Análisis de una campaña de redes de bots de fake news para la reelección de Donald Trump en las elecciones de 2020

Miguel Aguado Domínguez y Adrian Losada Casado

***Abstract*—La creciente tendencia de utilizar las redes sociales como fuente de información ha provocado la aparición de las denominadas *fake news*. Este tipo de bulos han sido usados en campañas políticas con el objetivo de afectar a la opinión de los votantes. En este artículo se va a investigar una campaña que trata de favorecer la victoria del actual Presidente de los Estados Unidos Donald Trump en las elecciones del 2020, utilizando una red periódicos locales falsos creados exclusivamente con este fin. Se analizará su descubrimiento y la recopilación de datos de las páginas web de los periódicos y sus cuentas de Twitter. Tras analizar los datos obtenidos se implementó un dataset para futuras investigaciones y se procesaron los datos para identificar otros nodos de la red y redes de cuentas automatizadas.**

***Index Terms*—fake news, análisis de la red, periódicos locales.**

I. Introducción y motivación

Durante la era de la información en la que vivimos actualmente la tendencia de informarse a través de medios digitales ha crecido exponencialmente. Un ejemplo de esto es la pérdida de ventas que los periódicos físicos han sufrido en los últimos años y que les ha forzado a apostar por el uso de estos medios. Entre estas formas de informarse online, las redes sociales suponen un porcentaje considerable debido a la gran popularidad que han adquirido en los últimos años como herramienta para mantenerse conectado y al día de las novedades.

Tras esta digitalización, la facilidad de publicación de información combinada con la ingenuidad de los usuarios provocó la proliferación de las llamadas *fake news*. Las *fake news* son un tipo de bulo que consiste en la creación de un contenido pseudoperiodístico cuyos principales objetivos son la desinformación y la manipulación de los lectores.

Sin embargo, este artículo se centra en las situaciones en las que esta práctica se lleva a cabo a gran escala, con objetivos que van más lejos que los anteriormente mencionados. En estas ocasiones, grandes organizaciones utilizan este tipo de bulos con fines políticos o económicos entre otros. Desacreditar a un rival en campaña, dañar la imagen de una marca de la competencia o generar malestar entre la población respecto a un cierto tema son algunos de los motivos por los que se organizan este tipo de campañas.

Con el paso del tiempo estas campañas se han modernizado y adaptado cada vez más a las posibilidades que ofrecen las redes sociales. En lugar de difundir bulos manualmente publicándolos en cuentas de páginas como Twitter o Facebook, han automatizado este proceso con el objetivo de crear una gran cantidad de cuentas conocidas comúnmente como ”bots”. Estas cuentas permiten publicar noticias falsas de forma automática y simultánea, generando de esta forma un mayor ruido. Alimentando, por ejemplo, determinados *hashtags* para que alcancen posiciones de *trending topic*, alcanzando de esta forma una mayor cantidad de gente.

Un ejemplo del alcance que las *fake news* pueden tener, es su impacto en las elecciones de Estados Unidos de 2016 en las que Donald Trump alcanzó el Despacho Oval. Estas elecciones marcaron el inicio de una nueva etapa de comunicación y campaña política a través de las redes sociales. Estudios posteriores tanto de organismos nacionales como de entidades privadas han demostrado que las *fake news* afectaron en la opinión de los votantes. En un estudio realizado por el **Pew Research Center** concluyó que un 23% de los americanos compartieron noticias falsas durante las elecciones [3].

Si nos situamos en un ejemplo más reciente, la actual crisis del coronavirus ha generado una oleada de *fake news* que ha obligado a empresas como Facebook a establecer un límite de reenvíos en su aplicación *WhatsApp* para evitar que estos bulos se propaguen. Como se ha indicado, se trata de un problema en alza que requiere soluciones y herramientas novedosas para ser resuelto.

En este artículo se va a realizar una investigación sobre una de estas campañas multimillonarias de *fake news*. Su principal objetivo es la reelección del Presidente Donald Trump mediante la manipulación de la comprensión de los procesos políticos del pueblo estadounidense y la explotación de la credibilidad que acompaña al periodismo local. Este tipo de periódicos suelen suscitar una mayor credibilidad entre la población local, debido entre otros factores a que se es más propenso a confiar en lo que se conoce y es más cercano.

El resto del documento se ha organizado como se detalla en este párrafo. En el apartado *II* se realiza un repaso al estado del arte que rodea el proyecto, incluyendo artículos que han tratado temas similares. A continuación, en el apartado *III* se detallará la metodología seguida durante el desarrollo de la investigación con el objetivo de dar a conocer tanto los métodos como los criterios utilizados. Después, en el apartado *IV* se detallarán los resultados obtenidos y se analizarán el el apartado *V*. Finalmente, en el apartado *VI* se detallan las conclusiones obtenidas y se darán posibles pautas futuras que se podrían seguir para ampliar este trabajo.

II. Estado del arte

Una de las principales influencias a la hora de escribir el presente artículo ha sido el trabajo desarrollado por Zannettou et al. en [8]. En este artículo se realiza un estudio sobre como las comunidades online se influencian a través de fuentes de noticias alternativas como las redes sociales como Twitter, Reddit o 4chan. Además, analiza como las *fake news* se propagan rápidamente en estas plataformas influenciando a la gente. También profundiza en lo poderosas que son estas herramientas en el ecosistema de información moderno. Esta primera aproximación al problema ha permitido establecer el importante papel que juegan estas plataformas en el sistema de información del mundo moderno.

