ÚNICO EN FACEBOOK: FORMULACIÓN Y EVIDENCIA DE LA PUBLICIDAD (NANO)DIRIGIDA A USUARIOS CON DATOS NO-PII

JOSÉ GONZÁLEZ-CABAÑAS, Universidad Carlos de III Madrid

ÁNGEL CUEVAS, Universidad Carlos III de Madrid,

UC3M-Santander Big Data Institute

RUBÉN CUEVAS, Universidad Carlos III de Madrid

UC3M-Santander Big Data Institute

JUAN LÓPEZ-FERNÁNDEZ, GTD System & Software Engineering

DAVID GARCÍA, Graz University of Technology

RESUMEN

La privacidad de un individuo está limitada por la capacidad de un tercero para revelar su identidad. Ciertos tipos de datos como el número de pasaporte o el número de teléfono móvil se pueden usar para identificar de manera unívoca a una persona. Además, trabajos previos en la literatura han reportado que, en un dataset con millones de usuarios, la combinación de algunos elementos, que de manera individual no son suficiente para identificar a un individuo, pueden identificar de manera unívoca a un usuario en ese dataset. En este trabajo, definimos un modelo basado en datos para cuantificar el número de intereses de un usuario que lo hacen único en Facebook. Hasta donde nosotros conocemos, este trabajo representa el primer estudio sobre la unicidad de un individuo a la escala de la población mundial. Además, los intereses de los usuarios se pueden activar para generar campañas de publicidad dirigida en Facebook. Hemos realizado un experimento mediante 21 campañas publicitarias en Facebook dirigidas a tres de los autores de este trabajo para probar que, si un anunciante conoce suficientes intereses de un usuario, la plataforma de publicidad de Facebook se puede usar de manera sistemática para enviar anuncios de manera exclusiva a un usuario específico. Definimos esta práctica como publicidad nanodirigida. Finalmente, discutimos riesgos dañinos asociados con la publicidad nanodirigida tales como la persuasión psicológica, la manipulación del usuario, el chantaje, y proporcionamos contramedidas que pueden ser fácilmente implementadas para evitar ataques con publicidad nanodirigida en Facebook.

Este trabajo es una traducción del trabajo referenciado a continuación:

José González-Cabañas, Ángel Cuevas, Rubén Cuevas, Juan López-Fernández, and David García. 2021. ***Unique on Facebook: Formulation and Evidence of (Nano)targeting Individual Users with non-PII Data.*** *In Proceedings of the 21st ACM Internet Measurement Conference (IMC '21). Association for Computing Machinery, New York, NY, USA, 464–479. https://doi.org/10.1145/3487552.3487861*

Índice de Contenidos

Índice de Figuras

Índice de Tablas

1 INTRODUCCIÓN

En el mundo hiper-conectado actual, la privacidad de un individuo está limitada por la cantidad de información que un tercero necesita conocer para identificarlo. Además de la información que directamente identifica a un individuo (p.ej. dirección de email, número de teléfono, dirección postal, número de pasaporte, etc.), un usuario puede ser identificado de manera unívoca mediante la combinación de un cierto número de elementos que individualmente no tienen esa capacidad. Definir cuántos de estos elementos se requiere para identificar de manera unívoca a un usuario es de gran importancia para entender los límites reales de nuestra privacidad. Estudios preliminares en el área de la unicidad de los usuarios han demostrado que conocer la información espacio-temporal asociada a 4 llamadas de teléfono móvil [3] o 4 compras con tarjeta de crédito[4] son suficientes para identificar de manera única a más del 90% de los usuarios en un dataset con más de 1.5 millones de usuarios. De manera similar, 3 elementos demográficos (género, código postal y fecha de nacimiento) son suficientes para identificar al 63% de los ciudadanos de Estados Unidos en el censo del año 2000 [19]. Estos estudios presentan limitaciones o porque trabajan con una base de usuarios relativamente pequeña o porque cubren un solo país.

En este trabajo, presentamos el primer estudio que analiza la unicidad de los usuarios teniendo en cuenta una base de usuarios en un orden de magnitud similar a la población mundial. El foco de nuestro estudio es Facebook (FB), una plataforma con más de 2.8B de usuarios activos al final de 2020 (el momento en que realizamos el estudio) [14]. Los elementos (que de manera individual no son válidos para identificar de manera unívoca a un usuario) que consideramos en nuestro análisis son los intereses que FB asigna a los usuarios en base a su actividad online y offline. Estos intereses son un importante activo de FB ya que su modelo de negocio se basa fundamentalmente en enviar anuncios relevantes a sus usuarios. Muchos anunciantes usan la plataforma de publicidad de FB para crear campañas y alcanzar a usuarios basados en sus intereses.

La primera contribución de este trabajo es un modelo basado en datos cuya salida es la métrica N_P, que se define como el número de intereses N que de manera unívoca identifican a un usuario en FB con una probabilidad P. Por ejemplo, $N_{50} = 12$ indica que la probabilidad de identificar a un usuario en FB con 12 intereses es de un 50%. Para obtener N_P, estudiamos el tamaño de miles de audiencias de FB formadas mediante la combinación de entre 1 y 25 intereses. Obtenemos el tamaño de las audiencias mediante el Administrador de Campañas de FB[13]. Para crear la combinación de intereses nos basamos en conjuntos reales de intereses de 2.4k usuarios de FB que han instalado la extensión para navegador FDVT [20], la cual recoge los intereses que FB ha asignado a dichos usuarios.

El resultado de nuestro modelo muestra que 4 intereses raros o 22 intereses aleatorios del conjunto total de intereses que FB asigna a un usuario lo hacen único en FB con una probabilidad de un 90%.

En contraposición a algunos elementos usados en estudios previos sobre la unicidad de usuarios (p.ej., transacciones con tarjeta de crédito o llamadas de teléfono), los intereses en Facebook han sido diseñados explícitamente para ser operables mediante campañas de publicad en FB. Por lo tanto, si un usuario puede ser identificado de manera unívoca mediante un conjunto e intereses en FB, puede que sea viable configurar una campaña publicitaria en FB usando un conjunto de intereses que alcance de manera exclusiva a un usuario único. En este trabajo definimos esta práctica como *publicidad nanodirigida* o *nanotargeting*.

La publicidad nanodirigida es una práctica potencialmente dañina. La literatura en el área de la *psicología de la persuasión* ha demostrado que persuadir a un individuo es más fácil si puedes crear mensajes a medida de las características psicológicas y motivaciones de esa persona [28]. En este contexto, la publicidad nanodirigida puede ser una herramienta muy potente para atacantes que persigan manipular a un individuo específico. Además, la publicidad nanodirigida podría usarse para chantajear a los usuarios.

Hasta donde nosotros conocemos, no existe evidencia previa de la posibilidad de explotar la plataforma de publicidad de FB de manera sistemática para implementar publicidad nanodirigida usando combinaciones de elementos (como los intereses de los usuarios) que de manera individual no permiten identificar a un usuario.

La segunda contribución consiste en obtener la primera evidencia empírica de que la publicidad nanodirigida puede ser implementada de manera sistemática en FB conociendo

simplemente un conjunto de intereses del usuario que se quiere alcanzar. En concreto, hemos configurado campañas de publicidad nanodirigida para tres de los autores de este trabajo. Hemos evaluado los resultados derivados de nuestro modelo creando anuncios específicos para cada uno de los autores usando combinaciones de 5, 7, 9, 12, 18, 20, y 22 intereses aleatorios seleccionados de la lista de intereses que FB les había asignado. En total, lanzamos 21 campañas de publicidad entre Octubre y Noviembre de 2020 para demostrar que la publicidad nanodirigida en FB es posible.

Nuestro experimento ha validado los resultados de nuestro modelo. Los resultados del experimento muestran que si un atacante conoce 19 interés o más de un usuario podrá tener éxito en un campaña de publicidad nanodirigida con alta probabilidad. En particular, 8 de las 9 campañas de publicidad configuradas con 18 o más intereses en nuestro experimento fueron exitosas en mostrar anuncios exclusivamente al usuario objetivo.

Tras demostrar la posibilidad de realizar campañas de publicidad nanodirigida de manera sistemática en FB, la última contribución de este trabajo se centra en discutir y proponer soluciones para proteger a los usuarios de los potenciales riesgos asociados a la publicidad nanodirigida (manipulación, chantaje, etc.). En primer lugar, añadimos una funcionalidad a la extensión de navegador FDVT para informar a los usuarios qué intereses de los que FB les ha asignado pueden ser más problemáticos para su privacidad (aquellos asociados con un audiencia de menor tamaño) usando una escala de color simple. Nuestra solución también permite que los usuarios eliminen los intereses que consideren oportuno con un solo clic. En segundo lugar, proponemos soluciones de fácil implantación que FB podría adoptar para evitar ataques de publicidad nanodirigida en su plataforma de publicidad.

2 CONTEXTO

En esta sección, describimos los escenarios tecnológicos que utilizamos en este trabajo. En particular, describimos en primer lugar la plataforma publicitaria de FB, que cumple una doble función en nuestra investigación. La utilizamos para (i) recuperar los tamaños de audiencia que sirven de entrada a nuestro modelo de unicidad de los usuarios en FB, y (ii) configurar las campañas publicitarias de nuestro experimento de publicidad nanodirigida. En segundo lugar, describimos la extensión del navegador FDVT, que proporciona el conjunto de intereses que utilizamos como entrada en nuestro análisis de la unicidad de los usuarios en FB.

2.1 Resumen del Administrador de Campañas de Publicidad de FB

Los anunciantes configuran sus campañas publicitarias en FB a través del *Administrador de Campañas de FB*[13]. Se trata de un panel de control en el que los anunciantes definen la audiencia (o perfil de usuario) a la que quieren dirigirse. El Administrador de Campañas de FB ofrece a los anunciantes una amplia gama de parámetros de configuración, tales como (pero no limitados a) *ubicación* (país, región, etc.), *parámetros demográficos* (sexo, edad, etc.), *comportamientos* (dispositivo móvil, SO y/o navegador web utilizado, etc.), y *intereses* (deportes, comida, etc.). En principio, todos estos atributos se consideran datos que no pueden utilizarse por sí solos de manera individual para identificar a un usuario.

Además, el Administrador de Campañas de FB publica el tamaño de la audiencia configurada en el panel de control a través del parámetro denominado *Alcance potencial*. Este parámetro informa del número de Usuarios Activos Mensuales (MAU por sus siglas en inglés) en FB que coinciden con la audiencia definida, que por definición es el tamaño de la audiencia. Además del panel de control, el Administrador de campañas de FB ofrece a los anunciantes una API para obtener automáticamente el alcance potencial de cualquier audiencia. Aprovechamos esa API para obtener el *Alcance potencial* asociado a las audiencias utilizadas para construir nuestro modelo.

Por otro lado, FB asigna a cada usuario un conjunto de intereses, denominados *preferencias publicitarias*. Las preferencias publicitarias de un usuario se infieren a partir de los datos y la actividad del usuario en FB y otros sitios web y servicios online en los que FB está presente. Estas preferencias publicitarias son, de hecho, los intereses que se ofrecen a los anunciantes en el Administrador de publicidad de FB. Por lo tanto, si un usuario tiene asignado *'comida italiana"* dentro de su lista de intereses, será un objetivo potencial de cualquier campaña publicitaria de FB configurada para llegar a usuarios interesados en *'comida italiana"*. Es importante señalar que los intereses en el ecosistema publicitario de FB son globales, por lo que no existen intereses específicos por país.

El único parámetro obligatorio para definir una audiencia en FB es la ubicación. Un anunciante puede combinar esa ubicación con cualquiera de los otros atributos disponibles. De este modo, podemos obtener el número de usuarios de FB en una ubicación concreta (por ejemplo, código postal, ciudad, país, etc.) o grupo de ubicaciones a las que se ha asignado un interés concreto (o grupo de intereses). Por motivos de privacidad, el valor mínimo de *Alcance Potencial* que FB devuelve para cualquier audiencia desde 2018 es de 1000. Anteriormente, este límite era solo de 20. En este trabajo, utilizamos un conjunto de datos recogidos en enero de 2017, por lo que nuestros datos están acotados por una limitación de tamaño de audiencia de 20 usuarios.

La limitación del alcance potencial no impide que los anunciantes realicen campañas para audiencias que incluyan menos de 1000 usuarios, sino que sólo impide que el anunciante conozca el tamaño real de la audiencia objetivo. Es más, un trabajo anterior [18] describe un mecanismo para reducir el límite actual del *Alcance potencial* mínimo de 1000 a 100. Este método puede ayudar a los anunciantes a inferir el tamaño de su público objetivo aunque éste sea inferior a 1000. Por lo tanto, cualquier autor que quiera replicar nuestro análisis de unicidad estaría limitado por un umbral de 100 en lugar de 20.

Es importante señalar que, en el momento en que recopilamos el conjunto de datos, el Administrador de Campañas de FB tenía dos limitaciones. En primer lugar, no podíamos crear consultas que incluyeran más de 25 intereses (esta limitación se mantiene hoy en día). En segundo lugar, el Administrador de Campañas de FB no incluía todo el mundo como posible ubicación (esta opción está disponible hoy en día). En su lugar, solicitaba introducir una ubicación específica (país, región, ciudad, código postal, etc.) o un grupo de ubicaciones. El número máximo de localizaciones permitidas en una consulta era de 50. Por lo tanto, para maximizar el número de usuarios a los que se dirigía nuestra investigación, realizamos nuestras consultas para un conjunto de ubicaciones que incluía los 50 países con el mayor número de usuarios de FB (véase el apéndice A). Estos países

representaban 1.500 millones de usuarios activos, lo que correspondía al 81% del total de FB cuando recogimos los datos [10].

