Medición de la Privacidad de Perfiles de Usuario mediante un Add-on de Navegador

José Estrada-Jiménez, Ana Rodríguez, Javier Parra-Arnau, Jordi Forné, David Rebollo-Monedero
Departamento de Ingeniería Telemática
Universidad Politécnica de Catalunya (UPC)
C. Jordi Girona 1-3, 08034 Barcelona, España
{jose.estrada, ana.rodriguez, javier.parra, jforne, david.rebollo}@entel.upc.edu

***Resumen*- Actualmente, la monitorización de los usuarios en Internet es permanente, y la información obtenida en este proceso es de enorme interés para grandes compañías de publicidad e incluso gobiernos. Además, la gran cantidad de datos susceptibles de recopilarse por los sistemas de información personalizados representa un grave riesgo para la privacidad del usuario en Internet. Quizá aún más crítico es que muchos usuarios no son conscientes de este riesgo, ya que éste no es tan manifiesto como en el mundo físico.**

En este artículo presentamos un *add-on* de navegador que estima el riesgo de privacidad del perfil de un usuario, quien por sus hábitos de navegación, está expuesto a mecanismos de *profiling* en Internet. El nivel de riesgo se muestra, de manera comprensible y accesible en la interfaz gráfica del navegador y se calcula tomando en cuenta diferentes modelos de adversario.

***Palabras Clave*- perfil de usuario, métricas de privacidad, entropía de Shannon, divergencia de Kullback-Leibler, extensión de navegador, *add-on de navegador*.**

I. Introducción

Actualmente en Internet, y gracias a los evidentes avances en las técnicas de análisis de datos, el *profiling* y la clasificación de usuarios son una práctica común, llevada a cabo por sistemas de personalización de contenido que se alimentan de toda la información que entrega el usuario aunque éste, en la mayoría de casos, no sea consciente de su magnitud.

La creación de perfiles de usuario, a partir sus patrones de navegación, permite la recomendación personalizada de contenido y, especialmente, de publicidad, pero a un precio bastante alto: la privacidad del usuario. Los rastros que deja un usuario, aunque sean dispersos o incluso perturbados, combinados con otros de distintas fuentes, podrían revelar información potencialmente sensible relacionada con preferencias personales [1], [2].

La información sujeta a análisis va desde el contenido de las páginas visitadas, el tiempo consumido en un sitio web, el número de clics, las consultas a un motor de búsqueda, los datos entregados en formularios, y las cookies, hasta la configuración particular del navegador [17].

En ese entorno existe, por lo tanto, una amplia gama de posibles atacantes: motores de búsqueda, sistemas de recomendación, redes sociales, sistemas de etiquetado, etc. Sin embargo, los proveedores de servicios de Internet son entidades que tienen acceso a toda esa información sobre la actividad del usuario y, en muchos casos, ésta es también comercializada con compañías de publicidad o directamente utilizada para alimentar una plataforma de anuncios, sin considerar la privacidad de los dueños de esos datos [3].

Existe mucha presión sobre estas empresas que manejan información personal ([4] y [5]) para que apliquen fuertes políticas de privacidad, con el fin de proteger los datos sensibles. Parece, sin embargo, que la presión externa (desde gobiernos por ejemplo) por revelar este rastro digital puede resultar mayor.

Además, las políticas de privacidad que aplican los proveedores de servicios se comunican de una manera tan deficiente que los usuarios apenas las leen y difícilmente las comprenden, por lo que las aceptan rápidamente sin reflexionar, con el único fin de hacer uso inmediato de algún servicio "gratuito".

Esto demuestra que existe una generalizada falta de consciencia respecto a los riesgos a los que están expuestos los usuarios en Internet y de la consecuente vulneración de su derecho a la privacidad.

Estos intereses económicos y políticos sobre la información en la Web y las prácticas inadecuadas respecto a la privacidad de los usuarios no parece que vayan a cambiar con el tiempo. Sin embargo, el comportamiento del usuario respecto a su información sí puede modificarse, si se logra evidenciar las debilidades de su conducta. El problema radica en que no existen herramientas que informen al usuario sobre su nivel de privacidad. Existen herramientas que implementan medidas de ofuscación o bloqueo, pero ninguna que permita al menos determinar su efectividad. Es que el nivel de privacidad (o riesgo de privacidad) puede ser relativo al entorno del usuario (la población) y, por lo tanto, las medidas de protección podrían depender incluso de sus intereses individuales.

A nuestro juicio, es imprescindible medir el nivel de privacidad para aplicar, en función de éste, un mecanismo de protección que se ajuste a las necesidades del usuario, especialmente cuando el enorme éxito de la publicidad dirigida incentiva a todos los proveedores de contenido a aplicar avanzadas técnicas de elaboración de perfiles para modelar el comportamiento de los usuarios.

A. Contribución

Considerando el riesgo para la privacidad que representan las actividades de *profiling*, y la inexistencia de mecanismos para poder evidenciarlo, proponemos la implementación de un *add-on* para el navegador Mozilla Firefox que permite medir la privacidad del usuario a partir de su perfil, que se obtiene de las consultas a motores de búsqueda, del contenido de las páginas web desplegadas, del número de clics sobre las páginas y del tiempo que el usuario permanece

en ellas. Este *add-on* captura todo este rastro dejado por el usuario y en base a éste último muestra, de una manera comprensible, varias mediciones de su privacidad, teniendo en cuenta distintos modelos de adversario.

El conocimiento y la interpretación de estas medidas de privacidad podrían ayudar a los usuarios a tomar una decisión informada respecto de su actividad en la Web y permitirían evaluar tecnologías de mejoramiento de la privacidad para determinar su utilidad real. Esto facilitaría, además, la comparación y la optimización de estas tecnologías.

Nuestro *add-on* reutiliza el módulo de *profiling* de otro *add-on* de Mozilla Firefox llamado Adnostic [5]. En concreto, dicho módulo nos permite obtener un perfil de usuario en base al cual determinamos varios niveles de riesgo de privacidad, mediante la utilización de métricas justificadas en conceptos de teoría de la información.

