre de 2018

CONTENIDO

1	Introducción	2
1.1	Principales Contribuciones y Organización de este Documento	3
2	Fundamentos de la Publicidad en Línea	4
2.1	Principales Actores	5
2.2	Proceso de Entrega de Anuncios	5
2.3	Tipos de Anuncios	7
2.4	Coincidencia de Cookie y Real-Time Bidding	8
3	Detección de Anuncios basados en el Perfil de Usuario y su Singularidad	9
3.1	Conceptos Básicos sobre Estadística y Teoría de Información	10
3.2	Detección de Anuncios basados en el Perfil de Navegación	10
3.2.1	Modelo de Distribución de Anuncios basados en Intereses	10
3.2.2	Test de Hipótesis Binario	12
3.2.3	Perfiles de Intereses a Corto y Largo Plazo	13
3.2.4	Detección Óptima bajo Incertidumbre de Anuncios basados en Intereses	14
3.3	Detección de Perfiles Singulares	17
4	MyAdChoices — una Herramienta para la Transparencia y Bloqueo de Anuncios	18
4.1	Principales Funcionalidades	18
4.1.1	Ejemplos de Políticas de Bloqueo de Anuncios	19
4.2	Arquitectura de Sistema y Detalles Prácticos de Implementación	20
4.2.1	Suposiciones	21
4.2.2	Componentes	22
5	Evaluación	28
5.1	Datos	28
5.2	Resultados	28
5.2.1	Rendimiento	29
5.2.2	Anuncios basados en Intereses y de tipo Redirección	30
6	Proyectos Relacionados	34
6.1	Bloqueadores de Anuncios	34
6.2	Transparencia de Anuncios	34
7	Conclusiones	38
	Apéndice A: Problema de Factibilidad	39
	Apéndice B: Formulación LP del Detector Minimax Robusto	39
	Apéndice C: Manual de Instalación y Configuración de MyAdChoices	40
C.1	Instalación	40
C.2	Configuración	41
C.3	Descripción de la información facilitada por MyAdChoices	42
	Referencias	43

1 INTRODUCCIÓN

Durante las últimas dos décadas, Internet y la Web se han ido integrando gradualmente en la vida cotidiana de las personas, permitiendo nuevas formas de comunicación, como el correo electrónico y la mensajería instantánea. La llamada red de redes se ha convertido en un canal de comunicación esencial no solo entre las personas, sino también entre las empresas y sus clientes.

La dinamización de las actividades comerciales tradicionales es precisamente una de las influencias más relevantes de Internet. La Web ha dado lugar a cambios empresariales clave que abarcan toda la cadena de valor en casi todos los sectores y empresas. Estos cambios han tenido un impacto en la forma en la que se venden los productos y también, y lo que es más importante, en cómo las empresas se acercan a sus clientes de manera personalizada, teniendo en cuenta sus preferencias únicas.

La industria de la publicidad, encabezada por Google DoubleClick y la tecnología real-time bidding (RTB), es un claro ejemplo de la transformación impulsada por la creciente sofisticación de las tecnologías Web. En el pasado, los propietarios de las páginas Web eran quienes acordaban directamente con los anunciantes la publicidad que se mostraría en sus páginas. Sin embargo, debido a la introducción gradual de compañías intermediarias con amplias capacidades para rastrear usuarios, la publicidad en Internet se ha vuelto cada vez más personalizada y omnipresente.

En los últimos años, además, hemos sido testigo de cómo la publicidad, habitualmente orientada a la venta de productos y servicios, ha ganado especial relevancia en el ámbito político. El caso de Cambridge Analytica, en el que los datos de millones de perfiles de Facebook fueron utilizados para enviar publicidad personalizada y manipular psicológicamente a los votantes en las elecciones de EE.UU. de 2016, es un claro ejemplo. En Europa, la campaña a favor del Brexit para el referéndum de 2016 invirtió casi todo su presupuesto en anuncios de Facebook, algo que se tradujo en unos mil millones de anuncios dirigidos a audiencias específicas de votantes. En España, la nueva ley de protección de datos que se ha debatido recientemente en el Senado permitirá a los partidos políticos rastrear datos de navegación sin que los usuarios tengan que dar su consentimiento expreso para hacer perfiles ideológicos y personalizar la propaganda en campaña electoral.

La capacidad de la industria del marketing para rastrear y perfilar la actividad de navegación Web de los usuarios es, sin duda, lo que permite servicios de publicidad más efectivos y personalizados, ya sea en el ámbito comercial como en la esfera política. La intrusión de estas prácticas y la creciente invasividad de la publicidad digital, sin embargo, han generado serias preocupaciones con respecto a la privacidad de usuario y la usabilidad de la Web. Según encuestas recientes, dos de cada tres usuarios de Internet están preocupados por el hecho de que su comportamiento en línea sea analizado sin su conocimiento y consentimiento [1]. Numerosos estudios en esta misma línea reflejan el creciente nivel de ubicuidad y abuso de la publicidad, que los usuarios perciben como una degradación significativa de su experiencia de navegación [2], [3].

En respuesta a estas preocupaciones, en los últimos años hemos asistido a la aparición de un gran número de herramientas de bloqueo de anuncios, cuyo objetivo principal es devolver a los usuarios el control sobre la publicidad. Esencialmente, los bloqueadores de anuncios supervisan todas las conexiones de red que pueden iniciarse cuando el navegador carga una página, e impiden aquellas realizadas con terceros que pueden corresponder a anuncios. Para llevar a cabo esta tarea, los bloqueadores de anuncios cuentan con listas negras de empresas que se dedican a la distribución de publicidad. Estas listas son mantenidas de forma manual por sus compañías de desarrollo y, en algunos casos, por comunidades de usuarios.

Al margen de la controversia generada por el uso de dichas listas, especialmente después de la revelación de que Adblock Plus [4], la más popular de estas tecnologías, recibía dinero de las empresas de publicidad para eliminarlas de dichas listas [5]—, el principal problema de estas herramientas es que fueron concebidas sin tener en cuenta dos aspectos clave: primero, el papel crucial de la publicidad en línea como el principal soporte de los servicios “gratuitos” de Internet; y en segundo lugar, el beneficio social y económico de la publicidad racional y no intrusiva.

Si bien los bloqueadores de anuncios pueden constituir un primer intento por recuperar el control sobre la publicidad, lo cierto es que su enfoque es limitado y radical: los usuarios solo pueden

elegir entre bloquear o permitir todos los anuncios de aquellas compañías presentes en dichas listas negras.

En un intento por abordar el problema de privacidad y de usabilidad que generan los anuncios en línea, la industria de la publicidad en Internet y el World Wide Web Consortium lanzaron dos iniciativas de autorregulación, Your Online Choices [6] y Do Not Track (DNT) [7]. Aunque estas dos iniciativas facilitan el denominado *opt-out* (la primera para dejar de recibir anuncios basados en los intereses de navegación Web, y la segunda para que los usuarios puedan dejar de ser rastreados a través de cookies de terceros), lo cierto es que los usuarios no pueden controlar si se respetan o no sus preferencias de publicidad y de rastreo en la Web.

Con alrededor de 200 millones de personas en todo el mundo que utilizan regularmente bloqueadores de anuncios, el modelo económico que sustenta la Web se está viendo amenazado [8]. La utilización de estas tecnologías ha desatado un fuerte debate sobre su ética y la necesidad de una solución que logre un mejor equilibrio entre el modelo empresarial dominante de Internet, la protección de los datos de navegación de usuario, y la usabilidad de la Web [9], [10], [11].

Creemos que la solución a este problema pasa por que los usuarios tengan un control real sobre la publicidad. Sin embargo, esto solo puede lograrse a través de tecnologías que empoderen al usuario e impongan sus preferencias, y no a través de los bloqueadores de anuncios actuales. De hecho, según una encuesta reciente, dos de cada tres usuarios de bloqueadores de anuncios no están en contra de los anuncios y aceptarían recibir publicidad a cambio de contenido gratuito [12]. Los usuarios, no obstante, ponen como condición que la publicidad sea un proceso transparente y tengan control sobre la información personal que se recopila [13]. La confianza, a través de la transparencia, parece ser clave en este sentido [14]. Sin embargo, como distintos usuarios pueden tener distintas motivaciones, necesitamos herramientas que permitan diferentes opciones con respecto al bloqueo de anuncios y transparencia.

En esta segunda edición del Premio de Investigación de Protección de Datos Personales Emilio Aced, presentamos *MyAdChoices*, una tecnología que tiene como objetivo, por un lado, proporcionar *transparencia* en el proceso de generación de los anuncios en línea, facilitando a los usuarios información sobre el uso que las empresas de Internet dan a sus datos de navegación; y por otro lado, permitir a los usuarios decidir qué tipo de anuncios desean recibir mientras navegan, ejerciendo un *control* preciso sobre la publicidad. El propósito último de *MyAdChoices* es dotar a los usuarios de ciertas garantías en términos de *protección de datos* y experiencia de navegación, respetando al mismo tiempo el modelo económico de la Web basado en anuncios.

La tecnología descrita en este documento ha sido diseñada y desarrollada en el marco del proyecto internacional *MyRealOnlineChoices* por investigadores de la Universitat Rovira i Virgili de Tarragona, de la École Polytechnique Fédérale de Lausanne de Suiza, y del Institut National de Recherche en Informatique et en Automatique de Francia. *MyAdChoices* está disponible públicamente¹ como prototipo de investigación desde diciembre de 2017.

El principal objetivo de *MyAdChoices* es dotar a los usuarios de Internet de herramientas que les permitan averiguar qué uso hacen las compañías de anuncios de sus datos de navegación, y al mismo tiempo puedan configurar preferencias de bloqueo de anuncios de una manera selectiva y respetuosa con el modelo económico actual de la Web.

1.1 Principales Contribuciones y Organización de este Documento

A continuación, resumimos las principales contribuciones técnicas de este trabajo:

- Proponemos un modelo teórico que investiga la *publicidad conductual* (“behavioral targeting”), una forma generalizada de publicidad que utiliza información del comportamiento de los

1. <https://sites.google.com/site/javierparraarnau/software>

usuarios en la Web (esencialmente, su historial de páginas Web visitadas) para publicar anuncios personalizados.

El modelo propuesto tiene como objetivo restar opacidad a este proceso de personalización de anuncios. Llevamos a cabo nuestro objetivo, primero, detectando los anuncios basados en la navegación, para cuantificar así hasta qué punto se utilizan estos datos; y segundo, estudiando la singularidad de los perfiles de navegación recopilados por las empresas que participan en este proceso. La clave del modelo propuesto está en su aproximación matemática al problema de la detección de publicidad conductual, a diferencia de trabajos anteriores basados en heurísticas y modelos simples que simplifican en exceso el proceso de personalización. La detección de dicha publicidad se formula en este trabajo como un problema de optimización que refleja la incertidumbre sobre la información que las plataformas publicitarias y rastreadores disponen de los usuarios. El modelo propuesto utiliza resultados fundamentales de los campos de estimación estadística y optimización robusta.

- En esta misma línea de transparencia, proponemos un segundo sistema de detección que arroja luz sobre la singularidad de los perfiles de navegación recopilados por las empresas de anuncios. Para ello, utilizamos una medida cuantificable de la singularidad del perfil del usuario: la divergencia Kullback-Leibler (KL) o *entropía relativa* entre la distribución de probabilidad de los intereses de navegación del usuario y la distribución de la población, una medida que justificamos y interpretamos en [15], [16] mediante los métodos de maximización de la entropía.
- Diseñamos una arquitectura de sistema que implementa los dos detectores mencionados anteriormente, y permite el bloqueo inteligente de anuncios a través de la especificación de políticas de control configurables por el usuario. Tanto el bloqueo de anuncios como la funcionalidad de transparencia se llevan a cabo en tiempo real, sin necesidad de comunicación con ninguna entidad externa, mediante un categorizador local de contenido Web y librerías de optimización de código abierto. La única excepción es el cálculo de la singularidad del perfil, que requiere de un servidor externo.

Un aspecto relevante de nuestro sistema es que ha sido concebido para funcionar en dos escenarios distintos en términos de rastreo, lo que permite a los usuarios configurar la funcionalidad de transparencia de acuerdo con sus propias percepciones al respecto. La arquitectura de sistema se ha desarrollado como una extensión de navegador para Google Chrome, está disponible públicamente como prototipo de investigación y ha recibido más de 2 000 descargas.

- Realizamos un análisis experimental a partir de los datos de usuario recopilados por nuestra extensión. Dicho análisis nos permite, primero, evaluar el sistema propuesto en un entorno real; y en segundo lugar, investigar varios aspectos relacionados con la publicidad conductual. Los experimentos realizados constituyen el primer estudio sobre este tipo de publicidad mediante perfiles de navegación de usuarios reales.

El contenido de este trabajo está organizado de la siguiente manera. La Sec. 2 proporciona los fundamentos básicos necesarios en publicidad online. Después, la Sec. 3 presenta el modelo teórico para la detección de anuncios basados en intereses y singularidad del perfil. La Sec. 4 describe los componentes principales de una arquitectura de sistema que tiene como objetivo proporcionar transparencia en el proceso de generación de anuncios, y funcionalidades avanzadas de bloqueo de éstos. La Sec. 5 analiza los datos recopilados por la herramienta propuesta mediante un experimento con 40 participantes. La Sec. 6 revisa el estado de la arte relevante para este trabajo. Las conclusiones se exponen en la Sec. 7. Los Apéndices B y A muestran, respectivamente, la formulación del programa lineal del detector de anuncios basado en intereses y la factibilidad de este problema de optimización. Por último, el manual de instalación y configuración de MyAdChoices se encuentran en el Apéndice C.

2 FUNDAMENTOS DE LA PUBLICIDAD EN LÍNEA

Esta sección describe el ecosistema de publicidad en línea, necesario para comprender las contribuciones técnicas de este trabajo. En primer lugar, la Sec. 2.1 ofrece una visión general de los principales

actores de este ecosistema. Posteriormente, la Sec. 2.2 detalla cómo se publican los anuncios en la Web, y después, la Sec. 2.3 proporciona una clasificación estándar de los tipos de anuncios existentes. Finalmente, la Sec. 2.4 presenta una de las tecnologías que permiten este proceso de publicación de anuncios. Para una explicación detallada y completa sobre el tema, se remite al lector a [17].

2.1 Principales Actores

La industria de la publicidad en línea está compuesta por un número considerable de entidades con roles muy específicos y complementarios, cuyo objetivo final es mostrar anuncios en sitios Web. Editores, anunciantes, plataformas publicitarias, agencias de anuncios, agregadores y optimizadores son algunas de las partes involucradas en el proceso de publicación de anuncios [18]. A pesar de la enorme complejidad² y la evolución constante del ecosistema publicitario, el proceso mediante el que los anuncios se muestran en la Web se caracteriza en términos únicamente de editores, anunciantes y plataformas publicitarias [20], [21], [22], [23], [24].

A continuación, describimos estos tres actores clave:

- Un *editor* es una entidad que posee una página Web (o un sitio Web) y que, a cambio de alguna compensación económica, está dispuesta a colocar anuncios de terceros en algunos espacios de su página (o sitio). Un ejemplo de editor es el propietario de la Web The New York Times.
- Un *anunciante* es una entidad que desea mostrar anuncios en uno de los espacios ofrecidos por un editor y está dispuesta a pagar por ello. Los anunciantes suelen contratar los servicios de una o varias *plataformas de anuncios* (se describen a continuación), que son los responsables de mostrar sus anuncios en los sitios de los editores. Como explicaremos más adelante en la Sec. 2.2, existen dos modelos de plataforma publicitaria, que permiten a los usuarios tener dos roles diferentes. En el enfoque tradicional (aunque prevalente), los anunciantes indican los tipos de anuncios más adecuados para sus campañas, es decir, a qué usuarios quieren que se muestren sus anuncios. Por ejemplo, un anunciante puede querer que la plataforma de anuncios sirva sus anuncios a una audiencia interesada en política o a personas que viven en Francia. Los anunciantes también deben especificar la cantidad de dinero que están dispuestos a pagar cada vez que se muestran sus anuncios, y cada vez que los usuarios hacen clic en ellos³. Por el contrario, en el modelo más reciente de *real-time bidding* (RTB), las plataformas publicitarias permiten a los anunciantes realizar ofertas por cada impresión de anuncio en el momento en que el navegador del usuario carga una página. Este modelo permite a los anunciantes tomar sus propias decisiones en lugar de confiar en un intermediario para tomar decisiones por ellos [17].
- Una *plataforma publicitaria* o *plataforma de anuncios* es un grupo de entidades que conecta a los anunciantes con los editores, es decir, recibe anuncios de los anunciantes y los coloca en los espacios disponibles de los editores. Para ello, las plataformas de anuncios rastrean y perfilan a los usuarios con el objetivo de adaptar los anuncios a sus intereses, ubicación y otros datos personales. Como describiremos con mayor detalle en la siguiente subsección, las plataformas de anuncios tradicionales realizan esta segmentación por su cuenta, de acuerdo con los requisitos y objetivos de la campaña especificados por los anunciantes. Las plataformas publicitarias basadas en RTB, por otro lado, comparten ciertos datos de rastreo de usuarios con los anunciantes, quienes deben decidir si alguno de esos usuarios encaja con una determinada campaña, y si es el caso, deben apostar para mostrarles el anuncio correspondiente. Algunos ejemplos de plataformas publicitarias incluyen DoubleClick, Gemini y Bing Ads, propiedad respectivamente de Google, Yahoo! y Microsoft.

2.2 Proceso de Entrega de Anuncios

Sin pérdida de rigor, a lo largo de este trabajo asumiremos un modelo de publicidad en línea compuesto principalmente por las tres entidades descritas en la subsección anterior. En estos términos,

2. La complejidad del ecosistema publicitario a menudo se ilustra en conferencias con el diagrama disponible en [19].

3. En la terminología de la publicidad en línea, estas cantidades se conocen como el coste por impresión (CPI) y el coste por clic (CPC), respectivamente.

el proceso de publicación de anuncios comienza con los editores, quienes incorporan en sus sitios un enlace a las plataformas de anuncios con las que desean trabajar. El resultado es el siguiente: cuando un usuario visita uno de esos sitios Web y carga el código de la página, su navegador se dirige inmediatamente a todos los enlaces embebidos. Luego, a través de las cookies de terceros, fingerprints del navegador u otras tecnologías de seguimiento, la plataforma publicitaria puede rastrear la visita del usuario a éste y cualquier otro sitio relacionado con él.

Como uno puede intuir, la capacidad de rastrear usuarios en la Web es de suma importancia para las plataformas publicitarias: les permite descubrir la página Web que se está visitando y, por lo tanto, su contenido; la ubicación del usuario a través de su dirección IP; y, lo que es más importante, sus intereses de navegación Web. Después, todos esos datos sobre el usuario son los que permiten a las plataformas publicitarias servir anuncios personalizados.

Para llevar a cabo esta tarea, la gran mayoría de las plataformas utilizan algoritmos personalización propietarios [17]. Los datos de usuario mencionados anteriormente, y los objetivos de las campañas y presupuestos de los anunciantes son las variables de entrada de estos algoritmos, que son responsables de seleccionar qué anuncio se mostrará en un espacio publicitario en concreto. Evidentemente, su objetivo principal es maximizar los ingresos de las plataformas publicitarias y satisfacer la demanda de los anunciantes.

Como anticipamos en la Sec. 2.1, recientemente ha aparecido una nueva clase de plataformas de anuncios que delegan este proceso de personalización a terceros partes, quienes compiten en subastas en tiempo real por la impresión de sus anuncios. Las plataformas de anuncios que se basan en este esquema generalmente comparten información sobre el usuario con estas partes, para que puedan decidir si apostar o no por una impresión de anuncio. Típicamente, las entidades que participan en estas subastas son grandes agencias publicitarias que representan pequeños y medianos anunciantes⁴, y plataformas publicitarias tradicionales que desean vender el inventario remanente. Este esquema de publicación de anuncios se llama RTB y su principal ventaja, en comparación con las plataformas publicitarias tradicionales, es permitir que los anunciantes (u otros que actúen en su nombre) compren impresiones individuales sin tener que confiar en los algoritmos de personalización de la plataforma publicitaria. En otras palabras, los anunciantes pueden decidir ahora si un usuario en particular es la persona adecuada para mostrarle sus anuncios.

Finalmente, independientemente del tipo de plataforma de anuncios involucrada (es decir, basada en RTB o no), el proceso de publicación de anuncios concluye cuando el anuncio seleccionado es mostrado en el navegador Web del usuario, un último paso que puede implicar una red de entrega de contenidos (“content-delivery network”).

Por último, aunque no menos importante, nos gustaría poder de manifiesto que el modelo de publicidad descrito aquí — y considerado en este trabajo — corresponde a *publicidad de venta indirecta*, también llamada publicidad basada en red o de terceros. Esto contrasta con el modelo de *venta directa*, donde editores y anunciantes negocian directamente, sin la mediación de plataformas publicitarias. En este último caso, en su mayoría encontramos sitios Web populares que venden espacio publicitario directamente a grandes anunciantes. Los anuncios que se publican de esta manera son esencialmente no personalizados y, a menudo, se muestran en sitios Web donde los productos y servicios anunciados están relacionados con sus contenidos. Esto se debe principalmente a que la capacidad de un editor para rastrear y perfilar a los usuarios está limitada solo a su propio sitio Web y quizás a algunos socios. Por ejemplo, el sitio Web del New York Times puede rastrear a los usuarios también a través del International Herald Tribune, propiedad de su grupo de medios. Sin embargo, esta capacidad de seguimiento es ridícula cuando la comparamos con los 2 millones de sitios a los que pueden acceder las plataformas de anuncios de Google [25].

4. Una clase especial de estas agencias son las *demand-side platforms* (DSP), que son sistemas que automatizan la compra de publicidad en línea en nombre de los anunciantes.

2.3 Tipos de Anuncios

Los anuncios entregados a través de anuncios de venta indirecta permiten a los anunciantes personalizar diferentes aspectos de un usuario Web. Los *objetivos de personalización* más populares incluyen la publicación de anuncios adaptados a la página Web visitada, su ubicación geográfica y sus intereses de navegación Web. Dependiendo del objetivo elegido por un anunciante, los anuncios se clasifican como anuncios *contextuales*, *basados en la ubicación*, *basados en intereses* y *no personalizados*. Ocasionalmente, nos referiremos a estos cuatro tipos de anuncios como *clases de anuncios*. A continuación, detallamos brevemente cada uno de ellos.

- Anuncios *contextuales*. Los anunciantes pueden llegar a su público a través de publicidad contextual y semántica, al personalizar anuncios relacionados con el contenido del sitio Web donde se mostrarán. Un ejemplo de dicha estrategia de personalización sería una compañía de seguros de salud que desea mostrar sus anuncios en sitios Web cuyo contenido se clasifica como “salud y fitness”.
- Anuncios basados en la *ubicación*. Se generan en función de la ubicación del usuario, por ejemplo, dada por el GPS de su smartphone o tableta, y también de acuerdo con los puntos de acceso Wi-Fi y la dirección IP de su ordenador o dispositivo del usuario. Los anuncios con personalización geográfica permiten a los anunciantes lanzar campañas dirigidas a usuarios dentro de una determinada ubicación geográfica.
- Anuncios basados en *intereses o perfil de navegación*. Los anunciantes también pueden personalizar a los usuarios en función de sus intereses de navegación Web. Por lo general, dichos intereses se deducen de las páginas rastreadas por las plataformas de anuncios y otras compañías de rastreo que pueden compartir esta información con las primeras. La secuencia de sitios Web explorados por un usuario y rastreados efectivamente por una plataforma publicitaria o rastreador se conoce como *clickstream* del usuario. En la práctica, esta es la información utilizada por la industria de la publicidad en línea para construir el perfil de intereses de un usuario [26], [27], [28], [21], [29], [17].
- *Anuncios genéricos*. Los anunciantes también pueden especificar secciones de los sitios Web de los editores (entre los que se asocian con la plataforma de anuncios) donde se mostrarán sus anuncios. Los anuncios servidos de esta manera no están necesariamente relacionados con el contenido del sitio Web, sino que simplemente pueden responder a alguna coincidencia entre los intereses de los visitantes y los productos anunciados. Estos anuncios también se denominan *anuncios genéricos* debido a que no utilizan ningún dato del usuario.