Por último, FB también permite a los anunciantes dirigirse a los usuarios basándose en elementos que los identifican unívocamente a través de la funcionalidad *Audiencia Personalizada* [11] en su plataforma publicitaria. Una audiencia personalizada se refiere a una lista de usuarios identificados por un elemento que los identifica unívocamente (por ejemplo, número de teléfono móvil, dirección de correo electrónico, etc.). Una campaña publicitaria de FB basada en una audiencia personalizada tiene como objetivo llegar a los usuarios incluidos en dicha lista de audiencia personalizada. Para ello, FB encuentra a los usuarios registrados que coinciden con alguno de los elementos personalizados incluidos. FB impone dos requisitos importantes para el uso de una audiencia personalizada: (i) Los anunciantes son responsables de obtener el consentimiento explícito de los usuarios incluidos en la audiencia para que se les envíen anuncios de las campañas publicitarias dirigidas a la audiencia personalizada. No hacerlo puede implicar que el anunciante/atacante esté infringiendo la normativa sobre datos personales, como el Reglamento General de Protección de Datos (RGPD) [7] en Europa. Este requisito parece no ser necesario cuando se utilizan atributos que no generan unicidad en los usuarios; (ii) El número mínimo de usuarios que forman una audiencia personalizada tiene que ser de 100. Aunque las audiencias personalizadas son de gran interés en el contexto de los estudios de privacidad, requieren datos que en sí mismo ya identifican unívocamente al usuario y, por lo tanto, están fuera del contexto de este trabajo.

2.2 Extensión de navegador FDVT

La lista de intereses considerada en este trabajo se obtiene a partir de 2.390 usuarios reales de FB que instalaron la extensión del navegador [20] antes de enero de 2017. Nuestra extensión fue lanzada públicamente en octubre de 2016 y nuestra base de usuarios está formada por usuarios que libremente decidieron instalarla. La principal funcionalidad de la extensión del navegador FDVT es proporcionar a los usuarios una estimación en tiempo real de los ingresos que generan para FB a partir de los anuncios que reciben mientras navegan en FB. Para calcular la estimación de ingresos, los usuarios proporcionan en un proceso de registro algunos parámetros demográficos que nos permiten construir una audiencia a partir de ellos: (i) País de residencia (obligatorio), (ii) Género (opcional), (iii) Edad (opcional), y (iv) Estado civil (opcional). La extensión del navegador FDVT recoge (entre otros datos) los intereses que FB asigna al usuario. Para ello, la extensión del navegador analiza la página de preferencias de anuncios del usuario [12].

2.3 Aspectos éticos y consentimiento de usuario

Nuestra investigación se compromete a cumplir con las normas éticas y legales. Desde el punto de vista legal, estamos sujetos al Reglamento General de Protección de Datos (RGPD) [7] que se aplica a todos los países de la UE. Para cumplir con el RGPD, en el momento de instalar la extensión del navegador FDVT (es decir, el proceso de registro), todos los usuarios tienen que: (i) aceptar de forma proactiva (opt-in) las condiciones de uso [16] y la política de privacidad [15], y (ii) darnos permiso explícito (opt-in) para utilizar

la información recogida de forma anónima con fines de investigación. Desde el punto de vista ético, el comité de ética de la institución de los autores proporcionó una aprobación para desarrollar la extensión del navegador FDVT como parte de un proyecto europeo H2020 y las actividades de investigación derivadas del mismo.

En la Sección 5, realizamos campañas publicitarias reales con el objetivo de enviar publicidad nanodirigida a 3 de los autores de este trabajo utilizando un número diferente de intereses. En algunos casos, estas campañas también llegan a otros usuarios distintos de los autores a los que van dirigidas. Hemos utilizado anuncios (véase la Figura 6) que simplemente promocionan la extensión de navegador FDVT en todas nuestras campañas publicitarias. Además, nuestros anuncios no hacen un seguimiento de las impresiones publicitarias. Por lo tanto, no hay forma de obtener la identidad de los usuarios que reciben el anuncio. En caso de que algún usuario haga clic en el anuncio, se le remite al sitio web de FDVT [17] (el sitio web del proyecto FDVT). Como en la mayoría de los servicios web estándar, recogemos la dirección IP del dispositivo que ha abierto la conexión. Esta es una práctica común por múltiples razones, de las cuales la más importante es la seguridad. Para proteger aún más la privacidad de los usuarios, en el contexto de este experimento, hemos convertido la dirección IP en un seudónimo utilizando una función hash (secreta) y una clave. Además, los únicos usuarios a los que se dirigen las campañas publicitarias de la sección 5 son los autores del artículo que conocen y aceptan el propósito de los experimentos de publicidad nanodirigida.

3 DATASET

Nuestro conjunto de datos se creó a partir de 2.390 usuarios reales que instalaron nuestra extensión de navegador FDVT entre octubre de 2016 (lanzamiento público) y enero de 2017. De estos usuarios, 1.949 declararon ser hombres, 347 mujeres y 94 no revelaron su sexo. Además, siguiendo la clasificación por grupos de edad propuesta en [6], 117 usuarios son adolescentes (de 13 a 19 años), 1374 adultos en edad temprana (de 20 a 39 años), 578 adultos (de 40 a 64 años), 19 mayores (mayores de 65 años), y 302 no facilitaron su edad. Por último, nuestros 2.390 usuarios declararon estar ubicados en 80 países diferentes. En el Apéndice B se puede encontrar un desglose detallado del número de usuarios por país en nuestro conjunto de datos.

Hemos obtenido 1,5 millones de ocurrencias de 99 mil intereses únicos de FB asignados a los 2.390 usuarios. La figura 1 muestra la CDF del número de intereses por usuario. El número de intereses de FB asignados a un usuario individual en nuestro conjunto de datos oscila entre 1 y 8.950, con una mediana de 426 intereses.

Para comprender la distribución de la popularidad de estos intereses, hemos extraído el tamaño de la audiencia que indica la API del gestor de publicidad de FB para cada uno de ellos. La figura 2 representa la CDF del tamaño de la audiencia para los 99k intereses únicos de nuestro conjunto de datos. Los resultados muestran una gran variabilidad en la popularidad de los intereses. En concreto, los percentiles 25, 50 y 75 de la distribución son 113.193, 418.530 y 1.719.925, respectivamente.

Somos conscientes de que nuestro conjunto de datos puede no ser una muestra estadísticamente representativa de todo el ecosistema de intereses de FB; sin embargo,

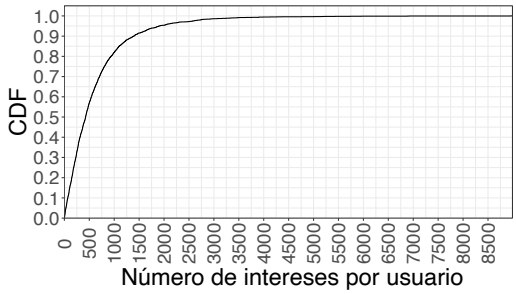

Fig. 1. CDF que muestra la distribución del número de intereses asignados a los 2.390 usuarios de nuestro conjunto de datos.

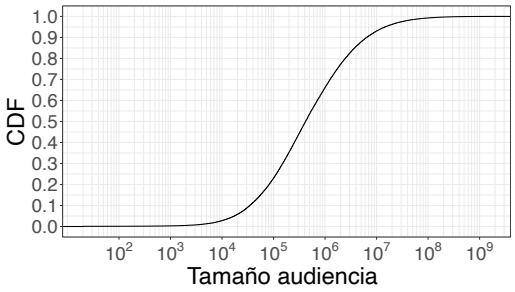

Fig. 2. CDF que muestra la distribución del tamaño de la audiencia para los 98.982 intereses asignados a los 2.390 usuarios de nuestro conjunto de datos.

incluye un número muy grande de intereses que cubren un rango de popularidad muy amplio, que es lo que necesitamos para el propósito de nuestro trabajo. Tal y como validan los resultados del artículo, el conjunto de datos recogidos es apropiado para (i) cuantificar cuántos intereses hacen que un usuario sea único en FB, y (ii) demostrar que la publicidad nanodirigida puede implementarse sistemáticamente en FB.

4 ANÁLISIS DE LA UNICIDAD DE LOS USUARIOS EN FACEBOOK

Dedicamos esta sección a analizar la unicidad de los usuarios en FB en base a sus intereses. El resultado de esta sección servirá para dos propósitos diferentes: (i) responderemos a la primera pregunta de investigación abordada en este trabajo: *¿Cuántos intereses son necesarios para identificar a un usuario en FB?* (ii) la respuesta a esta pregunta se utilizará como referencia para el número de intereses que debemos considerar en nuestro experimento de publicidad nanodirigida (véase la Sección 5.1).

4.1 Metodología

Definimos la variable N_P como el número de intereses que identifican de forma unívoca a un usuario con una probabilidad P en FB. Por ejemplo, si con 9 (18) intereses un usuario puede ser identificado de forma única en FB con una probabilidad de 0,3 (0,8), entonces $N_{0,3}$ = 9 ($N_{0,8}$=18).

Nuestro objetivo es proponer un modelo que defina N_P para cualquier valor de P. Utilizamos como fuente de datos los 99k intereses únicos asignados a los 2.390 usuarios de la extensión del navegador FDVT.

Consideremos un usuario en nuestro conjunto de datos u_i (i $\in$ [1, 2390]) y un número determinado de intereses N ($N \in$ [1,25]).[1] Para cada par (u_i, N), seleccionamos un conjunto de N intereses de la lista de intereses que FB asignó a u_i y recogemos el tamaño

[1]Este rango se debe a la limitación impuesta por la API de FB que permite recuperar los tamaños de audiencia para una combinación de un máximo de 25 intereses.

de la audiencia de FB asociado a esa combinación de intereses aprovechando la API de FB Ads Manager. Después de hacer esto para todas las combinaciones de u_i y N, obtenemos 25 vectores, uno por cada valor de N, que incluyen 2.390 muestras de tamaño de audiencia.[2] Por ejemplo, en el caso de $N = 5$, creamos un vector con 2.390 valores de tamaño de audiencia de 2.390 combinaciones diferentes de 5 intereses (uno por usuario en nuestro conjunto de datos).

Utilizando estos vectores, podemos producir una distribución del tamaño de la audiencia para cada valor de N y calcular los diferentes percentiles de la distribución. Basándonos en esto, definimos $AS(Q, N)$ como el tamaño de la audiencia para el percentil Q y el número de intereses N. Por ejemplo, un $AS(50, 5)$ = 500 significa que con una probabilidad del 50% el tamaño de una audiencia definida con 5 intereses es ≤ 500. Obsérvese que, dado que el tamaño mínimo de audiencia indicado por FB es 20, $AS(Q, N) \geq 20$ por definición.

A continuación, creamos un vector $V_{AS}(Q)$ que incluye los valores de $AS(Q, N)$ para un valor fijo de Qy todos los valores de N (de 1 a 25). $V_{AS}(Q)$ se define como

$$V_{AS}(Q) = [AS(Q, 1),\ AS(Q, 2),\ ...,\ AS(Q, 24),\ AS(Q, 25)]$$

Como $AS(Q, N) \geq AS(Q, N + 1)$, $V_{AS}(Q)$ presenta una tendencia decreciente. La Figura 3 muestra ejemplos de $V_{AS}(Q)$ para Q = 50 y Q = 90, donde el eje Y representa el tamaño de la audiencia y el eje X representa el número de intereses N.

En el modelo descrito, N_P *se define como el punto de corte donde* $V_{AS}(Q)$ *corta un tamaño de audiencia igual a 1*. Desgraciadamente, como podemos observar, $V_{AS}(Q)$ tiene una asíntota en 20, ya que éste es el tamaño mínimo de audiencia comunicado por FB.

Para superar este problema, ajustamos $V_{AS}(Q)$ utilizando el siguiente modelo logarítmico:

$$log(V_{AS}(Q)) \sim -Alog(N + 1) + B$$

A partir de este ajuste calculamos el punto de corte del número de intereses N en el que la recta de regresión intercepta un tamaño de audiencia de 1, es decir, $V_{AS}(Q)$ = 1. Dado que estamos utilizando un modelo logarítmico, el punto de corte se produce realmente cuando $log(V_{AS}(Q))$ = 0. Por lo tanto, N_P se define de la siguiente manera

$$N_P \geq 10^{B/A} - 1$$

Para evaluar la incertidumbre de esta estimación, repetimos la agregación de datos y el ajuste del modelo en 10.000 muestras bootstrap, calculando así el Intervalo de Confianza (IC) del 95% del punto de corte para cada valor de N. Nótese que no truncamos los datos para audiencias de tamaño 20, e incluimos el primer $AS(Q, N)$ = 20 en nuestra estimación. De este modo, nuestra estimación del punto de corte es conservadora pero robusta para el tamaño mínimo de 20 y nuestro método puede seguir aplicándose para el actual límite superior de 1.000 usuarios.

[2]Tenemos en cuenta que algunos de los vectores incluyen menos de 2.390 muestras porque en nuestro conjunto de datos encontramos usuarios que tienen asignados menos de 25 intereses. El vector más corto es el asociado a $N = 25$ que incluye 2.286 muestras.

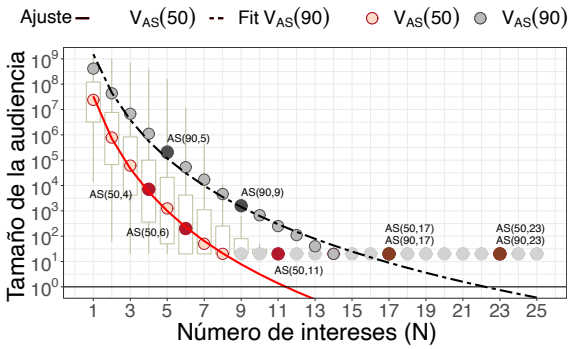

Fig. 3. Esta Figura representa una ilustración de nuestro modelo para calcular N_P. En primer lugar, muestra un ejemplo de las variables $V_{AS}(Q)$ y $AS(Q,N)$ para Q = 50 (puntos rojos) y Q = 90 (puntos negros). $AS(Q,N)$ colisiona para ambos casos a partir de N=14 cuando el valor del tamaño de la audiencia pasa a ser 20 (el límite impuesto por FB). En segundo lugar, la Figura ilustra el modelo de ajuste logarítmico utilizado para estimar el valor de N_P para $V_{AS}(50)$ (línea roja) y $V_{AS}(90)$ (línea negra discontinua) como punto de corte de las líneas con el valor y=1 (tamaño de audiencia igual a 1).

La Figura 3 muestra el resultado del proceso de ajuste de $V_{AS}(Q)$ para Q = 50 (línea discontinua roja) y Q = 90 (línea discontinua negra).