B. Organización

El artículo se ha organizado de la siguiente manera. La Sec. II explora algunas de las herramientas y mecanismos existentes orientados a la protección de privacidad. La Sec. III resume los modelos de atacante y las métricas utilizadas para determinar los niveles de riesgo de privacidad. La Sec. IV describe la arquitectura y los módulos de nuestro *add-on*. Finalmente en la Sec. V se mencionan las conclusiones.

II. Estado del arte

Actualmente existen algunas herramientas que tratan de proteger la privacidad del usuario en Internet, esencialmente mediante el bloqueo de funciones que facilitan la entrega de información personal. Estos mecanismos, generalmente basados en la heurística, no miden el riesgo de privacidad del usuario ni evalúan el nivel de protección que ofrecen.

Adnostic [6] es un *add-on* desarrollado para Mozilla Firefox que implementa una arquitectura para desplegar publicidad personalizada, sin comprometer la privacidad del usuario, ya que se decide en el navegador qué anuncios mostrar, en función de un perfil calculado localmente. Este perfil se obtiene a partir de un procesamiento de las consultas que realiza el usuario y del contenido de las páginas que visita. Luego, esta información es clasificada utilizando procesamiento natural de lenguaje dentro del navegador. Los anuncios, parte de un conjunto previamente descargado, se despliegan dependiendo de los intereses del usuario.

REPRIV [7] es otro sistema propuesto para trabajar en el navegador que ofrece una personalización mejorada de contenido y un mecanismo de control del usuario sobre la información que entrega a terceros. Usa la información de navegación del usuario para descubrir cuáles son sus intereses, y comunicarlos a terceros para que estos últimos puedan ajustar el contenido en base a esas preferencias. Propone interfaces para sitios web de terceros para los protocolos de comunicación de información personal que funcionan sobre HTTP. Promete una mejora importante en la provisión de contenido a medida, gracias al gran detalle de la información del navegador, pero el control de privacidad podría verse afectado por la falta de usabilidad de las políticas de protección que se implementen y que un usuario promedio tendría que gestionar. Además, no implementa ninguna métrica que indique al usuario el nivel de privacidad que posee.

En relación específica con la medición de privacidad, existen un par de estudios ([8] y [9]) sobre herramientas para redes sociales (Facebook en los dos casos) que determinan el riesgo de privacidad del usuario en función de la cantidad de información que de éste se puede inferir a partir de sus relaciones con otros usuarios. También implementan acciones de protección de privacidad bloqueando estos usuarios, analizando la configuración de privacidad de la cuenta o detectando y eliminando aplicaciones que poseen demasiados permisos.

TrackMeNot [10] es otra herramienta para protección de privacidad a nivel de navegador que propone ofuscar el flujo de consultas del usuario a motores de búsqueda mediante la generación de consultas falsas. Ha recibido muchas críticas respecto de su eficacia, aunque no se han propuesto muchos mecanismos para evaluar sus bondades. En [11] se muestra que estas consultas falsas podrían ser identificadas con relativa facilidad utilizando clasificadores basados en inteligencia artificial. Sin duda, la falta de una herramienta de medida de privacidad le impide al usuario valorar su condición de riesgo antes y después de aplicar una estrategia de protección como ésta.

Google Sharing [20] es otra herramienta que implementa un mecanismo de protección de privacidad al prevenir el rastreo del usuario realizado por Google mediante las consultas al motor de búsqueda. El mecanismo consiste en que el usuario envía sus peticiones a un proxy externo que gestiona un grupo de identidades asociadas a *cookies*. Estas *cookies* reemplazan las *cookies* de las peticiones, enmascarando la identidad del usuario, y luego son reenviadas con la petición original a Google. Aun cuando permite enviar peticiones cifradas desde el usuario, su privacidad puede comprometerse si hay colusión entre Google y el servidor proxy.

Ghostery y *Collusion* son *add-ons* que enfrentan el problema de privacidad del usuario mediante la identificación de *trackers* o entidades que rastrean los movimientos del usuario, generalmente a través de una *cookie* de terceros. *Ghostery*, en especial, es muy completa ya que detecta y muestra información sobre estos *trackers*, y además bloquea los elementos de ejecución dinámica no confiables que se cargan en el navegador, mediante los cuales es posible este rastreo. Estas herramientas dejan de lado, sin embargo, la información que el usuario entrega en Internet y que podría fácilmente revelar su identidad.

El modo de "navegación privada" es también una opción de protección de privacidad en los navegadores más conocidos. Ésta opción deshabilita el almacenamiento local de información (historial, imágenes, videos, *cookies*, etc.) durante la navegación web. Esto complica significativamente el acceso del usuario a muchos sitios en Internet, por lo que quienes usan este modo lo hacen durante intervalos de tiempo muy cortos. El nivel de protección se limita al ámbito local pues externamente existen otros mecanismos para identificar y clasificar al perfil del usuario.

El bloqueo o desactivación de ciertas características del navegador web es una medida común implementada por varias soluciones en forma de *plug-ins* de navegador (NoScript [18], AdBlock Plus [17], DoNotTrackMe [21]), y evitan que se libere información que pueda usarse para identificar al usuario.

Sin embargo, ninguna de estas herramientas o mecanismos evalúa el nivel de privacidad del usuario. Se considera como posibles adversarios únicamente a los anunciantes o a los servicios de redes sociales pero no a los proveedores de servicios de Internet (ISP, *Internet Service Provider*) que son las entidades que más información poseen sobre los usuarios, tomando en cuenta que es muy habitual que los usuarios naveguen utilizando conexiones sin cifrar (sin usar HTTPS). Con base a la última consideración, el ISP tiene acceso a mucha información, y el gran detalle de la misma representa un enorme incentivo para su comercialización, por lo que además existen muchos potenciales compradores.