Un aspecto importante de las clases de anuncios descritos anteriormente es que los tres primeros no son mutuamente excluyentes. En otras palabras, a excepción de los anuncios genéricos, los anuncios pueden personalizarse simultáneamente en función del contenido, la ubicación y los intereses. Por tanto, cuando nos refiramos a anuncios basados en intereses, queremos decir que están personalizados *al menos* teniendo en cuenta los datos de navegación. Nos referiremos a los anuncios basados en contenido y ubicación de manera análoga.

En la industria de la publicidad en línea, una estrategia de personalización muy popular es la redirección (“retargeting”), que permite a los anunciantes mostrar anuncios a aquellos usuarios que previamente visitaron su página Web. Por ejemplo, después de haber navegado por el sitio Web de Apple, a un usuario se le pueden mostrar anuncios sobre una nueva versión de iPhone cuando visita otros sitios, en un intento por recuperarlos.

Concluimos esta subsección con un ejemplo real de cómo los anunciantes pueden enviar su publicidad. La Fig. 1 muestra el panel de configuración disponible en la plataforma de anuncios de Yahoo!, mediante el cual los anunciantes pueden definir sus audiencias objetivo según la ubicación, edad, sexo, intereses⁵ y contexto (no se muestra en esta figura). Para cada campaña, el anunciante

5. Otras plataformas como la de Google, que permiten a los anunciantes especificar Se mostrarán restricciones adicionales, como la hora del día en que se mostrarán los anuncios, su frecuencia de aparición a un mismo usuario y ubicaciones de anuncios específicas.

The screenshot shows the 'Define your audience' interface. It is divided into two main sections: 'Define your audience' and 'Settings and budget'.

Define your audience:

- Locations:** A dropdown menu showing 'France' with a minus sign, and a text input field below it with the placeholder 'Add a country, state or city'.
- Age:** A row of buttons: 'All' (selected), '18-24', '25-34', '35-44', '45-54', '55-64', and '65+'.
- Gender:** A row of buttons: 'All' (selected), 'Female', and 'Male'.
- Interests:** A list of interest categories: 'Credit and Debt', 'Chronic Pain', and 'Dating', each with a minus sign. Below the list is a 'Browse' button and a 'Browse' link.

Settings and budget:

- Cost per click (CPC):** A dropdown menu showing 'EUR' and a text input field with the value '0.79'.
- Campaign budget:** A dropdown menu showing 'EUR' and a text input field with the value '25000'. To the right, there are two buttons: 'Per day' and 'In total' (selected).
- Schedule:** Two radio buttons: 'Start running ads immediately' (unselected) and 'Set a start and end date' (selected). Below the radio buttons are two date pickers: 'Start' with the date 'Nov 15, 2018' and 'End' with the date 'Nov 30, 2018'.

Fig. 1: Gemini, la plataforma publicitaria de Yahoo!, ofrece a los anunciantes la posibilidad de segmentar anuncios en función de una serie de parámetros, incluidos los intereses de navegación del usuario, que se eligen de un conjunto predefinido de 281 categorías. Las categorías seleccionadas en este ejemplo muestran simplemente el tipo de información personal involucrada en estas prácticas y, por lo tanto, no reflejan una campaña de marketing real.

debe configurar todas estas variables de manera adecuada, evidentemente con una restricción en el presupuesto de publicidad.

2.4 Coincidencia de Cookie y Real-Time Bidding

Esta última subsección explica con mayor detalle algunos aspectos operativos de RTB, un esquema de publicación de anuncios que representa el 20% de las ventas de anuncios digitales [17], aunque se espera que sea el paradigma publicitario dominante en los próximos años [30].

En la Sec. 2.2, mencionamos que las plataformas de anuncios RTB comparten información del usuario con ciertas entidades, que luego podrán ofertar por la impresión de sus anuncios. Los participantes de dichas subastas son típicamente agencias de anunciantes, DSPs y plataformas publicitarias tradicionales. Para facilitar el intercambio de información con estos apostantes, RTB utilizan un protocolo de *coincidencia de cookie* ("cookie matching").

En general, la coincidencia de cookies es un proceso mediante el cual dos dominios diferentes vinculan los IDs (identificadores) de usuario que han asignado a un mismo usuario y que almacenan en sus respectivas cookies. Típicamente, el proceso se realiza de la siguiente manera. Cuando un usuario visita uno de los dos dominios, este dominio redirige su navegador al otro, e incluye su ID de usuario como un parámetro en la URL. Luego, al recibir la solicitud, este último dominio vincula el ID con su propio ID para este usuario [31].

La coincidencia de cookies se aplica habitualmente en RTB, donde permite que la plataforma publicitaria y el apostante emparejen sus cookies para un usuario en particular [32]. Generalmente, el protocolo se ejecuta solo si el apostante gana una subasta y entrega su anuncio a ese usuario en particular. El protocolo permite al ganador buscar al usuario (si está presente) en su propia base de datos. También, si se producen más subastas de anuncios para este usuario, el apostante sabrá que la información del usuario proporcionada en esas subastas se refiere al mismo usuario. Debemos enfatizar que esto es bajo el supuesto de que este apostante se encuentre entre los destinatarios de las solicitudes de subastas enviadas por la plataforma publicitaria.

Después de describir la tecnología subyacente de RTB, a continuación examinamos brevemente el funcionamiento general del esquema de Google, probablemente el más representativo. La descripción siguiente, sin embargo, también es válida para otras plataformas de anuncios basadas en RTB.

Cuando un usuario visita un sitio Web con un espacio publicitario servido a través de RTB, se envía una solicitud HTTP a la plataforma publicitaria, que posteriormente envía *solicitudes de subastas* a los posibles participantes. Es importante destacar que la cantidad y el tipo de participantes involucrados pueden variar por subasta, a discreción de la plataforma publicitaria. Dentro de la solicitud de subasta, la plataforma publicitaria generalmente incluye los siguientes datos: la URL de la página visitada por el usuario; la categoría temática de la página; la dirección IP del usuario o partes de ella; y otra información relacionada con su navegador Web [33]. Acompañando a esta información, la plataforma de anuncios de Google incorpora un identificador de usuario específico suyo.

Al recibir la solicitud de subasta, el potencial apostante puede identificar al usuario dentro de su propia base de datos a través de la cookie o el identificador. Esto sucede cuando el protocolo de coincidencia de cookies se ha ejecutado previamente para este usuario. Gracias a dicha cookie o identificador, el apostante puede rastrear al usuario en las páginas Web en las que ha sido invitado a participar. Por lo tanto, a partir de esas páginas rastreadas, el apostante puede crear un perfil⁶, que tal vez complementa la información de rastreo u otros datos personales que pueda tener sobre el usuario.

El precio de la oferta se establece luego en función de los objetivos de campaña del apostante, es decir, si se dirige a los usuarios que visitan determinadas categorías del sitio, navegan desde una ubicación determinada y/o tienen algún perfil específico. Para evaluar si la impresión del anuncio cumple con dichos objetivos, el apostante se basa en el perfil mencionado y en la información incluida en la solicitud de subasta. Si está interesado, el apostante envía una puja a la plataforma de anuncios, quien, en un último paso, permite al ganador de la subasta entregar el anuncio al usuario. Conviene destacar que todo este proceso de recopilación de datos del usuario, y la posterior subasta y entrega del anuncio se realiza en solo decenas de milisegundos.

3 DETECCIÓN DE ANUNCIOS BASADOS EN EL PERFIL DE USUARIO Y SU SINGULARIDAD

Como describimos en la sección anterior, las plataformas de anuncios, las empresas de rastreo y también los anunciantes recopilan información sobre los usuarios (por ejemplo, las páginas visitadas y su ubicación) mientras navegan por la Web. Posteriormente, estos y otros datos se utilizan para mostrar anuncios personalizados al contenido de las páginas visitadas, su ubicación geográfica actual y/o sus intereses. También mencionamos que las plataformas publicitarias pueden ofrecer anuncios genéricos.

Esta sección propone un modelo matemático que tiene como objetivo cuantificar hasta qué punto la información recopilada sobre los *intereses de navegación* de un usuario es explotada posteriormente por la industria de la publicidad en línea para ofrecerles anuncios. El modelo propuesto se centra en la detección de anuncios basados en intereses, ya que son el resultado, y probablemente la causa, de rastrear y perfilar los hábitos de navegación de los usuarios en Internet, a menudo sin su conocimiento [35] y consentimiento. Es importante señalar que el modelo está restringido a la publicidad basada en la red, ya que la capacidad de los editores para rastrear y perfilar a los usuarios, en general, se limita a sus sitios.

Además de determinar si los anuncios mostrados pueden haber sido personalizados de acuerdo con un perfil de navegación, esta sección aborda otra pregunta ineludible relacionada con la segmentación de perfiles: *¿qué tan único* somos vistos a través de los ojos de las compañías que nos muestran anuncios? Como veremos en la Sec. 3.3, el riesgo de perfilado, así como la singularidad de los perfiles contruidos por estas compañías, están estrechamente relacionados con el riesgo de reidentificación.

6. DoubleClick especifica que, a menos que un apostante gane una impresión determinada, no debe usar los datos de esa impresión para perfilar a los usuarios [34]. Sin embargo, debido a que no se habilita ningún mecanismo activo de prevención, nada impide que un participante a dicha subasta haga un uso indebido de los datos del usuario.

En las siguientes secciones, proporcionaremos la base conceptual y la estructura operativa fundamental de dos detectores que tienen como objetivo (i) identificar anuncios basados en perfiles; y (ii) arrojar luz sobre la singularidad de los perfiles recopilados por las entidades que participan en el proceso de entrega de anuncios. Con ello, pretendemos dar un primer paso hacia el estudio de la relevancia comercial de nuestro historial de navegación, cuantificando su impacto real en la privacidad del usuario. Más adelante, en la Sec. 4 presentaremos *MyAdChoices*, una extensión de navegador Web que utiliza estos dos detectores para proporcionar transparencia en dicho proceso y habilitar el bloqueo de anuncios inteligente y selectivo.

3.1 Conceptos Básicos sobre Estadística y Teoría de Información

Esta sección establece aspectos de notación de teoría de la información.

El espacio medible en el que una *variable aleatoria* (r.v.) toma valores se llamará *alfabeto*. Sin pérdida de generalidad, asumiremos que el alfabeto es discreto. Seguiremos la convención de usar letras mayúsculas para las v.a.s, y letras minúsculas para los valores particulares que tomen. La función de masa de probabilidad (PMF) p de una v.a X es una función que mapea los valores tomados por X a sus probabilidades. Conceptualmente, una PMF es un histograma de frecuencias relativo.

A lo largo de este trabajo, las PMFs serán subindexadas por sus correspondientes v.a.s en caso de ambigüedad. En consecuencia, tanto $p(x)$ como $p_X(x)$ denotan el valor de la función p_X en x . Ocasionalmente, nos referiremos a la función p por su valor $p(x)$. Usamos las notaciones $p_{X|Y}$ y $p(x|y)$ equivalentemente.

Adoptamos la misma notación para las cantidades de teoría de información utilizadas en [36]. De acuerdo con ello, el símbolo D denotará la entropía relativa o divergencia KL. Recordamos brevemente este concepto. Todos los logaritmos son base 2.

Dadas dos distribuciones de probabilidad $p(x)$ y $q(x)$ sobre el mismo alfabeto, la *divergencia KL* $D(p \parallel q)$ se define como

$$D(p \parallel q) = \sum_x p(x) \log \frac{p(x)}{q(x)}.$$

La divergencia KL a menudo se denomina *entropía relativa*, ya que puede considerarse como una generalización de la entropía de Shannon de una distribución, relativa a otra.

Aunque la divergencia KL no es una distancia en el sentido matemático del término, porque ni es simétrica ni satisface la desigualdad triangular, sí que proporciona una medida de discrepancia entre distribuciones, en el sentido de que $D(p \parallel q) \geq 0$, con igualdad si, y solo si, $p = q$.

3.2 Detección de Anuncios basados en el Perfil de Navegación

Una de las funcionalidades clave de nuestro sistema es la detección de anuncios basados en perfiles, es decir, anuncios que han sido entregados a un usuario debido a sus intereses de navegación y, además, pero no necesariamente, a su ubicación y la última página Web visitada. Esta sección propone un modelo matemático para la identificación de estos anuncios, que parte de resultados fundamentales de estimación estadística y optimización robusta.

3.2.1 Modelo de Distribución de Anuncios basados en Intereses

Modelamos los anuncios entregados por una plataforma de anuncios (basados en RTB o no) a un usuario particular como v.a. independiente que toma valores en un alfabeto finito común de categorías o temas, a saber, el conjunto $\mathcal{X} = \{1, \dots, n\}$ para algún entero $n > 1$. Conviene destacar que nuestro modelo abarca las cuatro clases de anuncios descritos en la Sec. 2.3. El hecho de que cada anuncio esté asociado a una categoría de interés no significa que estemos considerando solo anuncios basados en intereses. Por ejemplo, un anuncio basado en contenido que se muestra en el sitio Web www.webmd.com se clasificará necesariamente en una categoría de interés relacionada con salud. Evidentemente, los anuncios basados en la ubicación y en la ubicación pueden asignarse a cualquiera de las n categorías asumidas en este trabajo.

Como se comentó en la Sec. 2.2, el proceso de publicación de anuncios tiene en cuenta una amplia gama de variables. Estas variables incluyen el rastreo y la creación de perfiles sobre el usuario en cuestión, el editor que se visita, los anunciantes y sus campañas correspondientes y, según el tipo de plataforma publicitaria, las ofertas de los participantes en la subasta o los criterios de la plataforma para maximizar sus ingresos.

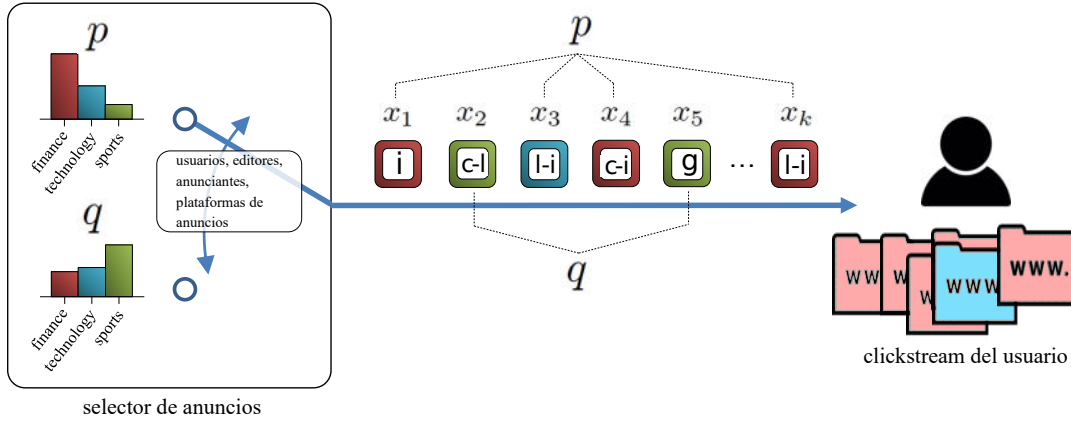


Fig. 2: Un selector de anuncios (por ejemplo, una plataforma de anuncios tradicional) muestra k anuncios en el navegador del usuario cuando navega por la Web. Las categorías de interés de los anuncios entregados se modelan como una secuencia de v.a.s independientes en las $n = 3$ categorías. Las categorías observadas, es decir, $(x_i)_{i=1}^k$, pueden verse como generadas por una fuente que conmuta entre las PMF p y q . El cambio entre anuncios basados en intereses (es decir, "i", "c-i", "l-i" y "cl-i" por un lado, y anuncios no basados en intereses (es decir, "c", "l", "g" y "c-l" en la otra, está determinado por una serie de parámetros relacionados con el usuario, los editores, los anunciantes y la plataforma de anuncios.

En nuestro modelo matemático, caracterizamos el proceso de publicación de anuncios llevado a cabo por una plataforma como una caja negra, cuyas entradas son las variables mencionadas anteriormente y cuyas salidas son los anuncios seleccionados. Explicamos en la sección de antecedentes que las plataformas de anuncios tradicionales son las que seleccionan el anuncio que se va a mostrar, mientras que en la publicidad basada en RTB la elección la realiza quien gana la apuesta, ya sea una agencia de publicidad o una plataforma de anuncios tradicional. Por simplicidad, en adelante usaremos el término *selector de anuncios* para referirnos de manera genérica a la entidad particular que impone la selección de un anuncio.

Para cada usuario y para cada espacio publicitario, los anuncios publicados pueden clasificarse como genéricos, basados en contenido, ubicación y interés. De acuerdo con los objetivos de las campañas de los anunciantes correspondientes, apreciamos que, de estas cuatro clases de anuncios, solo podemos tener ocho combinaciones posibles de esas clases. Si indicamos cada una de las clases de anuncios por su primera letra, el conjunto de todas estas combinaciones resulta ser

$$\mathcal{G} = \{c, l, i, g, cl, ci, li, cli\},$$

donde el elemento "c-l" representa un anuncio que se ha personalizado según el contenido y la ubicación. En otras palabras, \mathcal{G} incluye todas las combinaciones de objetivos de segmentación que un anunciante puede elegir.

Mencionamos en la Sec. 2.2 que los perfiles de usuario se crean esencialmente a partir del clickstream, es decir, de las páginas Web rastreadas. Para $k \gg 1$, sea $(X_i)_{i=1}^k$ la secuencia de anuncios que un selector de anuncios (por ejemplo, una plataforma de anuncios tradicional) entrega a un usuario en particular durante varias sesiones de navegación. Nuestra caracterización de este proceso de entrega de anuncios se deriva de la observación intuitiva de que, si pudiéramos descartar de dicha secuencia todos los anuncios excepto los basados en intereses, la distribución empírica [36] de las categorías de interés observadas se asemejaría en gran medida a los intereses de navegación del usuario, o equivalentemente, a su clickstream.

De acuerdo con esta observación y sin pérdida de generalidad, modelamos la secuencia de anuncios salientes, clasificados en categorías de interés, como la salida de una fuente de anuncios que *alterna* entre dos distribuciones de probabilidad, a saber:

- una distribución de categoría de intereses p que refleja el conocimiento que tiene el selector de anuncios de los *intereses del usuario*;
- y otra distribución de categorías de intereses q que representa el complemento de la distribución anterior y, por lo tanto, corresponde a (las categorías de intereses de) aquellos anuncios clasificados como no basados en intereses, es decir, *contextual, basado en la ubicación y genérico*.

Naturalmente, el modelo descrito anteriormente captura solo un aspecto del proceso de publicación de anuncios: refleja la selección de las categorías de interés de los anuncios dentro del conjunto \mathcal{X} , un paso que modelamos a través de las distribuciones p y q cuando la clase de anuncios se basa respectivamente en intereses y no en intereses. El modelo propuesto está respaldado por la suposición razonable de que las categorías de intereses acumulados de los anuncios basados en intereses se aproximarán, muy probablemente, a los intereses del usuario, o más precisamente, el clickstream que posee el selector de anuncios.

Por lo tanto, nuestro modelo no captura otros aspectos del proceso de publicación de anuncios, como la forma en que se elige una combinación particular de clase de anuncios de \mathcal{G} . Sin embargo, con esto reflejamos el simple hecho de que las categorías de interés de los anuncios salientes pueden distribuirse de acuerdo con los datos de navegación del usuario, ya sean parciales (o completos), o cualquier otra información que no incluya esos datos de navegación. Este modelo simplificado de publicación de anuncios basado en categorías de intereses nos permitirá en la siguiente subsección estimar la clase de anuncios elegida por el selector de anuncios, o más exactamente, si los anuncios entregados se clasifican como basados en intereses o no. La Figura 2 ilustra cómo modelamos este aspecto del proceso de publicación de anuncios.

3.2.2 Test de Hipótesis Binario

Asumiendo dicho modelo en el lado de la plataforma publicitaria, en el lado del usuario, nuestro objetivo es determinar si un anuncio, previamente clasificado en una categoría de interés, se ha mostrado al usuario según sus intereses pasados de navegación Web o no. Formalmente, podemos considerar esto como un problema binario de *test de hipótesis* [36] entre dos hipótesis, en concreto, si los datos (es decir, la categoría del anuncio mostrado) se han generado de acuerdo con la distribución p o q . A continuación damos más detalles sobre estas dos distribuciones.

Recuerde que, para un usuario y espacio publicitario en particular, el *selector de anuncios* es la entidad que finalmente decide qué anuncio se muestra a ese usuario en ese espacio publicitario. En el caso de las plataformas publicitarias tradicionales, el selector de anuncios es la propia plataforma publicitaria. En RTB, por el contrario, el selector de anuncios es el apostador que gana la subasta para mostrar su anuncio.

Como se describe en la Sec. 3.2.1, la PMF p representa el conocimiento que dicho selector de anuncios tiene sobre los intereses de navegación del usuario. A partir de ahora, nos referiremos a esta distribución como *perfil de interés* del usuario, teniendo en cuenta que es específico del selector de anuncios en cuestión.

En la práctica, estos perfiles generalmente se crean a partir de los sitios Web rastreados o *clickstream observado* [26], [27], [28], [21], [29], [17]. El clickstream disponible para un selector de anuncios, sin embargo, no necesariamente debe ser el resultado de un rastreo directo del usuario. Por ejemplo, las plataformas publicitarias pueden rastrear a los usuarios por su cuenta a través de sus cookies; y no satisfechos con eso, es posible que también deseen complementar dicha información con datos de rastreo de otras plataformas publicitarias o rastreadores. Por el momento, no especificaremos cómo, en nuestro modelo, el selector de anuncios perfila a un usuario a partir de su clickstream. Solo supondremos que los perfiles están representados como PMFs, como se suele hacer en la literatura [27], [37], [21], [23].