Utilizando como referencia el resultado del modelo presentado, hemos implementado un experimento ejecutando campañas reales de FB para enviar publicidad nanodirigida a tres de los autores de este trabajo para validar: (i) si es factible implementar un ataque sistemático de nanotargeting en FB basado en los intereses de los usuarios; (ii) si los valores de N_P derivados de nuestra metodología pueden ser utilizados como una buena referencia de la probabilidad de éxito de un potencial ataque de publicidad nanodirigida. Describimos el experimento en detalle y presentamos los resultados obtenidos en la Sección 5.1.

4.2 Estrategia de selección de intereses

El valor de N_P, es decir, el número de intereses que hacen que un usuario sea único en FB con probabilidad P, depende en gran medida de la estrategia utilizada para seleccionar los intereses.

La popularidad, es decir, el tamaño de la audiencia, de los intereses de FB es muy diversa y también lo es la popularidad de los intereses de un usuario individual. Nuestro conjunto de datos revela que, en general, entre los cientos de intereses que se suelen asignar a un usuario individual en FB, algunos son muy populares (con tamaños de audiencia del orden de decenas o cientos de millones de usuarios) mientras que otros son impopulares (con tamaños de audiencia del orden de decenas o algunos cientos de usuarios).

La mejor alternativa para conseguir enviar publicidad nanordirigida a un usuario consiste en realizar una campaña publicitaria seleccionando los intereses menos populares de

ese usuario como perfil objetivo, lo que se espera que conduzca a valores pequeños de N_P (incluso para valores altos de P como 0,9 o 0,95). Sin embargo, la aplicación de este ataque requeriría tener un conocimiento completo de la lista de intereses del usuario objetivo, lo que en la práctica es muy poco probable. En cambio, es más probable que un atacante conozca un subconjunto de los intereses del usuario objetivo, pero no todos.

Teniendo en cuenta que FB impone una limitación de 25 intereses en la definición de la audiencia, en este trabajo aplicamos dos estrategias diferentes para la selección de intereses basados en la discusión anterior:

- *Selección de intereses menos populares (MP):* Recuperamos el tamaño de la audiencia de todos los intereses asignados a un usuario y seleccionamos los 25 menos populares. Comenzamos obteniendo el tamaño de la audiencia para el interés menos popular y seguimos añadiendo los siguientes intereses menos populares uno a uno para obtener todos los tamaños de audiencia asociados hasta completar la combinación más larga de 25 intereses. A partir de ahora, utilizaremos la variable $N(MP)_P$ cuando N_P se calcule seleccionando los intereses menos populares de los usuarios.
- *Selección aleatoria de intereses (A):* Seleccionamos 25 intereses al azar entre los intereses asignados a un usuario. Empezamos a obtener el tamaño de la audiencia para un interés aleatorio y seguimos añadiendo intereses aleatorios secuencialmente uno a uno para obtener todos los tamaños de audiencia asociados hasta completar la combinación más larga de 25 intereses. A partir de ahora en el documento, utilizaremos la variable $N(A)_P$ cuando se calcule N_P seleccionando los intereses de los usuarios de forma aleatoria.

El valor de $N(MP)_P$ tiene una importante relevancia teórica, ya que establece un límite inferior teórico en términos de privacidad basado en el número de intereses que hacen que un usuario sea único entre 1500 millones de usuarios de FB (aproximadamente 1/5 de la población mundial) considerados en nuestro análisis de unicidad. Por lo tanto, $N(MP)_P$ es, hasta donde sabemos, el cálculo más aproximado realizado hasta ahora sobre el número de elementos (que inicialmente no identifican a un usuario de forma unívoca) que hacen que un individuo sea único entre toda la humanidad.

Sin embargo, como se ha comentado anteriormente, $N(MP)_P$ sólo sirve como referencia para fines de publicidad nanodirigida en aquellos casos en los que el atacante conoce la lista completa de intereses de un usuario, lo que se espera que no sea una situación habitual. Por lo tanto, utilizaremos $N(A)_P$ como referencia para nuestro experimento de publicidad nanodirigida presentado en la Sección 5.1.

4.3 Resultados

En esta sección, aplicamos el modelo desarrollado para calcular N_P, el número de intereses que hacen que un usuario sea único en FB con una probabilidad P. En particular, para realizar un análisis exhaustivo, consideramos P = 0,5, 0,8, 0,9 y 0,95 y las dos estrategia de selección de intereses definidos: intereses *menos populares ($N(MP)_P$)* y *aleatorios ($N(A)_P$)*.

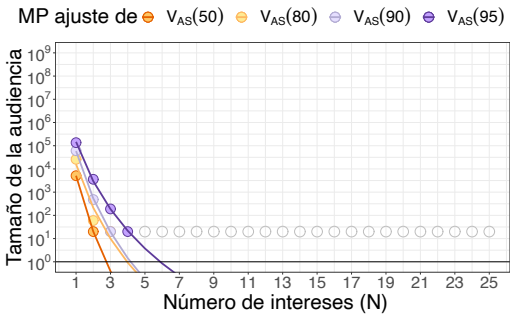

Fig. 4. La Figura ilustra los resultados de nuestro modelo para calcular el número de intereses que hacen único a un usuario en FB utilizando sus intereses menos populares. En particular, la Figura muestra los resultados para $N(MP)_{0,5}$, $N(MP)_{0,8}$, $N(MP)_{0,9}$ y $N(MP)_{0,95}$ aplicando nuestro modelo de ajuste a los vectores $V_{AS}(50)$, $V_{AS}(80)$, $V_{AS}(90)$ y $V_{AS}(95)$.

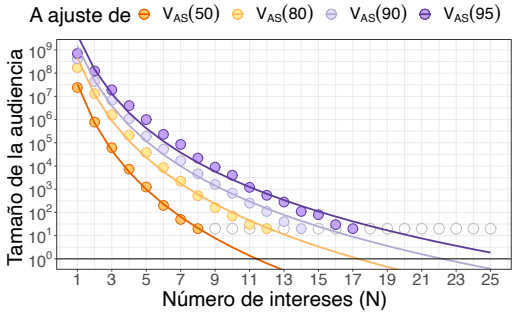

Fig. 5. La Figura ilustra los resultados de nuestro modelo para calcular el número de intereses que hacen único a un usuario en FB combinando intereses al azar. En concreto, la Figura muestra los resultados para $N(A)_{0,5}$, $N(A)_{0,8}$, $N(A)_{0,9}$ y $N(A)_{0,95}$ aplicando nuestro modelo de ajuste a los vectores $V_{AS}(50)$, $V_{AS}(80)$, $V_{AS}(90)$ y $V_{AS}(95)$.

También hemos calculado $N(MP)_P$ y $N(A)_P$ en diferentes grupos demográficos basados en el género, la edad y la ubicación (país) para explorar las diferencias en el número de intereses que hacen que un usuario sea único en estos grupos. Los resultados del análisis demográfico se describen en el Apéndice C.

4.3.1 $N(MP)_P$: Estrategia de selección de intereses menos populares

La Figura 4 muestra $V_{AS}(Q)$ para la selección de los intereses menos populares de los usuarios en nuestro conjunto de datos y Q = 50, 80, 90 y 95 junto con su correspondiente curva de ajuste lineal.

Además, la Tabla 1 presenta el valor estimado de $N(MP)_{0,5}$, $N(MP)_{0,8}$, $N(MP)_{0,9}$ y $N(MP)_{0,95}$ junto con el intervalo de confianza del 95% y el valor de R-cuadrado (R^2). Ambas métricas de calidad del modelo sugieren que el modelo de ajuste es muy preciso.

Como se ha comentado anteriormente, los valores de $N(MP)_P$ obtenidos ofrecen un límite inferior sobre el número de elementos que hacen que un usuario sea único entre 1500 millones de usuarios de FB, aproximadamente 1/5 de la población mundial. $N(MP)_{0,95}$ = 5,89, indica que un usuario puede ser identificado de forma única en FB basándose en sus 6 intereses menos populares con una probabilidad del 95%. Del mismo modo, $N(MP)_{0,9}$ = 4,16 y $N(MP)_{0,5}$ = 2,74 muestran que con aproximadamente los 4 y 3 intereses menos populares se puede identificar de forma única a un individuo entre 1500 millones de usuarios con una probabilidad del 90% y 50%, respectivamente.

Los resultados indican que el número de elementos (que individualmente no son suficiente para identificar a un usuario de manera única) que hacen único a un usuario en

N_P	P=0.5	95% CI	R^2	P=0.8	95% CI	R^2	P=0.9	95% CI	R^2	P=0.95	95% CI	R^2
$N(MP)_P$	2.74	(2.72,2.75)	1.00	3.96	(3.91,4.02)	0.92	4.16	(4.09,4.37)	1.00	5.89	(5.62,6.15)	1.00
$N(A)_P$	11.41	(11.21,11.6)	1.00	17.31	(16.98,17.6)	0.99	22.21	(21.73,22.69)	0.99	26.98	(26.34,27.68)	0.98

Tabla 1. Número de intereses necesarios para que un usuario sea único en FB con probabilidad 0,5, 0,8, 0,9 y 0,95 ($N_{0,5}$, $N_{0,8}$, $N_{0,9}$ y $N_{0,95}$). La primera fila revela los resultados para el caso en el que seleccionamos los intereses de los usuarios menos populares (es decir, $N(MP)_P$). La segunda fila expone los resultados para una selección aleatoria de los intereses de los usuarios (es decir, $N(A)_P$). Proporcionamos los resultados junto con el Intervalo de Confianza (IC) del 95% y el R-cuadrado (R^2) asociado con el modelo de ajuste utilizado para obtener $N(MP)_P$ y $N(A)_P$.

un conjunto de datos a escala de la población mundial es realmente pequeño (4 con una probabilidad del 90%). En otras palabras, la privacidad de un usuario sólo está limitada por un puñado de elementos que individualmente no pueden hacerlo.

Por último, en el contexto de la publicidad nanodirigida, nuestro resultado sugiere que un atacante que tenga acceso completo a la lista de intereses de un usuario puede enviarle un anuncio nanodirigido con una probabilidad del 90% realizando una campaña publicitaria con sólo 4 intereses. La probabilidad de éxito aumenta al 95% si el atacante utiliza los 6 intereses menos populares.

4.3.2 $N(A)_P$: Estrategia de selección de intereses aleatorios

En esta subsección, presentamos un análisis para $N(A)_P$ similar al presentado anteriormente para $N(MP)_P$. La Figura 5 muestra $V_{AS}(Q)$ basada en la selección aleatoria de los intereses de los usuarios y Q = 50, 80, 90 y 95 junto con su correspondiente curva de ajuste lineal.

Además, la Tabla 1 presenta las estimaciones obtenidas para $N(A)_{0,5}$, $N(A)_{0,8}$, $N(A)_{0,9}$, y $N(A)_{0,95}$ junto con sus intervalos de confianza asociados y el valor de R-cuadrado (R^2). Los intervalos de confianza y los valores de R-cuadrado indican de nuevo una buena precisión del modelo propuesto.

Los resultados obtenidos revelan que 12, 18, 22 y 27 intereses aleatorios hacen que un usuario sea único en FB con una probabilidad del 50%, 80%, 90% y 95%, respectivamente. Estos resultados tienen dos implicaciones prácticas principales: 1) Dado que FB suele asignar cientos de intereses a los usuarios, es probable que un atacante pueda inferir algunas decenas de esos intereses que le permitirían enviar publicidad nanodirigida a la víctima; 2) Realizar un ataque con un 95% de probabilidad de éxito es imposible en la práctica, ya que requiere dirigirse a un público que combine 27 intereses cuando FB impone un máximo de 25 intereses para un público objetivo. Obsérvese que utilizamos estos valores de $N(A)_P$ como referencia para realizar el experimento de publicidad dirigida de la Sección 5.1.

Resumen de los resultados del análisis de unicidad.

Nuestros resultados revelan que: (i) los 4 intereses más raros de un usuario lo hacen

único dentro de una base de usuarios del mismo orden de magnitud que la población mundial. Esto indica que la unicidad de un individuo está definida por la combinación de un pequeño número de elementos que individualmente no hacen al usuario único, por lo que la privacidad de los usuarios está muy comprometida en la actual sociedad híperconectada; (ii) *En comparación con estudios anteriores, nuestro análisis revela que el número de elementos (que individualmente no tienen capacidad de unicidad) que hacen único a un usuario en un conjunto de datos con millones de usuarios (4 compras con tarjeta de crédito o 4 llamadas de teléfono móvil) o miles de millones de usuarios (4 intereses más raros o 22 intereses aleatorios) está en el mismo orden de magnitud. Esto significa que pertenecer a grupos humanos de mayor escala no parece contribuir a mejorar significativamente los límites de la privacidad de los individuos.*

5 EXPERIMENTO DE PUBLICIDAD NANODIRIGIDA

En esta sección, presentamos un experimento que se basa en los resultados derivados de nuestro análisis de la sección anterior. Nuestro objetivo es proporcionar pruebas de que la plataforma publicitaria de FB puede ser explotada sistemáticamente para implementar campañas de publicidad dirigida con la combinación de elementos que individualmente no identifican a un usuario.

Recordamos que nuestra definición de publicidad nanodirigida requiere que el anuncio se envíe exclusivamente al usuario objetivo. Por lo tanto, cuando indicamos que una campaña de publicidad nanodirigida ha fallado no significa que la campaña no haya llegado al usuario objetivo. Significa que el anuncio ha sido entregado a más de un usuario, que puede incluir o no al usuario objetivo.

5.1 Descripción del experimento

El experimento consistió en crear campañas publicitarias en FB para llegar a tres autores de este trabajo utilizando conjuntos de intereses aleatorios obtenidos de la lista completa de intereses que FB les había asignado. Como se ha comentado anteriormente, decidimos centrar nuestro experimento en el uso de intereses aleatorios, ya que en la práctica es mucho más probable que un atacante conozca una lista aleatoria de intereses de un usuario que sus intereses menos populares.

Selección de intereses Basándonos en los resultados presentados en la Sección 4, hemos lanzado 7 campañas por usuario configuradas con 5, 7, 9, 12, 18, 20 y 22 intereses seleccionados aleatoriamente. Dado que nos dirigimos a 3 usuarios independientes, ejecutamos 21 campañas publicitarias de FB en total en nuestro experimento.

