The role of online attention in the supply of disinformation in Wikipedia

Anis Elebiary
University of South Florida
Tampa, FL, USA
anis@usf.edu

Giovanni Luca Ciampaglia
University of South Florida
Tampa, FL, USA
glc3@usf.edu

ABSTRACT

Online communities like Wikipedia have been extremely successful at producing large amounts of high-quality content. At the same time, the adoption of User-Generated Content (UGC) principles makes such communities vulnerable to false or misleading information. Previous work, for example, has shown that hoaxes on Wikipedia can often go undetected for extended periods of time. However, little is known about the circumstances leading to the creation of intentionally false or misleading information online. Does collective attention toward a topic increase its likelihood of containing more disinformation? In this work, we measure the relationship between allocation of attention and the production of hoax articles on the English Wikipedia. We analyze the traffic patterns surrounding the creation of a set of known hoax articles and their respective topics. We found that, compared to legitimate articles created on the same day, hoaxes tend to be more associated with the attention spikes preceding their creation. This observation is consistent with the idea that the supply of false or misleading information on a topic is driven by the attention it receives. These findings improve our comprehension of the determinants of disinformation in UGC communities and could help preserve the integrity of knowledge on Wikipedia.

CCS CONCEPTS

• **Information systems** → *Web mining*; *Wikis.*

KEYWORDS

Wikipedia, Disinformation, Hoaxes, Collective attention, Traffic volume, Computational Social Science, Social network analysis

1 INTRODUCTION

The Web distributes information worldwide without the need for a centralized entity vetting content [5]. This design choice has enabled several websites to become the hubs for communities where users can produce, consume, and disseminate content without central oversight. Examples of these *user-generated content* (UGC) websites include major social media platforms, like Facebook or Twitter, or global online knowledge production communities like Wikipedia. Nowadays, many of the top 50 most visited websites worldwide are UGC websites [30].

This new radical approach to online communication has had both positive and negative ramifications. On the positive side, exposure to novel concepts and ideas has never been easier with the rise of UGC websites. With a few clicks of a computer or a smartphone, the Web allows one to access large amounts of knowledge, to follow in real time current events, and to engage with other points of view about any given topic of conversation. However, a negative consequence of the popularity of UGC websites is that their low barriers to access, combined with the lack of supervision from experts or other gatekeepers, allow for the proliferation of false or misleading information on the Web [23, 36].

Social networking services, in particular, are vulnerable to the spread of false or misleading content [3, 4, 6, 14–16, 47], but there are growing concerns that other UGC communities like Wikipedia may be vulnerable to these threats too [29]. This is especially worrisome since Wikipedia is one of top most visited internet websites [30] and a popular source of knowledge [27]. Wikipedia contains over 50 million articles in more than 300 languages; in February 2022, the English language edition of Wikipedia alone received 781M visits (from unique devices) and was edited over 5M times [39, 43]. Hence, preserving the integrity of Wikipedia is of paramount importance for the Web as a whole [29].

There are many potential threats to the integrity of knowledge in Wikipedia [29]. One common threat comes from vandalism, which is “a deliberate attempt to compromise the integrity of the encyclopedia, often through the insertion of obscenities, insults, nonsense or crude humour, or by page blanking” [40]. Over the years Wikipedia has developed an impressive array of socio-technical solutions to the problem of vandalism. Early work on Wikipedia has shown that most acts of vandalism are repaired manually by the crowd of contributors within a matter of minutes [35]. In addition to human interventions, automated tools like ClueBot NG play a crucial role in maintaining encyclopedic entries clear from damage [11, 18]. On top of these, there are other methods such as patrolling recent changes, creating watchlists, blocking frequent vandalism creators, and using editorial filters. Finally, multiple research attempts have been conducted to aid in both the manual and the automatic detection of vandalism [2, 19, 28, 32].

Vandalism, however, is not the only threat to the integrity of Wikipedia’s content. Whereas vandalism focuses on defacing existing entries, there exists evidence showing that Wikipedia is also targeted by *hoaxes*, whose aim is to create whole new entries about fake, fictitious topics. An example of a famous Wikipedia hoax is the entry *Jar’Edo Wens*, a fake Australian aboriginal deity, which went undetected for almost 10 years before being debunked and deleted [8]. Hoaxes are thus not to be confused with vandalism; although vandalism is a much bigger threat in scope and size compared to hoax articles, hoaxes constitute a more subtle threat, which has received less attention compared to vandalism.