En [1], Badawy et al. estudiaron el uso de bots de Twitter durante la interferencia rusa en las elecciones de 2016, comentada en la introducción del presente artículo. Indican que, hasta estas elecciones, se habían usado las redes sociales para promocionar discursos políticos mientras que en estas, otros actores externos las usaron para explotar la discusión política online. Destacan el uso de de trolls y bots para propiciar la desinformación, difundir información políticamente sesgada y afectar la opinión política de la gente. En total recopilaron 43 millones de tweets, de los cuales obtuvieron que, aproximadamente, el 4,9% de los usuarios liberales y el 6,2% de los conservadores fueron bots automatizados. Además, este artículo recopila y describe las técnicas que han usado para analizar las cuentas de Twitter, lo que que nos han servido de gran ayuda para nuestro proyecto.

En Zannettou, Sirivianos et al. [9] también se profundiza en las redes de *fake news*. Comentan que en la actualidad se está produciendo una guerra de la información continua. En esta batalla los actores, tanto espontáneos como financiados y organizados, llevan a cabo campañas de información falsa para manipular la información pública respecto a diversos temas que pueden tener graves impactos, especialmente en las elecciones. Entre los principales objetivos que destacan de este artículo se puede encontrar 1) cómo el público percibe esta falsa información, 2) cómo se llega a propagar, 3) cómo detectar y contener los bulos, 4) y la información falsa en el ámbito político, siendo este uno de los puntos más importantes al poder influir directamente en la sociedad.

En *On the Origins of Memes by Means of Fringe Web Communities* de Zannettou et al. [10] comentan el uso de los memes para manipular la opinión pública política. Este estudio analiza como se propagan, evolucionan e influyen estos memes a través de varias comunidades web como Twitter, Reddit, 4chan. Encontraron un número sustancial de memes relacionados con política, apoyando los informes de los medios de comunicación de que los memes podrían ser utilizados para mejorar o dañar la imagen de los políticos. Finalmente, describen los procesos que han utilizado para medir la interacción y cuantificar la influencia de los memes entre las plataformas y los diferentes ecosistemas, lo cual nos ha sido de gran utilidad.

III. Metodología

A. Investigación previa

Esta investigación se originó tras la publicación de un artículo publicado en la edición impresa de marzo del *The Atlantic* llamado *The Billion-Dollar Disinformation Campaign to Reelect the President*[2]. En este artículo se denunciaba una campaña electoral de propaganda encubierta por para lograr la reelección del Presidente de los Estados Unidos, Donald Trump, en las elecciones de 2020. Esta campaña contaba con una gran financiación y iba desde mensajes de texto hasta la creación de periódicos de noticias falsas en masa.

Este artículo provocó algunos usuarios de la plataforma Reddit se organizasen para intentar hacer frente a esta campaña, bajo el nombre **MassMove**[12]. Con el paso del tiempo compartieron sus ideas para localizar la red de periódicos de *fake news* y las diferentes herramientas y recursos que pueden automatizar este proceso. Además, agruparon todos los conocimientos y resultados obtenidos en un *GitHub* colaborativo[4].

En este punto comenzó la primera etapa de este movimiento denominada con el nombre técnico de *Primer Hackaton*. Esta fase tenía como objetivo la identificación de los vectores de ataque de esta campaña mediante diferentes técnicas. Una de ellas consistía en al identificación de otras páginas web mediante la búsqueda de las noticias de una página ya encontrada en cualquier buscador. Al estar estos sitios web automatizados, las noticias que compartían estaban formadas por el mismo contenido, la misma estructura y las mismas imágenes. Otra técnica utilizada consiste en estudiar el servidor de Amazon AWS en el que están hosteados y buscar el resto de los dominios hosteados en esos ellos mediante *Reverse IP lookup"*.

Durante esta fase se identificaron más de 700 páginas web que simulaban ser periódicos locales de ciudades estadounidenses. El hecho de que sean periódicos locales es muy significativo, primero porque buscan no llamar la atención como lo haría un periódico a nivel nacional y segundo porque utilizan esa predisposición a creerse sus noticias que ha sido comentada en la introducción del presente documento. Se observó que este conjunto de páginas se trataba de una granja de bots dado que su estructura y diseño eran iguales y consistentes entre ellas. Además, las noticias utilizadas se repetían idénticas en todos los dominios. Los vectores de ataque identificados pueden ser observados en la siguiente imagen:

Una vez se identificaron los dominios y sus respectivas cuentas en Facebook y Twitter, se inició el llamado *Segundo Hackaton*. En esta fase los usuarios dividieron sus esfuerzos en diferentes frentes:

- Se monitorizaban las páginas web para registrar sus actualizaciones.
- Se analizan los seguidores de Facebook y quien comparte sus publicaciones con el objetivo de identificar patrones.
- Se analizan los nombres de las páginas web de los periódicos con el fin de encontrar patrones que permitan identificar otras nuevas de la misma campaña que aun se mantienen sin identificar.