Dividimos los experimentos en dos grupos en función de la probabilidad de éxito esperada. El primer grupo incluye los experimentos que utilizan 12, 18, 20 y 22 intereses. Nos referimos a él como “Grupo de éxito” porque nuestros resultados en la sección de análisis de unicidad indican que la probabilidad de éxito de una campaña de publicidad nanoridirigida para el número de intereses considerado en este grupo oscila entre el 50 y el 90 por ciento. Por lo tanto, esperamos que muchas de estas campañas envíen el anuncio exclusivamente al usuario objetivo.

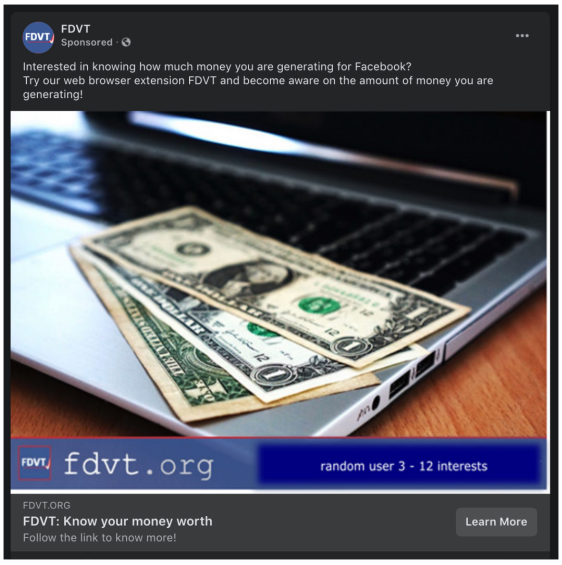

Fig. 6. Imagen del anuncio utilizado en la campaña dirigida al usuario 3 con 12 intereses. Cada anuncio incluía un texto que identificaba la campaña en la esquina inferior derecha.

El segundo grupo configura las campañas utilizando 5, 7 y 9 intereses y lo denominamos "Grupo de Fallo" ya que los resultados de nuestro modelo manifiestan una probabilidad de éxito del 2,5%, 15% y 30% para 5, 7 y 9 intereses, respectivamente. Basándonos en esto, esperamos que la mayoría de estas campañas de publicidad nanodirigida fallen y entreguen el anuncio a múltiples usuarios, aunque probablemente incluyan al usuario objetivo.

Para realizar los experimentos, seleccionamos un conjunto aleatorio de 22 intereses de cada usuario que se utilizan directamente en las campañas configuradas con 22 intereses. Para crear la campaña con 20 intereses, eliminamos 2 del conjunto inicial. Del mismo modo, para crear la campaña con 18 intereses, eliminamos 2 del conjunto utilizado en la campaña de 20 intereses. Siguiendo la misma estrategia, eliminamos 6 intereses de la campaña de 18 intereses y utilizamos los restantes para definir una campaña de 12 intereses. Seguimos el mismo proceso para definir las campañas restantes.

Rango geográfico. El alcance geográfico de las campañas publicitarias se define como "mundial". Esto hace que nuestras campañas no filtren a los usuarios en función de su ubicación y, por tanto, pueden llegar potencialmente a cualquier usuario de FB. Hay que tener en cuenta que FB informó que tenía 2.8B usuarios activos mensuales en el último trimestre de 2020 [14] cuando realizamos nuestro experimento.

Creatividad de los anuncios. Hemos creado una creatividad publicitaria específica para cada una de las 21 campañas configuradas. Cada creatividad publicitaria identifica tanto

el usuario al que se dirige (Usuario 1, Usuario 2 o Usuario 3) como el número de intereses utilizados en su campaña asociada (5, 7, 9, 12, 18, 20, 22). Por ejemplo, la Figura 6 presenta el anuncio recibido por el Usuario 3 en su muro de FB asociado a la campaña de publicidad nanodirigida configurada con 12 intereses aleatorios. Además, cada creatividad publicitaria está vinculada a una página de destino diferente alojada en nuestro servidor web.

Tiempo y duración. Cada campaña se ejecutó durante un tiempo total de 33 horas divididas en 4 ventanas de tiempo. En particular, las campañas del grupo de éxito (con los intereses 12, 18, 20 y 22) se desarrollaron en paralelo el jue. 29 de octubre de 2020, de 19h a 21h CET, el vie. 30 de octubre de 9h a 21h CET, lun. 2 de noviembre de 9h a 21h CET y el mar. Nov 3 de 9h a 16h CET. Detuvimos todas las campañas del *Grupo de Éxito* al mismo tiempo una vez que los tres usuarios habían recibido al menos una vez el anuncio dirigido de cada campaña. Las campañas del *Grupo de Fallo* (con 5, 7 y 9 intereses) se ejecutaron en paralelo exactamente a las mismas horas y días que las campañas del *Grupo de Éxito* en la semana siguiente.

Utilizamos la misma duración y los mismos días y horas de la semana en todas las campañas publicitarias, para garantizar que todas ellas tuvieran el mismo tiempo de entrega de impresiones publicitarias y para evitar posibles sesgos en nuestros resultados debidos a condiciones especiales que afectaran a días de la semana u horas concretas dentro de un día.

Presupuesto. Asignamos un presupuesto diario inicial de 70 euros a cada campaña publicitaria para un periodo de una semana. FB distribuye el presupuesto a lo largo de la ventana de tiempo proporcionada, por lo que nuestro gasto previsto era de aproximadamente 10€/día por campaña. Dado que la duración real de nuestras campañas fue inferior a la ventana de tiempo proporcionada de una semana, ninguna de las campañas consumió los 70€ asignados. El coste total de nuestro experimento fue de 305,36 euros.

Validación del éxito de las campañas de publicidad nanodirigidas. Utilizamos tres elementos para validar el éxito (o el fracaso) de una campaña publicitaria dirigida a un usuario específico en nuestro experimento:

(1) Facebook ofrece a los anunciantes un panel de control en el que pueden supervisar el progreso de sus campañas publicitarias. Este panel informa de cada campaña (entre otras cosas): el número de impresiones entregadas, el número de usuarios únicos alcanzados, el número de clics en el anuncio y el presupuesto gastado en la campaña publicitaria.
(2) Mantenemos un registro de cada impresión de anuncio entregada a los usuarios objetivo. Para ello, al recibir un anuncio dirigido, se indicaba al usuario que hiciera clic en él. Dado que cada creatividad publicitaria tiene una única página de destino asociada, este clic creaba una entrada de registro en nuestro servidor web que registraba los detalles de la campaña publicitaria (usuario objetivo y número de intereses) y la marca de tiempo.

(3) También se le indicó a cada usuario que tomara una instantánea del anuncio recibido junto con la información incluida en la opción *"¿Por qué estoy viendo este anuncio?"* que Facebook ofrece a los usuarios. Cuando un usuario hace clic en la opción *"¿Por qué estoy viendo este anuncio?"*, aparece una nueva ventana que muestra los parámetros utilizados por el anunciante para definir el público objetivo en la campaña publicitaria asociada al anuncio. En nuestro experimento, esos parámetros se refieren a la lista de intereses utilizada en la campaña publicitaria. Las Figuras 11 y 12 del Apéndice D ilustran un ejemplo de la opción *"¿Por qué estoy viendo este anuncio?"* captada por uno de los autores. Hemos comprobado que, para cada anuncio dirigido identificado por los autores, los parámetros incluidos en *"¿Por qué estoy viendo este anuncio?"* coincidían exactamente con la audiencia configurada asociada al anuncio recibido.

Combinando las tres informaciones descritas, podíamos identificar fácilmente si una campaña de publicidad nanodirigida había tenido éxito o no. En concreto, podíamos concluir con seguridad que una campaña de publicidad nanodirigida había tenido éxito si se daban las tres condiciones siguientes: (i) FB informaba de que sólo se había llegado a un usuario, (ii) teníamos un registro en nuestro servidor web generado por el clic del usuario en el anuncio, y (iii) el usuario objetivo recogía una instantánea del anuncio y su opción *"¿Por qué estoy viendo este anuncio?"*

5.2 Resultados

En la Tabla 2 se presentan los resultados del experimento de publicidad nanodirigida. Para cada usuario y campaña publicitaria (es decir número de intereses utilizados), la Tabla muestra las cinco métricas siguientes: (i) *Recibido*: es una métrica binaria que indica si el usuario ha recibido el anuncio o no; (ii) *Alcance*: informa del número total de usuarios únicos a los que ha llegado la campaña, según la información facilitada por el panel del Administrador de Campañas de FB; (iii) *Impresiones*: indica el número total de impresiones realizadas por la campaña, según el panel del gestor de campañas de FB. Tenga en cuenta que el número de impresiones suele ser mayor que el número de usuarios alcanzados porque el anuncio puede ser entregado varias veces al mismo usuario; (iv) *Tiempo hasta la primera impresión (TFI)*: muestra el tiempo transcurrido desde que se lanzó la campaña hasta que el usuario objetivo recibió la primera impresión del anuncio. Para calcular esta métrica sólo consideramos los periodos en los que la campaña estuvo activa; (v) *Coste*: informa del importe que nos facturó FB en euros. (v) *Clics*: Muestra el número total de clics recibidos en la campaña, y entre paréntesis el número de direcciones IPs (seudonimizadas) únicas que han llevado a cabo esos clics. Por último, la Tabla destaca en negrita las campañas de publicidad dirigida que han tenido éxito, es decir, las que han llegado exclusivamente al usuario objetivo.

En el resto de la sección, comentamos los aspectos más relevantes de los resultados obtenidos.

Viabilidad de la publicidad nanodirigida. 9 de las 21 campañas alcanzaron con éxito al usuario correspondiente. Estas campañas son todas de 20 y 22 intereses, dos (de las

	Usuario 1					
	Recibido	Alcance	Impresiones	TFI	Coste	Clics
5 intereses	No	9,824	42,273	-	€28.58	40 (38)
7 intereses	No	2,992	14,774	-	€29.47	14 (13)
9 intereses	Sí	743	4,883	2h 11’	€28,74	17(14)
12 intereses	Sí	152	1,110	9h 8’	€19.28	9 (4)
18 intereses	**Sí**	**1**	**1**	**3h 31’**	**€0.01**	**1 (1)**
20 intereses	**Sí**	**1**	**1**	**47’**	**Free**	**1 (1)**
22 intereses	**Sí**	**1**	**1**	**28h 40’**	**Free**	**1 (1)**

	Usuario 2					
	Recibido	Alcance	Impresiones	TFI	Coste	Clics
5 intereses	No	89,328	251,379	-	€28.97	94 (94)
7 intereses	Sí	1,843	10,004	2h 9’	€29.30	23 (22)
9 intereses	Sí	1,152	7,175	1h 47’	€29.00	19 (19)
12 intereses	Sí	201	970	4h 22’	€18.68	11 (6)
18 intereses	Sí	92	263	27h 57’	€4.15	6 (3)
20 intereses	**Sí**	**1**	**1**	**44’**	**€0.01**	**1 (1)**
22 intereses	**Sí**	**1**	**1**	**32h 10’**	**€0.01**	**1 (1)**

	Usuario 3					
	Recibido	Alcance	Impresiones	TFI	Coste	Clics
5 intereses	No	39,520	100,106	-	€30.05	93 (90)
7 intereses	No	2,221	11,248	-	€30.83	26 (25)
9 intereses	Sí	749	4,356	1h 50’	€28,19	22(15)
12 intereses	**Sí**	**1**	**1**	**12h 22’**	**€0.01**	**1 (1)**
18 intereses	**Sí**	**1**	**2**	**6h 19’**	**€0.02**	**2 (2)**
20 intereses	**Sí**	**1**	**5**	**3h 32’**	**€0.06**	**5 (3)**
22 intereses	**Sí**	**1**	**1**	**48’**	**Free**	**1 (1)**

Tabla 2. Resultados del experimento de publicidad dirigida para tres autores del artículo. Las filas indican el número de intereses utilizados en cada una de las 7 campañas publicitarias lanzadas por usuario. Las columnas representan el rendimiento: *Recibido* (si el usuario objetivo recibió el anuncio o no); *Alcance* (el número de usuarios alcanzados por la campaña); *Impresiones* (el número total de impresiones realizadas en la campaña); *TFI* (tiempo transcurrido hasta la primera impresión entregada al usuario objetivo); *Coste* (coste de la campaña); *Clics* (número de clics en la campaña y número de direcciones IP únicas seudonimizadas que generan esos clics, entre paréntesis).

tres) de 18 intereses y una (de las tres) de 12 intereses. Hay otras seis campañas (dos de 12 intereses, las tres de 9 intereses y una de 7 intereses) que también llegaron al usuario objetivo junto con otros pocos cientos o miles de usuarios. Por lo tanto, estas campañas no consiguieron realizar con éxito publicidad nanodirigida (es decir, alcanzar exclusivamente) a un usuario específico. Por último, hay cinco campañas (las tres de 5 intereses y las dos de 7 intereses) que no llegaron al usuario objetivo.

En resumen, nuestro experimento demuestra que un atacante puede enviar publicidad nanodirigida sistemáticamente a un solo usuario en FB si puede inferir un número suficiente de intereses del individuo objetivo.

Coste de la publicidad nanodirigida. Una cuestión importante es cuál es el coste real asociado a una campaña de publicidad nanodirigida. Hay que tener en cuenta que un coste muy elevado puede ser un factor disuasorio en la práctica. Lamentablemente, los resultados extraídos del gestor de campañas publicitarias de FB y recogidos en la Tabla 2 demuestran que realizar una campaña de publicidad nanodirigida a un usuario es bastante barato. De hecho, el coste global de las 9 campañas de publicidad nanodirigida realizadas con éxito fue de sólo 0,12 euros. Sorprendentemente, FB no nos cobró nada en tres de las campañas de las campañas que tuvieron éxito y que proporcionaron sólo una impresión de anuncio al usuario objetivo. Por lo tanto, más que un factor desalentador, el coste extremadamente bajo puede animar a los atacantes a aprovechar esta práctica.

Tiempo hasta la primera impresión (TFI). Los resultados exponen una amplia variabilidad del TFI, que oscila entre 44m y 32h10m en las 9 campañas realizadas con éxito. En particular, 3 de estas campañas muestran un TFI inferior a una hora, mientras que 3 de ellas presentan un TFI superior a 10h.