En [22], [23], [24] y [25] se abordan en detalle algunos mecanismos que podrían emplearse para la protección de la privacidad del usuario en entornos donde éste hace consultas o etiqueta contenido; considerando también el costo de estas estrategias que se refleja en la pérdida de utilidad de los datos, la pérdida de funcionalidad de un servicio o el consumo adicional de recursos. Se incluye entre estos mecanismos la falsificación de consultas o la supresión de etiquetas con el fin de mostrar una versión distorsionada del perfil del usuario que el atacante no pueda explotar. La optimización de estos mecanismos así como su impacto son también sujetos de estudio.

III. Modelos de Atacante y Métricas de Privacidad

En esta sección presentamos los dos modelos de atacante considerados en este artículo, así como las dos métricas que nos permiten evaluar el nivel de privacidad de un usuario. Los modelos y métricas de usuario son ampliamente justificados en [12].

A. Modelos de Atacante

Los criterios de privacidad se plantean inicialmente asumiendo que el perfil del usuario es modelado como una función de masa de probabilidad o un histograma de frecuencias relativas de datos de usuario a lo largo de un conjunto preestablecido de categorías de interés. Este modelo supone una representación muy habitual en los servicios de información personalizada.

El modelo de adversario permite definir las propiedades del atacante, considerando como tal a cualquier entidad capaz de tener acceso a información de usuario con el objetivo de obtener su perfil, con el riesgo a la privacidad que esto implica.

Conocer al adversario es importante ya que la privacidad del usuario se mide respecto a éste. En función de las propiedades del adversario, el usuario podría implementar medidas de protección de su privacidad que, por ejemplo, modificasen su perfil de intereses.

Esencialmente, se contemplan dos objetivos del atacante, en función de sus capacidades y que definen el modelo de adversario; *identificación* y *clasificación*.

- *Identificación*, cuando el atacante intenta distinguir al usuario del resto de la población, detectando desviaciones de sus intereses respecto del perfil promedio de la población.

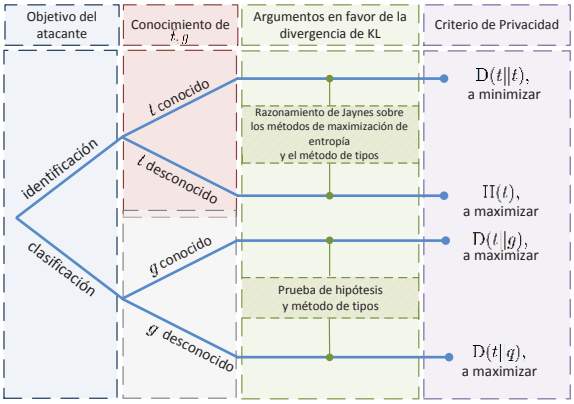

Fig. 1. Resumen de las interpretaciones de la entropía de Shannon y de la divergencia de KL (entropía relativa) como métricas de privacidad, de acuerdo con la justificación presentada en [12].

- *Clasificación*, cuando el atacante intenta clasificar al usuario en un grupo de población, comparando el perfil del usuario con el perfil representativo del grupo.

B. Métricas de Privacidad

En [12] se justifica la entropía de Shannon y la divergencia de Kullback-Leibler [28] (conocida también como divergencia de KL o entropía relativa) como medidas de privacidad. Las interpretaciones de estas medidas dependerán de las hipótesis que se hagan, fundamentalmente respecto del modelo de adversario.

Otra métrica más general, no limitada a la privacidad de perfiles, es la propuesta en [26]. En este trabajo, los autores proponen medir la privacidad como el error de estimación de un adversario, e interpretan, mediante argumentos de teoría de la información y teoría de decisión Bayesiana, otras métricas del estado del arte como casos particulares de la suya.

Para facilitar la comprensión, además, se resume a continuación las principales definiciones propuestas para la justificación de estas métricas de privacidad. Se revisa además la interpretación de estas dos cantidades de teoría de la información como métricas de privacidad de perfiles de usuario.

El símbolo H denotará la entropía de Shannon y D denotará la divergencia de KL. La entropía $H(p)$ de una variable aleatoria discreta X con distribución de probabilidad p es una medida de su incertidumbre, definida como

$$H(X) = -E \log p(X) = -\sum_x p(x) \log p(x).$$

La divergencia de KL o entropía relativa D ($p \parallel q$) entre dos distribuciones de probabilidad $p(x)$ y $q(x)$ sobre el mismo alfabeto se define como

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

La divergencia de KL es una medida de discrepancia entre distribuciones de probabilidad, garantizando que $D(p \parallel q) \geq 0$, con igualdad si, y sólo si, $p=q$. Consecuentemente se deduce que la entropía $H(p)$ alcanza su valor máximo en $H(u)=\log n$, siendo n la cardinalidad del alfabeto finito sobre el que se calcula $D(p \parallel u)$, para una distribución uniforme u:

$$D(p \parallel u) = \log n - H(p).$$

En concreto, acorde con el análisis en [12] tenemos que la maximización de la entropía resulta ser un caso especial de la

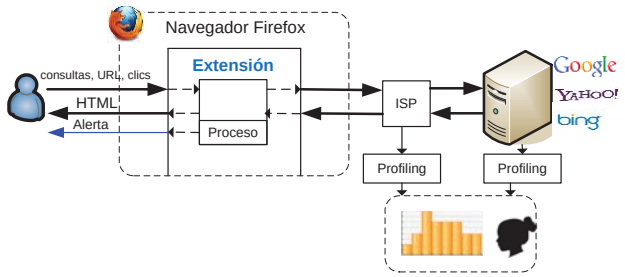

Fig. 2. Esquema de navegador y un *add-on* como intermediarios entre usuario y servicios de Internet (motores de búsqueda y proveedor de servicios) y el inherente riesgo de *profiling*.

minimización de la divergencia, alcanzada idealmente cuando la distribución a optimizar es idéntica a la de referencia.

Sea q el perfil de interés de un usuario, t una versión perturbada o modificada del mismo y $\bar{t}$ la distribución del perfil de la población. En la Fig. 1 se muestran las interpretaciones de la entropía de Shannon y de la divergencia de KL como medidas de privacidad. Éstas se explican a continuación, de acuerdo al objetivo del atacante.