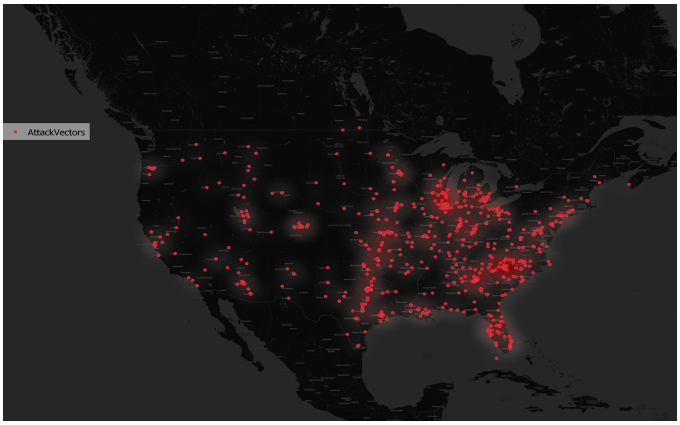

Fig. 1. Vectores de ataque identificados

- Se descarga la información y las publicaciones de Facebook con el objetivo de almacenarla para futuros análisis.

Finalmente, en el *Tercer Hackaton* se pasó a una fase activa en la que se buscaba provocar un impacto significativo en la red de periódicos falsos. Esta fase se llevó a cabo de forma contigua al *Segundo Hackaton* y aparte de utilizar los resultados del primero, también tenían en cuenta la información que se detectaba. Algunos usuarios que disponían de contactos dentro de Facebook, les comentaron el problema y solicitaron el cierre las páginas que estos periódicos locales tenían en la plataforma. Otros usuarios contactaron con periodistas con el objetivo de proporcionales información para denunciar este caso a escala global y darlo a conocer al pueblo americano. Mientras que algunos usuarios siguen buscando nuevas cuentas de esta campaña, otros han comenzado la búsqueda de otras redes de cuentas falsas que puedan estar actuando sin ser detectadas. Otras aproximaciones más manuales consisten en denunciar las páginas y las cuentas en redes sociales a nivel de usuario a los sitios donde están hosteadas.

Estas dos últimas fases son las que se encuentran activas durante la realización del presente documento. El número total de dominios identificados hasta ahora asciende a un total de 1083 periódicos locales falsos. La investigación realizada se ha centrado en las cuentas de Twitter de los periódicos identificados y de su dominio web asociado. Hemos seleccionado Twitter como red social de investigación por las siguientes razones:

- Las cuentas de Facebook asociadas a los dominios web han sido ampliamente investigadas y analizadas por los usuarios del movimiento *MassMove*.
- Twitter se ha convertido en la red social central del debate político y es la principal herramienta social en las campañas de los políticos de todo el mundo. A excepción de Estados Unidos, Facebook ha quedado relegado a un segundo plano en este ámbito. Los políticos prefieren publicar comunicados, hacer declaraciones e interactuar con sus votantes a partir de Twitter. Además, el debate y la crítica política también se ha convertido en una de las señas de identidad de esta red social.

Como punto de partida, hemos utilizado las cuentas de Twitter asociadas a los periódicos, que fueron identificadas por los miembros del movimiento *MassMove* dentro de los ficheros *html* de los dominios web encontrados. Estas cuentas junto con sus datos se detallan en un fichero *JSON* en su GitHub[4]. El número de cuentas que se van a analizar es de 37. En el resto del documento, estas cuentas serán denominadas *"cuentas principales"*.

B. Investigación de cuentas relacionadas

Para cada una de las cuentas principales, primero se han recuperado el nombre de todas las cuentas que siguen y las que las siguen. De esta forma es posible conocer la magnitud global de la red de cuentas automatizadas, comprobar que están relacionadas, realizar un análisis de las cuentas cercanas y como se interconectan entre ellas. Este proceso, al igual que todas las recopilaciones de datos de Twitter, se han realizado con scripts en Python que utilizan la API [11] proporcionada por la plataforma y la libería Tweepy[7].

Con el objetivo de encontrar nuevas cuentas pertenecientes a esta red de bots se ha analizado la lista de seguidores y seguidos de las diferentes cuentas principales. Esto permite detectar otras redes del mismo autor o organización puesto que comúnmente se siguen entre si para aumentar sus números iniciales y darse a conocer. Este hecho se puede observar en la Figura 2, donde excepto unas cuentas más distantes todas se siguen entre si.

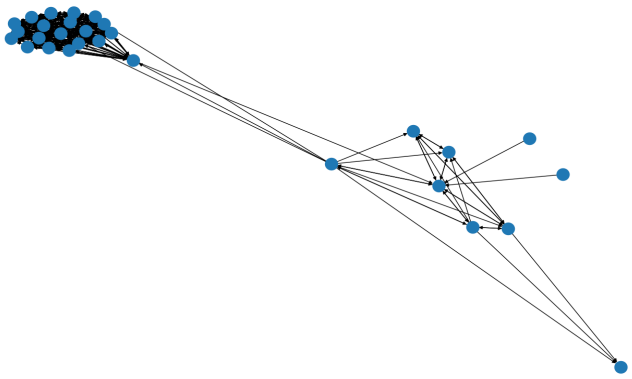

Fig. 2. Relaciones entre las cuentas principales

Primero, se han recopilado los seguidores de las cuentas principales que siguen a más de dos, descartando las cuentas principales que se siguen entre si. Este criterio se ha tenido en cuenta debido a que al ser periódicos locales no es común que un usuario siga a otras cuentas de noticias locales a parte de la de su ciudad y alguna cercana o en la que tenga interés. Identificar cuentas que siguen a muchas de las cuentas principales permitiría descubrir otras cuentas de periódicos falsos locales o de otras redes de bots relacionadas.