Clics. Mostramos el número de clics obtenidos en cada campaña publicitaria, así como el número de direcciones IP únicas (utilizando la información de la dirección IP seudonimizada que almacenamos) que generan esos clics. Este último representa un límite superior del número de usuarios que hicieron clic en el anuncio. Como sólo almacenamos la dirección IP pública (seudonimizada) del dispositivo que genera el clic, no podemos validar si el mismo usuario recibe nuestro anuncio varias veces y hace clic en él desde diferentes direcciones IP (por ejemplo, diferentes dispositivos, el mismo dispositivo móvil conectado a diferentes puntos de acceso, etc.). Por ejemplo, el usuario 3 hizo clic en las 5 impresiones de anuncios de la campaña de 20 intereses desde 3 direcciones IP diferentes. Como era de esperar, las campañas de publicidada nanodirigida que tuvieron éxito sólo recibieron clics de los usuarios a los que iban dirigidas, ya que el número de clics e impresiones coinciden uno a uno.

Resumen de los resultados de los experimentos de publicidad nanodirigida.
Las principales conclusiones derivadas de nuestro experimento son las siguientes (i) realizar una campaña de publicidad nanodirigida con éxito a un usuario en FB es altamente probable si un atacante puede inferir más de 18 intereses del usuario objetivo; (ii) la publicidad nanodirigida es extremadamente barata, y (iii) en base a nuestros experimentos, se espera que 2/3 de los anuncios de una campaña de publicidad nanodirigida sean entregados al usuario objetivo en menos de 7 horas efectivas de campaña.

6 EXTENSIÓN DEL FDVT PARA ELIMINAR INTERESES ESPECÍFICOS

Hemos desarrollado una solución que muestra una lista de los intereses que Facebook ha asignado a un usuario ordenados en función del tamaño de su audiencia, de menor

a mayor valor. Esta solución: (i) informa a los usuarios de que algunos de los intereses de su conjunto pueden ser demasiado específicos y pueden ser utilizados para prácticas abusivas de privacidad inapropiadas, como el uso de publicidad nanodirigida, (ii) permite a los usuarios eliminar fácilmente cualquiera de los intereses de la lista con sólo pulsar un botón. Por lo tanto, la solución ofrece un mecanismo sencillo y guiado para que los usuarios puedan eliminar los intereses menos populares de su lista para proteger su privacidad.

Hemos implementado esta solución como una nueva característica en la extensión de navegador FDVT que utilizamos para recoger el conjunto de datos utilizados para nuestro análisis de unicidad en la Sección 4. Para obtener el tamaño de la audiencia de los intereses asignados al usuario, cada vez que un usuario inicia una sesión en FB, la extensión del navegador FDVT obtiene su conjunto actualizado de preferencias publicitarias (es decir, los intereses) y el tamaño de la audiencia de cada interés de la API de FB Ads Manager. A partir de esta información, la extensión del navegador calcula una lista ordenada de los intereses asignados al usuario de menor a mayor popularidad. La interfaz gráfica de la extensión del navegador FDVT añade un nuevo botón con la etiqueta *"Riesgos de mis intereses en FB"*. Cuando un usuario hace clic en ese botón, la extensión muestra una página web con la lista de intereses ordenada. Hemos definido un código de colores para facilitar a los usuarios la comprensión de qué intereses pueden suponer un riesgo importante para la privacidad en función del tamaño de la audiencia asociada. Utilizamos la siguiente clasificación: Rojo (riesgo alto) para tamaños de audiencia mundiales $\leq 10k$ usuarios; Naranja (riesgo medio) para tamaños de audiencia entre 10k y 100k usuarios; Amarillo (riesgo bajo) para tamaños de audiencia entre 100k y 1M usuarios; Verde (sin riesgo) para tamaños de audiencia $\geq 1M$ usuarios. Hay que tener en cuenta que el umbral para cada categoría de riesgo puede modificarse fácilmente si otros trabajos científicos o expertos recomiendan utilizar valores diferentes.

Por último, la información que muestra esta nueva funcionalidad de la extensión del navegador FDVT es: (i) Nombre del interés, (ii) Nivel de riesgo (basado en el código de colores descrito), (iii) Tamaño de la audiencia, (iv) Botón de eliminar, que permite borrar el interés asociado del perfil del usuario, (v) Botón de más información, que muestra información histórica y la razón por la que ese interés aparece/aparece en el perfil del usuario, y (vi) Estado, ya sea activo (actualmente en el conjunto de preferencias de anuncios del usuario) o inactivo. La Figura 7 muestra una instantánea de la solución descrita.

7 TRABAJOS RELACIONADOS

En esta sección, presentamos la bibliografía más relevante para nuestro trabajo en el contexto de los análisis de unicidad basados en datos que individualmente no permiten identificar a un usuario y en experimentos de publicidad dirigida. Por simplicidad, en esta sección nos referiremos por las siglas en inglés PII (Personal Identifiable Information) a los elementos que individualmente permiten identifica a un usuario, y no-PII (non-Personal Identifiable Information) a los elementos que individualmente no permiten identificar a un usuario.

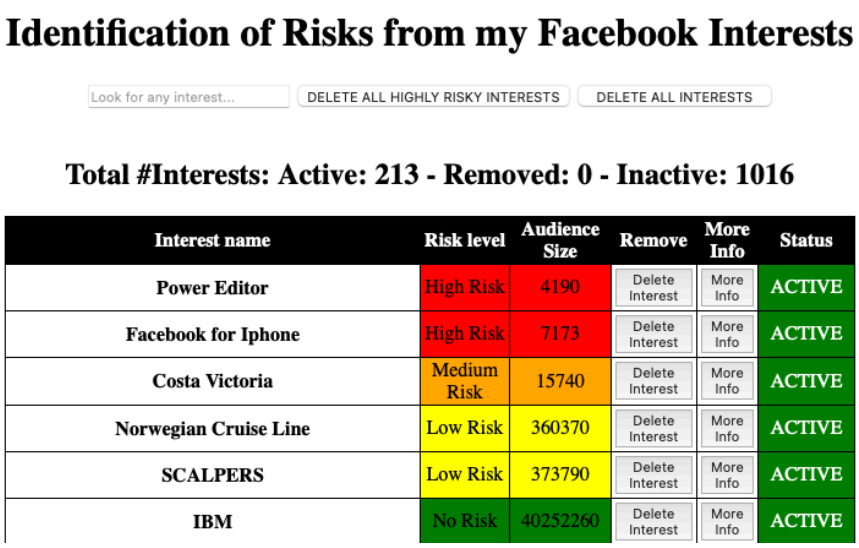

Fig. 7. Interfaz de la nueva funcionalidad de la extensión del navegador el FDVT. Informa del posible riesgo para la privacidad asociado a cada interés de FB mediante un código de colores. También permite a los usuarios eliminar cualquier interés con un clic.

7.1 Unicidad basada en elementos no-PII

Existe un cuerpo de literatura que ha explorado el número de elementos no-PII requeridos para identificar de forma única a una persona dentro de una gran base de usuarios. Sweeney [34] informó de que el acceso al género, el código postal y la fecha de nacimiento de los usuarios permitía revelar la identidad del 87% de los ciudadanos dentro de los datos del censo de Estados Unidos de 1990, que incluía a 248M de personas. Más recientemente, Golle et al. [19] reprodujeron el análisis de Sweeney utilizando el censo estadounidense del 2000, que incluía 281M de personas. Los resultados muestran un descenso del 87% al 63% en un periodo de 10 años. De Montjoye et al. [3] expusieron que conocer la hora y la ubicación asociadas a sólo cuatro llamadas de móvil era suficiente para identificar de forma exclusiva al 95% de los individuos en un conjunto de datos que incluía 1,5M de usuarios. Del mismo modo, De Montjoye et al. [4] estudiaron 3 meses de registros de tarjetas de crédito de 1,1 millones de personas y revelaron que cuatro puntos espacio-temporales de las compras con tarjeta de crédito son suficientes para identificar de forma exclusiva al 90% de los individuos de dicho conjunto de datos. Su et al. [33] demostraron la posibilidad de identificar de forma única a un usuario de Twitter basándose en su historial de navegación. Realizaron un experimento utilizando el historial de navegación de usuarios reales y lograron desanonimizar 268 de 374 cuentas reales de Twitter (72%). Por último, Narayanan et al. [31] trataron de desanonimizar el conjunto de datos Netflix Prize [1] que incluía más de 100M de valoraciones de películas de 480k suscriptores de Netflix entre diciembre de 1999 y diciembre de 2005. Los autores demostraron que 8 valoraciones de películas junto con sus fechas (que pueden tener un error de 14

días) son suficientes para identificar de forma única al 99% de los usuarios del conjunto de datos. Además, utilizaron una pequeña muestra de 50 usuarios (conocidos) de la base de datos de películas de Internet (IMDb) [25] que habían valorado públicamente películas y pudieron identificar a dos de ellos en la base de datos de Netflix. Esto demuestra que la información obtenida de un sistema online A puede utilizarse para desvelar la identidad de los usuarios de un sistema online B.

Nuestro trabajo contribuye a esta literatura de varias maneras:

(i) Nuestro estudio es el primero que analiza la unicidad dentro de una base de usuarios a escala de la población mundial. La base de usuarios de nuestro estudio representa alrededor de 1/5 de la población mundial. Los trabajos anteriores han utilizado conjuntos de datos de empresas privadas que incluyen como máximo 1.5M o el censo de Estados Unidos con hasta 281M de personas. Sin embargo, la capacidad de reidentificación en el conjunto de datos del censo de EE.UU. de 2000 (es decir, el análisis más reciente) es significativamente menor que en nuestro estudio.

(ii) Todos los trabajos anteriores se basan en la información de localización y/o información temporal de los usuarios. En cambio, nuestro estudio considera un tipo de información no-PII completamente diferente, representada por los intereses de los usuarios en las redes sociales.

(iii) Los elementos de información utilizados en nuestro estudio, es decir, los intereses de los usuarios, pueden utilizarse directamente para definir campañas publicitarias en plataformas como FB con tasas de reidentificación muy elevadas (el 90%). En cambio, los trabajos anteriores o bien se basan en información de empresas privadas (registros de llamadas o transacciones con tarjetas de crédito) que no es directamente procesable para llegar hasta un individuo con mensajes dirigidos, o bien alcanzan una tasa de reidentificación bastante baja (63% en el estudio del censo de EE.UU. de 2000).

7.2 Publicidad nanodirigida

Investigadores y profesionales han explorado la posibilidad de implementar la publicidad nanodirigida en FB. La literatura existente puede clasificarse en dos grupos según las herramientas de FB utilizadas para implementar campañas de publicidad nanodirigida. Por un lado, y de forma similar a nuestro enfoque, un par de estudios preliminares abordan la implementación de campañas de publicidad nanodirigida utilizando el Administrador de Anuncios de FB y datos no-PII. Por otro lado, encontramos varios trabajos que proponen campañas de publicidad nanodirigida utilizando la herramienta Audiencias Personalizadas de FB que requiere datos PII (por ejemplo, correo electrónico, número de teléfono móvil, etc.). A continuación, analizamos en detalle estos dos grupos de trabajos.

7.2.1 Publicidad nanordirigida basada en elementos no-PII

Dave Kerpen [26] explica en un libro cómo en 2009 realizó un experimento intentando llegar a su mujer mediante un anuncio en Facebook. Para ello, configuró una campaña con los siguientes parámetros *<de 31 años, casada, mujer, empleada de Likeable Media, que vive en la ciudad de Nueva York>*. El anuncio incluía el mensaje ”Te quiero y te echo de menos Carrie. Vuelve pronto de Texas”. El anuncio llegó a la esposa de Kerpen.

En 2010, Korolova et al. [27] aprovecharon el Administrador de Anuncios de FB para lanzar anuncios nanodirigidos a dos personas concretas. En primer lugar, eligieron a una amiga y utilizaron el sexo (femenino), el lugar de trabajo y la universidad a la que asistía esta persona para configurar una campaña de anuncios de Facebook dirigida a ella. Los autores sabían de antemano que estos parámetros sólo podían identificar a la persona a la que querían dirigirse porque era la única persona de su lugar de trabajo que había asistido a la universidad referida. Además, sabían que la usuaria a la que se dirigía había introducido la información sobre su género, la universidad y el trabajo en su perfil de FB. Como era de esperar, el anuncio dirigido se entregó exclusivamente a la amiga referida. Repitieron el experimento con un segundo individuo, pero esta vez obtuvieron información de su perfil público de FB. En concreto, lanzaron una campaña utilizando su género, edad, ubicación y algunos intereses. Una vez más, consiguieron entregar el anuncio sólo al usuario objetivo. El objetivo final de este trabajo no era la publicidad nanodirigida a los usuarios, sino demostrar que si se puede identificar de forma exclusiva a un individuo utilizando una configuración de audiencia determinada, se puede utilizar el ecosistema publicitario de FB para desvelar otra información personal de ese usuario. Por ejemplo, revelaron la edad de la mujer a la que se dirigieron en el primer experimento. Para obtener la edad, ampliaron la definición original de la audiencia añadiendo un valor de edad. Lanzaron múltiples campañas utilizando en cada una de ellas un valor de edad diferente. Entre estas campañas, sólo una entregó impresiones (el resto de las campañas no entregaron ni una sola impresión). La edad utilizada en esta campaña revelaba la edad del usuario objetivo, que efectivamente se validaba como la edad real de la persona en su perfil de FB. En resumen, este trabajo reveló una vulnerabilidad de privacidad del ecosistema publicitario de FB. Según los autores, tras su trabajo, FB actualizó su plataforma publicitaria y no permitió realizar campañas para las que el tamaño real de la audiencia era inferior a 20. Nuestros resultados sugieren que este límite no existe en la actualidad, ya que hemos sido capaces de de realizar camapñas de nanotargeting exitosas a usuarios en FB.

Aunque estos trabajos preliminares ofrecen ejemplos sobre la posibilidad de realizar camapañas nanodirigidas de usuarios en FB con datos no-PII, sólo presentan experimentos ad-hoc que pueden considerarse pruebas anecdóticas. En cambio, nuestro trabajo ofrece la primera formulación sistemática que proporciona directrices claras y específicas para realizar ataques de publicidad nanodirigida con éxito.