To date, little is known about Wikipedia hoaxes. Kumar et al. [22] collected a sample of known hoaxes from resources compiled by

Wiki Workshop ‘22, April 25, 2022, Online
2022.

the Wikipedia community, and studied their longevity, along with other characteristics. They found that one in a hundred hoaxes remain undetected for more than a year, with 92% of the cases detected within the first day. They also observed that, although only 1% of all hoaxes remain undetected for more than a year, those that go undetected have a higher chance over time of remaining so. Finally, they showed that, on average, hoaxes have a lower density of internal links and receive less traffic than legitimate (i.e., non-hoax) articles [22].

A crucial question that remains unresolved is what drives the creation of hoaxes on Wikipedia. Because their original authors are aware that these articles are false, hoax articles are different from mere *misinformation*, but should rather be considered instances of *disinformation* [23, 36]. As such, understanding the factors that determine the supply of hoaxes on Wikipedia could shed light on disinformation in general, including broader threats to the integrity of the Web, like state-sponsored propaganda [13, 21, 46] and conspiracy theories [33].

To bridge this gap, in this paper, we study the role of online attention in the supply of disinformation in Wikipedia. The idea of an economy of attention was first introduced by Simon [31], who observed that attention is a limited resource that needs to be allocated [12]. Here, to quantify the flow of collective attention to individual topics of knowledge, we take advantage of the unique Wikipedia traffic dataset and API. Traffic to Wikipedia has been used before to study collective attention. García-Gavilanes et al. [10] studied the patterns of attention to Wikipedia in the wake of airplane crashes. They found that the traffic to entries about *previous* airplane crashes was greater than that of the current crash, i.e. the one that triggered the attention surge [10].

Ciampaglia et al. [7] studied the creation of new Wikipedia entries (i.e., not just hoaxes) and observed that the creation of new information about a topic is preceded by spikes of attention toward it, as measured by traffic to neighboring entries [7]. This is consistent with a model in which the demand for information on a topic drives the supply of novel information about it. Consequently, measuring traffic to Wikipedia entries can help us get a step closer to understanding why and when hoaxes are more likely to be produced. Specifically, in this work we seek to answer the following questions:

- Q1. Does online attention toward a topic increase the likelihood of disinformation being created about it?
- Q2. Is there a relationship between traffic to Wikipedia and the production of hoax articles?
- Q3. How can we quantify shifts of attention to a topic around the creation time of a hoax about it?

To answer these questions, we compiled a list of known hoax articles [41] along with their creation timestamps and content. To control for potential confounding factors in the distribution of traffic to Wikipedia over time, for each hoax, we considered a cohort consisting of all the legitimate (i.e. non-hoax) Wikipedia articles that were created on the same day as that of the hoax. Similar to Kumar et al. [22], we find that hoaxes differ from legitimate articles in key appearance features, but do not strongly differ in the number of hyperlinks they contain. For each article (either hoax or non-hoax), we parsed its content and extracted all the out-links, i.e. its neighbors in the Wikipedia hyperlink network. The presence of a link between two Wikipedia entries is an indication that they are semantically related. Therefore, traffic to these neighbors gives us a rough measure of the level of online attention to a topic *before* a new piece of information (in this case an entry in the encyclopedia) is created.

Finally, we measure the relative change of traffic in the 7-day period before and after the creation of a hoax and compare this change to that of the articles in its cohort. To preview our results, we find that, on average, online attention tends to precede the creation of hoaxes more than it does for legitimate articles. This observation is consistent with the idea that the supply of false and misleading information on a topic is driven by the attention it receives.

The rest of the paper is structured as follows. Section 2 describes the details of the data collection process for the list of hoaxes and their cohorts. It also gives details about the comparison between hoaxes and legitimate articles in terms of their appearance features, and describes in detail the pre-processing of the Wikipedia traffic data — how it was cleaned and filtered of unwanted entries and titles. Section 3 discusses the techniques used to quantify online attention and its relationship to the hoax creation. It also describes the statistical procedures performed to asses the results. Finally, section 4 summarizes our findings and future directions.