- *Métricas contra identificación*

En caso de que el objetivo del atacante sea identificar al usuario, el razonamiento de Jaynes acerca de los métodos de maximización de la entropía permite justificar la divergencia y la entropía como medidas de privacidad.

La entropía del perfil aparente del usuario, que es el perfil observado por el atacante, es justificada en [12] como una medida de la probabilidad de este perfil perturbado, en el sentido de frecuencia de aparición de dicho perfil en la población de usuarios. Considerando esta probabilidad del perfil de usuario como una medida razonable de su anonimato (o privacidad), en [12] se justifica también la entropía como una métrica de privacidad. En concreto, mientras mayor sea la entropía de este perfil, mayor es su probabilidad, y por tanto mayor es el número de usuarios que se comportan de acuerdo con este perfil, haciéndolo más privado.

Además, como se puede observar en la primera rama de la Fig. 1, si la distribución del perfil de la población $\bar{t}$ es conocida, se utiliza la divergencia entre el perfil del usuario t y el perfil de la población como métrica de privacidad, de manera que, cuanto más pequeña sea esa divergencia, más privado se puede considerar el perfil.

En definitiva, la elección de perfiles aparentes que conduzcan a la minimización de la divergencia de KL mejora el anonimato. En términos más simples, una menor divergencia corresponde a una mayor frecuencia de ocurrencia de dicho perfil, permitiendo al usuario pasar más desapercibido. En el caso de un perfil de referencia de la población uniforme, esto equivale a la maximización de la entropía de Shannon.

- *Métricas contra clasificación*

Si el objetivo del atacante es clasificar al usuario como miembro de un grupo en particular, se utiliza la divergencia como métrica de privacidad, de acuerdo al análisis realizado en [12], a partir del *test* de hipótesis y el método de tipos. Como se indica en la Fig. 1, en la segunda rama, si el perfil del grupo g es desconocido en el lado del usuario, la opción es maximizar la divergencia entre el perfil real q y el perfil observado (aparente) t, con el fin de evitar ser clasificado de acuerdo a su perfil original.

Nótese que en el problema de clasificación, al contrario de lo que ocurre en el problema de identificación, buscamos maximizar la divergencia de KL, en lugar de minimizarla. La intuición subyacente al análisis citado es que se desea agrandar la distancia entre el perfil aparente del usuario y el perfil real, o el representativo del grupo en el que deseamos evitar la categorización.

IV. Medición de la privacidad en el navegador

En este artículo presentamos un *add-on* para el navegador Mozilla Firefox, que mide la privacidad del usuario en términos de riesgo o ganancia de privacidad.

Tal como se mencionó en la Sec. II, casi no existen herramientas que muestren, en tiempo real, el estado de privacidad del usuario. Desde un principio, por tanto, el usuario no es realmente consciente del peligro que representa para su privacidad todo el rastro digital que va dejando en los distintos servicios que utiliza en Internet. Esto constituye ya un grave problema, pues una percepción clara del riesgo al que se enfrenta el usuario derivaría en sospecha y ésta, muy probablemente, conduciría a un comportamiento más activo (i.e. defensivo) respecto al manejo de la información [13].

En este trabajo, proponemos una herramienta que presenta información comprensible al usuario, referente a sus niveles de privacidad, que le permiten asimilar el riesgo y probablemente, desde su perspectiva e interés, tomar una decisión para protegerse.

La información relativa a los niveles de privacidad, como se mencionó previamente, se determina en base a los conceptos de teoría de la información que se resumen en la Sec. III y consta, básicamente, de una alerta de riesgo de identificación y el resultado de la clasificación del perfil del usuario entre varios grupos definidos.

A. *Consideraciones de Diseño*

Partimos de la premisa de que el usuario no confía en ningún agente externo distinto de su dispositivo local de comunicación por lo que no está interesado en ceder información de su perfil.

Adicionalmente, consideramos dos atacantes posibles cuyas actividades de *profiling* podrían representar un grave riesgo a la privacidad, dada la gran cantidad de información a la que tienen acceso. Estos atacantes son, como se muestra en la Fig. 2, los motores de búsqueda y el proveedor de servicios de Internet. Los motores de búsqueda pueden recopilar todas las consultas que realizan sus usuarios y en las que se revela información detallada de sus intereses. Los ISPs tienen acceso a la mayor parte del contenido que el usuario genera desde su navegador; es decir, tiene acceso a contenido de sitios web, clics sobre enlaces, tiempo de permanencia en páginas web (dependiendo del sitio web) y también las consultas realizadas a motores de búsqueda. Esta afirmación se cumple siempre y cuando la conexión del usuario no esté cifrada (sin utilizar HTTPS); de manera que si la conexión está cifrada, el ISP dispondrá solamente de información sobre qué sitios visita el usuario y no de los contenidos de las páginas. Muy pocos sitios implementan mecanismos de cifrado.

El patrón de navegación formado por la información descrita podría permitir a los atacantes obtener un perfil muy detallado de los usuarios. Obtener este perfil es también crucial para el usuario para que pueda comprender el riesgo al