7.2.2 Publicidad nanodirigda basada en elementos PII

Además del panel de control habitual de Administrador de Anuncios de FB, FB ofrece a los anunciantes una herramienta alternativa para configurar las campañas publicitarias denominada Audiencias Personalizadas [11]. Tal y como describimos en la Sección 2, un anunciante puede definir una audiencia personalizado utilizando una lista de elementos PII como correos electrónicos, teléfonos móviles, etc. Facebook utiliza esa lista para identificar a los usuarios que están registrados en la plataforma utilizando el elemento PII proporcionado. Los anunciantes pueden crear audiencias personalizados combinando la lista de PII con otros parámetros de personalización. Por ejemplo, pueden crear una audiencia que incluya a los usuarios de la lista PII que sean hombres, o a los usuarios de

la lista PII que estén interesados en el fútbol. La función de Audiencias Personalizadas se ha utilizado en múltiples ocasiones para crear campañas de publicidad nanodirigida. Facebook intentó aumentar las garantías de privacidad de los usuarios estableciendo un tamaño mínimo de 20 usuarios verificados en una lista de Audiencia Personalizada. Este límite se incrementó posteriormente hasta 100, que es el límite actual. Ninguno de estos límites fue suficiente para evitar la publicidad nanodirigda.

Encontramos varios ejemplos en la literatura en los que la herramienta de Audiencia Personalizada se combina con otros criterios de segmentación para mostrar un anuncio exclusivamente a un individuo [35][21][22][37][23][8]. Por ejemplo, en [21] el objetivo era dirigirse a un usuario masculino específico. Para ello, los autores utilizaron una lista de correos electrónicos pertenecientes exclusivamente a mujeres, excepto uno que pertenecía al usuario masculino objetivo. Configuraron una Audiencia Personalizada que tenía como objetivo los hombres dentro de la lista de usuarios proporcionada. De este modo, se garantizaba que el anuncio se entregaría exclusivamente al usuario masculino objetivo. Del mismo modo, Korolova et al. [9] utilizaron la función de Audiencia Personalizada para demostrar que era posible entregar un anuncio a un usuario específico. Pasaron por alto el umbral de Audiencias Personalizada (20 en el momento en que se publicó este trabajo) incluyendo en la lista 19 cuentas de FB no alcanzables (por ejemplo, usuarios que tienen un bloqueador de anuncios instalado o que no son activos en FB) y sólo 1 cuenta activa. De este modo, se podía garantizar la entrega del anuncio exclusivamente al usuario objetivo. Los autores propusieron que el límite de tamaño de la audiencia personalizada se aumentara a 1000, pero como ya hemos mencionado ese límite es actualmente de 100.

En resumen, algunos trabajos demuestran que las Audiencias Personalizadas de FB pueden ser explotadas para implementar publicidad nanodirigida. Sin embargo, requieren conocer y utilizar elementos PII del usuario objetivo. En cambio, nuestra investigación pretende demostrar que la publicidad nanodirigida puede implementarse de forma sistemática utilizando datos no-PII.

8 DISCUSIÓN

En esta sección, primero discutimos los riesgos potenciales de la publicidad nanodirigida. A continuación, describimos cuáles son las medidas actuales que aplica FB y por qué son ineficaces. Por último, proponemos varias contramedidas que pueden ser fácilmente adoptadas por FB (y otros participantes del ecosistema publicitario) para evitar eficazmente la publicidad nanodirigida.

8.1 Riesgos asociados a la publicidad nanodirigida

Existe un cuerpo de literatura denominado *persuasión psicológica* que demuestra que es más fácil persuadir a un individuo si se crean mensajes adaptados a las características psicológicas y a las motivaciones de esa persona[2][38][29][24][5][28]. Algunos estudios han demostrado que los anuncios personalizados tienen una capacidad de atracción mucho mayor, lo que lleva, por ejemplo, a un aumento de la tasa de clics (CTR) de hasta el 670% [30]. En el contexto de FB, Matz et. al [28] realizaron un experimento con 3,5

millones de usuarios de FB utilizando anuncios personalizados junto con técnicas de comunicación de persuasión psicológica. Los autores informan de que fueron capaces de aumentar el número de clics hasta un 40% y las compras hasta un 50% en comparación con las campañas no personalizadas. Además, hemos encontrado algunas historias en la web que explican cómo se utilizó el ecosistema de anuncios de FB para persuadir a una persona concreta para que realizara una acción. Como ejemplo, Michael Harf [21] explica cómo utilizó la herramienta Custom Audience de FB para ofrecer anuncios nanodirigidos con el fin de persuadir a un cliente potencial, que había expresado previamente su interés en cambiar de agencia digital, para que se uniera a la agencia de Harf.

Hay otra historia interesante que, aunque no utiliza explícitamente la publicidad nanodirigida tal como se define en nuestro documento, es válida para ilustrar algunos otros riesgos potenciales asociados a ésta [22][37]. En la campaña de 2017 en el Reino Unido, el líder del partido laborista, Jeremy Corbyn, quiso invertir fuertemente en anuncios digitales que fomentaran el registro de votantes. Sin embargo, los jefes de la campaña del Partido Laborista pensaron que era una mala idea. Para contentar a Corbyn y, al mismo tiempo, gastar el dinero de la campaña en otros objetivos, los jefes de campaña invirtieron 5.000 £en una campaña de Facebook que explotaba la herramienta Custom Audience para llegar sólo a Corbyn, a sus colaboradores y a unos pocos periodistas afines. De este modo, Corbyn estaba convencido de que la campaña se llevaba a cabo siguiendo sus instrucciones.

Todos los ejemplos anteriores ilustran claramente los riesgos potenciales de la publicidad nanodirigida. En primer lugar, puede utilizarse eficazmente para manipular a un usuario con el fin de persuadirle de que compre un producto o convencerle de que cambie de opinión respecto a un tema concreto. Además, la publicidad nanodirigida podría utilizarse para crear una percepción falsa en la que se exponga al usuario a una realidad que difiera de lo que ven el resto de los usuarios (como ocurrió en el caso de Corbyn). Por último, ésta podría ser explotada para llevar a cabo otras prácticas dañinas como el chantaje.

Cualquiera de las prácticas presentadas representa una manipulación muy preocupante de los seres humanos. Para llevar a cabo dicha manipulación, los atacantes pueden aprovechar plataformas como FB que les permiten ofrecer anuncios exclusivamente a un usuario objetivo. Esto representa una vulnerabilidad para la privacidad de los usuarios de FB que urge a que FB adopte e implemente contramedidas eficientes.

8.2 Actuales medidas (ineficientes) contra la publicidad nanodirigida

La contramedida más importante que aplica Facebook para evitar que los anunciantes se dirijan a públicos muy reducidos son los límites impuestos al número mínimo de usuarios que pueden formar un público. Sin embargo, se ha demostrado que esos límites son completamente ineficaces. Por un lado, Korolova et. al [27] afirman que, motivados por los resultados de su trabajo, Facebook desautorizó la configuración de públicos de tamaño inferior a 20 mediante el Administrador de Anuncios de FB. Nuestra investigación muestra que este límite no se aplica actualmente. Por otra parte, FB impone un tamaño mínimo de Audiencias Personalizadas de 100 usuarios. Como se presenta en la sección 7.2.2,

varios trabajos en la literatura mostraron diferentes formas de superar este límite e implementar campañas publicitarias nanodirigidas utilizando la herramienta Audiencias Personalizadas.

Es relevante mencionar que, en el proceso de configuración de una de las 21 campañas utilizadas en nuestro experimento de publicidad nanodirigida, Facebook advirtió que nuestro público era demasiado estrecho y nos recomendó ampliarlo para poder ejecutar la campaña asociada. Sin embargo, sólo tuvimos que sustituir un interés en la lista de intereses y la advertencia desapareció. Efectivamente, la campaña referida consiguió llegar exclusivamente al usuario asociado. Hemos buscado en la documentación pública de Facebook y no hemos podido encontrar ningún límite especificado oficialmente para el tamaño mínimo de la audiencia asociada a una campaña publicitaria.

Por último, cabe mencionar que unos días después de que terminara nuestro experimento de publicidad nanodirigida, Facebook cerró la cuenta que utilizamos para realizar las campañas publicitarias. También utilizamos esa misma cuenta recientemente en otros trabajos de investigación que consultaron intensamente la API de FB Ads Manager. Facebook no nos dio ninguna explicación sobre los motivos que llevaron a la eliminación de la cuenta, por lo que no podemos confirmar si se debió al experimento de publicidad nanodirigida o no.

Incluso si asumimos que la cuenta fue eliminada debido a nuestro experimento de publicidad nanodirigida, esto sólo ocurrió días después de que la última campaña hubiera terminado. Esto representaría una medida reactiva, la cual es ineficiente, ya que no nos impidió realizar con éxito nuestras campañas de publicidad nanodirigida y llegar a los usuarios solicitados varias veces.

8.3 Medidas eficientes contra la publicidad nanodirigida

La publicidad nanodirigida basada en los intereses de los usuarios podría evitarse aplicando una actualización muy sencilla en la plataforma publicitaria de FB. Facebook debería reducir el número máximo de intereses permitidos en la definición de un público del límite actual (25) a menos de 9. Nuestro análisis en la Sección 4 indica que esto reduciría drásticamente la posibilidad de ejecutar eficazmente campañas publicitarias nanodirigidas. Además, hemos consultado a TAPTAP Digital [36], una empresa que gestiona Sonata [32], un DSP de tamaño medio. Nos han confirmado que es muy extraño configurar audiencias dirigidas con más de 9 intereses. De hecho, menos del 1% de las campañas tiene una configuración de este tipo.[3] Esto sugiere que se espera que la contramedida propuesta tenga un impacto muy limitado en los ingresos de FB.

[3]Un experimento para demostrar que esta afirmación es correcta requiere acceder a la configuración de cientos o miles de campañas publicitarias. La configuración de las campañas publicitarias es un activo muy sensible para los anunciantes, las agencias de marketing y otros agentes del mercado de la publicidad online, por lo que no podríamos acceder a esa información. En este contexto, al menos pudimos obtener una confirmación informal de dos expertos en AdTech.

La medida propuesta es eficaz para proteger a los usuarios de la publicidad nanodirigida basada en los intereses, sin embargo, no funciona para evitarla a través de la herramienta Facebook Custom Audience. Por lo tanto, para este asunto, proponemos una segunda medida (simple) que evitaría cualquier tipo de publicidad nanodirigida. FB no debería permitir la ejecución de ninguna campaña publicitaria cuyo tamaño de público objetivo sea inferior a un límite determinado de usuarios activos. Es importante señalar que sólo los usuarios activos (por ejemplo, en el último mes) deberían contar para calcular el tamaño de la audiencia.

El límite referido no debe ser inferior a 100 y nuestra recomendación es fijarlo en 1000. Esta solución invalidaría trucos como el citado anteriormente en el que un Custom Audience estaba integrado por un solo hombre, y la campaña publicitaria estaba configurada para dirigirse a hombres dentro de la lista de Custom Audience. Si nuestra solución estuviera en marcha, FB identificaría que el tamaño de la audiencia activa para dicha campaña es realmente uno y, como resultado, la campaña no sería aceptada.

En resumen, soluciones muy básicas como las descritas anteriormente proporcionarían una fuerte protección contra las prácticas de publicidad nanodirigida.

9 CONCLUSIÓN

Este trabajo presenta dos contribuciones fundamentales. En primer lugar, proporcionamos una metodología analítica para estudiar el número de elementos no PII, es decir, los intereses, que hacen único a un usuario en una base de usuarios que incluye 1,5 mil millones de individuos registrados en FB. Se trata del primer análisis de la unicidad de un usuario en una base de usuarios de la misma escala que la población mundial. Nuestros resultados indican que los 4 intereses más raros de un usuario en FB lo hacen único en la mencionada base de usuarios con una probabilidad del 90%. Si, en cambio, consideramos una selección aleatoria de intereses, se necesitarían 22 intereses para hacer único a un usuario con una probabilidad del 90%.

En segundo lugar, dado que los intereses de los usuarios son operables en FB para configurar campañas publicitarias dirigidas, aprovechamos los resultados de nuestro análisis de la unicidad de los usuarios para realizar experimentos reales en el último trimestre de 2020 con el fin de implementar campañas publicitarias nanodirigidas en FB, es decir, campañas que lleguen exclusivamente al usuario seleccionado. Nuestros experimentos demuestran que es posible hacer publicidad nanodirigida sistemáticamente a un usuario en FB en función de sus intereses. También proponemos medidas para evitar ataques de publicidad nanodirigida potencialmente dañinos que exploten la plataforma publicitaria de FB.

Por último, cabe señalar que nuestro trabajo sólo ha revelado la punta del iceberg en lo que respecta a la utilización de datos no PII con fines de nanosegmentación. Nuestro trabajo se basa exclusivamente en los intereses de los usuarios, pero un anunciante puede utilizar otros parámetros sociodemográficos disponibles para configurar audiencias en el FB Ads Manager, como la ubicación del domicilio (país, ciudad, código postal, etc.), el lugar de trabajo, la universidad, el número de hijos, el dispositivo móvil utilizado (iOS,

Android), etc., para reducir rápidamente el tamaño de la audiencia para hacer llegar publicidad nanodirigida a un usuario. Por lo tanto, la combinación de los parámetros sociodemográficos con los intereses puede implicar que el número de elementos no PII necesarios para implementar con éxito un ataque de publicidad nanodirigida sea inferior a lo que hemos informado en este documento. Tenemos previsto abordar esta cuestión en nuestro futuro trabajo y queremos estudiar la singularidad de los usuarios, así como la probabilidad de llevar a cabo con éxito ataques de publicidad nanodirigida en FB cuando se considera una combinación de parámetros sociodemográficos e intereses en la configuración de las audiencias.

AGRADECIMIENTOS

Esta investigación ha recibido financiación del programa de acción de innovación Horizonte 2020 de la Unión Europea en el marco del proyecto PIMCITY (Grant 871370) y del proyecto TESTABLE (Grant 101019206); del Ministerio de Economía, Industria y Competitividad, España, y del Fondo Social Europeo(UE), en el marco del programa Ramón y Cajal (Grant RyC-2015-17732); el Ministerio de Educación, Cultura y Deporte, España, a través del programa FPU (Subvención FPU16/05852); la Agencia Estatal de Investigación (AEI) en el marco del proyecto ACHILLES (Subvención PID2019-104207RB-I00/AEI/10.13039/501100011033); el proyecto sinérgico de la Comunidad de Madrid EMPATIA-CM (Subvención Y2018/TCS-5046); la Fundación BBVA en el marco del proyecto AERIS; y el Vienna Science and Technology Fund a través del proyecto “Emotional Well-Being in the Digital Society” (Grant VRG16-005).