Claramente, dependiendo de la capacidad del selector de anuncios para rastrear a los usuarios a través de la Web (de forma independiente o no), el perfil p se asemejará, en mayor o menor medida, a sus intereses reales. Denotamos por t el perfil de interés resultante de *real clickstream*, es decir, todos los sitios Web visitados por un usuario. Ocasionalmente nos referiremos a p y t como *perfiles observado* y *real*, respectivamente. La Fig. 3 amplía el modelo de segmentación de anuncios que se

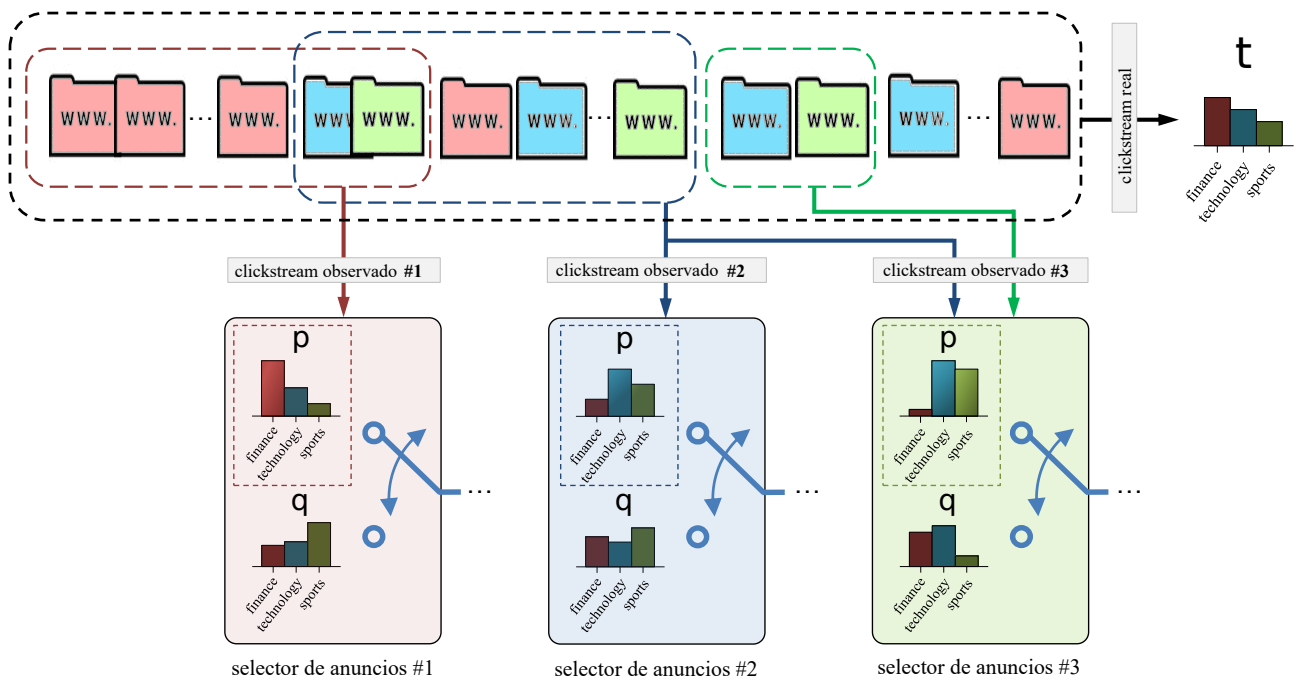


Fig. 3: Mostramos cómo tres selectores de anuncios rastrean a un usuario a través de diferentes sitios Web. Los selectores de anuncios 1 y 2 podrían representar dos plataformas de anuncios que se superponen con sus secuencias de clics observadas. Esto reflejaría una situación común para grandes plataformas de anuncios como Google AdSense y OpenX. El selector de anuncios 3, por otro lado, podría ser una pequeña empresa de publicidad. Debido a su capacidad limitada para rastrear a los usuarios por su propia cuenta, este último selector de anuncios puede decidir adquirir datos de rastreo del selector de anuncios 2. Sin embargo, independientemente de los datos intercambiados, ninguno de los tres selectores de anuncios podrá obtener el clickstream real.

muestra en la Fig. 2, para reflejar el hecho de que p se construye a partir de la secuencia de clics observada y , y por lo tanto, no puede capturar el usuario perfil de intereses reales t .

La distinción entre estos dos perfiles también se empleará más adelante en la Sec. 4 para modelar dos escenarios posibles con respecto al seguimiento e intercambio de datos del clickstream: por un lado, un escenario paranoico en el que se realiza un seguimiento de los usuarios en cada página visitan y dichos datos de seguimiento se intercambian entre todas las entidades que sirven anuncios. Y, por otro lado, un escenario de referencia donde p se crea fundamentalmente a partir de la secuencia de clics que un selector de anuncios puede obtener por sí mismo, a través de cookies u otras tecnologías de seguimiento, sin depender de los datos de seguimiento de otras fuentes.

Para llevar a cabo nuestras pruebas de hipótesis, también deberemos estimar la distribución q . Con este fin, consideramos un entorno donde no se realiza ningún seguimiento, de manera similar a cuando los usuarios habilitan el modo privado del navegador Web. Recuerde que esta PMF es la distribución de categorías de interés de aquellos anuncios que no están basados en perfiles, es decir, aquellos clasificados como “c”, “l”, “g” y “c-l”. Como, a excepción de los anuncios genéricos, estos anuncios dependerán de la ubicación del usuario y de las páginas visitadas durante esta sesión de seguimiento gratuito, q será específico para cada usuario en particular. Para estimar esta distribución en el lado del usuario, será necesario capturar la categoría de todos los anuncios recibidos, bajo la suposición razonable de que, cuando los usuarios navegan en modo privado, no se utilizan las páginas visitadas para personalizar anuncios. En la Sec. 4.2.2.1, describiremos más específicamente cómo nuestro detector estimará esta PMF.

3.2.3 Perfiles de Intereses a Corto y Largo Plazo

En secciones anteriores, señalamos que los perfiles de interés de los usuarios se crean principalmente a partir de la categorización de los sitios Web visitados. También comentamos que los perfiles se modelan esencialmente como PMFs, es decir, como histogramas de frecuencias de sitios visitados a lo largo en un conjunto de categorías de interés. En esta subsección, examinamos brevemente un aspecto crucial de dicho modelo de usuario: la importancia que los selectores de anuncios pueden

otorgar a los intereses recientes en comparación con los acumulados durante un largo período de tiempo.

Desde la perspectiva de la segmentación basada en perfil, la necesidad de ponderar el clickstream es evidente. Un historial de navegación con pocas páginas aunque visitadas recientemente puede ser suficiente para mostrar anuncios de productos que no requieren mucha reflexión, como comprar una película en Google Play. Pero otro tipo de transacciones, como matricularse en un posgrado en línea, pueden necesitar un historial de navegación más prolongado para garantizar una cierta probabilidad de conversión. En terminología de marketing en línea, la conversión generalmente significa el acto de convertir a los visitantes del sitio Web en clientes de pago [29].

Dependiendo de la ventana de tiempo elegida, los perfiles de usuario pueden clasificarse como perfiles *a corto plazo* y *a largo plazo*. Los primeros representan los intereses actuales e inmediatos del usuario, mientras que los últimos capturan intereses que no están sujetos a cambios frecuentes [38]. En general, diferentes sistemas de mercadotecnia basados en intereses pueden contemplar diferentes ventanas de tiempo para construir perfiles. Muchos sistemas comerciales optan por perfiles a largo plazo, mientras que otros utilizan clickstreams cortos y recientes. Algunos estudios parecen no ponerse de acuerdo en cuál es la mejor estrategia. Por ejemplo, [29] proporciona evidencias de que los historiales largos de navegación dan mejores resultados, mientras que otros muestran lo contrario.

Como veremos en la Sec. 3.2.4.1, nuestro sistema de detección capturará la incertidumbre asociada con la ventana de tiempo utilizada por un selector de anuncios. Como en la práctica es imposible determinar este parámetro, consideraremos las clases de incertidumbre de los perfiles de usuario. Estas clases nos permitirán caracterizar las distintas opciones que un selector de anuncios podría haber elegido para crear un perfil, y nos conducirán al diseño de un detector robusto óptimo.

3.2.4 Detección Óptima bajo Incertidumbre de Anuncios basados en Intereses

En esta sección, formulamos el problema de diseñar un detector de anuncios basado en intereses como un problema de optimización minimax robusto. Para ello, seguiremos la metodología desarrollada por [39], [40].

Sea X una v.a. que modela la categoría a la que pertenece un anuncio. Denota por H la v.a. que representa las dos posibles hipótesis sobre la distribución de la categoría observada X . En nuestra notación, $H = 1$ indica que el anuncio está basado en el perfil (primera hipótesis), y $H = 2$ que no está basado en el perfil (segunda hipótesis). Dicho de otro modo, X condicionado en H tiene PMF p cuando $H = 1$ y q cuando $H = 2$. Para compactar nuestra notación, denotamos por $P \in \mathbb{R}^{n \times 2}$ la matriz que tiene p y q como columnas.

Un *estimador aleatorio* o *detector* \hat{H} de H es una regla de decisión probabilística determinada por la probabilidad condicional de \hat{H} dada X , es decir, $p_{\hat{H}|X}$. La interpretación de dicho estimador es la siguiente: si se observa que X tiene un valor j , el detector concluye $H = 1$ con probabilidad $p_{\hat{H}|X}(1|j)$, y $H = 2$ con el complemento de esa probabilidad.

Un detector aleatorio también admite una interpretación en términos matriciales, en particular como una matriz $\mathbb{R}^{2 \times n}$, donde la columna j -th corresponde a la distribución de probabilidad de \hat{H} cuando recibimos un anuncio que pertenece a la categoría de interés j . A lo largo de esta sección, usaremos convenientemente esta notación matricial para estimadores, y denotaremos por D la matriz que los define.

El desempeño de una regla de decisión generalmente se caracteriza en términos de sus probabilidades de detección y error. Es posible capturar este desempeño de forma compacta mediante la matriz $M = DP$, cuyo elemento M_{ij} nos da la probabilidad de decidir $\hat{H} = i$ cuando en realidad $H = j$, es decir, $p_{\hat{H}|H}(i|j)$. Los elementos diagonales de esta matriz 2×2 son las probabilidades de acertar. Las probabilidades de error están representadas por los elementos fuera de la diagonal M_{21} y M_{12} , que dan las probabilidades de un falso negativo y un falso positivo, respectivamente. En nuestro contexto, el primer elemento es la probabilidad de concluir que el anuncio no está basado en perfiles cuando en realidad lo está; y el segundo elemento es la probabilidad de decidir que el anuncio se basa en intereses cuando no lo está.

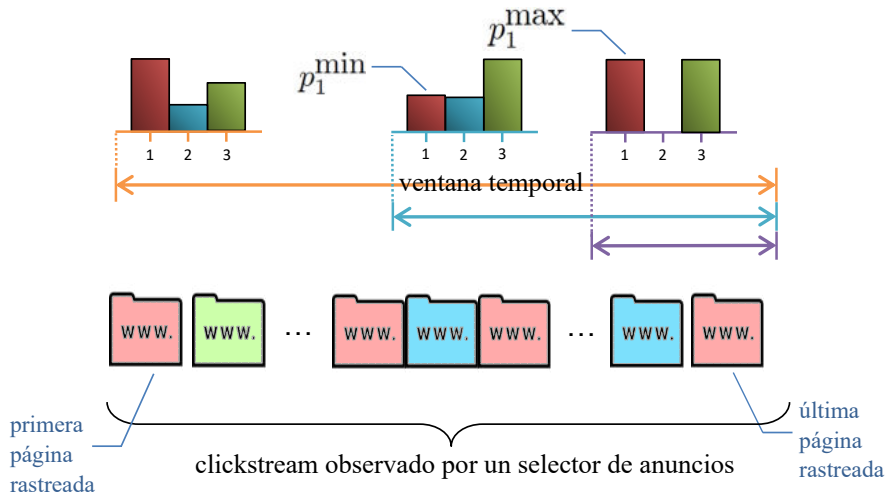


Fig. 4: Los selectores de anuncios pueden crear perfiles de interés basados en las páginas Web rastreadas. Nuestro detector captura todas las opciones posibles que un selector de anuncios puede considerar para calcular esos perfiles de páginas rastreadas. Todas estas opciones están directamente relacionadas con la/s ventana/s de tiempo elegidas, o de manera equivalente, con el número de páginas tomadas de la secuencia de clics observada. Modelamos estas posibles elecciones como intervalos entre los valores de interés mínimo y máximo por categoría.

Nuestro objetivo es diseñar la matriz D que define el detector de anuncios basado en intereses, para que se cumplan ciertos criterios de rendimiento. Entre otros requisitos, podríamos estar interesados en minimizar (maximizar) una de las probabilidades de error (detección), con una restricción en el complemento de la probabilidad objetiva. Además, podríamos considerar minimizar las probabilidades de error o una combinación convexa de ellas, si se dispone de información previa sobre p_H .

3.2.4.1 Estimación Robusta: Independientemente de los criterios elegidos, el problema de este diseño es que requiere el conocimiento completo de las distribuciones de probabilidad definidas por P . Como describimos en la sección anterior, podemos obtener una estimación fiable de q de manera local (es decir, en el lado del usuario), pero no podemos saber cómo los selectores de anuncios construyen el perfil p a partir de su clickstream observado. Algunos selectores de anuncios pueden querer representar los perfiles en base a intereses a corto plazo. Algunos pueden querer trabajar con perfiles más largos y relativamente estables, y otros pueden optar por ambos tipos de modelos. En cualquier caso, la ventana de tiempo empleada por un selector de anuncios es lo que determina los perfiles que se usarán para la personalización de anuncios. Debido a que esta información es desconocida, no es posible tener una especificación precisa de la distribución p , o estimarla de manera fiable.

El problema de estimar una distribución bajo incertidumbre también se ha dado en otros campos y aplicaciones tales como procesado de señal [41], optimización de carteras [42] y redes de comunicaciones [43]. En todos estos casos, las distribuciones de probabilidad se especifican como pertenecientes a conjuntos de distribuciones, típicamente llamados *clases de incertidumbre*. En nuestro caso, la clase de incertidumbre de p viene dada por la longitud mínima y máxima de las ventanas de tiempo que un selector de anuncios puede definir para modelar intereses a corto y largo plazo. En la práctica, la longitud máxima podría corresponder a todo el clickstream, mientras que una ventana de tiempo mínima para perfiles a corto plazo podría ser un día [29], [23].

Para $i = 1, \dots, n$, denotamos por p_i^{\max} el valor de interés máximo p_i estimado por el selector de anuncios, sobre *todas* las ventanas de tiempo posibles que van desde un día hasta la totalidad del clickstream observado⁷. Definimos p_i^{\min} de manera análoga, y modelamos intuitivamente la incertidumbre sobre la distribución p como intervalos entre estos límites superior e inferior. Más específicamente, definimos el conjunto de posibles perfiles de interés como

$$\mathcal{P} = \{p : p^{\min} \preceq p \preceq p^{\max}, \mathbf{1}^T p = 1, p \succeq 0\}, \quad (1)$$

7. En la Sec. 4.2.2.1, veremos que una ventana de tiempo máxima de 1.5 meses puede ser suficiente.

donde el símbolo “ \preceq ” indica desigualdad por componentes, y la última desigualdad y igualdad reflejan el hecho de que p debe ser una PMF.

A nivel conceptual, el poliedro \mathcal{P} captura todos los perfiles posibles que un selector de anuncios puede haber construido al agregar observaciones incrementales de un sitio Web al modelo de intereses. Al calcular los intereses máximos y mínimos observados en todos estos modelos incrementales, y definir intervalos de valores de interés entre estos dos extremos, obtenemos una clase de incertidumbre que refleja *cualquier* posible decisión tomada por el selector de anuncios con respecto a la ventana de tiempo. Nos gustaría poner de relieve que la clase de incertidumbre \mathcal{P} también incluye la posibilidad de que un selector de anuncios esté utilizando más de un perfil — con diferentes ventanas de tiempo — para un mismo usuario. La Fig. 4 ilustra la incertidumbre alrededor de la/s ventana/s de tiempo seleccionada/s.

Una forma posible de diseñar un estimador cuando una distribución de probabilidad debe pertenecer a una clase de incertidumbre es contemplar el rendimiento del *peor caso* en esta clase. En este caso, se dice que la regla de decisión resultante es *robusta* a las incertidumbres en la distribución de probabilidad [44]. Siguiendo la notación de [39], definimos la matriz de desempeño en el peor de los casos M^w asociada con un detector robusto como

$$M_{ij}^w = \sup_{p \in \mathcal{P}} M_{ij},$$

para $i, j = 1, 2$, con $i \neq j$, y

$$M_{ii}^w = \inf_{p \in \mathcal{P}} M_{ii},$$

para $i = 1, 2$. En términos generales, los elementos fuera de la diagonal de esta matriz nos dan la mayor probabilidad de errores en todo $p \in \mathcal{P}$. Las entradas diagonales, por otro lado, dan las probabilidades de detección más pequeñas posibles. En base a estas últimas probabilidades, podemos definir la *probabilidad de error de peor caso* como $P_i^w = 1 - M_{ii}^w$, que representa la mayor probabilidad de error sobre la clase de incertidumbre cuando $H = i$. Claramente, notamos que $M_{12}^w = M_{12}$ y $M_{22}^w = M_{22}$, ya que en nuestro caso la incertidumbre solo es en p .

3.2.4.2 Diseño Minimax: En esta subsección, especificamos el diseño de un detector robusto de anuncios basado en intereses, y formulamos el problema de test hipótesis entre H_1 y H_2 como un programa lineal (LP).

En base a las probabilidades de error y detección mostradas en la subsección anterior, se pueden realizar varios diseños. Algunos criterios clásicos de optimalidad son los diseños de Bayes, Neyman-Pearson y minimax [40]. En este trabajo, consideramos un enfoque minimax robusto que minimiza la probabilidad de error de peor caso, en las dos hipótesis. Adoptamos este enfoque porque, en nuestro intento por detectar anuncios basados en intereses, ambas probabilidades de error son igualmente importantes.

De acuerdo con este criterio de diseño, el detector minimax robusto propuesto viene dado por la matriz D que resuelve el problema de optimización

$$\min \max_{i=1,2} P_i^w. \quad (2)$$

Sea \tilde{d}^T la primera fila de D , es decir, las probabilidades condicionales $p_{\hat{H}|X}(1|j)$ para $j = 1, \dots, n$. En el Apéndice B mostramos que (2) es equivalente al siguiente problema de optimización en las variables $\lambda, \mu, \tilde{d} \in \mathbb{R}^n$ y $\nu \in \mathbb{R}$:

$$\begin{aligned} &\text{maximizar} && \zeta \\ &\text{sujeto a} && \mu^T p^{\min} - \lambda^T p^{\max} + \nu \geq \zeta, \\ &&& 1 - \tilde{d}^T q \geq \zeta, \\ &&& \mu - \lambda + \nu \mathbf{1} \preceq \tilde{d}, \\ &&& \lambda \succeq 0, \mu \succeq 0, \\ &&& 0 \preceq \tilde{d} \preceq \mathbf{1}. \end{aligned} \quad (3)$$

La importancia de formular (2) como LP reside en que nos permite recurrir a métodos extremadamente eficientes y potentes para calcular el detector óptimo. Esto tiene una gran relevancia práctica ya que nuestro objetivo es proporcionar dicha funcionalidad de detección localmente (en el lado del usuario), mediante un software independiente que opere en tiempo real, es decir, mientras el usuario navega por la Web. La Sec. 4 describe la librería de optimización utilizada para este cálculo. La factibilidad de este problema de optimización se muestra en el Apéndice A.

3.3 Detección de Perfiles Singulares

En la subsección anterior, proporcionamos el diseño de un detector robusto basado mediante el que los usuarios pueden averiguar hasta qué punto se explotan sus perfiles de navegación para servirles anuncios. Esta subsección investiga otro aspecto crucial relacionado con la publicidad conductual, en concreto, si los perfiles construidos por las empresas de publicidad pueden revelar patrones de navegación *únicos*.

La importancia de este aspecto radica en el riesgo potencial de reidentificación a partir de datos *únicos*, aunque no identificables personalmente, como demostró el escándalo de AOL [45]. En nuestro caso, el riesgo de perfilar va de la mano del riesgo de reidentificación, especialmente cuando se considera en el contexto de información adicional que se puede obtener de un usuario, como su ubicación y aspectos relacionados con el navegador Web y el sistema operativo. Cuando el perfil se agrega además a la gran cantidad de datos compartidos a través de numerosos servicios de información, que un atacante de privacidad podría observar y cruzar, dicho atacante podría eventualmente descubrir, incluso en un sentido estadístico, la identidad real del usuario.

Habiendo motivado el riesgo de la singularidad del perfil, esta subsección describe cómo detectar si los anuncios entregados a un usuario pueden haberse generado como resultado de un patrón de navegación común, o por el contrario, un historial de navegación que se desvía de un comportamiento típico. Con este fin, primero proporcionamos una breve justificación de la divergencia KL como medida de la singularidad de un perfil. En [15], [16], describimos en mayor detalle por qué la divergencia KL es una métrica apropiada de unicidad. Posteriormente explicamos cómo estimar esta cantidad de información teórica.

Aunque mencionamos en la Sec. 3.1 que la divergencia KL no es una métrica adecuada, su sentido de discrepancia entre distribuciones permite una justificación intuitiva de cuán popular es un perfil. Particularmente, siempre que el perfil observado por un selector de anuncios difiera demasiado del perfil promedio de todos los usuarios rastreados, el selector de anuncios podrá determinar si los intereses del usuario en cuestión son atípicos, en contraste con las estadísticas de la población general.

Una justificación más elaborada surge de la célebre justificación de Jaynes de los métodos de maximización de la entropía [46], [47], basada en el método de los tipos [36, S 11], una técnica de la teoría de grandes desviaciones. Partiendo de este razonamiento, la entropía relativa entre un perfil observado y el perfil de la población puede considerarse como una medida de la singularidad de la primera distribución dentro de dicha población. La idea principal es que el método de tipos establece una relación monótona aproximada entre la *probabilidad* de una PMF en un sistema estocástico, y su divergencia con respecto a una distribución de referencia, la de la población. En términos generales, y en nuestro contexto, cuanto menor sea la divergencia de un perfil con respecto al perfil promedio, más probable será, y más se comportarán los usuarios de acuerdo con él. Bajo esta interpretación, la divergencia KL se interpreta, por lo tanto, como un indicador (inverso) de lo común que es un perfil dentro de dicha población.

Habiendo justificado el uso de la divergencia KL como una medida de frecuencia de un perfil, a continuación exploramos la incertidumbre que plantea la estimación de este valor de divergencia. Recuerde de Sec. 3.2.3 que los selectores de anuncios pueden construir perfiles de varias maneras a partir del clickstream observado. Al igual que hicimos con el diseño del estimador de anuncios basado en intereses, continuamos considerando una estimación de unicidad de caso más desfavorable en el espacio de los posibles perfiles creados por un selector de anuncios. Indique por \bar{p} el

perfil de interés de la población. Formalmente, para cada usuario y selector de anuncios, definimos la *unicidad mínima* sobre todos los perfiles como

$$u_{\min} = \inf_{p \in \mathcal{P}} D(p \parallel \bar{p}), \quad (4)$$

lo que da una medida de la similitud de perfil, que tiene en cuenta a su vez la incertidumbre inherente en la ventana de tiempo utilizada por un selector de anuncios.

El problema de minimización de divergencias descrito anteriormente capta el peor de los casos con respecto a la frecuencia de perfiles. En particular, nos dice cuán peculiares pueden ser nuestros intereses, a los ojos de un selector de anuncios. Para cualquier selector de anuncios, el valor u_{\min} (en el intervalo $[0, \infty)$ bits) variará claramente con el tiempo a medida que el usuario navega por la Web. Desde el punto de vista más práctico, sin embargo, la información transmitida cada vez por este valor *absoluto* de singularidad puede no ser lo suficientemente informativa para el usuario.

Para ayudar al usuario a interpretar un valor dado de u_{\min} , consideramos hacerlo *relativo* a una población de usuarios. Con ello, los usuarios pueden comparar sus valores de singularidad de perfil con los de otros usuarios de nuestra extensión de navegador Web, y así obtener una perspectiva más amplia de cómo son perfilados. Además, los usuarios pueden utilizar esta información para definir las políticas de bloqueo de anuncios correspondientes. Más adelante, en la Sec. 4, describiremos el intercambio de información entre los usuarios de nuestro sistema y un repositorio central para estimar esos valores relativos de unicidad de perfil.

4 MYADCHOICES — UNA HERRAMIENTA PARA LA TRANSPARENCIA Y BLOQUEO DE ANUNCIOS

Esta sección describe *MyAdChoices*, un sistema prototipo que proporciona transparencia en el proceso de entrega de anuncios, para que los usuarios puedan tomar una decisión informada y equitativa con respecto al bloqueo de anuncios. El sistema propuesto proporciona dos funcionalidades principales. Habilitado por el detector de anuncios basado en intereses y el estimador de singularidad de perfil diseñado en la Sec. 3, la funcionalidad de transparencia de anuncios permite a los usuarios entender lo que está sucediendo con sus datos de navegación Web. La funcionalidad de bloqueo de anuncios, por otro lado, permite a los usuarios reaccionar en consecuencia, de una manera flexible y no radical. Esto difiere sustancialmente de las tecnologías actuales de bloqueo de anuncios, que simplemente bloquean o permiten todos los anuncios. *MyAdChoices* no solo considera estos dos extremos, sino cualquier posibilidad intermedia. Con esta última funcionalidad, los usuarios pueden indicar el tipo de anuncios que desean recibir o, en caso contrario, aquellos que desean bloquear. Al combinar ambas funcionalidades y, por lo tanto, dotar de transparencia y control preciso sobre la publicidad en línea, el sistema propuesto puede ayudar a preservar el modelo económico dominante de Internet, actualmente amenazado por la aparición de bloqueadores de anuncios.