Code and data needed to replicate the findings of this study are available on Github at github.com/CSDL-USF/wiki-workshop-2022-elebiary-ciampaglia.

2 DATA AND METHODS

This section shows how the dataset of hoaxes was collected. It also shows the process of building the cohort of each hoax. The cohort of a hoax is defined as all the legitimate Wikipedia articles that were created on the same day as their respective hoax. We show how the pageview dataset [38] was utilized to find the hourly traffic count. This section also details the pre-processing of the pageview dataset. Finally, we present how the difference of relative volume change between each hoax and its cohort was calculated, as well as the process of sampling the distribution of means for those differences.

2.1 Data Collection

Wikipedia editors maintain a list of known detected hoaxes [41]. To be included in this list, hoaxes meet either of the following two characteristics: (i) they have gone undetected for more than a month after patrolling [22], or (ii) they were discussed by reliable media sources.

To collect this list, we queried the Wikipedia API using the 'prefix search' endpoint [25]. The total number of titles retrieved was $N_h = 190$. We then used the Toolforge [44] to query the database replica of the English Wikipedia for the creation date of each hoax article, defined as the timestamp of the first revision recorded in the database. Figure 1 shows a summary of the number hoaxes created over time, with the majority of hoaxes appearing in the period 2005–2007, and a decline starting in 2008. This observed behavior can be in part explained by the fact that the Wikipedia community started patrolling new pages in November of 2007 [22, 42] and is also consistent with the well-known peak of activity of the English Wikipedia community [17].

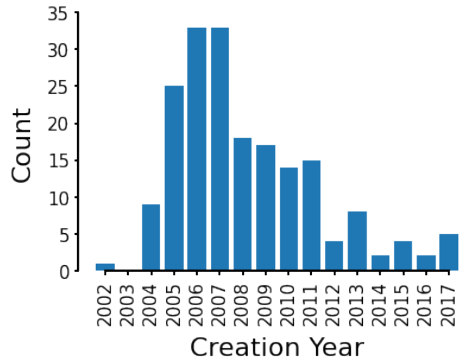


Figure 1: The number of detected hoaxes in the English Wikipedia over time [41].

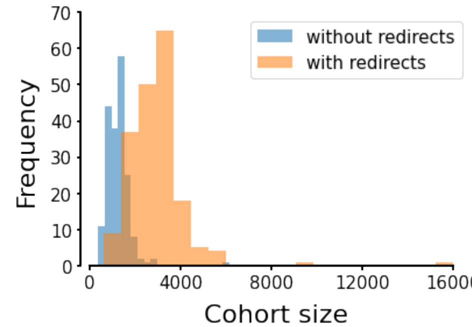


Figure 2: Histogram of the number of legitimate articles created on the same day (cohort) as hoaxes in our dataset. Cohort size was computed with and without resolving redirects.

Finally, to build the cohort of each hoax, we queried the Wikipedia database replica for all legitimate articles created on the same day. Since Wikipedia entries are often accessible through different titles, in collecting the cohort, we resolved all redirects page created in the same day of the hoax. Treating these redirects as separate entries would inflate the cohort size and could skew traffic statistics used later for estimating the level of online attention. Figure 2 shows the effect that redirects have on the size of each cohort. In some cases, failing to account for redirects can increase the size of cohorts to up to 16, 000 articles.

2.2 Appearance Characteristics Analysis

To understand the differences between each hoax and its cohort members, we analyzed their appearance features, inspired by the work of Kumar et al. [22] who, in addition to appearance features, studied network, support, and editor features for both hoax and legitimate articles [22]. We considered the following features: plain text length, ratio of plain to markup text, density of wiki-links, and density of external links. The plain text length is the number of words in an article after removing all the wiki markup. The ratio of plain to markup text is the number of words obtained after removing all markup in the body of the article, divided by the number of words before removal. The density of wiki-links is the number of wiki-links per 100 words, counted before markup removal. Finally, the density of external links is defined similarly as the density of wiki-links, but for links to external Web pages.