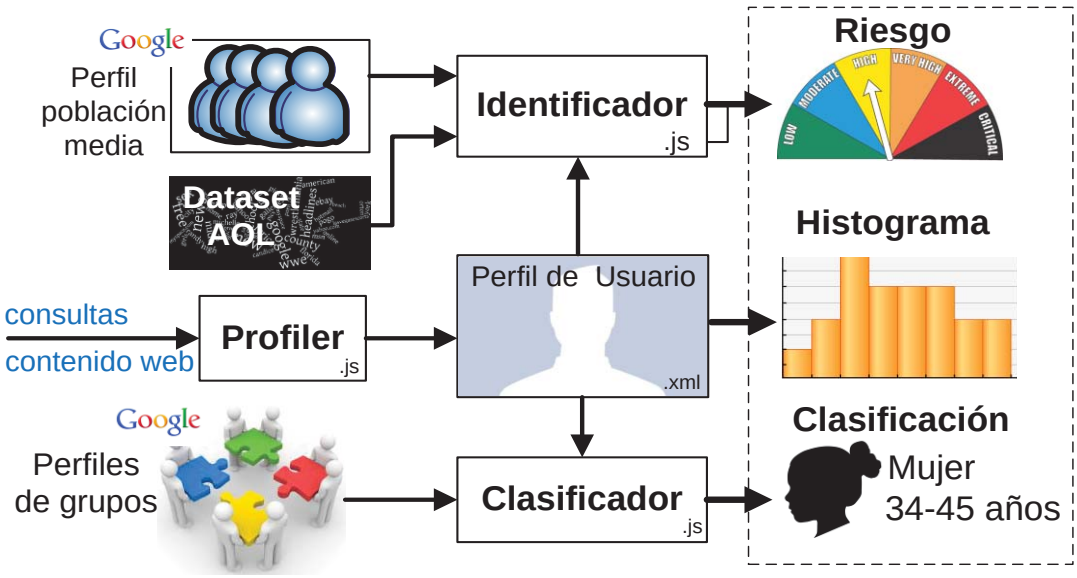

Fig. 3. Arquitectura para el cálculo del nivel de privacidad.

que se enfrenta cuando navega en Internet. Este perfil es también importante para generar las alertas que le ayuden a tomar una decisión respecto de su privacidad.

El navegador web se adapta muy bien a estas premisas, pues, dado que actúa como un intermediario entre el usuario e Internet (Fig. 2), es el encargado de gestionar todas las peticiones que el usuario realiza y todas las respuestas que recibe de Internet, y que se despliegan generalmente como páginas web. La información de la actividad del usuario, obtenida a partir del navegador, es muy detallada y por lo tanto muy útil para obtener su perfil, tal como lo modelarían los atacantes mencionados.

Con el fin de considerar el entorno del usuario en el proceso de evaluación de su privacidad, especialmente cuando se intenta clasificarlo, se requiere la información de los perfiles de varios grupos de población.

Finalmente, cumpliendo con el antecedente de un usuario que no confía en terceros, la información del perfil de usuario, como el resultado de su procesamiento se mantendrá siempre en el ámbito local del navegador.

B. Arquitectura

En esta sección se explica el funcionamiento de los componentes principales de la arquitectura para la medición de la ganancia y riesgo de privacidad del usuario. La estructura se ilustra en la Fig. 3, en donde se refleja el flujo de procesos y los resultados obtenidos por cada módulo funcional.

Navegador web. En este caso el navegador Mozilla Firefox, es el encargado de recibir las órdenes de navegación del usuario, generalmente en forma de palabras, recogidas mediante formularios que se traducen en peticiones HTTP hacia servidores web o motores de búsqueda. Están disponibles varias interfaces necesarias para acceder a gran parte de la información que envía y recibe el navegador en nombre del usuario, mediante los *add-ons*. El uso de Firefox en este trabajo se justifica en el aprovechamiento que se da de la extensión *Adnostic*, también desarrollada para este navegador, de la que se reutiliza el módulo de *profiling* para construir el perfil de usuario cuya privacidad se mide en nuestra herramienta. Sin embargo, este proceso de medición de privacidad puede implementarse en otros navegadores como Chrome o Internet Explorer, por ejemplo. Dado que estos ofrecen distintas interfaces de desarrollo, nuestra herramienta debería ser adaptada para que use los componentes de la interfaz de los distintos navegadores a los que se desea portarla.

Profiler. El proceso de *profiling* o establecimiento del perfil del usuario consiste en obtener una tabla de frecuencias de un conjunto de categorías pre-establecido. La "puntuación" de cada categoría irá incrementándose conforme se vayan revelando las preferencias del usuario a través de sus consultas y la información que recibe de los servidores web.

Histograma. Es nuestro modelo de perfil de usuario. Está formado por barras cuyo tamaño representa la popularidad de cada categoría en este perfil. El esquema de categorización, que heredamos de Adnostic, y que se basa en la representación que hacía[1] *Google Ad Preferences*, usa 3 niveles de categorías con 602 categorías en total. El primer nivel de la jerarquía se compone de 27 categorías. En el histograma se muestran las 8 categorías más representativas de este primer nivel de jerarquía. Esta representación gráfica le da al usuario una impresión básica de su perfil.

Identificador. Tal como se muestra en la Fig. 4, este módulo determina el nivel de privacidad del usuario ante un ataque de identificación. Este nivel se muestra de 3 maneras distintas, tal como se explicó en la Sec. III.

La primera forma en que mostramos al usuario su nivel de privacidad es mediante la entropía de su perfil, que sirve como una métrica de anonimidad o ganancia de privacidad.

La segunda forma está relacionada con la misma entropía del perfil de usuario, pero ahora usando como referencia los

[1] Actualmente *Google Ad Preferences* utiliza un número mayor de categorías.

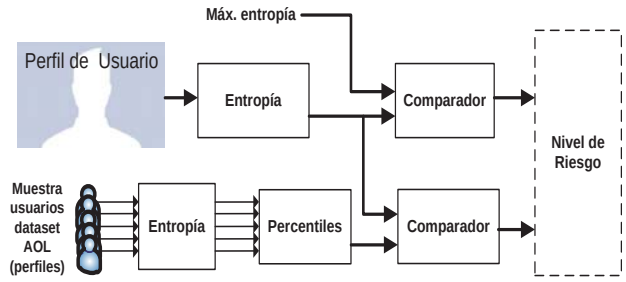

Fig. 4. Arquitectura del proceso de Identificación.

valores de entropía de una población real. Los perfiles de esta población de referencia se obtienen a partir de un subconjunto de un *dataset* de consultas de AOL liberado hace varios años.