REFERENCIAS

[1] J. Bennett and S. Lanning. 2007. The Netflix Prize. In *Proceedings of the KDD Cup Workshop 2007*. ACM, New York, 3--6. http://www.cs.uic.edu/~liub/KDD-cup-2007/NetflixPrize-description.pdf

[2] Joseph Cesario, E. Tory Higgins, and Abigail A. Scholer. 2008. Regulatory Fit and Persuasion: Basic Principles and Remaining Questions. *Social and Personality Psychology Compass* 2, 1 (2008), 444--463. https://doi.org/10.1111/j.1751-9004.2007.00055.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1751-9004.2007.00055.x

[3] Yves-Alexandre de Montjoye, César A Hidalgo, Michel Verleysen, and Vincent D Blondel. 2013. Unique in the Crowd: The privacy bounds of human mobility. *Scientific reports* 3, 1 (2013), 1376. https://www.nature.com/articles/srep01376

[4] Yves-Alexandre De Montjoye, Laura Radaelli, Vivek Kumar Singh, et al. 2015. Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science* 347, 6221 (2015), 536--539. https://science.sciencemag.org/content/347/6221/536

[5] David Dubois, Derek D. Rucker, and Adam D. Galinsky. 2016. Dynamics of Communicator and Audience Power: The Persuasiveness of Competence versus Warmth. *Journal of Consumer Research* 43, 1 (Feb. 2016), 68--85. https://doi.org/10.1093/jcr/ucw006 arXiv:https://academic.oup.com/jcr/article-pdf/43/1/68/7049938/ucw006.pdf

[6] Erik H Erikson and Joan M Erikson. 1998. *The life cycle completed (extended version)*. WW Norton & Company.
[7] EU. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). European Union. http://eur-lex.europa.eu/eli/reg/2016/679/oj accessed on 21 September, 2021.
[8] Marc Faddoul, Rohan Kapuria, and Lily Lin. 2019. SNIPER AD TARGETING. *Berkeley School of Information* (May 2019). https://www.ischool.berkeley.edu/projects/2019/sniper-ad-targeting
[9] Irfan Faizullabhoy and Aleksandra Korolova. 2018. Facebook's Advertising Platform: New Attack Vectors and the Need for Interventions. In *Workshop on Technology and Consumer Protection (ConPro 2018)*. arXiv:1803.10099 [cs.CY] https://arxiv.org/abs/1803.10099
[10] FB. 2017. Facebook Reports Fourth Quarter and Full Year 2016 Results. Facebook Inc.. https://investor.fb.com/investor-news/press-release-details/2017/Facebook-Reports-Fourth-Quarter-and-Full-Year-2016-Results/default.aspx accessed on 21 September, 2021.
[11] FB. 2021. About Custom Audiences. Facebook Inc.. https://www.facebook.com/business/help/744354708981227 accessed 11 July, 2021.
[12] FB. 2021. Facebook Ad Preferences. Facebook Inc.. https://www.facebook.com/adpreferences/ accessed on 21 September, 2021.
[13] FB. 2021. Facebook Ads Manager. Facebook Inc.. https://www.facebook.com/ads/manager accessed on 21 September, 2021.
[14] FB. 2021. Facebook Reports Fourth Quarter and Full Year 2020 Results. Facebook Inc.. https://investor.fb.com/investor-news/press-release-details/2021/Facebook-Reports-Fourth-Quarter-and-Full-Year-2020-Results/default.aspx accessed on 21 September, 2021.
[15] FDVT. 2016. FDVT: Privacy Agreement. https://www.fdvt.org/privacy_agreement.html accessed on 21 September, 2021.
[16] FDVT. 2016. FDVT: Terms of Use. https://www.fdvt.org/terms_of_use accessed on 21 September, 2021.
[17] FDVT. 2021. FDVT: Data Valuation Tool for Facebook™ Users Website. https://fdvt.org/ accessed on 21 September, 2021.
[18] Cloé Gendronneau, Dilek Yıldız, Yuan Hsiao, Martin Stepanek, Guy Abel, Stijn Hoorens, Arkadiusz Wiśniowski, Emilio Zagheni, Lee Fiorio, and Ingmar Weber. 2019. Measuring Labour Mobility and Migration Using Big Data - Exploring the potential of social-media data for measuring EU mobility flows and stocks of EU movers. https://ec.europa.eu/social/BlobServlet?docId=22084
[19] Philippe Golle. 2006. Revisiting the Uniqueness of Simple Demographics in the US Population. In *Proceedings of the 5th ACM Workshop on Privacy in Electronic Society* (Alexandria, Virginia, USA) *(WPES '06)*. ACM, New York, NY, USA, 77--80. https://

doi.org/10.1145/1179601.1179615
[20] José González Cabañas, Ángel Cuevas, and Rubén Cuevas. 2017. *FDVT: Data Valuation Tool for Facebook Users*. Association for Computing Machinery, New York, NY, USA, 3799–3809. https://doi.org/10.1145/3025453.3025903
[21] Michael Harf. 2017. Sniper Targeting on Facebook: How to Target ONE specific person with super targeted ads. Medium. https://medium.com/@MichaelH_3009/sniper-targeting-on-facebook-how-to-target-one-specific-person-with-super-targeted-ads-515ba6e068f6 accessed on 21 September, 2021.
[22] Caroline Haskins. 2018. Facebook ad micro-targeting can manipulate individual politicians. The Outline. https://theoutline.com/post/5411/facebook-ad-micro-targeting-can-manipulate-individual-politicians accessed on 21 September, 2021.
[23] Jonathan Hawkins. 2019. Facebook Ads Sniper Method: How to Put Your Ad in front of ONE Specific Person. Jonathan Hawkins. https://jonathanhawkinsofficial.com/blog/facebook-ads-sniper-method-how-to-put-your-ad-in-front-of-one-specific-person accessed on 21 September, 2021.
[24] Jacob B. Hirsh, Sonia K. Kang, and Galen V. Bodenhausen. 2012. Personalized Persuasion: Tailoring Persuasive Appeals to Recipients' Personality Traits. *Psychological Science* 23, 6 (2012), 578--581. https://doi.org/10.1177/0956797611436349 arXiv:https://doi.org/10.1177/0956797611436349 PMID: 22547658.
[25] IMDb. 2021. The Internet Movie Database. https://www.imdb.com/ accessed on 21 September, 2021.
[26] Dave Kerpen. 2011. *Likeable social media : how to delight your customers, create an irresistible brand, and be generally amazing on Facebook (and other social networks)*. McGraw-Hill.
[27] Aleksandra Korolova. 2010. Privacy Violations Using Microtargeted Ads: A Case Study. In *ICDMW 2010, The 10th IEEE International Conference on Data Mining Workshops*, Wei Fan, Wynne Hsu, Geoffrey I. Webb, Bing Liu, Chengqi Zhang, Dimitrios Gunopulos, and Xindong Wu (Eds.). IEEE Computer Society, Sydney, Australia, 474--482. https://doi.org/10.1109/ICDMW.2010.137
[28] S. C. Matz, M. Kosinski, G. Nave, and D. J. Stillwell. 2017. Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences* 114, 48 (2017), 12714--12719. https://doi.org/10.1073/pnas.1710966114 arXiv:https://www.pnas.org/content/114/48/12714.full.pdf
[29] Youngme Moon. 2002. Personalization and Personality: Some Effects of Customizing Message Style Based on Consumer Personality. *Journal of Consumer Psychology* 12, 4 (2002), 313--325. https://doi.org/10.1016/S1057-7408(16)30083-3 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1016/S1057-7408%2816%2930083-3
[30] J. Mullock, S. Groom, , and P. Lee. 2010. International online behavioural advertising survey 2010. Osborne Clarke.
[31] Arvind Narayanan and Vitaly Shmatikov. 2008. Robust De-Anonymization of Large Sparse Datasets. In *Proceedings of the 2008 IEEE Symposium on Security and*

Privacy (SP '08). IEEE Computer Society, USA, 111–125. https://doi.org/10.1109/SP.2008.33
[32] Sonata DSP. 2021. Global Platform for Mobile-Centric Audience Engagement. https://www.sonataplatform.com accessed on 21 September, 2021.
[33] Jessica Su, Ansh Shukla, Sharad Goel, and Arvind Narayanan. 2017. De-Anonymizing Web Browsing Data with Social Networks. In *Proceedings of the 26th International Conference on World Wide Web* (Perth, Australia) *(WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1261–1269. https://doi.org/10.1145/3038912.3052714
[34] Latanya Sweeney. 2000. Simple demographics often identify people uniquely. *Health (San Francisco)* 671, 2000 (2000), 1--34.
[35] Brian Swichkow. 2014. The Ultimate Retaliation: Pranking My Roommate With Targeted Facebook Ads. Ghost Influence. http://ghostinfluence.com/the-ultimate-retaliation-pranking-my-roommate-with-targeted-facebook-ads/ accessed on 21 September, 2021.
[36] TAPTAP Digital. 2021. Omnichannel advertising and marketing intelligence powered by location. https://www.taptapdigital.com/ accessed on 21 September, 2021.
[37] Political Editor Tim Shipman. 2018. Labour HQ used Facebook ads to deceive Jeremy Corbyn. The Sunday Times. https://www.thetimes.co.uk/article/labour-hq-used-facebook-ads-to-deceive-corbyn-3hvn0jzr8 accessed on 21 September, 2021.
[38] S. Wheeler, Richard Petty, and George Bizer. 2005. Self-Schema Matching and Attitude Change: Situational and Dispositional Determinants of Message Elaboration. *Journal of Consumer Research* 31 (March 2005), 787--797. https://doi.org/10.1086/426613

APÉNDICE

A BASE DE USUARIOS DE FB PARA EL ANÁLISIS DE LA UNICIDAD

En el momento en que recogimos los datos (enero de 2017), la API de Facebook no permitía definir una audiencia de todo el mundo. En su lugar, era obligatorio insertar una ubicación o grupo de ubicaciones para definir la cobertura geográfica de la audiencia definida. El número máximo de ubicaciones permitido en una campaña publicitaria era de 50. Para maximizar la base de usuarios de nuestro análisis de unicidad en FB, seleccionamos los 50 principales países en términos de usuarios activos. Estos 50 países incluían 1.500 millones de usuarios activos de FB, lo que correspondía al 81% del total de usuarios de Facebook. La Tabla 3 enumera los 50 países considerados junto con el número de usuarios de FB.

B DESGLOSE DE LA UBICACIÓN DE LOS USUARIOS

El único parámetro obligatorio para definir una audiencia en el Administrador de Anuncios de FB es una ubicación (por ejemplo, país, región, código postal, etc.). Esto significa que un público puede ser configurado por una sola ubicación, pero si se desea utilizar múltiples atributos al menos uno de ellos debe ser una ubicación. Basándose en esta restricción, en

code	country	users (M)	code	country	users (M)
US	Estados Unidos	203	DZ	Algeria	16
IN	India	161	NG	Nigeria	16
BR	Brasil	114	AU	Australia	15
ID	Indonesia	91	IQ	Irak	14
MX	México	70	PL	Polonia	14
PH	Filipinas	56	SA	Arabia Saudí	14
TR	Turquía	46	ZA	Sudáfrica	14
TH	Tailandia	42	MA	Marruecos	13
VN	Vietnam	42	VE	Venezuela	13
GB	Reino Unido	39	CL	Chile	12
EG	Egipto	33	MM	Myanmar	12
FR	Francia	33	RU	Rusia	12
DE	Alemania	30	NL	Países Bajos	10
IT	Italia	30	EC	Ecuador	9.80
AR	Argentina	29	RO	Rumanía	8.60
PK	Paquistán	28	AE	Emiratos Árabes	7.70
CO	Colombia	26	NP	Nepal	6.70
JP	Japón	26	BE	Bélgica	6.50
BD	Bangladesh	23	SE	Suecia	6.20
ES	España	23	TN	Túnez	6.10
CA	Canadá	22	KE	Kenia	6
MY	Malasia	20	PT	Portugal	5.90
PE	Perú	19	UA	Ucrania	5.90
KR	Corea del Sur	18	GT	Guatemala	5.50
TW	Taiwan	18	HU	Hungría	5.30

Tabla 3. Lista de los 50 países incluidos en nuestras consultas al FB Ads Manager y su número asociado de usuarios en millones.

el proceso de registro de la extensión del navegador FDVT, los usuarios debían rellenar obligatoriamente su ubicación (es decir, el país de residencia). De lo contrario, la extensión del navegador no podía recuperar ninguna información de la API de FB y, por tanto, no podía proporcionar a los usuarios los ingresos estimados que generan para FB. Nuestra base de 2.390 usuarios estaba formada por usuarios de 80 ubicaciones diferentes. La Tabla 4 muestra el número de usuarios por país.

C ANÁLISIS DEMOGRÁFICO

Una pregunta interesante es si el número de intereses que hacen que un usuario sea único en FB muestra diferencias significativas entre los distintos grupos demográficos. Para responder a esta pregunta, obtenemos el valor de $N(MP)_{0,9}$ y $N(A)_{0,9}$ a través de tres parámetros demográficos: género, edad y ubicación. Este análisis demográfico sólo pretende ilustrar que puede haber diferencias en la unicidad de los usuarios en función de los parámetros demográficos.