Después, como segundo criterio y con el objetivo de comprobar si una cuenta está automatizada para compartir noticias falsas, se ha seleccionado únicamente las cuentas que no han dado "me gustas". En las cuentas principales, se ha observado que al estar automatizadas no marcan ningún otro tweet como favorito. Como tercer criterio se ha comprobado que la cuenta únicamente haya publicado tweets que contengan enlaces. Esto se debe a que al estar automatizadas, se ha observado que las

cuentas identificadas con anterioridad únicamente publican las noticias falsas con un enlace a su página web.

Finalmente, las cuentas que hayan superado estos filtros se someterán a una revisión manual para confirmar que son cuentas automatizadas.

C. Estudio de las cuentas principales

Para poder analizar más en detalle estas cuentas, se procederá con la identificación de todas las URL que cada una de ellas ha escrito en un tweet alguna vez. Para recuperar estos tweets se ha utilizado la API de Twitter. Estas direcciones se almacenarán en un diccionario en formato JSON con clave el nombre del usuario y valor una lista con todas las URL recopiladas.

Adicionalmente, para un mejor análisis y para evitar la perdida de toda esta información si las páginas web fueran eliminadas, se procederá con la creación de un crawler, un programa que analiza documentos de sitios web, mediante Python y Selenium [5], con el que descargar todos los archivos HTML asociados a estas direcciones mediante el navegador Mozilla Firefox.

Cabe destacar que, además, en todas las URL de los usuarios identificados, se puede apreciar la siguiente estructura: *https://nombredelperiodico.com/stories/012345678-nombre-del-articulo*. En dicha estructura, aparece un identificador numérico de nueve dígitos único para cada artículo, por lo que se procederá a aislarlo dentro de la URL para utilizarlo como nombre de los archivos descargados.

Sin embargo, entre todos estos archivos descargados podría haber errores, es decir, ficheros que habría descartar previo análisis. Entre estos posibles errores se podrían destacar 1) páginas de error con un peso de archivo muy pequeño, y 2) archivos con un nombre que no siguiera la estructura de los identificadores anterior. Una vez obtenidos los archivos HTML se procederá con un primer análisis con el que poder descartar todos estos errores.

Una vez recopilados todos los archivos HTML de las diferentes cuentas, será necesario analizarlos con el objetivo de comprobar si las estructuras de las paginas web de los diferentes periódicos son similares entre sí. Para ello se escogerá ejemplo de entre todos los archivos correctos de cada periódico, es decir, aquellos que no fueran ni páginas de error ni tuvieran un nombre no ajustado al identificador, con el objetivo de compararlo con los representantes del resto de los periódicos.

Como primera forma de comparación se utilizará el método *ratio* del módulo de Python *difflib*, con el que comparar cada par de archivos obteniendo como resultado su porcentaje de similitud. Cabe destacar que, dependiendo del orden de los parámetros, el resultado podría ser diferente, por lo que es necesario comparar todos con todos [6] para luego obtener unos resultados óptimos.

La segunda forma de comparación se realizará obteniendo aquellas líneas que son diferentes entre los diferentes archivos. Para ello, en primer lugar, se eliminarán todos aquellos aspectos que puedan afectar a la comparación, entre los que destacan saltos de línea, espacios en blanco y líneas duplicadas. En segundo lugar, cada línea del archivo que este siendo analizado se buscará en su par correspondiente y, en caso de que no existiera, se añadirá dicha línea a una lista encargada de almacenar todas las diferentes encontradas. Por último, se obtendrá el tamaño del resultado obtenido, significando por lo tanto que, cuanto mayor sea este tamaño, más diferente es el archivo analizado con respecto al que ha sido comparado.

D. Limitaciones

Una de las principales limitaciones técnicas del estudio ha sido el conjunto de restricciones de la API de Twitter. Debido a la cantidad de datos que pueden ser manejados por la API, Twitter ha impuesto una serie de restricciones que evitan la sobrecarga en sus sistemas. Una de las principales restricciones es el límite de tweets máximos por día que pueden ser descargados. La API dispone de una suscripción de pago que durante este estudio no ha sido abonada. Debido a esto, los procesos de recogida de datos de Twitter han necesitado de largos tiempos de recopilación. Otra desventaja de la API es que no permite por restricción de Twitter recopilar las cuentas que han dado ”me gusta” a un tweet. Esto se hace por motivos de privacidad, pero impide poder conocer las cuentas que más interactúan con los periódicos falsos.

En cuanto a las desventajas que impone la forma en la que están automatizadas las cuentas podemos encontrar que no disponen de tweets marcados como ”me gusta”. Esto permite identificar estas cuentas fácilmente pero impide obtener información muy útil.