Además, y dado que se dispone de una distribución aproximada del perfil de la población media (obtenido de la herramienta Ad Planner de Google [14]), tenemos una tercera forma de mostrar el nivel privacidad dada por la divergencia de KL del perfil de usuario relativo al perfil de la población media. Un valor de divergencia de KL de 0 indicaría una distribución del perfil de usuario equivalente a la de la población media, lo que representaría el nivel más bajo de riesgo de privacidad. Sin embargo, este valor no se puede normalizar respecto de un máximo, para determinar otros niveles de riesgo, pues este máximo no está acotado superiormente. Posteriormente este valor podría ser utilizado para medir la ganancia de privacidad, luego de aplicar alguna medida de protección.

Clasificador. Este módulo emplea la divergencia de KL entre la distribución del perfil de usuario y la del perfil de varios grupos predefinidos (ver Fig. 5). Se intenta recrear el ataque que haría un adversario con la intención de clasificar al usuario como parte de un grupo. Con ese objetivo, se calcula la divergencia de KL entre la distribución del perfil de usuario y la promedio de cada grupo de población en los que Google clasifica a sus usuarios de acuerdo a sus preferencias (datos obtenidos de la herramienta Ad Planner de Google).

El menor valor de divergencia es el que identifica al grupo con el cual el perfil del usuario tiene la menor discrepancia y, por lo tanto, el grupo al que el usuario tiene mayor probabilidad de pertenecer.

Desde el punto de vista del usuario, ésta es información muy ilustrativa ya que le da una idea bastante clara de lo previsible que es su perfil en Internet y, especialmente, lo mucho que se puede inferir a partir de su rastro digital.

Este método de clasificación es consistente con la métrica y el ataque correspondientes en la Sección III.B, resumidos en la Fig. 1, en el que el perfil representativo del grupo se escoge como el promedio de los perfiles pertenecientes.

Como nota marginal, cualquier método de clasificación supervisada, pongamos por ejemplo máquinas de vectores de soporte, podría ser utilizado por el atacante o la arquitectura para clasificar un perfil en uno de los subgrupos de población predeterminados. El método elegido en esta arquitectura es conceptual y computacionalmente simple, además de consistente con la métrica de privacidad propuesta en [12]. Una observación adicional es que la divergencia de KL es un caso particular de divergencia de Bregman, y por ello el promedio de un grupo es su centroide. Así, el método empleado corresponde al establecimiento de celdas de Voronoi en cuantificación con divergencias de Bregman [27].

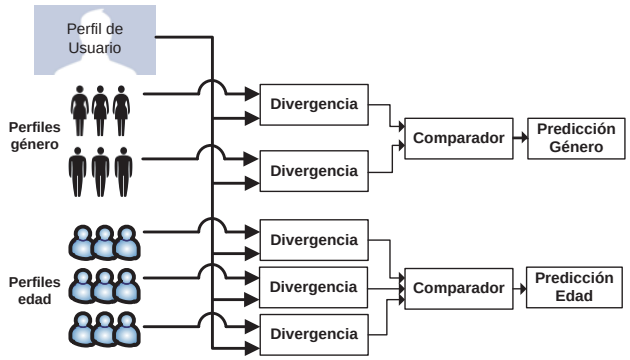

Fig. 5. Arquitectura del proceso de Clasificación.

C. Aspectos relevantes de la implementación

Este *add-on* para el navegador Firefox se ha programado utilizando los lenguajes Javascript y XUL.

Profiling. El perfil de usuario se obtiene mediante el módulo *profiler*, tomado del *add-on* Adnostic. Aquí se detectan los eventos de despliegue de las páginas Web en Firefox cuando el usuario navega y, al hacerlo, se recupera la información del usuario que da forma a su perfil. Esta información se reduce a: las consultas realizadas en motores de búsqueda, las *keywords* del código html perteneciente a las páginas que el usuario visita, los clics sobre estas páginas y el tiempo que permanece en ellas. Estos elementos permiten una clasificación bajo el esquema de 602 categorías utilizado por *Google Ad Preferences* y conforme a los resultados de la clasificación se actualizan los "puntajes" de las categorías en el perfil. Esta información tabulada de categorías es utilizada luego por nuestros módulos para evaluar el nivel de privacidad del usuario.

Modificamos este módulo para que el puntaje absoluto de cada categoría en el perfil del usuario no estuviese limitado a un valor de 500. Del mismo modo, realizamos cambios para que permitiese el *profiling* cuando el usuario se conecta a sitios con https.

Recolección de datos de población. Para obtener las métricas de privacidad relacionadas con la divergencia de KL, especialmente cuando el objetivo es clasificar al usuario en un grupo de varios predefinidos, es necesario disponer de la información del perfil de cada uno de estos grupos. Para poder comparar los distintos perfiles, desde luego, deben basarse en el mismo alfabeto de categorías.

Ventajosamente, Google posee una herramienta de publicidad en línea llamada Ad Planner [14] que dispone de mucha información sobre la distribución de los intereses de usuarios, clasificados utilizando la misma jerarquía que se utiliza para definir las preferencias de los usuarios de Google y Gmail. Esta información está tabulada de diversas formas: geográficamente, por grupos de edad y por grupos de género.

En números absolutos, en esta herramienta de análisis de intereses, se indica la cantidad de personas que estarían interesadas en cada una de las categorías tomadas como referencia para la clasificación.

Los datos recopilados para el proceso de clasificación del perfil del usuario pertenecen a los siguientes grupos;

edad (en años):

- 18 a 24,
- 25 a 34,

Fig. 7. Ventana de Información de métricas de privacidad (riesgo de privacidad) para el usuario.

- 35 a 44,
- 45 a 54,
- 55 a 64,
- 65 y más,

y género.

Estos datos, agrupados en correspondientes archivos, se incluyen en el *add-on* para recuperarse el momento de calcular la divergencia del perfil del usuario.

Training de datos de AOL. Se utilizó un conjunto de datos de consultas de usuarios liberado por AOL en 2006 [15] para encontrar la distribución de los valores de entropía en una población real.
Este *conjunto* está formado por cerca de 20 millones de consultas Web realizadas por alrededor de 650.000 usuarios, en un período de 3 meses.