Esta sección está organizada de la siguiente manera. La Sec. 4.1 primero describe las funcionalidades de transparencia de anuncios y de bloqueo provistas por nuestro sistema. Luego, la Sec. 4.2 detalla los componentes de una arquitectura de sistema que implementa estas dos funcionalidades.

4.1 Principales Funcionalidades

Nuestro sistema aporta *transparencia* en dos aspectos centrales de la publicidad de comportamiento. Por un lado, permite a los usuarios saber si la información recopilada sobre sus intereses de navegación podría haber sido utilizada por la industria de la publicidad para enviarles los anuncios. Específicamente, nuestro sistema permite al usuario saber si los anuncios recibidos pueden haber sido generados de acuerdo con sus intereses de navegación o, más precisamente, en base a los perfiles que los selectores de anuncios pueden tener de ellos. Por otro lado, proporciona información sobre los perfiles de navegación que los selectores de anuncios pueden haber deducido de las páginas rastreadas. En particular, *MyAdChoices* muestra el valor de unicidad de perfil en el peor de los casos para cada selector de anuncios, además de la categoría de interés de los anuncios recibidos.

Con respecto al servicio *anuncio de bloqueo*, nuestro sistema contempla los siguientes parámetros configurables por el usuario:

- **Categoría de interés del anuncio.** Ofrecemos a los usuarios la posibilidad de filtrar anuncios por categoría de interés. Por ejemplo, un usuario podría bloquear los anuncios que pertenecen a ciertas categorías sensibles como pornografía y salud.
- **Clase de anuncio.** Este parámetro permite a los usuarios bloquear los anuncios basados en intereses o los no basados en intereses, para todas las categorías de interés del anuncio o para un subconjunto de ellas.
- **Unicidad de perfil.** Los usuarios pueden decidir bloquear los anuncios entregados por aquellos selectores de anuncios que pueden haber compilado perfiles atípicos, y por lo tanto potencialmente identificables, de sus hábitos de navegación.
- **Redirección.** Por último, los usuarios pueden decidir bloquear los anuncios redirigidos, es decir, los anuncios procedentes de anunciantes que han sido visitados anteriormente por el usuario (véase la Sec. 2.3).

4.1.1 Ejemplos de Políticas de Bloqueo de Anuncios

Esta subsección proporciona un par de políticas de control de anuncios sencillas, que pretenden ilustrar los parámetros descritos en la subsección anterior. Estos ejemplos están precedidos por una definición general de política de filtrado de anuncios, inspirada en el campo del control de acceso.

Definición 1 (Política de bloqueo de anuncios). *Una política pol es un par $(AC, sign)$, donde AC es una restricción de anuncio, y $sign \in \{+, -\}$ modela una acción a llevar a cabo cuando un anuncio cumple con esa restricción. Una restricción de anuncio se representa con un triple (I, i, u) , donde $I \in \{0, 1\}$ indica si un anuncio está basado en intereses o no, $i \in \mathcal{X}$ es una categoría de interés, y u_{\min} denota un requisito de unicidad de perfil mínima.*

Una restricción de anuncios representa el conjunto de anuncios que pertenecen a una categoría de interés i , que se clasifican como basados en el interés (o no), y que se han entregado según un perfil con una unicidad mínima dada por u_{\min} . Por otro lado, $sign$ denota si el anuncio debe ser bloqueado (-), o mostrado en el navegador del usuario (+).

Debido a que el soporte para políticas positivas y negativas puede causar conflictos (es decir, podemos tener una restricción de anuncios que satisfaga tanto las políticas positivas como las negativas), debemos contar con un mecanismo de resolución de conflictos. La literatura de control de acceso proporciona varios enfoques para abordar tales situaciones. Aquí, por simplicidad, asumimos que las políticas negativas dominan, ya que este enfoque ofrece mayores garantías con respecto al riesgo de mostrar anuncios inapropiados. Sin embargo, otras políticas de resolución de conflictos también podrían integrarse fácilmente.

A continuación se dan dos ejemplos de políticas. En estos ejemplos, nos referimos a algunas de las categorías de interés utilizadas por el sistema propuesto (véase la Sec. 4 para obtener más detalles). Para mayor brevedad, en esta sección indicaremos las categorías relevantes por su nombre. Además, por simplicidad y claridad, en los ejemplos seguiremos usando la notación formal de políticas introducida en Definición 1. Sin embargo, observamos que esta notación, que describe cómo se implementan realmente las políticas en el sistema, debe ser transparente desde el principio para mejorar la facilidad de uso y para ayudar a los usuarios a especificar políticas que reflejen en la medida de lo posible sus preferencias. Como explicaremos en la Sec. 4.2.2.5, se diseñarán varias estrategias para este propósito, por ejemplo, el uso de etiquetas textuales en lugar de valores numéricos.

Ejemplo 2 (Políticas para permitir determinados anuncios personalizados.). *Sofía había planeado visitar Nueva York en vacaciones. Hace algunos días compró los billetes de avión y reservó un hotel, todo a través de Internet. Durante los días siguientes a la reserva, visitó varios sitios Web en busca de visitas turísticas y excursiones de un día. A medida que navegaba por la Web, Sofía veía cómo los anuncios que le aparecían en el navegador estaban cada vez más relacionados con su próximo viaje. Sofía ahora está cansada de recibir*

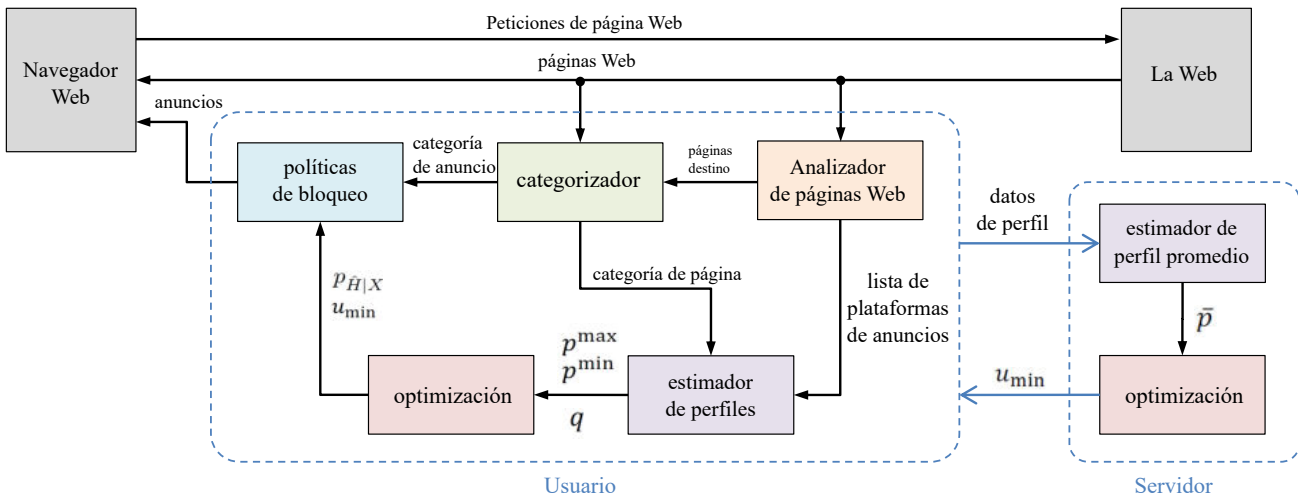


Fig. 5: Componentes internos de la arquitectura propuesta.

anuncios de hoteles en Nueva York, por lo que está considerando instalar AdBlock Plus para bloquearlos todos. Sin embargo, ella aprecia el valor y la utilidad de la publicidad conductual, y debido a que aún no ha decidido su itinerario, todavía desea recibir anuncios personalizados asociados con las categorías 1 (“Viajes-Trenes”) y 2 (“Viajes-Parques Temáticos”). En consecuencia, Sofía especifica las siguientes políticas:

- $pol_1 = ((c_1, 1, \cdot), +)$,
- $pol_2 = ((c_2, 1, \cdot), +)$,

donde el símbolo “.” significa que el valor del parámetro en cuestión no está especificado.

Ejemplo 3 (Política de equilibrio entre personalización y privacidad.). Juan trabaja en una tienda de dietética y nutrición. Como parte de su trabajo, a veces consulta páginas sobre salud y estado físico. Ocasionalmente, y cuando nadie le ve, pasa algún tiempo visitando páginas Web relacionadas con la fibromialgia, enfermedad de la que ha sido diagnosticado recientemente. Hace unos días se sorprendió cuando un par de anuncios sobre tratamientos biológicos para su enfermedad aparecieron mientras navegaba por la Web. Desde entonces, Juan está preocupado con que se puedan mostrar anuncios de ese tipo cuando sus compañeros de trabajo miren a su monitor. Sin embargo, a pesar de sus preocupaciones, no desea recurrir a las herramientas de bloqueo de anuncios típicas, ya que dichos servicios de publicidad personalizados le ayudan a mantenerse al día de los últimos productos y tendencias en su trabajo. Para lograr un equilibrio entre la privacidad y la personalización, Juan define un filtro que bloquea los anuncios de salud basados en su perfil de navegación solo cuando dicho perfil refleja intereses relativamente atípicos. En particular, define la siguiente política:

- $pol_1 = ((c_3, 1, \pi_{u_{\min}} \geq 25\%), -)$,

donde la categoría 3 corresponde a “salud y fitness”, y $\pi_{u_{\min}}$ denota el valor percentil de u_{\min} .

4.2 Arquitectura de Sistema y Detalles Prácticos de Implementación

En esta sección describimos los componentes de una arquitectura de sistema que implementa las dos funcionalidades especificadas en la Sec. 3. El sistema propuesto se ha desarrollado como una extensión de navegador Web y está disponible para Google Chrome. Cabe destacar que esta extensión no solo proporciona servicios de transparencia y bloqueo de anuncios en tiempo real, sino que también funciona como un sistema independiente, es decir, realiza todos los cálculos y operaciones localmente, sin necesidad de contar con ninguna infraestructura o entidad externa. La única excepción es el cálculo del valor mínimo de unicidad de perfil, que no se realiza en el lado del usuario, ya que requiere el perfil promedio de la población \bar{p} . Como veremos más adelante en la Sec. 4.2.2.4, este servicio en particular se proporciona solo si el usuario acepta compartir datos de perfil con el servidor *MyAdChoices*.

4.2.1 Suposiciones

Antes de continuar con la descripción de la arquitectura del sistema, examinamos las suposiciones asumidas al implementar el detector de anuncios basado en intereses y el estimador de unicidad de perfil diseñado en las Secciones. 3.2 y 3.3.

Nuestra primera suposición está relacionada con la imposibilidad de saber, con absoluta certeza, la información de navegación que los selectores de anuncios tienen sobre los usuarios. En la Sec. 3.3 llamamos a esta información el clickstream observado, y la definimos más precisamente como la secuencia de páginas Web que el selector de anuncios sabe que el usuario visitó. Al observar las solicitudes de red de terceros, nuestra extensión de navegador puede capturar las páginas que las plataformas de anuncios pueden rastrear a través de cookies HTTP u otros métodos más sofisticados, como los fingerprints del navegador. Sin embargo, no podemos saber si ésta es toda la información de la que disponen, es decir, si esas páginas representan sus clickstreams observados o no. — los selectores de anuncios y los rastreadores Web también pueden intercambiar sus datos de seguimiento, por ejemplo, a través de la coincidencia de cookies, una práctica que parece ser mucho más común que los métodos de seguimiento directo [48], [35], [49]. El hecho de que un protocolo de coincidencia de cookies se ejecute entre dos entidades no implica, sin embargo, que terminen intercambiando sus datos de rastreo. Existe un incentivo obvio para agregar información y obtener más datos sobre el historial de navegación de un usuario, pero como este intercambio no se realiza por medio del navegador del usuario, no podemos concluir con seguridad que se haya efectuado.

En el caso de RTB, las solicitudes de oferta enviadas por una plataforma publicitaria pueden permitir a los participantes de la subasta rastrear a un usuario determinado. Dado que el apostante ganador (es decir, el selector de anuncios) es el que sirve el anuncio, nuestro sistema puede marcar fácilmente la página correspondiente como este postor. Sin embargo, el problema es que no podemos determinar si este selector de anuncios podría haber recibido otras solicitudes de oferta para este usuario (mientras visitaba otras páginas) y, por lo tanto, podría haberlas rastreado en esas páginas. Las plataformas de anuncios generalmente permiten que los apostantes creen perfiles *solo* de las subastas que ganan, pero, en realidad, nada les impide utilizar tales datos de seguimiento. En resumen, debido a que no hay forma de conocer a los destinatarios de esas solicitudes y el uso que hacen de dichos datos, nuestro conocimiento de los sitios rastreados a través de RTB se limita a aquellos sitios donde el selector de anuncios sirve un anuncio.

En este trabajo, abordamos todas estas limitaciones considerando dos escenarios en términos de seguimiento y uso compartido de datos de clickstream:

- Escenario de *baseline*, donde el sistema opera con el clickstream que, según nuestras observaciones (i.e., las que realiza la herramienta MyAdChoices), puede tener el selector de anuncios. Es decir, asumimos que el clickstream observado de un selector de anuncios coincide con los datos de rastreo que disponemos, y, por lo tanto, ignoramos cualquier posible intercambio de información de seguimiento con otras entidades. En términos prácticos, nuestra extensión de navegador Web compilará este clickstream examinando si el selector de anuncios está presente, como un dominio de terceros, en las páginas visitadas por el usuario. En otras palabras, asumiremos que todos los dominios de terceros presentes en una página pueden rastrear la visita de un usuario a dicha página. Al hacerlo, podremos capturar los sitios donde un selector de anuncios ha incorporado un enlace (a través de los editores correspondientes), y aquellas páginas donde ha ganado el derecho a publicar un anuncio a través de RTB.
- un escenario *paranoico* en el que asumimos que el seguimiento Web es ubicuo y la información del clickstream se comparte entre todas las entidades que participan en el proceso de publicación de anuncios. En este caso, consideramos que el clickstream observado coincide con el clickstream real, es decir, con la secuencia de todas las páginas que un usuario ha visitado. Sin embargo, reconocemos que puede que no haya empresas publicitarias y rastreadores en ciertas páginas y, por lo tanto, no se podría obtener en la práctica un perfil real completo y preciso.

Nos gustaría subrayar que los dos escenarios descritos anteriormente se refieren *únicamente* a los datos de seguimiento de usuarios disponibles para los selectores de anuncios. Dicho de otra

manera, nuestro sistema no tiene en cuenta los datos de intereses ni la información personal que los usuarios podrían haber comunicado *explícitamente* a estas entidades, y podría utilizarse para fines de personalización de anuncios.

Habiendo especificado los dos modos de operación de nuestro sistema, a continuación exponemos nuestra segunda suposición, que se refiere a la forma en que los selectores de anuncios construyen perfiles de usuario a partir de los clickstream observados. En la Sec. 3.2.2 asumimos que los selectores de anuncios modelan los perfiles como PMFs, esencialmente en línea con gran parte de la literatura del área. Para computar tales distribuciones en la práctica, nuestro sistema asume, con una ligera pérdida de generalidad, que los selectores de anuncios emplean la estimación de probabilidad máxima (ML) [50]. Nos gustaría enfatizar que este es, con mucho, el método más popular de estimación de parámetros en estadística.

Nuestra tercera y última suposición tiene que ver con la categorización temática del contenido Web. Consideraremos que el categorizador utilizado por nuestro sistema coincide, en gran medida, con el empleado por las plataformas publicitarias⁸. Esto implica que tanto nuestra extensión como las plataformas de anuncios dependen del mismo conjunto predefinido de categorías de interés y el mismo algoritmo de categorización, de modo que cualquier página visitada por el usuario se clasificará en la misma categoría tanto por el sistema propuesto como por las plataformas publicitarias que rastrean esta visita. Creemos que este es una suposición razonable ya que nuestro algoritmo de categorización se basa en una taxonomía de temas estándar desarrollada por el Interactive Advertising Bureau [51], una organización que representa la gran mayoría de las empresas de publicidad en línea en los EE.UU.

4.2.2 Componentes

Esta sección describe la funcionalidad de los componentes principales de nuestra arquitectura, justifica los criterios de diseño, y proporciona *algunos* detalles clave de implementación de bajo nivel. La Fig. 5 representa la arquitectura implementada, que consta de dos partes principales, el lado del usuario y el lado del servidor. Este último se encarga de calcular los valores de unicidad mínima por cada uno de los selectores de anuncios. Como esto requiere la obtención del perfil \bar{p} , dicho cálculo se realiza solo si el usuario acepta compartir sus datos de perfil con nuestro servidor. El resto de funcionalidades y el procesado correspondiente se realizan íntegramente en el lado del usuario. A continuación analizamos los componentes de ambos lados.

4.2.2.1 Estimador de Perfiles: En el lado del usuario, este módulo tiene como objetivo estimar (1) el conjunto \mathcal{P} de posibles perfiles de usuario que un selector de anuncios puede haber asignado a un usuario; y (2) la distribución q de las categorías de intereses de aquellos anuncios clasificados como no basados en intereses. Es importante destacar que, independientemente del escenario asumido (es decir, baseline o paranoico), la estimación de q debe realizarse para *cada* selector de anuncios. En el escenario anterior, el cálculo de p también es necesario por cada selector de anuncios. Sin embargo, dado que el último escenario considera que las secuencias de clics observadas de todos los selectores de anuncios coinciden con el clickstream real del usuario, asumimos que $p = t$.

Como se explica en la Sec. 3.2.2, la estimación de la PMF q requiere una sesión de navegación en la que no se rastree al usuario. Nuestra versión actual del complemento implementa esta sesión libre de seguimiento mediante el modo privado o incógnito del navegador, una característica del navegador que, entre otras funcionalidades, evita el seguimiento a través de cookies HTTP y Flash. Sin embargo, reconocemos que los selectores de anuncios también pueden seguir las visitas de los usuarios como resultado del uso de súper cookies, respawning [52], [53], canvas fingerprints [54] o simplemente sus direcciones IP. No obstante, dado que estos mecanismos de seguimiento son muy infrecuentes o bastante inexactos, podemos suponer razonablemente que el modo de incógnito del navegador coincide aproximadamente con una sesión sin seguimiento, si no completamente. De hecho, estudios recientes indican que la prevalencia de estos métodos de seguimiento más sofisticados es solo del

8. Las plataformas publicitarias son las que clasifican el contenido de una página. En la publicidad de RTB, normalmente incluyen la categoría de la página del editor en las solicitudes de subasta.

Tabla 1: Categorías de interés de alto nivel.

adultos	economía	pasatiempos y aficiones	política
agricultura	educación	hogar	inmobiliario
animales	familia	derecho	religión
arquitectura	moda	militar	ciencia
arte y entretenimiento	folklore	noticias	sociedad
automoción	alimentos y bebidas	finanzas personales	deportes
negocios	salud y fitness	mascotas	tecnología e informática
ámbito profesional	historia	filosofía	viajes

Tabla 2: Subcategorías correspondientes a tres categorías de primer nivel.

Categoría de alto nivel	Bottom-level category
Arte y entretenimiento	animación, celebridades, cómics, diseño, bellas artes, humor, literatura, películas, música, ópera, poesía, radio, televisión, teatro y videojuegos.
Salud y bienestar	medicina alternativa, anatomía, asma, autismo, incontinencia intestinal, tumor cerebral, cáncer, paro cardíaco, dolor crónico, resfriado y gripe, sordera, cuidado dental, dermatología, diabetes, dieta, epilepsia, ejercicio, cuidado de los ojos, primeros auxilios, enfermedad cardíaca, VIH / SIDA, medicina, salud masculina, depresión mental, nutrición, ortopedia, pediatría, fisioterapia, psicología y psiquiatría, salud de personas mayores, sexualidad, trastornos del sueño, dejar de fumar, estrés, abuso de sustancias, tiroides Enfermedad, vitaminas, pérdida de peso y salud de la mujer.
Finanzas personales	banca, crédito, deuda y préstamos, criptomonedas, noticias financieras, planificación financiera, seguros, inversiones, plan de pensiones, acciones y planificación fiscal.

5% en Alexa 100 000 sitios [49]. En resumen, consideraremos que la PMF q estimada de esta manera refleja efectivamente la distribución temática del anuncio cuando el selector de anuncios considera al usuario como un nuevo usuario, por lo que los anuncios solo pueden basarse en la ubicación, ser contextuales y genéricos.

En términos prácticos, hay una diferencia entre la estimación de p y q . En este último caso, se realiza a partir de los anuncios que recibe el navegador durante dicho modo incógnito. En el caso de p , o equivalentemente \mathcal{P} , la estimación se realiza a partir las páginas que el selector de anuncios puede rastrear, por su cuenta y/o a través de otras fuentes de datos.

Una de las dificultades para estimar estas dos distribuciones es que, mientras que q requiere navegar en dicha sesión de libre de rastreo, la PMF p debe reflejar las páginas rastreadas por cualquier posible selector de anuncios. Una posible solución para lidiar con esta incompatibilidad consiste en alternar entre los modos de incógnito y normal de manera regular. Sin embargo, el problema con este enfoque es que los usuarios pueden ser reacios a navegar en el modo privado durante el tiempo necesario para calcular y actualizar las PMF q de un número suficiente de selectores de anuncios.

Motivado por esto, la arquitectura del lado del usuario estima simultáneamente ambas distribuciones volviendo a visitar, en el modo incógnito y de forma completamente automatizada, una fracción ρ de las páginas visitadas por el usuario. En términos prácticos, cada nueva visita se realiza abriendo una nueva ventana minimizada en modo privado. Procedemos de esta manera porque queremos evitar el seguimiento entre diferentes pestañas en el mismo modo incógnito. Sin embargo, admitimos que esta solución podría tener un impacto no despreciable en estos dos aspectos: primero, en términos de la sobrecarga de tráfico incurrida; y en segundo lugar, puede penalizar a los anunciantes en cierta medida, ya que los anuncios recibidos en la sesión de rastreo obviamente no se presentarán al usuario. Actualmente, el sistema propuesto opera con una fracción de re-visitadas de $\rho = 25\%$. Aunque esta reducción en el número de re-visitadas tiene un impacto en la exactitud de la estimación de q , creemos que puede representar un coste aceptable en términos de sobrecarga de tráfico e impacto publicitario. Como nota al margen, nos gustaría enfatizar que el impacto de tales revisiones es, desde una perspectiva de usabilidad, casi imperceptible.

Después de examinar las condiciones de navegación Web en las que se obtienen p y q , a continuación describimos aspectos más concretos relacionados con el estimador de estas distribuciones.

Como se menciona en la Sec. 4.2.1, este trabajo asume que los selectores de anuncios dependen de la estimación de ML, un método de estimación ampliamente utilizado en muchos campos de la ingeniería. Sea m la cantidad total de anuncios recibidos (páginas visitadas), y m_i el número de esos anuncios (páginas) que pertenecen a la categoría de interés i . Recuerde que la estimación ML de una PMF se define como

$$q_i = \frac{m_i}{m},$$

para $i = 1, \dots, n$.

Para tomar una decisión sobre si los anuncios mostrados están basados en intereses o no, nuestro estimador ML requiere observar el mismo número mínimo de páginas w_{\min} que necesita un selector de anuncios para modelar los intereses a corto plazo. Varios estudios señalan que la ventana de tiempo más pequeña que las empresas de publicidad podrían usar para este tipo de modelado es un día (véase la Sec. 3.2.3. De acuerdo con estos estudios y al número promedio de páginas exploradas por un usuario por día [55], configuramos $w_{\min} = 87$. En el otro extremo, en línea con los trabajos citados en esa sección, consideramos que el mayor clickstream utilizado para modelar intereses a largo plazo es de 8 semanas. Por tanto, hemos establecido $w_{\max} = 3915$. Para estimar q , procedemos de forma análoga, al establecer una ventana deslizante de la misma longitud.

Finalmente, en el lado del servidor, nuestra arquitectura tiene como objeto calcular, para cada usuario que quiera compartir datos de perfil con el servidor, el perfil promedio y la clase de incertidumbre de cada selector de anuncios.