Debido a la gran cantidad de ficheros HTML descargados no es posible compararlos todos entre si para analizar su similitud, ya que el tiempo requerido para esta tarea sería mucho mayor que del que se dispone para la realización del artículo. Debido a esto se va a escoger una única página representativa del total de las obtenidas de cada periódico local falso. De esta forma se podrá seguir comparando la similitud entre los distintos periódicos para garantizar que corresponden al mismo autor sin necesitar un tiempo excesivo.

Mediante una revisión manual, se ha comprobado que los datos de las cuentas de una red se crean con la misma plantilla gráfica variando, únicamente, el nombre. Al no disponer de una red convolucional entrenada para este propósito, no se ha podido realizar un análisis más exhaustivo de la plantilla de estas cuentas.

Los filtros diseñados para detectar cuentas de bots de la misma red o del mismo autor tienen criterios muy específicos para este caso. Unos bots más complejos y que traten de simular un comportamiento más común superarán este filtro. Aunque es una limitación, estos filtros funcionan a la perfección para este estudio.

IV. Resultados

En este apartado se van a detallar los resultados de los procesos de investigación realizados. Con el objetivo de proteger la privacidad de estas cuentas y debido a que no tenemos una sentencia que confirme que estas cuentas son redes de bots automatizadas, se va a ocultar su nombre de usuario único. Sin embargo, el nombre de las cuentas y sus imágenes de perfil

al no ser identificativos debido a que no son únicos y pueden ser modificados en cualquier momento, van a ser usados para representar un ejemplo de los resultados de la investigación de posibles cuentas automatizadas.

A. *Investigación de cuentas relacionadas*

Primero, de cada una de las cuentas se ha recopilado la lista de seguidores y seguidos. Como se puede observar en la Figura 3 el número de seguidores y seguidos es constante para la mayoría de las cuentas principales. Sin embargo, para dos de las cuentas podemos encontrar una diferencia destacable. La primera es la cuenta 24, cuyo número de seguidores es de 3.991. Sin embargo, este número se puede explicar como algo casual y puede deberse a muchas razones como por ejemplo, que una cuenta con muchos seguidores le hayan dado *retweet* a una de sus publicaciones mientras que el resto de cuentas no han tenido esa ”fortuna”. Además, el número de cuentas seguidas es similar al del resto. La segunda cuenta es la 32 que tiene un total de 27.490 seguidores y 23.013 seguidos. Estos números son demasiado elevados para deberse a una casualidad. Tras una revisión manual de la cuenta no se ha observado nada que explique su popularidad. Esta cuenta se dedica al área judicial de una ciudad considerablemente grande de Estados Unidos, por lo que podría deberse, por ejemplo, a que un caso de esta localidad ganó popularidad a nivel nacional y al simular ser un periódico local de la ciudad aumentó su actividad.

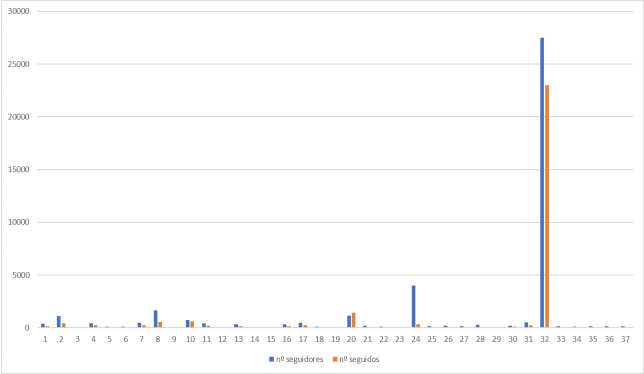

Fig. 3. Número de seguidores y seguidos de las cuentas principales

Tras la realización de este primer filtro, se ha obtenido que el número de cuentas que siguen a más de una de las cuentas principales es de 541. Como se puede observar en la Figura 4, el número de cuentas principales seguidas por estas cuentas es muy variado. La mayoría de las cuentas siguen a 2 o 3 usuarios pero otras alcanzan la increíble cifra de 26, lo que indica que estas cuentas pueden estar automatizadas y pertenecer al mismo autor.

A las cuentas resultantes de este primer filtro pasaron a la segunda fase. En este segundo filtro, se contabilizó el número de likes de cada cuenta. Como es común, la mayoría de las cuentas tenían un número alto y diferenciado de likes. Sin embargo, se pudo observar que 31 de ellas no tenían ningún tweet marcado como ”me gusta”, lo cual es muy inusual. Además, tras observar que 17 cuentas tenían un único ”me

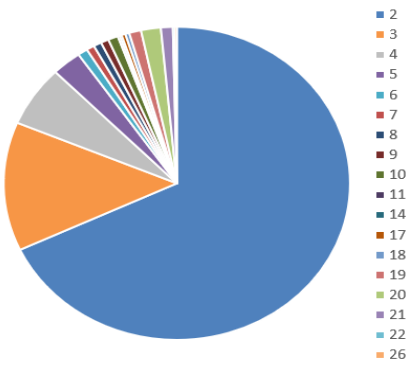

Fig. 4. Frecuencia del número de seguidores

gusta” se decidió relajar este criterio a ningún o un tweet marcado como ”me gusta” debido a que este valor también es inusual. Pese a que esto puede deberse a cuentas con poca actividad, se quiso minimizar las posibilidades de que alguna cuenta automatizada pudiese saltarse este filtro. Debido a esto se han tenido en cuenta las 48 para la siguiente fase.