La cantidad de consultas por usuario oscila entre 1 y varios cientos de miles. Con el fin de obtener una muestra de usuarios con una cantidad de consultas suficiente para obtener perfiles representativos, seleccionamos los usuarios con 501 a 1000 consultas, dando un total de 6674 usuarios.

Modificamos el módulo de *profiling* de Adnostic para que recibiera las consultas realizadas por los 6674 usuarios del grupo seleccionado y que por cada uno de ellos creara un perfil distinto. Luego, calculamos los valores de entropías de cada perfil, obteniendo una distribución que se puede observar en la Fig. 8.

Finalmente, en esta distribución se calcularon los percentiles de manera que podamos ubicar la entropía del perfil de usuario en alguno de los intervalos, para determinar su nivel de riesgo de privacidad. El *add-on* únicamente incluye estos valores de percentiles.

Interfaz gráfica. Para la creación de ventanas que muestren las distintas métricas que se han descrito utilizamos XUL, un lenguaje implementado como dialecto de XML orientado al desarrollo de interfaces de usuario.

En la ventana principal del navegador Firefox, el primer indicador de riesgo de privacidad que se muestra se ubica en la barra de complementos en la parte inferior. Tal como se observa en la Fig. 6, se detallan 3 elementos que le dan al usuario una idea rápida de su condición de privacidad: una escala de nivel que ilustra el riesgo de privacidad del usuario, el valor de la entropía del perfil de usuario y la última categoría en la que el usuario se ha clasificado. El nivel de riesgo se dibuja en función del percentil al que pertenece la entropía del usuario.

Fig. 6. Barra de información de privacidad del usuario ubicada en la parte inferior del navegador.

Adicionalmente, en el menú de herramientas del navegador se crea una opción para abrir una ventana en la que se muestran las distintas métricas que se han calculado en base al perfil de usuario y el de los distintos grupos poblacionales. En la Fig. 7 se puede observar esta información (usando los dos modelos de atacante mencionados): el histograma de categorías de interés, la entropía del perfil de usuario, un indicador de riesgo de privacidad y los grupos de edad y género en los que se ha clasificado al usuario.
Mediante dos botones, en esta ventana, se pueden abrir dos cuadros de diálogo adicionales. Uno para desplegar un histograma extendido con la información de hasta 19 categorías, y otro para mostrar una ventana con información un poco más detallada sobre las métricas de privacidad.

Esta última ventana incluye, además, los valores de las divergencias del perfil de usuario respecto de cada uno de los grupos de población. Dicha discrepancia se muestra también en una regla de nivel para fácil interpretación.

Importación de historial. Como parte de la implementación, se incorpora también un mecanismo para que el usuario, una vez que ha instalado el *add-on*, pueda alimentar su perfil, utilizando la información disponible en el historial de su navegador. Para eso se agrega otro elemento en el menú Herramientas que da acceso al proceso de categorización y *profiling* de los registros del historial de Firefox. En este caso, se utilizan como entradas al módulo de *profiling* los títulos de las páginas y las palabras ingresadas mediante formularios.

V. Conclusiones

Considerando el gran riesgo al que se enfrentan los usuarios cuando navegan en Internet debido a los agresivos mecanismos de *profiling* empleados por los principales sistemas de acceso a la Web (por ejemplo motores de búsqueda o ISPs) y la escasez de herramientas que evalúen este compromiso de seguridad, hemos propuesto un *addon* para Mozilla Firefox capaz de medir la ganancia y el riesgo de privacidad, poniendo estas magnitudes en un contexto determinado por los intereses del usuario y su entorno.

Una gran cantidad de usuarios se sienten anónimos en Internet y otros muchos apenas advierten la gravedad de las

amenazas contra su intimidad en el mundo digital. Por ello necesitan información accesible al respecto y fácil de interpretar. Con este *add-on* proveemos métricas interpretables como niveles de privacidad y un modelo de clasificación basado en conceptos justificados en la teoría de la información y el test de hipótesis, que el usuario puede emplear para tomar una decisión adecuada.

Estamos seguros de que la percepción que tiene el usuario respecto al riesgo puede ser un aporte muy valioso. Quién mejor que el usuario para escoger los intereses o los datos que debe proteger, dependiendo de su situación particular.

Del mismo modo, esta herramienta de medición e interpretación de la privacidad podría ser empleada para determinar la efectividad de otras herramientas que implementan mecanismos de protección (por ejemplo TrackMeNot [16]), así como también para comparar la utilidad entre ellas. El entorno de desarrollo para Firefox ofrece muchas facilidades para la implementación de esta herramienta para medir la privacidad (en especial la disponibilidad de la extensión *Adnostic* de la que se reutiliza el mecanismo de *profiling*), pero sería también conveniente portarla a Chrome o Internet Explorer, considerando la gran cantidad de usuarios que tienen estos navegadores actualmente. Aunque se debería tomar en cuenta que, con respecto a la privacidad, ahora se tiene muchas reservas respecto de Google y Microsoft, creadores de Chrome e Internet Explorer, respectivamente.

Aunque no se ha validado la utilidad real de la extensión con usuarios, se intuye que la misma va a depender mucho de la situación particular de cada usuario (intereses, valores, preocupaciones, actitudes, etc.) y por tanto de la visión individual que tenga cada usuario con respecto a su privacidad.

Finalmente, la estructura de categorías podría modificarse para incluir algunas categorías no tomadas en cuenta por *Google Ad Preferences* relacionadas con raza, religión, orientación sexual, o salud para disponer de un perfil de usuario mucho más completo.

AGRADECIMIENTOS

Este trabajo fue apoyado en parte por el Gobierno español a través de los proyectos Consolider Ingenio 2010 CSD2007-00004 “ARES”, TEC2010-20572-C02-02 Consequence y por el Gobierno de Catalunya a través de la subvención 2009 SGR 1362. Del mismo modo, se recibió el apoyo del Gobierno ecuatoriano a través de la Secretaría Nacional de Ciencia y Tecnología – SENESCYT, mediante la beca de estudios otorgada a Ana Rodríguez y José Estrada. D. Rebollo-Monedero disfruta de una beca postdoctoral Juan de la Cierva, ref. JCI-2009-05259, otorgada por el Ministerio de Ciencia e Innovación español.