4.2.2.2 Analizador de Páginas Web: Este bloque tiene como objetivo obtener cierta información sobre (1) las páginas Web exploradas por el usuario y (2) los anuncios mostrados dentro de esas páginas, *ambos* en las sesiones rastreadas y de incógnito. Específicamente, cuando el navegador descarga una página, ya sea en modo normal o privado, el módulo genera una lista de todas las entidades que rastrean esta página y publican anuncios en ellas.

Además, nuestro sistema intenta recuperar la *página de destino* de todos los anuncios mostrados en ambos modos, es decir, la página del anunciante a la que se redirige el navegador al hacer clic en su anuncio [56]. Recuerde que nuestro sistema necesita la categoría de interés de un anuncio para tomar una decisión sobre si está basado en el perfil o no. Para clasificar un anuncio en una categoría temática, el módulo de categorización (descrito más adelante en la Sec. 4.2.2.3) requiere su página de destino. Sin embargo, dado que al hacer clic en cada anuncio (para obtener esta información) nos llevaría a cometer el llamado fraude de clic [57], las funcionalidades proporcionadas por nuestra herramienta en términos de transparencia y bloqueo están limitadas a aquellos anuncios donde la información de la página de destino está disponible sin hacer clic en ellos. A pesar de esta limitación, algunos estudios recientes [56], [21] han reportado una disponibilidad de la página de destino por encima del 80%.

4.2.2.3 Categorizador: Este módulo clasifica las páginas visitadas por el usuario, así como las páginas de destino de los anuncios dirigidos a ellos, en un conjunto predefinido de intereses temáticos. El módulo emplea una taxonomía jerárquica de 2 niveles, compuesta por 32 categorías de nivel superior y 330 categorías o subcategorías de nivel inferior. Las Tablas 1 y 2 muestran las categorías de nivel superior y las subcategorías correspondientes a tres de estas categorías.

El algoritmo de categorización integrado en nuestro sistema está parcialmente inspirado por la metodología presentada en [56] para clasificar anuncios no textuales en categorías de interés. El algoritmo también se basa en la taxonomía disponible en el complemento Firefox Interest Dashboard [58] desarrollado por Mozilla.

Nuestro categorizador se basa en dos fuentes de datos previamente clasificados. Primero, una lista de URLs, o más específicamente, dominios y nombres de host, que son consultados para determinar la categoría de la página. En segundo lugar, una lista de unigramas y bigramas [59] que se usan cuando falla la búsqueda de URL. El primer tipo de datos se justifica por el hecho de que una parte relativamente pequeña de toda la Web representa la mayoría de las visitas. Además, es evidente que

Librería de optimización	Tiempo de ejecución [s]		
	promedio	varianza	máximo
Coin-OR Linear Programming (CLP), v.1.16.6 [66], [67]	0.0315505	0.0000010	0.0460014
GNU Linear Programming Kit (GNULPK), v.4.48 [68]	0.0337618	0.0000055	0.0681626
Object Oriented Quadratic Programming (OOQP), v.0.99.22 [69]	0.0401395	0.0000028	0.0805860
LPSolve, v.5.5.2.0 [70]	0.0645488	0.0000024	0.0808482
C Library for Semidefinite Programming (CDSP), v.6.1 [71]	0.5878725	0.0017888	1.1033131
Dual-Scaling Semidefinite Programming (DSDP), v.5.8 [72]	2.0933280	0.0137100	4.1946620
Coin-OR Interior Point OPTimizer (IPOPT), v.3.12.3 [73], [74]	0.2014676	0.1510803	5.7872007
Limited Memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS), v.3.0 [75]	0.2054921	0.1669853	6.1828331
NLopt, v.2.4.2 [76]	0.5781220	0.0010485	0.6520662

Tabla 3: Probamos 6 librerías de optimización para calcular la solución al problema LP (3), y otras 3 para el problema de minimización de divergencias (4). Esta figura muestra el promedio, la varianza y los valores máximos de tiempo de ejecución, obtenidos a partir de mil instancias de los dos problemas.

la búsqueda pre categorizada requiere pocos recursos computacionales en el navegador y puede ser más precisa. El último tipo de información, por otro lado, se justifica como una alternativa y nos permite aplicar heurísticas del lenguaje natural a las palabras disponibles en la URL, el título, las palabras clave y el contenido.

Para casi cada una de las categorías de nivel superior, la versión actual del complemento incorpora los 500 sitios Web principales de Alexa.com. Además, la lista de URL incluye las páginas clasificadas por el complemento de Mozilla (alrededor de siete mil). Por otro lado, el número de unigramas y bigramas en inglés es de aproximadamente 76 000. Tres listas adicionales, aunque con un número menor de entradas, también están disponibles en francés, español e italiano⁹. Para compilar todas estas listas de palabras, hemos utilizado los siguientes datos:

- una versión refinada de los datos de categorización provistos por la extensión Firefox Interest Dashboard;
- un subconjunto de los términos en inglés disponibles en WordNet 2.0 [60] para los cuales la jerarquía de dominios de WordNet [61], [62] proporciona una etiqueta de dominio;
- un subconjunto de los términos disponibles en el repositorio central multilingüe de WordNet 3.0 [63], para permitir la categorización de sitios Web escritos en los idiomas mencionados anteriormente;
- y los datos de mapeo de sincronización entre las versiones 2.0 y 3.0 de WordNet [64].

El módulo categorizador recurre a estas listas solo cuando el nombre de host y el dominio no se encuentran en la base de datos de URLs. Cuando esto sucede, el algoritmo trata de clasificar la página utilizando los unigramas y bigramas extraídos de los siguientes campos de datos: URL, título, palabras clave y contenido. Dependiendo del campo de datos en cuestión, el categorizador asigna diferentes ponderaciones a los unigramas y bigramas correspondientes. Al hacerlo, podemos reflejar el hecho de que los términos que aparecen en la URL, el título y, especialmente, las palabras clave especificadas por el editor (si están disponibles), suelen ser más descriptivas y explicativas que las que se incluyen en el cuerpo de la página.

Como se hace con frecuencia en el campo de la recuperación de información y en la minería de texto, nuestro clasificador de páginas Web también utiliza la frecuencia de término—frecuencia inversa de documento (TF-IDF) [65]. Dicho de otro modo, ponderamos las categorías resultantes en función de la frecuencia de aparición de los unigramas y bigramas correspondientes, y de una medida de su frecuencia dentro de toda la Web.

En aras de la eficiencia computacional, el algoritmo almacena las categorías derivadas de las últimas 500 páginas visitadas del usuario. De esta manera, cuando el usuario vuelve a visitar una de esas páginas, las categorías temáticas se obtienen directamente sin necesidad de pasar por todos los pasos anteriores.

9. Las próximas versiones de esta extensión del navegador Web incluirán más idiomas.

En términos de almacenamiento, la lista completa de unigramas, bigramas y sus correspondientes valores IDF ocupan aproximadamente 1 megabyte en formato comprimido. Creemos que esto es un coste aceptable para el tamaño de descarga del complemento.

Por último, una inspección manual de los resultados de la categorización de una gran colección de páginas Web y anuncios indica que el algoritmo es, en casi todos los casos, ciertamente preciso. Sin embargo, se requeriría más investigación para evaluar el desempeño del categorizador de una manera más rigurosa.

4.2.2.4 Módulos de Optimización: Los módulos de optimización incorporados en el lado del usuario y del servidor son responsables de calcular las soluciones a los problemas (3) y (4), y así obtener el detector minimax robusto y el de perfil de singularidad mínima, respectivamente. Los parámetros de entrada del módulo del lado del usuario son la distribución q y las tuplas p^{\min} y p^{\max} . En el lado del servidor, nuestro sistema requiere del clickstream observado de cada selector de anuncios para calcular el perfil promedio y la clase de incertidumbre asociada. Nos gustaría señalar que la transparencia de los anuncios y las funciones de bloqueo relacionadas con la singularidad del perfil solo se proporcionarán en caso de que el usuario dé su consentimiento para transmitir dichos datos de clickstream.

En la arquitectura implementada, ambos módulos se basan en librerías de optimización de código abierto. El diseño de tales módulos requirió el examen y la comparación de una variedad de softwares de optimización. Debido a que nuestro sistema puede necesitar calcular el detector robusto cada vez que se muestra un anuncio, priorizamos la eficiencia y la fiabilidad en el lado del usuario. Estos mismos requisitos también se tuvieron en cuenta en el lado del servidor. Sin embargo, debido a que los valores de unicidad mínima u_{\min} se deben calcular para *cada usuario*, optamos por aligerar el procesado y el cálculo en esta parte de la arquitectura. En particular, en lugar de procesar los datos del perfil cada vez que hay una actualización del lado del usuario, especificamos intervalos regulares de 1 día (desde el momento en que se instala el complemento) para el intercambio de información con el servidor. Reconocemos que, dependiendo de la actividad del usuario, esto podría tener un cierto impacto en la precisión de los datos de unicidad de perfil proporcionados.

Teniendo en cuenta todos estos requisitos, realizamos un benchmark para el problema LP y el problema de minimización de divergencias. Empleamos la herramienta de optimización de Matlab OPTI [77], y probamos mil instancias de problemas con valores aleatorios — aunque factibles — para las entradas mencionadas anteriormente. Para el problema (4), también proporcionamos el gradiente de las funciones de objetivo y restricción. Además, redujimos la complejidad de este último problema utilizando una representación de nivel superior de \bar{p} , p^{\min} y p^{\max} con solo 32 categorías.

Los resultados se muestran en la Tabla 3 para 9 softwares de optimización. En base a nuestro análisis de rendimiento, seleccionamos CLP [66], [67] y IPOPT [73], [74], que proporcionan un método de punto símplex y de punto interior [39], respectivamente. Los dos optimizadores mostraron el tiempo de ejecución promedio más bajo en nuestro análisis, con 32 y 201 milisegundos respectivamente, así como valores de varianza aceptables. Vale la pena mencionar que todas las instancias de problemas fueron resueltas satisfactoriamente por las librerías probadas, y que los dos optimizadores elegidos están disponibles bajo la Licencia Pública de Eclipse [78].

En nuestro sistema, ambos optimizadores se configuraron para tener un tiempo de ejecución máximo permitido. Cuando nuestra extensión se instala por primera vez, ésta ejecuta varias instancias del problema para establecer este parámetro; esto es para el caso del cálculo del detector minimax robusto de anuncios basado en intereses. En el lado del servidor, el cálculo del valor mínimo de singularidad del perfil de usuario se limita a 0.5 segundos.

4.2.2.5 Políticas de Bloqueo: La funcionalidad de este módulo es aplicar las políticas de bloqueo de anuncios definidas por el usuario. Su implementación actual simplifica la notación de política formal presentada en la Sec. 4.1.1, en un intento por proporcionar una interfaz fácil de usar y, por lo tanto, mejorar la usabilidad.

Con este objetivo, nuestra extensión permite a los usuarios definir las políticas *solo* con *sign* negativo. Es decir, en lugar de especificar qué anuncios deben mostrarse (+) y cuáles deben bloquearse (-), solo habilitamos la última declaración de bloqueo, que puede facilitar la definición de tales

	interest-based	non-interest-based	re-targeted
Select all	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
adult	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
agriculture	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
animals	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
architecture	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
arts & entertainment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
automotive	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
business	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
careers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fig. 6: El panel de configuración que se muestra en esta figura permite a los usuarios definir políticas de bloqueo de anuncios detalladas. Las opciones disponibles para los usuarios incluyen filtrar anuncios por categoría de interés, publicidad conductual y de redirección. Aunque no se muestra en esta figura, los usuarios también pueden indicar las condiciones de bloqueo de anuncios en función de la singularidad de los perfiles que los selectores de anuncios podrían crear.

políticas. Además, la especificación de los valores percentiles de la unicidad del perfil se reduce, en esta implementación, a una opción binaria: los usuarios solo pueden decidir si desean bloquear (o permitir) aquellas entidades que pueden haber compilado perfiles “muy únicos” de ellos, lo que significa que $\pi_{u_{\min}} \geq 90\%$. La Fig. 6 muestra el panel de configuración mediante el cual los usuarios pueden configurar las políticas de bloqueo, así como el escenario que desean asumir en términos de seguimiento Web.

El funcionamiento de este módulo se describe a continuación. Cuando un usuario visita una página, el módulo espera a que el categorizador envíe la categoría de cada anuncio. Luego, recibe los detectores de anuncios basados en intereses de cada una de las entidades que entregan esos anuncios. Y, finalmente, consulta una base de datos interna (es decir, del lado del usuario) para obtener los valores mínimos de unicidad asociados con dichas entidades. Con toda esta información, nuestro sistema solo necesita verificar si se cumple con cada restricción de anuncios y, en consecuencia, decidir si bloquear el anuncio o no.

Debemos resaltar que nuestro sistema no bloquea los anuncios en el mismo sentido que las tecnologías de bloqueo de anuncios actuales. Mientras que estas tecnologías evitan las peticiones de red de terceros¹⁰. Al ser enviadas, nuestra extensión de navegador Web los permite. Solo cuando la página está completamente cargada y, por lo tanto, se muestran los anuncios (si los hay), nuestro sistema decide ocultarlos o no, mediante la aplicación de una máscara negra encima de ellos¹¹. Para resaltar este aspecto en particular, nos referimos a la acción de bloquear de manera más precisa como *ofuscación*. La Fig. 7 muestra una captura de pantalla de los anuncios procesados por nuestra herramienta en una página Web concreta.

La herramienta notifica a los usuarios sobre el tipo de anuncios recibidos a través de un pequeño icono ubicado en la esquina izquierda de cada anuncio detectado. Los íconos indican si un anuncio está basado en intereses (rojo), redirigido (rojo), no está basado en intereses (verde), está bloqueado según la política del usuario (negro) o el sistema no puede tomar una decisión (naranja). Este último caso ocurre, por ejemplo, cuando la página de destino del anuncio no está disponible o el

10. Adblock Plus [4], por ejemplo, no bloquea *todas* las solicitudes de red de terceros, sino solo aquellas de la lista negra [79].

11. En una nota técnica, el sistema puede alternativamente eliminar la imagen del anuncio.

The screenshot shows the New York Times Technology section. At the top, there's a navigation bar with the site logo, a search icon, and buttons for 'SUBSCRIBE NOW', 'SIGN IN', and 'Register'. Below this is a large banner for Samsung Galaxy A5, advertising it for 'desde 8'5€/mes' with 'Unidades limitadas!' and a 'Lo quiero ya!' button. The main content area is divided into several sections: 'FEATURED' with an article titled 'F.A.A. Drone Laws Start to Clash With Stricter Local Rules' by Cecilia Kang, featuring a photo of a man flying a drone; 'Bits' section with a link to 'Continue to Bits'; 'PERSONAL TECHNOLOGY' section with a 'TECH FIX' article 'For Parental Controls, iPhones Beat Androids' by Brian X. Chen; 'THE GETAWAY' section with an article '12 Travel Apps Worth Keeping in 2016' by Stephanie Rosenbloom; and a 'Free O'Reilly Book' advertisement for 'Graph Databases New 2nd Edition!' with a 'download now' button. At the bottom right, there is a black box labeled 'Ad hidden'.

Fig. 7: Mostramos una captura de pantalla de los anuncios identificados por nuestro sistema en el sitio Web de The New York Times. Uno de estos anuncios se clasifica como redireccionado, otro no está basado en intereses y el de abajo a la derecha se oculta de acuerdo con la política de bloqueo del usuario.

categorizador no puede clasificarla; cuando no hay datos suficientes para entrenar los modelos PMF de p y q ; o cuando la ejecución del optimizador excede el tiempo de ejecución máximo permitido.

5 EVALUACIÓN

En esta sección, evaluamos empíricamente el sistema propuesto y analizamos varios aspectos de la publicidad basada en el comportamiento. El análisis de esta forma de publicidad se realiza a partir de los anuncios, así como los datos de navegación de 40 usuarios de *MyAdChoices*. A nuestro entender, este estudio constituye el primer intento, aunque preliminar, de investigar la publicidad conductual y la singularidad del perfil en un entorno real a partir de perfiles de navegación de usuarios reales.

5.1 Datos

Distribuimos *MyAdChoices* a colegas y amigos y les pedimos que lo instalaran y navegasen por Internet de manera habitual durante al menos un mes. El experimento se realizó entre diciembre de 2015 y enero de 2016, aunque se han llevado a cabo nuevos experimentos durante 2017 y 2018. Los datos recopilados por nuestra extensión de navegador Web se envían a nuestros servidores cada hora. Por otro lado, la extensión se configuró para una fracción de las páginas revisadas de 100%. Es decir, cada página explorada por un usuario fue revisitada por nuestro sistema en el modo incógnito.

Los participantes son en su mayoría investigadores y estudiantes en nuestros países de residencia, Francia, India y España. No se intentó relacionar los datos recopilados con las identidades personales de los voluntarios. Como preprocesado, eliminamos a aquellos usuarios que visitaron menos de 100 sitios, quedando un total de 40 usuarios.

5.2 Resultados

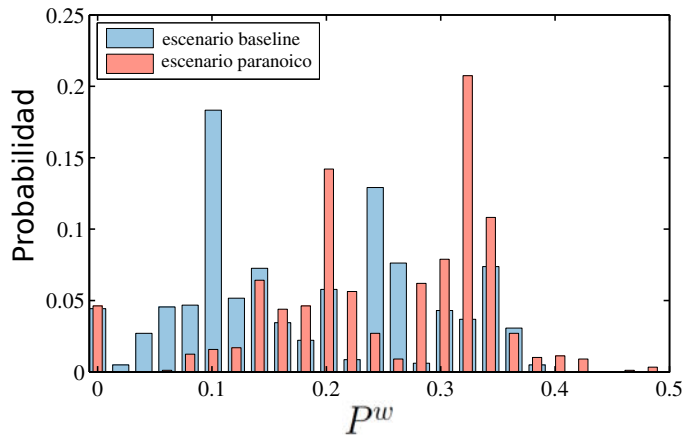


Fig. 8: PMF de la probabilidad de error en el peor casos para los dos escenarios asumidos en este proyecto.

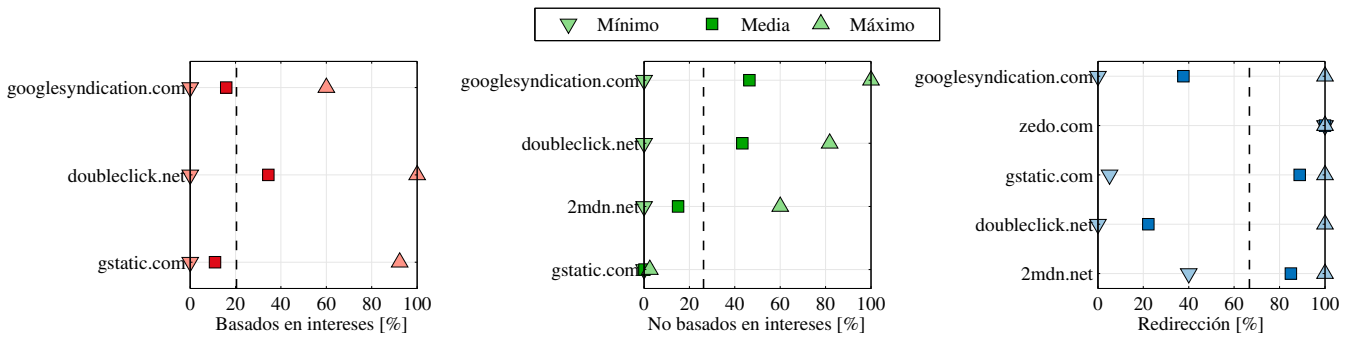


Fig. 9: Selectores de anuncio. Anuncios basados en intereses, no basados en intereses, y redirigidos para el escenario *baseline*.

5.2.1 Rendimiento

La evaluación de una herramienta de transparencia de anuncios es extremadamente complicada, ya que no es posible conocer los motivos (localización, perfil de intereses, contexto) por los que los selectores de anuncios decidieron enviar los anuncios a nuestros participantes. La eficacia de herramientas de transparencia se ha evaluado ocasionalmente a través de la inspección manual [80], [81]. Sin embargo, recientemente se ha demostrado que esta vía es extremadamente propensa a errores [82]. En esta sección, evaluamos la probabilidad de error del detector de anuncios basado en intereses, teniendo en cuenta la imposibilidad de verificar las decisiones de un detector con la verdadero estado de los anuncios analizados (es decir, si realmente están basados en intereses o no).

Antes de continuar con esta evaluación, primero nos centramos en la disponibilidad de los datos de categorización en nuestro conjunto de datos. Recuerde que nuestro sistema clasifica los anuncios en categorías temáticas de sus páginas de destino. Para ello, el módulo de categorización utiliza las palabras incluidas en la URL, las palabras clave, el título y el contenido de la página de destino. En nuestros experimentos, encontramos que solo el 0.60% de los anuncios no se pudieron clasificar utilizando esta información, lo que representa un buen índice de disponibilidad. En la mayoría de los casos, la razón fue la falta de apoyo lingüístico. Como se explica en la Sec. 4.2.2.3, actualmente nuestro módulo de categorización funciona solo para español, inglés, francés e italiano.

Habiendo comprobado el rendimiento de nuestro categorizador, ahora pasamos a analizar el detector minimax robusto. En todas las ejecuciones de la librería de optimización CLP (que incluye los dos escenarios *baseline* y *paranoico*), no se informó de ningún solo error a nuestros servidores. Es decir, nuestro sistema pudo computar con éxito dicho detector, sin exceder el tiempo de ejecución máximo permitido para este cálculo, establecido en 0.5 segundos en estos experimentos. Del mismo modo, el software IPOPT no informó de ningún error al calcular los valores de unicidad mínima.

La Fig. 8 muestra la PMF de la probabilidad de error del detector de anuncios basado en intereses. En el escenario *baseline*, observamos una media y una varianza de 0.1827 y 0.0105, respectivamente. En el caso *paranoico*, estos dos momentos dan 0.2504 y 0.0094. Hay dos observaciones importantes en estas figuras. Primero, ambos casos exhiben probabilidades de error relativamente bajas, con valores

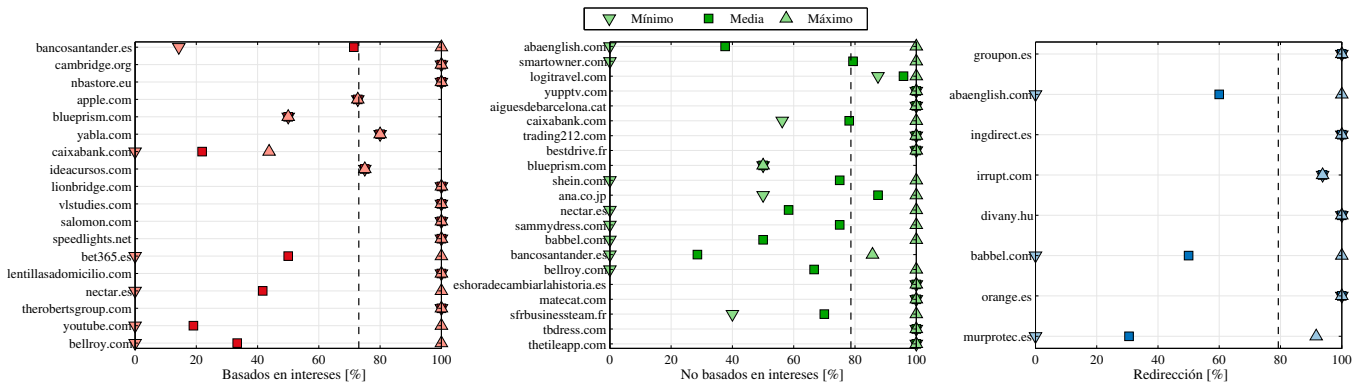


Fig. 10: *Anunciantes*. Anuncios basados en intereses, no basados en intereses, y redirigidos para el escenario *baseline*.

esperados aproximadamente por debajo de 1/4. En segundo lugar, el escenario paranoico parece ser un poco más propenso a errores en términos de detección de anuncios basados en intereses. Una posible explicación para esto es un mayor parecido entre las distribuciones p y q en este escenario. Intuitivamente, cuanto más distintas son estas distribuciones, menor es la probabilidad de identificar incorrectamente un anuncio basado en intereses.