El tercer y último filtro tenía como criterio que la cuenta únicamente hubiera publicado tweets con enlaces a una publicación de su página web. Tras realizar este proceso, se obtuvieron 19 cuentas que cumplían con los criterios establecidos.

Finalmente, estas cuentas fueron sometidas a una revisión manual para comprobar si se trataban de cuentas de noticias. De estas cuentas 2 eran de personas con poca actividad que únicamente habían publicado enlaces compartiendo notificaciones de aplicaciones enlazadas. Otras 13 fueron identificadas como cuentas de propósito no periodístico. Sin embargo, estas cuentas si eran activas, tenían una gran cantidad de tweets únicamente con enlaces y no habían marcado tweets como ”me gusta”, lo que resulta muy extraño y antinatural. Al no ser el objetivo principal de esta investigación, no fueron analizadas más en profundidad para demostrar que se trataban de cuentas automatizadas. Finalmente, 4 de las cuentas cumplían todos los requisitos al ser cuentas informativas de periódicos locales.

Tras la revisión manual se ha llegado a la conclusión de que estas cuentas se tratan de bots automatizados con el mismo propósito que los identificados hasta el momento. Además, al analizar sus seguidores se ha encontrado una gran cantidad de cuentas con imágenes que, al igual que las de las cuentas principales, han sido creadas con la misma plantilla y disponen de la misma descripción cambiando el nombre de la ciudad de la que simulan ser. En conjunto, se han identificado dos nuevas redes diferentes. Las cuentas miembro de las redes comparten imágenes y descripciones similares como se puede ver en la Figura 8 que muestra el ejemplo de una de las redes encontradas. Para la primera red, se han identificado 19 cuentas miembro automatizas. Por otro lado, para la segunda el número de cuentas miembro fue de 17, lo que da un total de 36 entre ambas redes. El resto de seguidores se han registrado en un fichero JSON junto con sus datos de cuenta.

B. Estudio de las cuentas principales

1) Descarga de archivos: En cuanto a la recopilación de URL de los usuarios analizados, en total, se han obtenido 91.889 tweets del total de las 37 cuentas diferentes. De estos tweets se han recopilado 91.858 direcciones web. De este total, mediante el uso del crawler desarrollado, se han podido descargar 91,589 archivos, es decir, un 99,7%, ocupando un espacio total de 14,9GB.

Tras esto, se ha procedido con el análisis de posibles errores en la descarga de los archivos y, al igual que se había supuesto, se han encontrado numerosos ficheros 1) con un tamaño idéntico de 3KB y 2) con nombres que no coincidían con la estructura de nueve dígitos identificada.

En cuanto a 1), al ver el contenido de estos HTML, se puede ver que son páginas de error 404, es decir, direcciones no encontradas de las que no se ha podido obtener ninguna información útil. Debido a esto se han descartado todas las páginas por debajo de los 5KB de tamaño, un total de 3.396.

En cuanto a 2), se han analizado algunas de los archivos que no siguen este patrón y en ellos se ha podido apreciar que, A) no se corresponden con el dominio del usuario identificado y que, B) algunos ejemplos se encuentran en otros idiomas, como, por ejemplo, el indonesio. Así, todas las páginas que no seguían la estructura identificada en el nombre han sido descartadas, un total de 1.043.

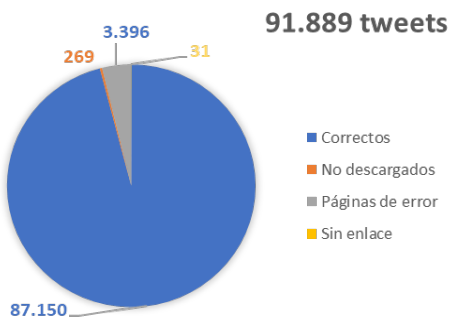

Fig. 5. Archivos HTML descargados respecto a los tweets obtenidos

Con esto, como puede apreciarse en la Figura 5, del total de 91.858 enlaces recopilados, se han acabado obteniendo un total de 87.150 archivos HTML, un 94,87% del total de enlaces y un 95,15% del total de los archivos descargados.

En definitiva, de los 37 usuarios analizados, se ha obtenido de cada uno de ellos una media de 2.482 direcciones web, con 7 enlaces no procesados, 92 páginas de error y 29 ficheros con nombre erróneo, es decir, una media de 2.355 páginas descargadas correctamente por usuario.

2) Comparación de archivos: Tras la obtención definitiva de los archivos se ha procedido con el estudio de la similitud entre la estructura de los diferentes periódicos. Una vez escogido el representante de cada periódico, se han ejecutado ambos programas de comparación destacando que, el primero de ellos, aquel que utiliza el método *ratio* del módulo *difflib* de Python, ha tardado aproximadamente 12 horas en completar el total del proceso, motivo principal por el que no ha sido posible realizar comparaciones entre más de un representante de cada periódico.