To be able to calculate these features for each hoax and its cohort, we consulted the API to extract their plain text. This was made possible by using the *TextExtracts* extension [24]. To get the wiki markup for each article, we used the revisions API, which is part of the MediaWiki action API [26]. After retrieving the content of all articles, the aforementioned features were calculated. A regular expression was used to count the number of words in plain and markup text. To find the wiki and external links within each article we used *wikitextparser* [1] as it made it easy to extract and count the number of links and hence calculate the densities.

Aside from the plain to markup ratio, the chosen appearance features have very skewed distributions. To illustrate this point, fig. 3 shows the distribution of each score for five manually sampled

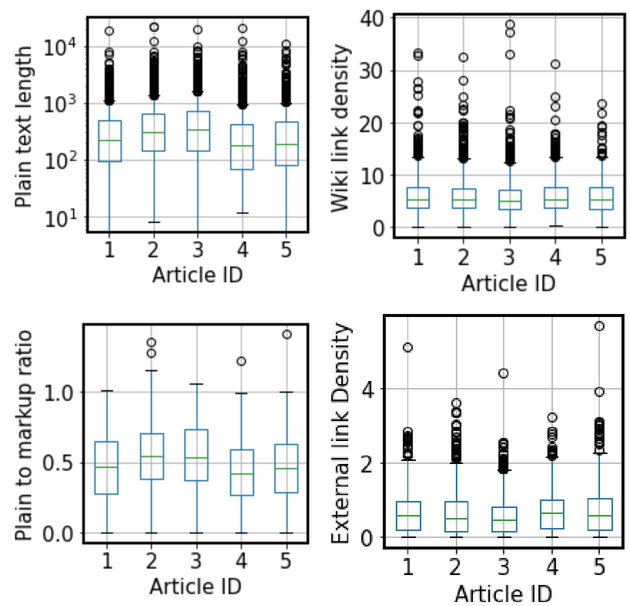


Figure 3: Distribution of appearance features for five manually sampled cohorts in our data.

cohorts in our data. For the plain text length, fig. 3 shows that the median is between 100 and 1, 000 words, yet there exist articles that reach and even exceed 10, 000 words. The same case persists in the wiki-link density — the median is under 10 links per 100 words, however some articles have up to 40 links, and similar for the other two features.

Thus, after collecting all the four features, we computed the modified z-score z' to compare different hoaxes together:

$$z' = \frac{x - \tilde{x}}{\mathrm{MAD}} \tag{1}$$

Where x is a feature measured on a hoax, $\tilde{x}$ the median value of the feature on the non-hoaxes, and MAD the median absolute deviation of x with respect to $\tilde{x}$. We chose to use z' instead of the regular z-score since it is more resilient to outliers [20].

2.3 Analyzing Wikipedia Traffic Data

To analyze the traffic that the articles in our dataset receive, we used the pageview dataset compiled by the Wikimedia foundation [38]. Since most of the hoaxes in our dataset were created in the period between 2005 and 2011, we have decided to use the pagecounts-raw dump, which spans from December 2007 to August 2016. This dataset contains the count of non-unique HTTP requests made for each article in an hourly time frame, collected by the proxy server of Wikipedia [7], along with request title and additional metadata. However, to effectively and readily analyse the view count for hoaxes and their cohorts, we pre-processed pagecounts-raw to resolve redirects, filter unwanted entries, and clean illegal titles.

Pre-processing the data was performed over the following three steps. First, the raw data was filtered. The filtration process selected only entries related to the English Wikipedia project while removing all pages from namespaces other than the 'main' MediaWiki namespace. Second, the filtered data was cleaned from illegal titles. Illegal titles were discarded by removing characters which are not allowed in Wikipedia page titles [37, cf. 'Page Restrictions']. The pound sign '#' is considered illegal only if it is the first character in a title; otherwise it indicates a subsection within a page. Hence, a title including '#' is discarded only in the former case. In addition to removing illegal characters, we decoded common URL-encoded characters (e.g. '%20') and replaced any space with an underscore character. Third, to resolve redirects, the Toolforge was consulted to extract all the redirects within the main namespace of the English Wikipedia. At the end of this procedure, each entry in the cleaned data contained only two fields — the page title and the view count for each hour. The result was a cleaned and filtered hourly dataset of the view count for pages within the main namespace of the English Wikipedia.