Tabla 4: Valores de porcentaje mínimo, medio y máximo de anuncios basados en intereses, no basados en intereses y redireccionados sobre todos los usuarios en nuestro conjunto de datos.

	Escenario <i>baseline</i> [%]			Escenario <i>paranoico</i> [%]		
	mín.	media	máx.	mín.	media	máx.
Basados en intereses	0	13.2	60.0	0	17.8	66.7
No basados en intereses	0	31.7	78.4	0	29.4	76.1
Redirección	0	55.1	100	0	52.8	100

5.2.2 Anuncios basados en Intereses y de tipo Redirección

En esta sección se examinan varios aspectos de la publicidad conductual y la redirección, incluyendo un análisis de las entidades que entregan tales formas de publicidad; las categorías temáticas más específicas en nuestros experimentos; las discrepancias entre los escenarios *baseline* y *paranoicos*; y un estudio preliminar de la relación entre la publicidad basada en intereses y la singularidad del perfil.

En la Tabla. 4 se muestran algunas cifras generales sobre publicidad conductual y de tipo redirección. Para obtener estas cifras, hemos calculado, para cada usuario con un mínimo de 10 anuncios recibidos, el porcentaje de anuncios basados en intereses, no basados en intereses y redirigidos. Los valores mínimo, medio y máximo de esos porcentajes sobre todos los usuarios son los valores representados en esta tabla.

Los resultados indican claramente que la redirección es la estrategia de segmentación de anuncios más común, seguida por la publicidad no basada en intereses y la segmentación por conducta. Este orden se observa tanto en ambos escenarios, con pequeñas diferencias en los valores porcentuales. Uno de los resultados más interesantes es la prevalencia relativamente pequeña de la publicidad basada en perfiles, que representa un tercio de los anuncios de redirección. Esto contrasta con trabajos anteriores que reportan porcentajes promedio más altos de este tipo de publicidad para perfiles falsos [83], pero en línea con los recientes estudios de marketing [84] que señalan que los anuncios de redirección se preferidos a los anuncios basados en intereses en una proporción 3:1.

5.2.2.1 Selectores de Anuncios y Anunciantes: En esta subsección, examinamos los selectores de anuncios que, en nuestro conjunto de datos, fueron responsables de la entrega de publicidad conductual, no basados en perfiles y de redirección. Para ello, calculamos el porcentaje de anuncios redirigidos y no basados en intereses y basados en intereses de cada una de estas entidades. La Fig. 9 muestra los valores mínimos, medios y máximos de dichos porcentajes para cada selector de anuncios con un mínimo de 10 anuncios. En cada uno de los tres diagramas, los selectores de

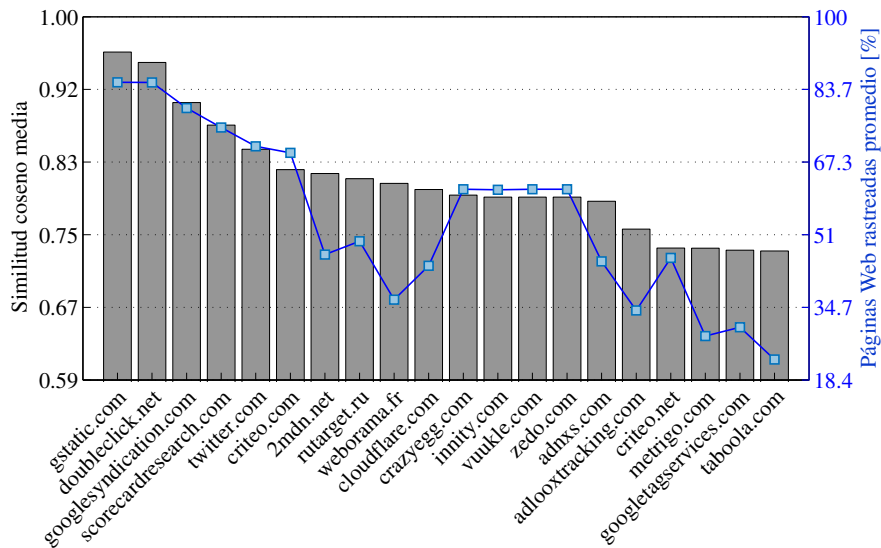


Fig. 11: Mostramos los valores de similitud coseno entre los perfiles real y observado, promediados para todos los usuarios y por entidad de rastreo.

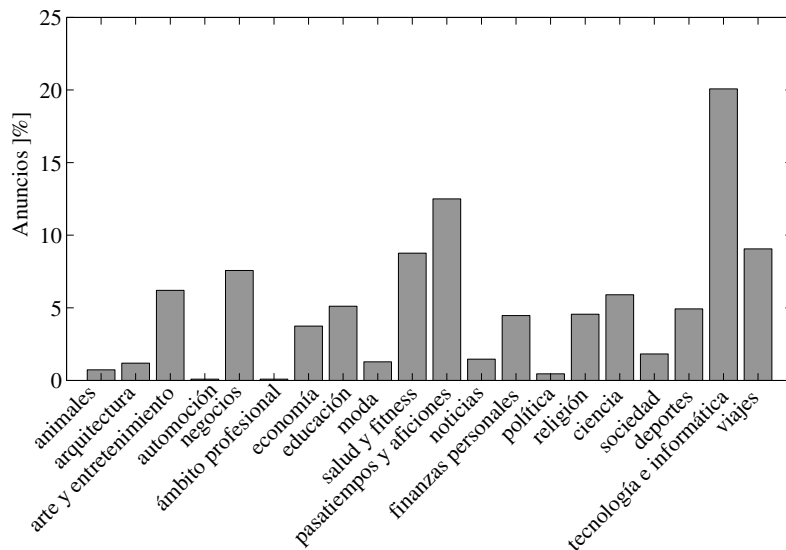


Fig. 12: Porcentaje de anuncios en las 20 categorías principales.

anuncios se clasificaron en orden decreciente del número total de anuncios publicados, de arriba a abajo. Las líneas verticales de puntos indican porcentajes promedio sobre los selectores de anuncios mostrados.

La figura en cuestión muestra solo cinco selectores de anuncios. En nuestro conjunto de datos, estas entidades fueron responsables del 98,99% del número total de anuncios. Aunque no fue inesperado, las empresas de publicidad de Google (*googlesyndication.com*, *doubleclick.net* y *gstatic.com*) fueron las que monopolizaron las tres clases de anuncios. Se observó que dicha plataforma de anuncios distribuyó principalmente anuncios no basados en intereses y redireccionados, mientras que DoubleClick y *gstatic.com* proporcionan en mayor medida publicidad conductual y de redirección, respectivamente. Los selectores de anuncios restantes fueron *zedo.com* y *2mdn.net*. La mayoría de los anuncios publicados por estas empresas publicitarias fueron de redirección. Por último, el caso paranoico muestra resultados similares.

Se utilizó la misma metodología para analizar los anunciantes de nuestro conjunto de datos y para generar la Fig. 10. Esta figura muestra a Banco Santander, Cambridge University Press, NBA Store y Apple como los anunciantes con las tasas más altas de publicidad conductual. SmartOwner, Logitravel.com, YuppTV y CaixaBank, por otro lado, lideran el ranking de anuncios no basados en

intereses, y Groupon, ABA English e Ing Direct son las compañías más interesadas en redirección. Si bien no podemos derivar una regla general de estos resultados, observamos que las grandes empresas parecen más interesadas en anuncios no basados en intereses. Esto podría ser una consecuencia inmediata de las mayores posibilidades de que esas empresas, por ejemplo, ganen subastas de anuncios en RTB, en comparación con las empresas con un poder de compra limitado.

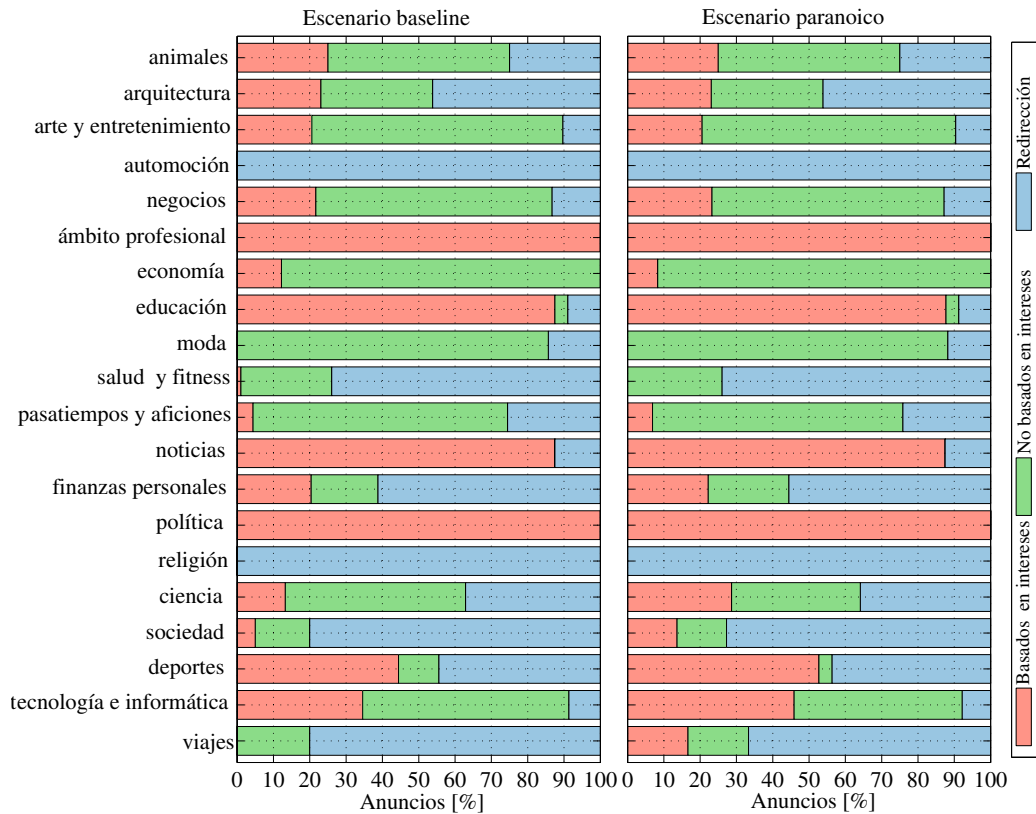


Fig. 13: Porcentaje de anuncios en algunas de las categorías de interés de nivel superior para los escenarios baseline y paranoico.

5.2.2.2 Escenarios Baseline y Paranoico: A continuación, analizamos el porcentaje general de coincidencia entre los dos escenarios en términos de detección de anuncios basados en intereses. Para ello, para cada anuncio verificamos si la decisión tomada por el detector en el modo baseline coincidía con la decisión tomada por el detector en el caso paranoico.

El porcentaje de coincidencia observado en nuestro conjunto de datos fue ciertamente alto, especialmente para la plataforma de anuncios `gstatic.com`, que obtuvo un 97.4%. Aunque más bajos, los porcentajes de coincidencia para `DoubleClick` (75.6%) y `googlesyndication.com` (87.0%) también fueron notables. Una explicación plausible de este comportamiento es la similitud del perfil p estimado en ambos escenarios, lo que podría indicar que `gstatic.com` se utilizó únicamente sus propios datos de seguimiento y, por lo tanto, no complementó esta información con los perfiles de navegación de otras fuentes.

Precisamente, la similitud de los perfiles p y t se investiga en nuestra siguiente figura, la Fig. 11. Recuerde que estos perfiles se estiman a partir de los clickstream observados y reales, respectivamente. Para calcular la Fig. 11, mantuvimos un registro de todas las entidades que rastreaban las visitas de los usuarios; estas entidades eran plataformas publicitarias, anunciantes y también rastreadores de datos analíticos. Luego, a partir de dichos registros, calculamos el porcentaje de páginas rastreadas por cada una de estas entidades, así como la similitud coseno entre los perfiles observados y los reales. La figura en cuestión muestra estos valores de porcentaje y similitud promediados para todos los usuarios.

Hay una serie de observaciones a destacar en esta figura. En primer lugar, las plataformas de anuncios de Google son las entidades con las capacidades de seguimiento más amplias. En particular,

gstatic.com, DoubleClick y googlesyndication.com rastrearon a los usuarios en 92.9%, 88.2% y 81.2% de las páginas visitadas. Una consecuencia inmediata de esto son los altos valores de similitud coseno observados. En segundo lugar, los resultados son consistentes con los porcentajes de coincidencia de escenario proporcionados al comienzo de esta subsección. En tercer lugar, se observó que los perfiles p de las empresas publicitarias con capacidades de seguimiento limitadas, como Metrigo y Taboola, son relativamente similares a los perfiles reales correspondientes. Si bien no es posible encontrar una respuesta precisa para este resultado, la razón podría encontrarse en el modelo de perfil de usuario basado en frecuencias *relativas*.

Finalmente, nos gustaría enfatizar lo apropiado de los escenarios propuestos para los selectores de anuncios examinados en estos experimentos. Recuerde que el escenario baseline no contempla el intercambio de información de seguimiento con otros selectores de anuncios y rastreadores, mientras que el caso paranoico sí lo hace; este último escenario también considera que el seguimiento es ubicuo. Los resultados proporcionados a lo largo de esta sección experimental se basan en el supuesto de que googlesyndication.com, DoubleClick y gstatic.com funcionan de manera independiente en el escenario baseline. Sin embargo, dado que todas son empresas de publicidad de Google, uno podría esperar que estas tres empresas hubieran intercambiado información entre sí. El escenario paranoico captura con precisión este posible intercambio de datos de seguimiento. Además, la ubicuidad del seguimiento se justifica por el hecho de que estas plataformas publicitarias se combinan para un total de 99.08% de páginas rastreadas (es decir, rastrean a los usuarios casi en todas las páginas que visitan).

5.2.2.3 Categorías de Interés Observadas: La Fig. 12 muestra la distribución de probabilidad de las categorías temáticas de anuncios. En esta figura, consideramos solo aquellos temas para los que recopilamos un mínimo de 5 anuncios. Los resultados indican que las categorías de interés más populares fueron “tecnología e informática”, “pasatiempos y aficiones”, “viajes” y “salud y fitness”, con porcentajes de 18.4%, 11.5%, 8.3% y 8.1%, respectivamente.

La Fig. 13 ilustra, por otra parte, las estrategias de personalización que se observaron en cada una de las 20 categorías representadas en la Fig. 12. Como puede verse, se obtuvieron resultados muy similares para baseline y el escenario paranoico. Nuestros hallazgos muestran que los anuncios redirigidos fueron más frecuentes en categorías como “automoción”, “religión”, “sociedad” y “viajes”, lo que parece estar en parte de acuerdo con algunas encuestas de marketing [85], [86]. Por otra parte, los anuncios basados en perfiles se observaron predominantemente en “ámbito profesional”, “educación”, “noticias” y “política”, y los anuncios no basados en intereses se orientaron en gran medida a “moda”, “economía” y “pasatiempos y aficiones”.

5.2.2.4 Anuncios basados en Intereses y Singularidad de Perfiles: En nuestros últimos experimentos, exploramos brevemente si es más probable (o no) que los perfiles de navegación comunes reciban anuncios basados en intereses. Con este propósito, para cada anuncio clasificado como basado en intereses y no basado en intereses, analizamos los valores mínimos de unicidad del selector de anuncios que lo sirve. Las distribuciones de probabilidad de tales valores se representan en la Fig. 14.

Como se puede observar, las dos PMF son muy similares, lo que significa que la probabilidad de entregar un anuncio basado en intereses puede no depender de la singularidad del perfil observado. De hecho, los valores esperados de estas distribuciones son 0.8949 bits para anuncios basados en perfiles y 0.8834 bits para anuncios no basados en intereses; y la divergencia KL (una medida de su discrepancia) es de 0.4344 bits. Sobre la evidencia actualmente disponible, parece justo sugerir que la singularidad o el carácter común de un perfil no es una característica que los selectores de anuncios usen en general para decidir sus estrategias de segmentación de usuarios. Sin embargo, se requeriría el análisis de grandes volúmenes de datos y pruebas adicionales para respaldar esta afirmación.

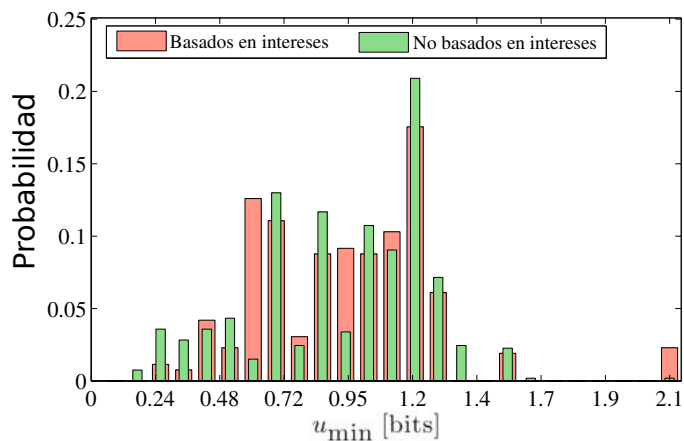


Fig. 14: Mostramos las PMFs de los valores de unicidad de perfil analizados, cuando los anuncios se clasifican como basados en intereses y cuando se consideran no basados en intereses.

6 PROYECTOS RELACIONADOS

Esta sección revisa el estado del arte relevante para este trabajo. Procedemos a explorar, primero, las tecnologías de software actuales destinadas a bloquear anuncios; y en segundo lugar, examinamos aquellas propuestas destinadas a proporcionar transparencia a la publicidad en línea.

6.1 Bloqueadores de Anuncios

En Internet existe una gran variedad de tecnologías de bloqueo de anuncios. En esencia, estas tecnologías actúan como cortafuegos entre el navegador Web del usuario, por una parte, y las plataformas publicitarias y las empresas de seguimiento, por otra. Específicamente, los bloqueadores de anuncios impiden aquellas solicitudes HTTP que se realizan cuando el navegador carga una página Web y que no son originadas por su editor. Estas solicitudes se conocen comúnmente como solicitudes de red de terceros, como se menciona en la sección introductoria de este trabajo.

La mayoría de estas herramientas se implementan como complementos de navegador de código abierto, y realizan dicho bloqueo con la ayuda de una base de datos o *lista negra* de plataformas de anuncios y rastreadores. Básicamente, estas listas incluyen expresiones y reglas regulares para filtrar las solicitudes de red de terceros que se consideran pertenecientes a anuncios o rastreadores. El mantenimiento de dichas listas negras se realiza manualmente por los desarrolladores de las tecnologías y, en algunos casos, por comunidades de usuarios. Algunos de los bloqueadores de anuncios más populares son Adblock Plus [4] y Adblock [87]. Dentro de esta lista de tecnologías de bloqueo, también incluimos herramientas anti-rastreo como Ghostery [88], Disconnect [89], Lightbeam [90] y Privacy Badger [91], que, desde un punto de vista técnico, funcionan exactamente como bloqueadores de anuncios y, por lo tanto, también pueden bloquear anuncios.

Recientemente ha surgido una iniciativa para el bloqueo de anuncios que utiliza listas blancas para permitir solo “anuncios aceptables”. Los criterios de aceptabilidad típicamente incluyen no invasividad, silencio y tamaño pequeño [92]. Sin embargo, debido a que estos criterios dependen en última instancia de los desarrolladores de los bloqueadores de anuncios, esta solución no representa ningún avance real en el retorno a los usuarios del control sobre la publicidad. De hecho, esta iniciativa de “anuncios aceptables” causó una gran controversia en la industria, cuando se descubrió que el bloqueador de anuncios más popular estaba aceptando dinero de algunas de las compañías incluidas en la lista blanca [5].

6.2 Transparencia de Anuncios

A nuestro entender, en términos de transparencia, nuestro trabajo es el primero en proporcionar a los *usuarios finales* información detallada sobre publicidad conductual en *tiempo real*. Como veremos a continuación, solo un par de trabajos abordan el problema de la detección de anuncios basada en intereses. Sin embargo, el principal inconveniente de estas propuestas es que no están pensadas para

los usuarios de la Web, es decir, no están diseñadas para ser utilizadas por un solo usuario que desee averiguar qué anuncios concretos han sido generados en base a sus datos de navegación. En cambio, estas propuestas consisten en plataformas destinadas a recopilar y analizar datos publicitarios a gran escala con fines de *investigación*. En general, permiten ejecutar experimentos en un entorno limitado y controlado, para estudiar los anuncios mostrados a perfiles de usuario muy específicos y generados artificialmente.

En esta subsección examinaremos estos trabajos, teniendo en cuenta que ninguno de ellos está concebido como una herramienta de la que los usuarios puedan beneficiarse directa y plenamente. Además, e igualmente importante, veremos que estas propuestas se basan en un modelo demasiado simplista y, en muchos casos, erróneo, del proceso de entrega de anuncios real. Además, muy a menudo recurren a heurísticas simples, no justificadas rigurosamente, para llevar a cabo sus estudios de medición de la publicidad conductual.

En contraste con estos trabajos, nosotros proponemos un estudio formal de esta forma de publicidad que se basa en un modelo más general y preciso del proceso de publicación de anuncios, que tiene en cuenta su complejidad y el nuevo paradigma de RTB, y que aborda los desafíos, mientras que simplemente los ignoran. Procedemos siguiendo una metodología formal fundamentada en los campos de la estimación estadística y la optimización robusta. Además, en comparación con estos trabajos, nuestro análisis no solo determina si un anuncio se basa en intereses o no, sino que también explora un aspecto crucial de los intereses rastreados y perfilados por las empresas publicitarias, en concreto, el carácter común de los perfiles de usuario. A continuación, describimos en más estas propuestas.

El primer intento en identificar los retos que pueden surgir al medir diferentes aspectos de la publicidad en línea se realizó en [93]. Aunque no estaban particularmente interesados en la publicidad conductual, los autores investigaron aspectos como el impacto de la recarga de páginas y las cookies en la publicidad, y destacaron las dificultades encontradas a través de algunos experimentos simples.

Después de este trabajo,[83], [94] propuso una plataforma que automatiza la recopilación de ciertas estadísticas sobre la publicidad basada en perfiles de navegación. La plataforma propuesta crea perfiles de usuario artificiales con categorías temáticas muy específicas y no superpuestas (es decir, perfiles con categorías activas solo en deportes, solo viajes, etc.) mediante la emulación de las visitas a páginas relacionadas con esos temas. La herramienta en cuestión alterna esta navegación de entrenamiento con visitas a páginas Web meteorológicas, donde verifican si las categorías de los anuncios recibidos coinciden con la categoría del perfil correspondiente; los autores justifican el uso de estas páginas relacionadas con el clima argumentando que, allí, los anuncios contextuales se detectan más fácilmente. Para llevar a cabo esta comprobación, la herramienta primero filtra las páginas de destino que pueden corresponder a anuncios genéricos y contextuales. Con este objetivo — y de manera similar a nuestra herramienta —, vuelve a visitar, en modo incógnito, cada visita a una página de meteorología y mantiene un registro de los anuncios entregados en esta sesión. Al eliminar las páginas de destino comunes a ambas sesiones, los autores afirman descartar la *mayoría* de los anuncios no basados en intereses y basados en contenido.

Aparte del hecho de que dicha plataforma no está destinada a usuarios finales ni proporciona funcionalidades de transparencia de anuncios en tiempo real, el inconveniente más importante es su alcance de aplicación extremadamente limitado. En primer lugar, solo funciona para perfiles de interés único, y en segundo lugar, la transparencia solo puede incluirse en dichas páginas meteorológicas, lo que proporciona una visión muy simplista y superficial de la publicidad conductual. No obstante, esta no es la única limitación. Para detectar anuncios basados en intereses, los autores cometen el error de simplificar en exceso el proceso de entrega de anuncios al asumir algún tipo de determinismo: consideran que *la mayoría* de los anuncios no basados en intereses que un usuario puede recibir en una sesión con rastreo y otra sin rastreo será exactamente lo mismo, lo que ignora la aleatoriedad inherente del proceso de publicación de anuncios.