REFERENCIAS

[1] Arvind Narayanan y Vitaly Shmatikov, "Robust De-anonymization of Large Sparse Datasets". En Security and Privacy, 2008. SP 2008. IEEE Symposium on, C1, 2008.

[2] Michael Barbaro y Tom Zeller Jr., "A Face Is Exposed for AOL Searcher No. 4417749". En The New York Times, Technology, URL http://www.nytimes.com/2006/08/09/technology/09aol.html?pagewanted=all, Agosto 2006.

[3] Eric Pfanner, "Internet Providers in Deal for Tailored Ads". En The New York Times, Technology, URL http://www.nytimes.com/2008/02/18/technology/18target.html?_r=2&oref=slogin&, Feb. 2008.

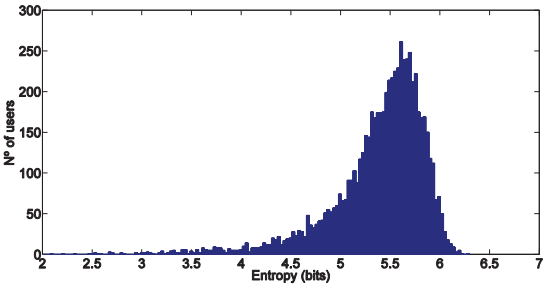

Fig. 8. Distribución de entropía de usuarios con 500 a 1000 consultas del *dataset* de consultas de AOL.

[4] Katy Hafner, “Google Resists U.S. Subpoena of Search Data”. En The New York Times, Technology, URL http://www.nytimes.com /2006/01/20/technology/20google.html?_r=1, Enero 2006.

[5] Russia Today, “Google se enfrenta al FBI para no revelar datos privados de los usuarios”, Abril 2013.

[6] V. Toubiana, D. Boneh, H. Nissenbaum, and S. Barocas, “Adnostic: Privacy Preserving Targeted Advertising ∗”. En Proceedings of the 17th Annual Network and Distributed System Security Symposium (NDSS), 2009.

[7] S. Vi-a, G. News, S. Vi-b, I. Browsing, and I. Explorer, “R E P RIV : Re-Envisioning In-Browser Privacy.”

[8] J. Becker y H. Chen, “Measuring Privacy Risk in Online Social Networks”. En Proceedings of W2SP 2009: Web 2.0 Security and Privacy, 2009.

[9] M. Fire, D. Kagan, A. Elishar, and Y. Elovici, “Social Privacy Protector - Protecting User' Privacy in Social Networks”.

[10] D. Howe and H. Nissenbaum. “TrackMeNot: resisting surveillance in web search”, 2006. mrl. nyu.edu/˜dhowe/trackmenot/.

[11] S. Teja Peddinti y N. Saxena, “On the Privacy of Web Search Based on Query Obfuscation: A Case Study of TrackMeNot”. En 10th International Symposium, PETS, 2010.

[12] J. Parra-Arnau, D. Rebollo-Monedero, J. Forné, “Measuring the Privacy of User Profiles in Personalized Information Systems”. En Future Generation Computer Systems, 2013.

[13] J. Drennan,G. Sullivan and J. Previte, “Privacy, Risk Perception, and Expert Online Behavior: An Exploratory Study of Household End Users”. En Journal of Organizational and End User Computing (JOEUC), 18(1), 1-22, 2006.

[14] Google Ad Planner, URL https://www.google.com/adplanner/ #audienceBuilder.

[15] G. Pass, A. Chowdhury, C. Torgeson, "A Picture of Search". En The First International Conference on Scalable Information Systems, Hong Kong, June, 2006.

[16] TrackMeNot, URL http://cs.nyu.edu/trackmenot/.

[17] Peter Eckersley, “How Unique Is Your Web Browser?”.

[18] Palant, Wladimir: Adblock Plus: Save your time and traffic, http://adblockplus.org/.

[19] Maone, Giorgio. NoScript. Online: http://noscript.net, 2009.

[20] Google Sharing, URL https://addons.mozilla.org/en-us/firefox/addon/googlesharing/

[21] DoNotTrackMe, URL https://addons.mozilla.org/en-US/firefox/addon/donottrackplus/?

[22] David Rebollo-Monedero, Jordi Forné, y Josep Domingo-Ferrer, “Query Profile Obfuscation by Means of Optimal Query Exchange between Users”. En IEEE Trans. Depend., Secure Comput., 2012.

[23] J. Parra-Arnau, D. Rebollo-Monedero and J. Forné, “A Privacy-Preserving Architecture for the Semantic Web based on Tag Suppression”. En Proc. Int. Conf. Trust, Priv., Secur., Digit. Bus., Bilbao, España, pp. 58-68, 2010.

[24] J. Parra-Arnau, D. Rebollo-Monedero, J. Forné, J. L. Muñoz y O. Esparza, “Optimal tag suppression for privacy protection in the semantic Web”. En Data, Knowl. Eng., vol. 81-82, pp. 46-66, 2012.

[25] J. Parra-Arnau, A. Perego, E. Ferrari, J. Forné and D. Rebollo-Monedero, “Privacy-Preserving Enhanced Collaborative Tagging”. En IEEE Trans. Knowl. Data Eng., 2012.

[26] D. Rebollo-Monedero, J. Parra-Arnau, Claudia Diaz and J. Forné, “On the Measurement of Privacy as an Attacker's Estimation Error”. En Springer, International Journal of Information Security, vol. 12, n. 2, pp. 129-149, 2013.

[27] D. Rebollo-Monedero, "Quantization and transforms for distributed source coding," Ph.D. dissertation, Stanford Univ., Dec. 2007.

[28] Wikipedia, “Divergencia de Kullback Leibler”, URL http://es.wikipedia.org/wiki/ Divergencia_ de_Kullback-Leibler