Hemos seleccionado P = 0,9 por dos razones: (i) revela el número de intereses que identifican de forma exclusiva a un usuario en FB con una probabilidad muy alta (0,9); (ii) $N(MP)_{0,9}$ y $N(A)_{0,9}$ son ambos inferiores a 25 (el número máximo de intereses que puede utilizarse para definir una audiencia en FB) y, por tanto, son operables en la práctica para realizar publicidad nanodirigida.

code	country	users	code	country	users
ES	España	1131	AU	Australia	2
FR	Francia	335	CY	Chipre	2
MX	México	122	DO	República Dominicana	2
AR	Argentina	115	GR	Grecia	2
EC	Ecuador	89	HK	Hong Kong SAR China	2
PE	Perú	78	ID	Indonesia	2
CA	Canadá	61	IE	Irlanda	2
CO	Colombia	48	LU	Luxemburgo	2
US	Estados Unidos	40	PL	Polonia	2
BE	Bélgica	36	RE	Reunión	2
UY	Uruguay	35	AL	Albania	1
GB	Reino Unido	26	AM	Armenia	1
CH	Suiza	24	AO	Angola	1
PT	Portugal	21	AX	Åland Islands	1
VE	Venezuela	18	BG	Bulgaria	1
SV	El Salvador	17	BT	Bután	1
CL	Chile	14	CI	Côte d'Ivoire	1
PY	Paraguay	13	CR	Costa Rica	1
DE	Alemania	11	CZ	República Checa	1
IT	Itala	11	DJ	Djibouti	1
BO	Bolivia	9	GI	Gibraltar	1
MA	Marruecos	8	GN	Guinea	1
BR	Brasil	6	IN	India	1
GT	Guatemala	6	IQ	Irak	1
HN	Honduras	6	LK	Sri Lanka	1
NI	Nicaragua	6	LT	Lituania	1
NL	Países Bajos	6	MG	Madagascar	1
PA	Panamá	6	MO	Macao SAR China	1
TN	Túnez	6	MU	Mauricio	1
BD	Bangladesh	5	NC	Nueva Caledonia	1
SE	Suecia	4	NP	Nepal	1
TH	Tailandia	4	NZ	Nueva Zelanda	1
AD	Andorra	3	PH	Filipinas	1
AT	Austria	3	PM	St. Pierre & Miquelon	1
DK	Dinamarca	3	PR	Puerto Rico	1
DZ	Algeria	3	RO	Rumanía	1
FI	Finlandia	3	RS	Serbia	1
PK	Paquistán	3	RU	Rusia	1
SN	Senegal	3	RW	Ruanda	1
AF	Afganistán	2	TW	Taiwan	1

Tabla 4. Desglose completo del número de usuarios por ubicación en nuestro conjunto de datos de 2.390 usuarios obtenidos con la extensión del navegador FDVT.

C.1 Análisis de género

Dividimos nuestro conjunto de datos en hombres (1.949 usuarios) y mujeres (347 usuarios) y calculamos $N(MP)_{0,9}$ y $N(A)_{0,9}$ para cada grupo. La Figura 8 muestra el resultado en forma de gráfico de barras. Obsérvese que presentamos el intervalo de confianza del 95% de nuestro modelo de ajuste en forma de barra de error en cada gráfico de barras.

Observamos que $N(MP)_{0,9}$ es casi el mismo para los hombres (4,16) y las mujeres (4,20), lo que indica que el número de intereses que hacen que un hombre o una mujer sean únicos dentro de una base de usuarios a escala de población mundial es similar y cercano a 4.

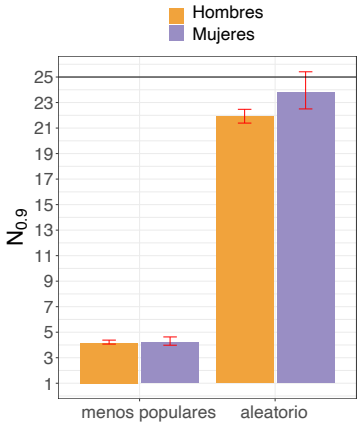

Fig. 8. Análisis de unicidad en función del género. $N(MP)_{0,9}$ (izquierda) y $N(A)_{0,9}$ (derecha) para hombres (amarillo) y mujeres (morado). La figura incluye el intervalo de confianza del 95% de los resultados.

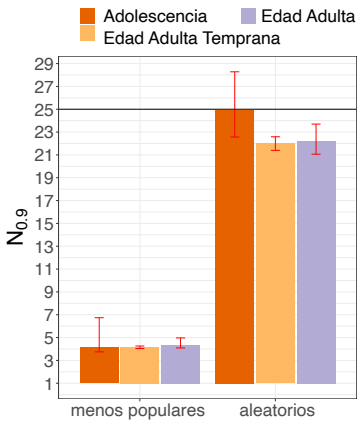

Fig. 9. Análisis de unicidad entre grupos de edad. $N(MP)_{0,9}$ (izquierda) y $N(A)_{0,9}$ (derecha) para los grupos de la adolescencia (naranja), la edad adulta temprana (amarillo) y la edad adulta (púrpura). La figura incluye el intervalo de confianza del 95% de los resultados.

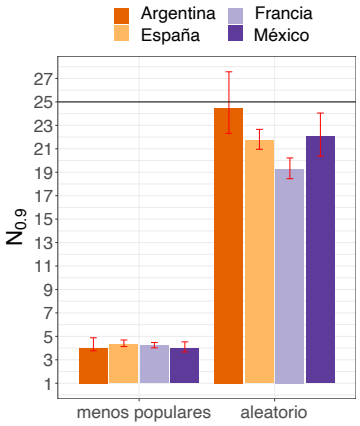

Fig. 10. Análisis de unicidad entre países. $N(MP)_{0,9}$ (izquierda) y $N(A)_{0,9}$ (derecha) para Argentina (naranja), España (amarillo), Francia (morado claro) y México (morado oscuro). La figura incluye el intervalo de confianza del 95% de los resultados.

$N(A)_{0,9}$ presenta una diferencia mayor, siendo de 23,80 para las mujeres y de 21,92 para los hombres. Este hallazgo indica que un atacante necesitaría inferir (aproximadamente) dos intereses más para hacer publicidad nanodirigida a una mujer en comparación con un hombre. Esto sugiere que los perfiles de intereses de las mujeres son ligeramente más privados que los de los hombres y, por tanto, son más difíciles de alcanzar con publicidad nanodirigida.

C.2 Análisis de edad

Dividimos a los usuarios de nuestro conjunto de datos en los siguientes grupos de edad basándonos en la división propuesta por Erikson et al. [6]: 13-19 (Adolescencia), 20-39 (Edad adulta temprana), 40-64 (Edad adulta) y 65+ (Mayores). El número de usuarios en los grupos de Adolescencia, Edad adulta temprana, Edad adulta y Mayores es de 117, 1.374, 578 y 19, respectivamente. Debido al bajo número de usuarios que forman el grupo de Mayores, decidimos excluirlo del análisis.

La Figura 9 muestra el valor de $N(MP)_{0,9}$ y $N(A)_{0,9}$ para los grupos de edad de Adolescencia, Edad Adulta temprana y Edad adulta junto con el intervalo de confianza del 95 de nuestro modelo.

Los valores de $N(MP)_{0,9}$ son muy similares en todos los grupos de edad considerados (4,11, 4,16 y 4,45 para los grupos de la adolescencia, la edad temprana y la edad adulta, respectivamente). Este resultado indica que la singularidad de un usuario en FB no parece estar correlacionada con su grupo de edad.

Si nos centramos ahora en los valores de $N(A)_{0,9}$, observamos que los usuarios de la edad adulta temprana y de la edad adulta pueden ser alcanzados con publicidad nanodirigida con una probabilidad de éxito del 90% con 22 intereses ($N(A)_{0,9}$ = 21,99 y 22,20 para la edad adulta temprana y la edad adulta, respectivamente). Para los usuarios en el grupo de Adolescencia es más difícil ya que requiere 25 intereses ($N(A)_{0.9}$ = 24.92).

C.3 Análisis de ubicación

Aunque nuestro conjunto de datos incluye usuarios de 80 países diferentes (véase la Tabla 4), la mayoría de ellos presentan un bajo número de usuarios. Por lo tanto, para obtener resultados significativos, seleccionamos aquellos países para los que tenemos más de 100 usuarios en nuestro conjunto de datos. Estos son: España (1131 usuarios), Francia (335), México (122) y Argentina (115).

Siguiendo el mismo análisis que realizamos para los grupos de género y edad, la figura muestra gráficos de barras que capturan los valores de $N(MP)_{0,9}$ y $N(A)_{0,9}$ para los países considerados junto con los intervalos de confianza del 95% proporcionados por el modelo de ajuste en forma de barras de error.

Al igual que en el caso del género y la edad, $N_{0.9}(MP)$ es muy similar para los cuatro países considerados (3,96; 4,03; 4,21 y 4,29 para México, Argentina, Francia y España, respectivamente), lo que confirma que ninguno de los parámetros demográficos considerados parece ser relevante para afectar a la unicidad del usuario en FB.

Los valores de $N_{0.9}(A)$ son 19,28; 21,7; 22,05 y 24,49 para Francia, España, México y Argentina, respectivamente. Esto indica que llevar a cabo una campaña publicitaria nanodirigida sería notablemente más fácil en Francia en comparación con Argentina, ya que un atacante necesitaría inferir 5 intereses menos en el primer país para realizar una campaña nanodirigida a un usuario con una probabilidad de éxito del 90%. Este resultado sugiere que la ubicación es un factor que puede ser relevante en el número de intereses requeridos para la unicidad de un usuario en FB.

En resumen, nuestro análisis demográfico revela que las mujeres, los adolescentes y los usuarios de Argentina (en comparación con los de Francia, España y México) están mejor protegidos de los ataques de publicidad nanodirigida basados en la selección aleatoria de intereses.

D PRUEBAS DE ANUNCIOS NANODIRIGIDOS

Cada campaña de publicidad nanodirigida de nuestro experimento utilizaba una creatividad publicitaria exclusiva, que incluía un texto para identificar al usuario al que se dirigía y el número de intereses utilizados en la campaña asociada. Con esta información, los autores a los que se dirigía el experimento podían identificar fácilmente los anuncios asociados a éste cuando los veían en su muro de FB.

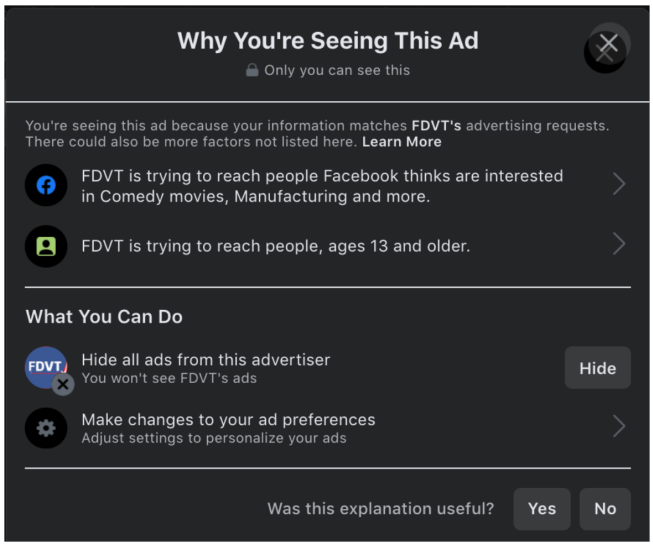

Fig. 11. Instantánea de la ventana ``*¿Por qué veo este anuncio?"* asociada a la impresión del anuncio de la campaña dirigida al usuario 3 con 12 intereses.

Además, cada creatividad publicitaria está vinculada a una página de destino diferente alojada en nuestro servidor web. Los autores fueron instruidos para hacer clic en los anuncios nanodirigidos cada vez que los vieran en el muro de FB. Esos clics se registraron en nuestro servidor web junto con su marca de tiempo.

Por último, se pidió a los usuarios que capturaran varias instantáneas asociadas a los anuncios recibidos de las campañas de publicidad nanodirigida:

(1) Los tres autores debían capturar una imagen de la impresión del anuncio recibido en su muro de FB. La Figura 6 ilustra el anuncio recibido por el usuario 3 asociado a la campaña de nanotargeting configurada con 12 intereses aleatorios.
(2) Los autores tuvieron que obtener una instantánea del campo *"Por qué estoy viendo este anuncio"* donde Facebook muestra al usuario cuáles son los motivos por los que recibe el anuncio asociado. La Figura 11 muestra una instantánea de los motivos *"Por qué estoy viendo este anuncio"* asociados a la impresión del anuncio de la campaña dirigida al usuario 3 con 12 intereses.
(3) También se pidió a los autores que hicieran clic en la opción ofrecida en la ventana *"¿Por qué veo este anuncio?"*, que proporciona los atributos específicos utilizados en la campaña publicitaria asociada al anuncio. Esto permite recuperar la lista real de intereses utilizada en la campaña publicitaria asociada al anuncio. La Figura 12 muestra una instantánea de la lista de intereses incluida en el campo *"Por qué estoy viendo este anuncio"* para la campaña publicitaria dirigida al usuario 3 con 12 intereses aleatorios.

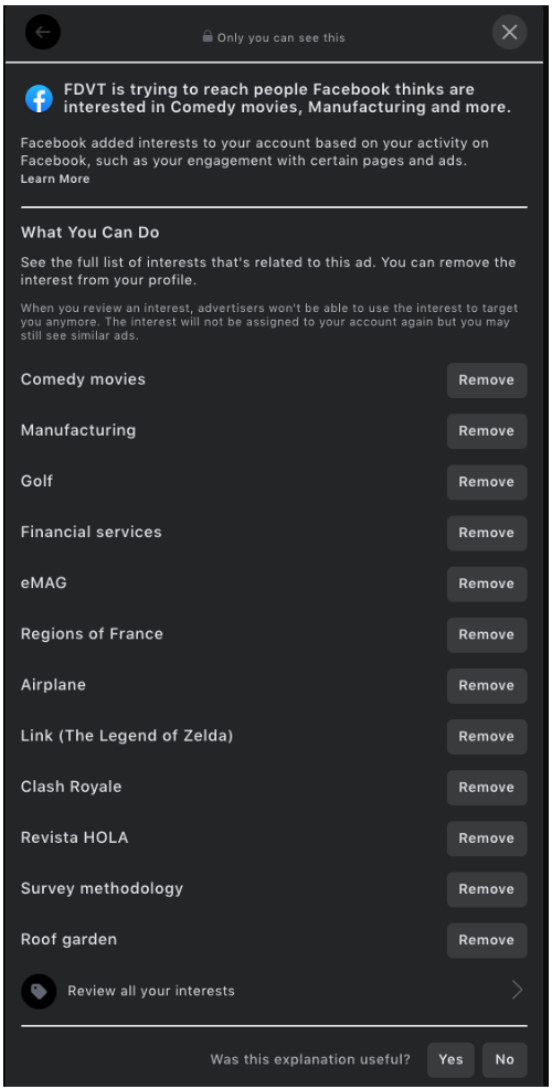

Fig. 12. Instantánea de la lista de intereses utilizada en la campaña publicitaria dirigida al usuario 3 con 12 intereses obtenidos a partir de la función ``*Por qué estoy viendo este anuncio"*.