Finalmente, los trabajos citados [83], [94] evalúan su propuesta utilizando una medida de distancia entre los términos que aparecen en las páginas de destino de los anuncios y los de las páginas de

Tabla 5: MyAdChoices versus otras herramientas de transparencia en publicidad conductual.

Enfoque	Tipo de herramienta	Inconvenientes
[83], [94]	plataforma de investigación	<ul style="list-style-type: none"> ○ válido para perfiles de una sola categoría, ○ funcionalidad de transparencia disponible solo en páginas relacionadas con la meteorología, ○ modelo impreciso del proceso de distribución de anuncios, ○ navegación paralela en modo incógnito, ○ solo escenario paranoico, ○ no se permite la detección de diferentes tipos de anuncios;
[95]	plataforma de investigación	<ul style="list-style-type: none"> ○ válido para perfiles de una sola categoría, ○ funcionalidad de transparencia limitada a usuarios que visitan las mismas páginas, ○ se ignoran los anuncios genéricos y los contextuales, ○ solo escenario paranoico;
[21]	plataforma de investigación	<ul style="list-style-type: none"> ○ sólo funciona para DoubleClick, ○ modelo simplificado del proceso de entrega de anuncios (e.g., se ignoran los anuncios genéricos y RTB, sólo para perfiles de intereses a largo plazo), ○ decisión binaria, i.e., los anuncios, o bien son contextuales, o bien están basados en intereses, ○ modelo inconsistente de rastreo y compartición de datos de usuario;
[80], [82], [81]	plataforma de investigación	<ul style="list-style-type: none"> ○ no escalable para navegación Web [80], [82], ○ tráfico de red inaceptable y sobrecarga de cómputo, si está destinado a usuarios finales [80], [82], ○ puede detectar redirección pero no anuncios basados en comportamiento;
MyAdChoices	herramienta para usuarios	<ul style="list-style-type: none"> ○ una fracción de visitas en modo incógnito.

entrenamiento. Si bien esto cuantifica la similitud entre las categorías de los anuncios y las categorías individuales de los perfiles, los autores no evalúan el método para detectar anuncios basados en perfiles. Una consecuencia importante de esta falta de evaluación es que los anuncios genéricos que pertenecen a la categoría activa del perfil siempre se clasificarán como basados en intereses (siempre que no se hayan entregado en las sesiones de incógnito), y la plataforma no informará de ningún error al respecto en dicha clasificación. Por el contrario, MyAdChoices proporciona, para cada anuncio, la probabilidad de error incurrido en la estimación de su clase.

Siguiendo el mismo espíritu,[95] presenta una infraestructura de rastreo de anuncios que no va dirigida a proporcionar transparencia, sino a analizar diferentes aspectos de la publicidad a gran escala. Entre otros aspectos, los autores estudian la tasa de llegada promedio de nuevos anuncios y la distribución del número de anuncios y anunciantes por página. Además, examinan brevemente la publicidad conductual siguiendo un enfoque similar al de [83], [94]. Emulan los hábitos de navegación de alrededor de 300 usuarios con intereses en una sola categoría e intentan ver qué anuncios están más orientados a qué perfiles cuando visitan un subconjunto de páginas Web seleccionadas. Su análisis de la orientación del perfil asume que, si un anuncio se muestra con más frecuencia en un perfil dado que en otros, entonces este anuncio se ha personalizado a dicho perfil. Sobre la base de ciertas heurísticas, los autores comparan la frecuencia de aparición de cada anuncio (para cada perfil) con un perfil uniforme, y concluyen que un anuncio está basado en intereses si el resultado de dicha comparación excede un cierto umbral.

El marco propuesto adolece de las mismas limitaciones de los dos trabajos mencionados anteriormente. Además, los autores ignoran que los anuncios también pueden ser contextuales y genéricos, y que la frecuencia de aparición de los anuncios depende de factores altamente dinámicos. Por otro lado, una implementación práctica de este marco en el lado del usuario sería inviable, ya

que requeriría que los usuarios visiten las mismas páginas (para habilitar la funcionalidad de transparencia) y exhiban perfiles de una sola categoría.

Una plataforma similar se propone en [21], que estudia los anuncios entregados a algunos perfiles artificiales, en este caso, a partir del conjunto de datos de búsquedas de AOL [45]. La herramienta no está destinada a usuarios finales y proporciona un marco que tiene como objetivo estudiar la publicidad contextual y basada en intereses a gran escala. La plataforma, que funciona en modo offline y está restringida a los anuncios de DoubleClick, analiza dos conjuntos de datos con este fin: las categorías de interés de los anuncios de *todos* los recibidos tanto en una sesión rastreada como en un modo de navegación incógnito. Luego, los autores utilizan un clasificador binario para decidir si un anuncio que pertenece a una determinada categoría está basado en el perfil de navegación o es contextual.

Las principales limitaciones de esta herramienta provienen del modelo simplificado e inexacto asumido para el proceso de entrega de anuncios. En primer lugar, no tiene en cuenta los anuncios genéricos o no personalizados. En segundo lugar, la decisión es binaria en el sentido de que el resultado de una clasificación de anuncios no puede ser contextual y basado en intereses, por lo que se pasa por alto que la gran mayoría de las plataformas de anuncios permiten la selección de múltiples objetivos de campaña. En tercer lugar, dicha clasificación se basa en el conjunto de datos *completos* de los anuncios recopilados en las sesiones rastreadas e de incógnito, lo que ignora el hecho de que DoubleClick (como cualquier selector de anuncios) puede crear perfiles a corto plazo o usar *cualquier* ventana temporal para perfilar los intereses de navegación de los usuarios, no necesariamente la que abarca todo el historial de navegación. Por último, pero no menos importante, la herramienta en cuestión no refleja el funcionamiento real de la plataforma publicitaria en la que se centra, es decir, que DoubleClick puede emplear tecnologías RTB modernas para publicar anuncios [48], [35]. Por un lado, los autores parecen asumir un escenario de referencia, ya que el perfil de usuario se crea solo a partir de las páginas rastreadas por esta plataforma publicitaria. Pero, por otro lado, ignoran completamente la tecnología de publicación de anuncios RTB, y el hecho de que los participantes de la subasta de anuncios de DoubleClick no compartan los mismos datos de perfiles. Es decir, los autores parecen asumir, al mismo tiempo, un escenario paranoico, lo cual es contradictorio. Nos gustaría enfatizar que nuestro trabajo aborda estos cuatro problemas, modelando la combinación de múltiples decisiones de segmentación de anuncios, basándonos en la noción de selector de anuncios, construyendo modelos de perfil de usuario independientes para elector de anuncios, y considerando cualquier posible ventana de tiempo elegida por tales entidades a través de la definición de clases de incertidumbre.

Otro trabajo más reciente que realiza experimentos a partir de perfiles artificiales es [80]. Este estudio rastrea los datos personales recopilados por varios servicios Web e intenta correlacionar entradas de datos (por ejemplo, correos electrónicos y consultas de búsqueda) con salidas de datos (por ejemplo, anuncios y enlaces recomendados). La plataforma propuesta aborda este problema de correlación en un sentido amplio, y ha sido testada para los anuncios que se muestran en Gmail. La plataforma se basa en el mantenimiento de un número de *cuentas ocultas*, es decir, réplicas de la cuenta original (por ejemplo, una cuenta de correo electrónico), pero que difieren en un subconjunto de entradas. Todas estas instancias de cuenta son operadas en paralelo por el sistema y se usan para comparar las salidas recibidas. Intuitivamente, si un anuncio se muestra con mayor frecuencia en aquellas cuentas que comparten una determinada entrada (por ejemplo, un correo electrónico), y este anuncio nunca aparece en el resto de instancias ocultas, entonces esta entrada probablemente sea la causa de dicha publicidad.

La plataforma en cuestión no requiere una cuenta oculta para cada combinación posible de datos de entrada, sino un número logarítmico de dichas cuentas en el número de entradas, lo que la hace adecuada para la aplicación donde se crea una instancia. Sin embargo, sería totalmente inviable extenderlo para analizar los anuncios recibidos de esta aplicación controlada, por ejemplo, mientras navega por la Web. Primero, en términos de escalabilidad. Los autores afirman que su plataforma soporta la correlación de cientos de entradas (correos electrónicos), con costos razonables en términos de cuentas ocultas. Esto puede funcionar para un proveedor de servicios *único*, pero claramente no

cuando se considera en el contexto más general de la publicidad Web, con miles de compañías publicitarias rastreando usuarios en toda la Web [88] y aproximadamente noventa páginas visitadas en promedio por día [55]. En segundo lugar, crear perfiles de navegación ocultos equivalentes en el lado del usuario no sería práctico en términos de tráfico de red y carga computacional. Por otro lado, la solución propuesta verifica qué datos de entrada o combinación en particular (con un tamaño de combinación reducido, para lograr la escalabilidad mencionada anteriormente) es responsable de los datos de salida dados (por ejemplo, un anuncio). Como resultado, dicha plataforma puede funcionar para formularios de publicidad como retargeting, donde una sola visita puede ser la causa de la visualización de un anuncio, y para anuncios contextuales, que dependen de la página que se está visitando actualmente. Sin embargo, no opera a un nivel de granularidad mucho más amplio y, por lo tanto, no es adecuado para estudiar la publicidad conductual, donde los anuncios suelen publicarse en función de los historiales de navegación acumulados durante largos períodos de tiempo.

Un par de mejoras de esta última propuesta han sido publicadas en [82], [81]. Estas propuestas proporcionan respectivamente cierta validación estadística de sus hallazgos e investigan la causalidad en anuncios basados en texto. No obstante, estos trabajos no dejan de ser plataformas de medición y tienen las mismas limitaciones en términos de detección de publicidad basada en perfiles de navegación. La Tabla 5 resume las principales conclusiones de esta sección.

7 CONCLUSIONES

En los últimos años, como resultado de la proliferación de anuncios intrusivos e invasivos, el uso de herramientas de bloqueo de anuncios y anti-rastreo se ha generalizado. El problema de estas tecnologías es que plantean una opción binaria al usuario y, por lo tanto, ignoran el papel crucial de la publicidad como el principal soporte del contenido gratuito de Internet.

Creemos que tales tecnologías son solo una solución a corto plazo y que se necesitan mejores herramientas para resolver este problema a largo plazo. La mayoría de los usuarios no está en contra de los anuncios y en realidad están dispuestos a aceptar algunos anuncios para ayudar a los sitios Web. Sin embargo, esto se pasa porque el proceso de publicación de anuncios sea transparente y los usuarios puedan controlar la información personal recopilada.

Dado que diferentes usuarios pueden tener diferentes motivaciones para usar bloqueadores de anuncios y anti-rastreadores, este documento propone una tecnología Web inteligente que aporta transparencia a la publicidad en línea y ayuda a los usuarios a imponer sus propias decisiones particulares sobre los anuncios. El objetivo principal de esta tecnología es, primero, informar a los usuarios de cómo las empresas de publicidad explotan sus datos de navegación; y, en segundo lugar, permitirles reaccionar en consecuencia otorgándoles un control flexible sobre la publicidad.

La tecnología propuesta proporciona transparencia sobre la publicidad conductual mediante dos estimadores aleatorios. El primero se basa en un modelo teórico del proceso de publicación de anuncios y utiliza la metodología de optimización robusta para abordar el problema de modelar los perfiles disponibles en las plataformas de anuncios. Este último arroja luz sobre estos perfiles al calcular la estimación de unicidad de perfiles en el caso más desfavorable.

Estos dos detectores se han integrado en una arquitectura de sistema que puede proporcionar transparencia de anuncios y servicios de bloqueo, todo en tiempo real y en el lado del usuario. En términos de transparencia, nuestra herramienta permite a los usuarios (1) saber si los anuncios que se les entregan pueden haber sido generados como respuesta a sus perfiles de navegación, y (2) averiguar si dichos perfiles pueden revelar patrones de navegación únicos. En términos de bloqueo de anuncios, el sistema propuesto permite a los usuarios filtrar anuncios basados en intereses, no basados en intereses y redirigidos por categoría temática, y especificar condiciones de bloqueo basadas en la exclusividad del perfil.

El sistema propuesto ha sido implementado como una extensión de navegador Web y se ha evaluado experimentalmente con 40 participantes. En términos de rendimiento, los dos estimadores mostraron tiempos de ejecución por debajo de 0.5 segundos y no reportaron errores. Además, casi

todas las páginas pudieron ser categorizadas. Realizamos un análisis de la publicidad basada en perfiles de navegación, en función de los anuncios y los datos de navegación de esos voluntarios. Entre otros resultados, nuestros hallazgos muestran que la estrategia de personalización de anuncios basada en la redirección es la más común; que las empresas publicitarias de Google son las que lideran la publicidad conductual y redirigida; que las grandes empresas podrían ser los responsables de los anuncios basados en perfiles; y que la singularidad del perfil puede no ser un criterio ampliamente utilizado para publicar anuncios.

A diferencia de algunos trabajos anteriores sobre transparencia Web, nuestra herramienta está dirigida a usuarios finales, parte de un modelo más fiel y preciso del proceso de publicación de anuncios, tiene en cuenta el esquema RTB recientemente establecido, y se cimienta sobre en una metodología formal basada en los campos de estimación estocástica y optimización robusta.

APÉNDICE A PROBLEMA DE FACTIBILIDAD

Este apéndice demuestra la factibilidad de los problemas de optimización (2) y (4). En particular, muestra que las restricciones dadas por el poliedro \mathcal{P} son consistentes, o dicho de otra manera, que el conjunto de puntos que los satisfacen es no vacío. Por simplicidad, cambiamos el nombre de las tuplas p^{\min} y p^{\max} simplemente con los símbolos r y s , respectivamente.

Para que (2) y (4) sea factible, necesitamos que $\sum_i r_i \leq 1$ y $\sum_i s_i \geq 1$. Para comprobar esto, considere lo contrario. Por un lado, tener $\sum_i r_i > 1$ y $\sum_i s_i < 1$ nos lleva a una contradicción, ya que por definición $r_i \leq s_i$. Por otro lado, es sencillo verificar que, si $\sum_i r_i > 1$, entonces $\sum_i p_i > 1$, y que, si $\sum_i s_i < 1$, entonces $\sum_i p_i < 1$, lo que contradice el hecho de que p es una PMF.

A continuación, comprobamos que se cumple el requisito $\sum_i s_i \geq 1$. La prueba de la condición $\sum_i r_i \leq 1$ continúa en la misma línea y se omite. Recuerde de la Sec. 3.2.4.1 que la clase de incertidumbre \mathcal{P} se calcula considerando un modelo incremental en la secuencia de clics. Es decir, cada vez que el usuario visita una página Web, se calcula una nueva estimación de p a partir de todas las páginas visitadas hasta el momento. Luego, en función de esta nueva distribución estimada, nuestro sistema actualiza r y s , si es necesario.

El sistema propuesto requiere un número mínimo de páginas visitadas w_{\min} para estimar p . Siguiendo la notación introducida en la Sec. 4.2.2.1, denotamos por m_i el número de páginas que están clasificadas en la categoría de tema i . Cuando se cumple dicho requisito, las tuplas r y s se inicializan a $r_i = s_i = \frac{m_i}{w_{\min}}$ para todo $i = 1, \dots, n$. En otras palabras, r y s se convierten en el MLE de p .

Sea s_i^m el componente i -th de la tupla s que se obtiene después de haber visitado m páginas. Es inmediato comprobar que $s_i^{w_{\min}} \leq \dots \leq s_i^m$ es una secuencia no decreciente para todos los i , lo que implica que $\sum_i s_i \geq 1$. Esto demuestra la factibilidad de los problemas (2) y (4).

APÉNDICE B FORMULACIÓN LP DEL DETECTOR MINIMAX ROBUSTO

Siguiendo la metodología desarrollada por [39], [40], este apéndice muestra la formulación de LP del diseño minimax robusto (2). A partir de las definiciones de P_i^w y M_{ii}^w , es fácil verificar que (2) es equivalente a

$$\max_{i=1,2} \min_{p \in \mathcal{P}} \inf M_{ii},$$

y por lo tanto equivalente al problema de optimización

$$\begin{aligned} & \text{maximize} && \zeta \\ & \text{subject to} && \inf\{\tilde{d}^T p : p \in \mathcal{P}\} \geq \zeta, \\ & && 1 - \tilde{d}^T q \geq \zeta, \\ & && 0 \preceq \tilde{d} \preceq \mathbf{1}. \end{aligned} \tag{5}$$

Debido a que el problema original (2) es factible, la calificación de restricción de Slater se cumple y, por lo tanto, se mantiene la denominada “fuerte dualidad” para el problema dual de Lagrange asociado al programa lineal (5). El problema dual en cuestión es

$$\begin{aligned} & \text{maximize} && \mu^T p^{\min} - \lambda^T p^{\max} + \nu \\ & \text{subject to} && \mu - \lambda + \nu \mathbf{1} \preceq \tilde{d}, \\ & && \lambda \succeq 0, \mu \succeq 0, \end{aligned}$$

donde λ, μ, ν son los vectores multiplicadores de Lagrange asociados con el problema de minimización (5), y p^{\min}, p^{\max} determina el poliedro \mathcal{P} definido en (1). A partir de este problema dual, es inmediato derivar la formulación LP de (3).

APÉNDICE C

MANUAL DE INSTALACIÓN Y CONFIGURACIÓN DE MYADCHOICES



C.1 Instalación

Para instalar MyAdChoices, abra Google Chrome. En el panel de opciones, seleccione “Más herramientas” y después “Extensiones” (Fig. 15a). A continuación, clique en “Cargar descomprimida” (Fig. 15b) e introduzca la ruta donde se encuentra el directorio de MyAdChoices.

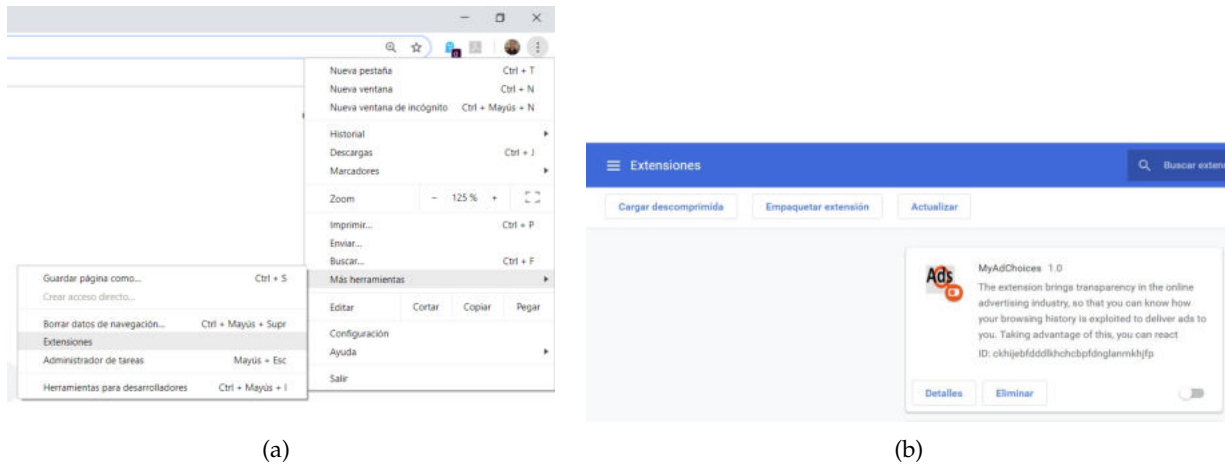


Fig. 15: Cómo instalar la extensión MyAdChoices.

Una vez instalada, debe activar el modo incógnito. Nuestra extensión requiere volver a visitar, en dicho modo, una fracción de las páginas visitadas por el usuario. En términos prácticos, cada nueva página visitada es cargada en una nueva ventana incógnito para evitar el rastreo (a través de las cookies del navegador) entre diferentes pestañas de una misma ventana del modo incógnito. Utilizamos el modo incógnito para determinar si los anuncios que se le envían a un usuario se basan en su perfil de navegación o no.

Para activar el modo incógnito, debe clicar la opción “Detalles” en el cuadro de la extensión MyAdChoices (Fig. 16a). Una vez dentro, haga clic en “Permitir en modo de incógnito” (Fig. 16b).

Para un correcto funcionamiento de nuestra herramienta, es necesario además que deshabilite todas las extensiones de bloqueo de anuncios y anti-rastreadores.

Por otro lado, si desea deshabilitar nuestra extensión durante un tiempo, es necesario que la desinstale en lugar de deshabilitarla. Nuestra extensión necesita construir no solo el perfil de usuario, sino también el perfil de los rastreadores. Cuando desactiva la extensión, MyAdChoices no puede

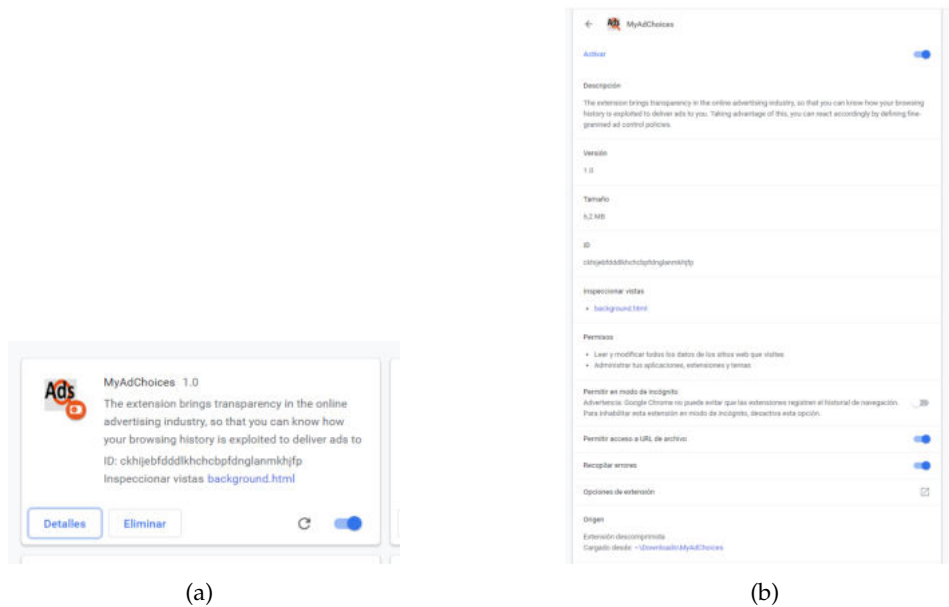


Fig. 16: Cómo habilitar el modo incógnito de Google Chrome para MyAdChoices.

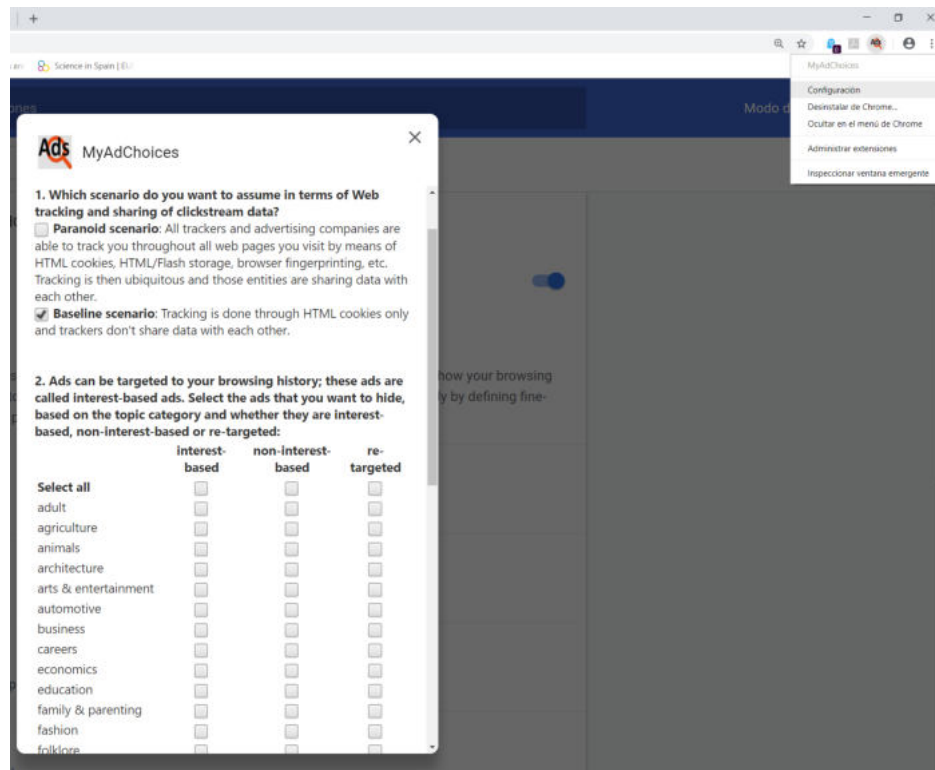


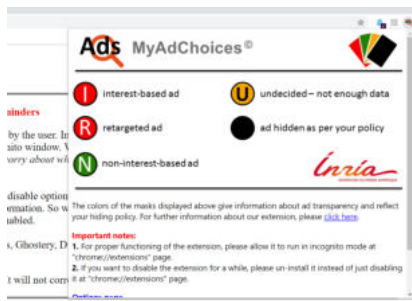
Fig. 17: Panel de configuración de MyAdChoices.

capturar esta información. Esto se debe a que Google Chrome no permite que una extensión sepa cuándo está deshabilitada o habilitada.