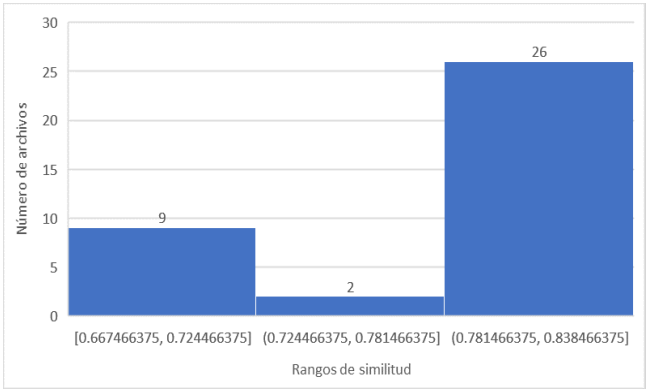

Fig. 6. Número de archivos por rango medio de similitud

Como se puede apreciar en la Figura 6, respecto a la obtención del porcentaje de similitud entre todos los archivos, estos se encuentran en una serie de rangos bastante definidos: 9 de los periódicos son similares al resto en una media de entre un 66,75% y un 72,45%; otros dos lo son entre un 72,45% y un 78,15%; y,los 26 periódicos restantes, conforman el ultimo rango entre el 78,15% y el 83,84%.

Así, se obtiene un ratio promedio entre todos los periódicos de un 78,65%, con el archivo más similar al resto con un porcentaje de 83,67, correspondiente a la cuenta número 1, y el más diferente con un 66,75%, correspondiente a la cuenta número 33.

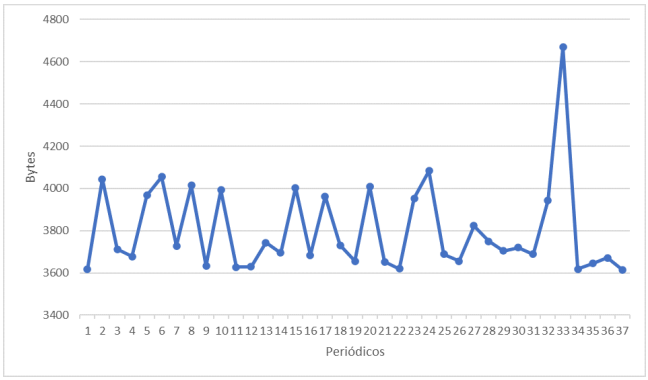

Fig. 7. Bytes medios de diferencia entre el total periódicos

En cuanto a la segunda forma de comparación, el tamaño de diferencia entre el total de los archivos, como se puede ver en la Figura 7, todos los valores oscilan en un rango similar, obteniendo una diferencia media entre el total de ellos de 3.802,52 bytes. Como se ha mencionado anteriormente, el que estos valores representen el tamaño que ocupan las lineas **diferentes**, hace que, cuanto menor sea el valor obtenido, más parecido es dicho archivo con respecto al resto, y viceversa. De esta forma, el archivo que más bytes ocupa, 4.668,44, correspondiente a la cuenta número 33, es el más original de todos; mientras que el que menos, con 3.614,89 bytes, correspondiente a la cuenta número 37, es el más similar,

seguido muy de cerca por los 3.618,44 bytes de la cuenta número 1.

V. Discusión

A. Investigación de cuentas relacionadas

De los resultados obtenidos tras la investigación realizada sobre los seguidores y seguidos de las cuentas principales se ha obtenido una serie de cuentas catalogadas como sospechosas de ser cuentas automatizadas. La aplicación de diferentes filtros y criterios a estas cuentas han reducido el número de sospechosos hasta alcanzar mediante revisión manual, 4 cuentas pertenecientes a redes de difusión de noticias falsas.

La revisión manual de estas cuentas y del resto de las cuentas miembro de las redes a las que pertenecen, ha demostrado con un alto índice de confianza que se tratan de redes de bots de periódicos locales falsos creados por el mismo autor. Estas 36 cuentas han sido compartidas con los miembros del movimiento **MassMove** para fomentar su denuncia y conseguir así acabar con su actividad en las redes sociales. Además, estas cuentas podrán aportar una mayor cantidad de datos para futuras investigaciones sobre *fake news*.

B. Estudio de las cuentas principales

En cuanto a la descarga de archivos HTML cabe destacar, en primer lugar, la estructura del identificador explicada. Y es que, aunque las páginas web de los periódicos deberían ser diferentes entre sí, como se ha podido ver, **todas** ellas siguen una misma forma en sus URL, obteniendo, como se ha visto en los resultados una media de 29 ficheros con nombre erróneo por periódico frente a las 2.355 páginas descargadas correctamente.

Asimismo, de la comparación entre los representantes de cada periódico se ha podido ver, en cuanto al ratio de similitud, que la media entre todos ellos es de un 78,65%. Este dato es bastante revelador ya que, obviamente, los textos de las noticias no serán iguales ni siquiera dentro de los artículos del propio periódico. Sin embargo, la estructura HTML, aquello que el usuario no ve, podría haber sido generado por la misma persona o programa, destacando en esta estructura, por lo tanto, gran parte del mérito del alto porcentaje obtenido.

En cuanto a la similitud del texto de las propias noticias cabe destacar que, si bien en un mismo periódico no deberían repetirse, esto sí que podría ocurrir entre los diferentes obtenidos ya que, al ser noticias falsas, muchas de ellas podrían tener el mismo contenido. Sin embargo, debido a las limitaciones al solo analizar un representante por periódico, esto no ha podido ser comprobado.