Finalmente, la extensión sólo es compatible con páginas Web en español, inglés, francés e italiano.

C.2 Configuración

Una vez instalada la herramienta, debe configurar sus preferencias de bloqueo de anuncios. Para ello, clique en “Configuración” en el icono de la extensión (Fig. 17). Aparecerá una ventana donde podrá elegir el escenario de rastreo que considere más oportuno (baseline o paranoico), así como



(a)

Ad selector	Advertiser	Ad-topic category	Percentile uniqueness	Transparency decision	Prob. of error	Hiding decision
googlesyndication.com	thegrid.io	technology & computing/general	81.82th	interest-based	N/A	hidden
zedo.com	thehindu.com	religion/hinduism	N/A	retargeted	N/A	hidden
googlesyndication.com	sammydress.com	fashion/general	81.82th	still learning your interests	N/A	not hidden
googlesyndication.com	sammydress.com	fashion/general	81.82th	still learning your interests	N/A	not hidden
googlesyndication.com	thegrid.io	technology & computing/general	81.82th	interest-based	N/A	hidden

(b)

Fig. 18: Panel de información de MyAdChoices.

las categorías temáticas de anuncios que desea ofuscar, en función de si los anuncios son basados en perfil, no lo son, o se trata de anuncios de redirección.

C.3 Descripción de la información facilitada por MyAdChoices

El panel de información de MyAdChoices contiene los siguientes datos (véase Fig. 18):

- **Selector de anuncios:** esta es la entidad que finalmente decide qué anuncio se muestra a un usuario en un espacio publicitario determinado. En el caso de las plataformas publicitarias tradicionales, el selector de anuncios es la propia plataforma publicitaria. En RTB, por el contrario, el selector de anuncios es el apostante que gana la subasta por mostrar su anuncio.
- **Anunciante:** entidad que desea mostrar sus anuncios en uno o varios de los espacios disponibles en una página web. La información mostrada por nuestra extensión corresponde a la página de destino del anuncio, es decir, la página del anunciante a la que se redirige el navegador al hacer clic en él. Tenga en cuenta que nuestra herramienta puede procesar un anuncio solo si se dispone de esta información de página de destino.
- **Categoría temática de anuncio:** nuestra extensión clasifica las páginas de destino de los anuncios mostrados a los usuarios, en un conjunto predefinido de temas o intereses. Utilizamos una taxonomía jerárquica de 2 niveles, compuesta por 32 categorías de nivel superior y 330 categorías o subcategorías de nivel inferior. En aras de la simplicidad, las políticas de bloqueo de anuncios solo se pueden definir para las categorías de nivel superior.
- **Unicidad de percentil:** los anuncios basados en intereses son el resultado de rastrear y perfilar nuestros intereses de navegación web. El parámetro de percentil-unicidad nos permite averiguar cuán únicos somos a través de los ojos de las compañías que nos muestran anuncios.
- **Decisión de transparencia:** para cada anuncio, nuestra extensión calcula la probabilidad que el anuncio se haya generado en base a intereses de navegación o no, y sea de tipo redirección. Si todavía no se tienen suficientes datos de navegación del usuario, el selector de anuncios y el anunciante son etiquetados como “información insuficiente”.
- **Probabilidad de error:** mostramos la probabilidad de error de nuestro detector de anuncios basados en intereses.
- **Decisión de ocultación:** cada anuncio puede ocultarse o no dependiendo de los criterios de bloqueo de anuncios.

REFERENCIAS

- [1] K. Purcell, J. Brenner, and L. Rainie, "Search engine use 2012," Pew Internet, Amer. Life Project, Res. Rep., Mar. 2012.
- [2] "The state of online advertising," Adobe, Tech. Rep., 2012, Accessed on 2015-09-11. [Online]. Available: http://www.adobe.com/aboutadobe/pressroom/pdfs/Adobe_State_of_Online_Advertising_Study.pdf
- [3] G. Marvin, "Consumers now notice retargeted ads," Marketing Land, Tech. Rep., Dec. 2013, Accessed on 2015-08-12. [Online]. Available: <http://marketingland.com/3-out-4-consumers-notice-retargeted-ads-67813>
- [4] "Adblock plus," Accessed on 2015-10-22. [Online]. Available: <https://adblockplus.org>
- [5] R. Cookson, "Google, Microsoft and Amazon pay to get around ad blocking tool," Feb. 2015, Accessed on 2014-03-10. [Online]. Available: <http://www.ft.com/cms/s/0/80a8ce54-a61d-11e4-9bd3-00144feab7de.html>
- [6] "YourOnlineChoices," European Interact. Digit. Advertising Alliance. [Online]. Available: <http://www.youronlinechoices.com/>
- [7] "Tracking preference expression (DNT)," Tech. Rep., Aug. 2015. [Online]. Available: <http://www.w3.org/TR/tracking-dnt/>
- [8] "The cost of ad blocking," PageFair, Res. Rep., Aug. 2015.
- [9] M. Arment, "The ethics of modern web ad-blocking," Aug. 2015, accessed on 2015-08-15. [Online]. Available: <http://www.marco.org/2015/08/11/ad-blocking-ethics>
- [10] W. Davis, "Ftc's julie brill tells ad tech companies to improve privacy protections," MediaPost Commun., Sep. 2015, accessed on 2015-10-01. [Online]. Available: <http://www.mediapost.com/publications/article/259210/ftcs-julie-brill-tells-ad-tech-companies-to-impro.html>
- [11] S. Thielman, "Rise of ad-blockers shows advertising does not understand mobile, say experts," Oct. 2015, accessed on 2015-10-05. [Online]. Available: <http://www.theguardian.com/technology/2015/oct/03/ad-blockers-advertising-mobile-apple>
- [12] "Adblock plus user survey results, part 3," Eyeo, Tech. Rep., Dec. 2011, Accessed on 2015-07-11. [Online]. Available: <https://adblockplus.org/blog/adblock-plus-user-survey-results-part-3>
- [13] D. Rogers, "How business can gain consumers' trust around data," Nov. 2015, Accessed on 2015-11-03. [Online]. Available: <http://www.forbes.com/sites/davidrogers/2015/11/02/how-business-can-gain-consumers-trust-around-data/>
- [14] T. Morey, T. Forbath, and A. Schoop, "Customer data: Designing for transparency and trust," Internet Draft, May 2015, accessed on 2014-02-18. [Online]. Available: <https://hbr.org/2015/05/customer-data-designing-for-transparency-and-trust>
- [15] J. Parra-Arnau, D. Rebollo-Monedero, and J. Forné, "Measuring the privacy of user profiles in personalized information systems," *Future Gen. Comput. Syst. (FGCS), Special Issue Data, Knowl. Eng.*, vol. 33, pp. 53–63, Apr. 2014. [Online]. Available: <http://dx.doi.org/10.1016/j.future.2013.01.001>
- [16] D. Rebollo-Monedero, J. Parra-Arnau, and J. Forné, "An information-theoretic privacy criterion for query forgery in information retrieval," in *Proc. Int. Conf. Secur. Technol.(SecTech)*, ser. Commun. Comput., Inform. Sci. (CCIS), vol. 259. Jeju Island, South Korea: Springer-Verlag, Dec. 2011, pp. 146–154.
- [17] M. Smith, *Targeted: How Technology Is Revolutionizing Advertising and the Way Companies Reach Consumers*, 1st ed. New York: AMACOM, Nov. 2014.
- [18] S. Yuan, A. Z. Abidin, M. Sloan, and J. Wang, "Internet advertising: An interplay among advertisers, online publishers, ad exchanges and web users," *arXiv: 1206.1754*, 2012, arXiv preprint.
- [19] T. Kawaja, "Display LUMAscape," Accessed on 2015-09-23. [Online]. Available: <http://www.lumapartners.com/lumascapes/display-ad-tech-lumascape>
- [20] V. Toubiana, "SquiggleSR," 2007. [Online]. Available: www.squigglesr.com
- [21] B. Liu, A. Sheth, U. Weinsberg, J. Chandrashekar, and R. Govindan, "Adreveal: Improving transparency into online targeted advertising," in *Proc. Hot Topics in Netw.* ACM, 2013, pp. 121–127.
- [22] J. Yan, N. Liu, G. Wang, W. Zhang, Y. Jiang, and Z. Chen, "How much can behavioral targeting help online advertising?" in *Proc. Int. WWW Conf.* ACM, 2009, pp. 261–270.
- [23] M. Aly, A. Hatch, V. Josifovski, and V. K. Narayanan, "Web-scale user modeling for targeting," in *Proc. Int. WWW Conf.* ACM, 2012, pp. 3–12.
- [24] M. M. Tsang, S. C. Ho, and T. P. Liang, "Consumer attitudes toward mobile advertising: An empirical study," *Int. J. Electron. Commer.*, vol. 8, no. 3, pp. 65–78, 2004.
- [25] "Audience Buying Guide 2011," Accessed on 2015-03-27. [Online]. Available: <http://brandedcontent.adage.com/audiencebuyingguide2011/network.php?id=12>
- [26] "Clickstream or clickpath analysis," Accessed on 2015-03-27. [Online]. Available: <http://www.opentracker.net/article/clickstream-or-clickpath-analysis>
- [27] V. Toubiana, A. Narayanan, D. Boneh, H. Nissenbaum, and S. Barocas, "Adnostic: Privacy preserving targeted advertising," in *Proc. Symp. Netw. Distrib. Syst. Secur. (SNDSS)*, Feb. 2010, pp. 1–21.
- [28] "Cisco service control online advertising solution guide," Cisco Syst., Tech. Rep., Jan. 2009.
- [29] S. Pandey, M. Aly, A. Bagherjeiran, A. Hatch, P. Ciccolo, A. Ratnaparkhi, and M. Zinkevich, "Learning to target: What works for behavioral targeting," in *Proc. Int. Conf. Inform., Knowl. Manage. (CIKM)*. ACM, 2011, pp. 1805–1814.
- [30] "US programmatic ad spend tops \$10 billion this year, to double by 2016," eMarketer, Tech. Rep., Oct. 2014. [Online]. Available: <http://www.emarketer.com/Article/US-Programmatic-Ad-Spend-Tops-10-Billion-This-Year-Double-by-2016-1011312>
- [31] S. Englehardt, "The hidden perils of cookie syncing," Aug. 2014, Accessed on 2014-12-10. [Online]. Available: <https://freedom-to-tinker.com/blog/englehardt/the-hidden-perils-of-cookie-syncing/>

- [32] “Real-time bidding protocol - cookie matching,” Accessed on 2015-10-07. [Online]. Available: <https://developers.google.com/ad-exchange/rtb/cookie-guide>
- [33] “Real-time bidding protocol - processing the request,” Accessed on 2015-10-07. [Online]. Available: <https://developers.google.com/ad-exchange/rtb/request-guide>
- [34] “Google DoubleClick ad exchange (AdX) buyer program guidelines,” Apr. 2015, Accessed on 2015-08-05. [Online]. Available: <http://www.google.com/doubleclick/adxbuyer/guidelines.html>
- [35] L. Olejnik, “Measuring the privacy risks and value of web tracking,” Ph.D. dissertation, Jan. 2015.
- [36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.
- [37] S. Puglisi, D. Rebollo-Monedero, and J. Forné, “You never surf alone. ubiquitous tracking of users’s browsing habits,” in *Proc. Int. Workshop Data Priv. Manage. (DPM)*, ser. Lecture Notes Comput. Sci. (LNCS), Vienna, Austria, Sep. 2015.
- [38] S. Gauch, M. Speretta, A. Chandramouli, and A. Micarelli. Springer-Verlag, 2007, ch. User profiles for personalized information access, pp. 54–89.
- [39] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [40] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*, 1st ed. Springer-Verlag, 2008.
- [41] A. M. Zoubir, V. Koivunen, and Y. C. M. Muma, “Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts,” *IEEE Signal Process. Mag.*, vol. 29, no. 4, pp. 61–80, Jul. 2012.
- [42] T.-D. Nguyen, “Robust estimation, regression and ranking with applications in portfolio optimization,” Ph.D. dissertation, MIT, Jun. 2009.
- [43] K. Yang, Y. Wu, J. Huang, X. Wang, and S. Verdu, “Distributed robust optimization for communication networks,” in *Proc. Joint Conf. IEEE Comput., Commun. Soc. (INFOCOM)*, 2008.
- [44] P. Kouvelis and G. Yu, *Robust Discrete Optimization and Its Applications*, 1st ed. Springer-Verlag, Nov. 1996.
- [45] “AOL search data scandal,” Aug. 2006, accessed on 2013-11-15. [Online]. Available: http://en.wikipedia.org/wiki/AOL_search_data_scandal
- [46] E. T. Jaynes, “On the rationale of maximum-entropy methods,” *Proc. IEEE*, vol. 70, no. 9, pp. 939–952, Sep. 1982.
- [47] —, “Information theory and statistical mechanics II,” *Phys. Review Ser. II*, vol. 108, no. 2, pp. 171–190, 1957.
- [48] L. Olejnik, T. Minh-Dung, and C. Castelluccia, “Selling off privacy at auction,” in *Proc. Symp. Netw. Distrib. Syst. Secur. (SNDSS)*. Internet. Soc., Feb. 2014.
- [49] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz, “The web never forgets: Persistent tracking mechanisms in the wild,” in *Proc. ACM Conf. Comput., Commun. Secur. (CCS)*, Washington, DC, Nov. 2014, pp. 674–689.
- [50] M. J. Schervish, *Theory of Statistics*. New York: Springer-Verlag, 1995.
- [51] “Iab quality assurance guidelines (qag) taxonomy,” Interact. Advertising Bureau, Accessed on 2015-09-11. [Online]. Available: <http://www.iab.com/guidelines/iab-quality-assurance-guidelines-qag-taxonomy/>
- [52] A. Soltani, S. Canty, Q. Mayo, L. Thomas, , and C. J. Hoofnagle, “Flash cookies and privacy,” in *Proc. AAAI Spring Symp. Intell. Inform. Priv. Manage.* Assoc. Adv. Artif. Intell., 2010.
- [53] “Evercookie - virtually irrevocable persistent cookies,” Oct. 2010. [Online]. Available: <http://samy.pl/evercookie>
- [54] K. Mowery and H. Shacham, “Pixel perfect: Fingerprinting canvas in HTML5,” in *Proc. IEEE Web 2.0 Workshop Secur., Priv. (W2SP)*. IEEE Comput. Soc., May 2012.
- [55] “Topline u.s. web data for march 2010,” Tech. Rep., Mar. 2010, Accessed on 2015-02-19. [Online]. Available: <http://www.nielsen.com/us/en/insights/news/2010/nielsen-provides-topline-u-s-web-data-for-march-2010.html>
- [56] A. Kae, K. Kan, V. K. Narayanan, and D. Yankov, “Categorization of display ads using image and landing page features,” in *Proc. ICDM Workshop Large-Scale Data Min.: Theory, Appl.* ACM, 2011, pp. 1–8. [Online]. Available: <http://doi.acm.org/10.1145/2002945.2002946>
- [57] B. J. Jansen, “Click fraud,” *IEEE Comput.*, vol. 40, no. 7, pp. 85–86, Jul. 2007.
- [58] “Firefox interest dashboard,” Nov. 2014, Accessed on 2015-05-02. [Online]. Available: <https://www.mozilla.org/en-US/firefox/interest-dashboard/>
- [59] C. D. Manning and H. Schütze, *Foundations of Statistical Natural Language Processing*. Cambridge, MA, USA: MIT Press, 1999.
- [60] G. A. Miller, “WordNet: A lexical database for english,” *Commun. ACM*, vol. 38, no. 11, pp. 39–41, 1995.
- [61] B. Magnini and G. Cavaglià, “Integrating subject field codes into wordnet,” in *Proc. Lang. Resource, Evaluation (LREC)*, Jun. 2000, pp. 1413–1418.
- [62] L. Bentivogli, P. Forner, B. Magnini, and E. Pianta, “Revising wordnet domains hierarchy: Semantics, coverage, and balancing,” in *Proc. PostCOLING Workshop Multiling. Ling. Resources*, Hangzhou, China, Aug. 2004, pp. 101–108.
- [63] A. Gonzalez-Agirre, E. Laparra, and G. Rigau, “Multilingual central repository version 3.0: upgrading a very large lexical knowledge base,” in *Proc. Global WordNet Conf.*, 2012.
- [64] J. Daudé, , L. Padró, and G. Rigau, “Validation and tuning of wordnet mapping techniques,” *Proc. Int. Conf. Recent Adv. Nat. Lang. Process. (RANLP)*, Sep. 2003.
- [65] G. Salton, A. Wong, and C. S. Yang, “A vector space model for automatic indexing,” *Commun. ACM*, vol. 18, no. 11, pp. 613–620, 1975.
- [66] “COIN-OR linear programming solver,” accessed on 2015-09-15. [Online]. Available: <https://projects.coin-or.org/Clp>
- [67] R. Lougee-Heimer, “The Common Optimization INterface for Operations Research: Promoting open-source software in the operations research community,” *IMB J. Res. Develop.*, vol. 47, no. 1, pp. 57–66, Jan. 2003.
- [68] “GNU linear programming kit, (GLPK),” accessed on 2015-09-17. [Online]. Available: <http://www.gnu.org/software/glpk>
- [69] E. M. Gertz and S. J. Wright, “Object-oriented software for quadratic programming,” *ACM Trans. Math. Softw.*, vol. 29, pp. 58–81, 2003.

- [70] M. Berkelaar, K. Eikland, and P. Notebaert, "Open source (mixed integer) linear programming system," May 2004. [Online]. Available: <http://lpsolve.sourceforge.net>
- [71] B. Borchers, "CSDP, a C library for semidefinite programming," *Optim. Method, Softw.*, vol. 11, no. 1, pp. 613–623, 1999.
- [72] S. J. Benson, Y. Ye, and X. Zhang, "Solving large-scale sparse semidefinite programs for combinatorial optimization," (*SIAM*) *J. Optim.*, vol. 10, no. 2, pp. 443–461, 2000.
- [73] "COIN-OR Interior Point OPTimizer," accessed on 2015-09-17. [Online]. Available: <https://projects.coin-or.org/Ipopt>
- [74] A. Wächter and L. T. Biegler, "On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming," *Math. Program.*, vol. 106, no. 1, pp. 25–57, 2006.
- [75] C. Zhu, R. H. Byrd, and J. Nocedal, "L-bfgs-b: Algorithm 778: L-bfgs-b fortran routines for large scale bound constrained optimization," *ACM Trans. Math. Softw.*, vol. 23, no. 4, pp. 550–560, 2007.
- [76] S. G. Johnson, "NLOPT nonlinear-optimization package," accessed on 2015-09-16. [Online]. Available: <http://ab-initio.mit.edu/nlopt>
- [77] J. Currie and D. I. Wilson, "OPTI: Lowering the barrier between open source optimizers and the industrial Matlab user," in *Proc. Found. Comput.-Aided Process Oper.*, Jan. 2012.
- [78] "Eclipse public license - version 1.0," Eclipse Found., Accessed on 2015-09-17. [Online]. Available: <https://www.eclipse.org/legal/epl-v10.html>
- [79] "The official easylist website," Accessed on 2015-10-22. [Online]. Available: <https://easylist.adblockplus.org>
- [80] M. Lecuyer, G. Ducoffe, F. Lan, A. Papancea, T. Petsios, R. Spahn, A. Chaintreau, and R. Geambasu, "XRay: Enhancing the web's transparency with differential correlation," in *Proc. Conf. USENIX Secur. Symp.*, Aug. 2014.
- [81] A. Datta, M. C. Tschantz, and A. Datta, "Automated experiments on ad privacy settings," in *Proc. Int. Symp. Priv. Enhanc. Technol. (PETS)*, Darmstadt, Germany, Jul. 2015.
- [82] M. Lecuyer, R. Spahn, Y. Spiliopoulos, A. Chaintreau, R. Geambasu, and D. Hsu, "Sunlight: finegrained targeting detection at scale with statistical confidence," in *Proc. ACM Conf. Comput., Commun. Secur. (CCS)*, Aug. 2015.
- [83] J. M. Carrascosa, J. Mikians, R. Cuevas, V. Erramilli, and N. Laoutaris, "I always feel like somebody's watching me. measuring online behavioural advertising," in *Proc. ACM Int. Emerg. Netw. Experiments, Technol. (CoNEXT)*, Heidelberg, Germany, Dec. 2015.
- [84] K. G. Allen, "Search marketing tops online retail customer acquisition tactics," NFR, Tech. Rep., Jul. 2014, Accessed on 2016-01-11. [Online]. Available: <https://nrf.com/media/press-releases/shoporgforrester-search-marketing-tops-online-retail-customer-acquisition>
- [85] J. Q. Freed, "Hoteliers rake in returns through retargeting," Hotel News Now, Tech. Rep., Mar. 2012, Accessed on 2016-01-14. [Online]. Available: <http://www.hotelnewsnow.com/Article/7710/Hoteliers-rake-in-returns-through-retargeting>
- [86] J. Butler, "Case study: How display ad remarketing works in travel," Tnooz, Tech. Rep., Sep. 2010, Accessed on 2016-01-14. [Online]. Available: <http://www.tnooz.com/article/case-study-how-display-ad-remarketing-works-in-travel/>
- [87] M. Gundlach, "AdBlock," Accessed on 2015-12-28. [Online]. Available: <https://getadblock.com/>
- [88] "Ghostery," Accessed on 2016-04-08. [Online]. Available: <https://www.ghostery.com>
- [89] "Disconnect," Accessed on 2015-11-19. [Online]. Available: <https://disconnect.me/>
- [90] "Lightbeam," Accessed on 2015-09-24. [Online]. Available: <https://www.mozilla.org/en-US/lightbeam/>
- [91] "Privacy badger," accessed on 2015-09-24. [Online]. Available: <https://www.eff.org/es/node/73969>
- [92] P. Sayer, "Adblock extension begins whitelisting "acceptable ads"," Oct. 2015. [Online]. Available: <http://www.pcworld.com/article/2988838>
- [93] S. Guha, B. Cheng, and P. Francis, "Challenges in measuring online advertising systems," in *Proc. ACM Internet Meas. Conf. (IMC)*, Melbourne, Australia, Nov. 2010.
- [94] J. M. Carrascosa, J. Mikians, R. Cuevas, V. Erramilli, and N. Laoutaris, "Understanding interest-based behavioural targeted advertising," in *arXiv: 1411.5281v1*, Nov. 2014.
- [95] P. Barford, I. Canadi, D. Krushevskaja, Q. Ma, and S. Muthukrishnan, "Adscope: Harvesting and analyzing online display ads," in *Proc. ACM Int. WWW Conf.* ACM, 2014, pp. 597–608.