Por último, cabe recordar que se ha realizado una segunda comparación pero, en este caso, obteniendo el tamaño de las líneas que son diferentes entre cada par de artículos. En este análisis destaca, en primer lugar, lo mencionado anteriormente, y es que, al comparar lineas enteras, aquellas que tengan el propio texto del artículo, deberían ser diferentes. Por lo tanto, aquellas líneas que hayan sido identificada como similares entre cada par de artículos deberían, de nuevo, ser parte de la estructura HTML de los mismos.

Además, comparando los resultados tanto del ratio como del tamaño se puede apreciar que la misma cuenta, la número 33, es aquella más diferente al resto, al tener el menor porcentaje con un 66,75% y el mayor tamaño de lineas diferentes con 4.668,44 bytes. La cuenta más parecida al resto, si bien no coincide en ambas comparaciones, podría considerarse la número 1, al tener un porcentaje de 83,67% de similitud y un tamaño de líneas de diferentes de 3.618,44 bytes, la segunda cuenta con menor tamaño obtenido.

Por lo tanto, debido al alto índice de similitud entre ellas como ha podido apreciarse, sobre todo en términos que deberían ser únicos, como la estructura HTML, se puede afirmar que con un alto de índice de confianza, que estas cuentas suponen una red automatizada de bots.

Todos los datos obtenidos, incluyendo los tweets publicados y los archivos HTML descargados, han sido publicados para su uso por parte de la comunidad científica en un repositorio público en GitHub *[próxima subida, en este borrador aun no está disponible]*.

VI. Conclusiones

En este artículo se ha llevado a cabo un estudio sobre la campaña para la reelección de Donald Trump que medios conservadores están realizando con la ayuda de las nuevas tecnologías. Esta campaña ha utilizado una red de periódicos locales falsos para alterar la opinión política de la ciudadanía estadounidense a su favor.

Tras el descubrimiento de esta red por el movimiento **MassMove**, este artículo ha tenido como objetivo identificar nuevos nodos de la red y otras redes de la misma autoría y objetivo a través de la plataforma Twitter. Una de las principales aportaciones ha sido la identificación de una gran cantidad de sitios web de noticias falsas de otras redes diferentes a la inicial y su denuncia para su bloqueo. Debido a esto, el acceso a muchas de las cuentas está ya siendo restringido por Twitter y sus dominios eliminados por los dueños de los servidores que los hospedaban.

Como segunda aportación se ha creado un conjunto de datos compuesto por toda la información de las cuentas de Twitter de estos falsos periódicos, incluyendo todos los tweets publicados durante su existencia, y las páginas de las noticias que publicaban en sus dominios web a través de estos tweets. Este dataset podría ser de gran utilidad para futuros análisis, dado que permitiría investigar sus orígenes aunque las cuentas sean bloqueadas y sus dominios web eliminados.

Como trabajos futuros se ha considerado el seguimiento en tiempo real de las cuentas de este tipo de campañas para analizar su propagación por la red. Además, el uso de una red de deep learning permitiría automatizar los procesos de revisión manual comentados en este artículo. De esta forma se podría rastrear activamente la plataforma en busca de este tipo de cuentas.

References

[1] Adam Badawy, E. F. and Lerman, K. (2018). Analyzing the digital traces of political manipulation: The 2016 russian

interference twitter campaign. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*.

[2] Coppins, M. (2020). The billion-dollar disinformation campaign to reelect the president. https://www.theatlantic.com/magazine/archive/2020/03/the-2020-disinformation-war/605530/. [Accedido el 15-04-2020].

[3] Michael Barthel, A. M. and Holcomb, J. (2016). Many americans believe fake news is sowing confusion. https://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/. [Accedido el 14-04-2020].

[4] Move, M. (2020). Mass move github. https://github.com/MassMove/AttackVectors. [Accedido el 15-04-2020].

[5] Postels, P. (2019). Build a scalable web crawler with selenium and python. https://towardsdatascience.com/. [Accedido el 15-04-2020].

[6] PythonSoftwareFoundation (2020). Difflib: Helpers for computing deltas. https://docs.python.org/3/library/difflib.html#difflib.SequenceMatcher.ratio. [Accedido el 16-04-2020].

[7] Rivera, P. (2020). Tweepy. https://www.tweepy.org/. [Accedido el 17-04-2020].

[8] Savvas Zannettou, Tristan Caulfield, E. D. C. et al. (2017). The web centipede: Understanding howweb communitiesinfluence each other through the lens of mainstream and alternative news sources. *Proceedings of IMC '17*.

[9] Savvas Zannettou, Michael Sirivianos, J. B. and Kourtellis, N. (2019). The web of false information: Rumors, fake news, hoaxes,clickbait, and various other shenanigans. *ACM Journal of Data and Information Quality, Vol. 11, No. 3*.

[10] Savvas Zannettou, Tristan Caulfield, J. B. et al. (2018). On the origins of memes by means of fringe web communities. *The 18th ACM Internet Measurement Conference 2018*.

[11] Twitter (2020). Twitter api developer guide. https://developer.twitter.com/en/docs/api-reference-index. [Accedido el 17-04-2020].

[12] u/mcoder (2020). Mass move. https://www.reddit.com/r/MassMove/. [Accedido el 15-04-2020].

Fig. 8. Cuentas automatizadas descubiertas