3 RESULTS

3.1 Appearance Features

We start by analyzing the appearance features of hoaxes relative to the non-hoaxes in their cohort. Figure 4a shows that most hoaxes have either similar or slightly smaller plain text length compared to that of their cohorts. We also observe the presence of several outliers, indicating that a subset of hoaxes in our sample tend to have unusually higher word counts. This is consistent with the results of Kumar et al. [22], who observed that 'successful' hoaxes (i.e., that have gone undetected for at least 1 month) have a median plain text length of 134 words — almost twice as large as that of legitimate articles. However, the analysis of Kumar et al. [22] differs from ours in multiple ways. First, they used a different, larger set of hoaxes. Second, they used a matching procedure to compare each hoax to only one legitimate article created on the same day. They also considered other types of articles, such as wrongly flagged articles and failed hoaxes. Another potential differentiating factor is the method of extraction for the plain text, markup content, and links for each page, which might contribute to not obtaining exactly the same results.

Figure 4b shows that hoaxes tend to have similar density of wiki-links than non-hoaxes. This is important, since to quantify online attention toward a topic we compute the volume of traffic to the wiki-link neighbors of an article. Thus, in the following analysis on traffic, we can safely exclude potential confounding factors due to different linking patterns between hoaxes and non-hoaxes.

Figures 4c and 4d show the distributions of the ratio of plain to markup text and of external link density, respectively. Aside from a few outliers, hoaxes almost always contain more plain text than markup text, compared with non-hoaxes. This is consistent with the findings of [22], who observed that, on average, 58% of a legitimate article, 71% of a successful hoax, and 92% of a failed hoax is just plain text.

In summary, hoaxes tend to have more plain text than legitimate articles and fewer links to external web pages outside of Wikipedia. This means that non-hoax articles, in general, contain more references to links residing outside Wikipedia. Such behavior is not unexpected, considering it would require a significant effort for the author of a hoax to craft external resources outside of Wikipedia that could be pointed at from it.

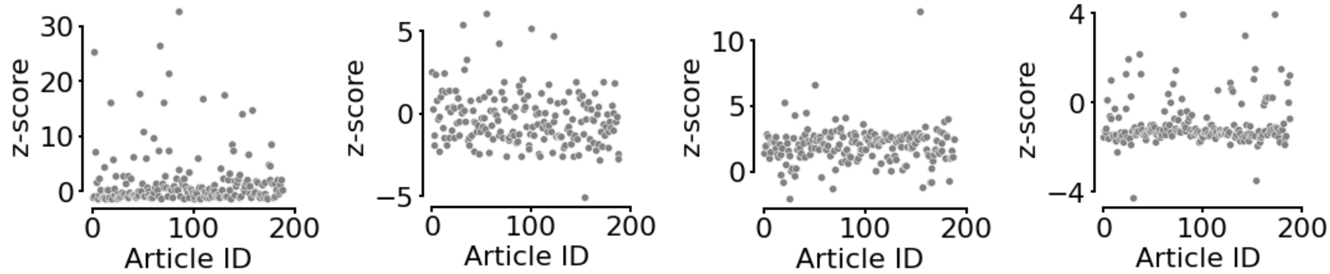
3.2 Traffic Analysis

Recall that the cohort of a hoax is defined as all the non-hoax articles created on the same day it was created. To understand the nature of the relationship between the creation of hoaxes and the attention their respective topics receive, we first seek to quantify the relative volume change before and after this creation day. Here, a *topic* is defined as all of the (non-hoax) neighbors linked within the contents of an article i.e., its (non-hoax) out-links. Traffic to Wikipedia is known to fluctuate following circadian and weekly patterns, and is likely to depend on a host of additional, unknown factors, such as the relative popularity of Wikipedia over the years, the total number and geographic distribution of web users [45], etc. To account for these potential confounding factors, [7] proposed to quantify the volume change in a way that controls for the circadian rhythm and the fluctuating nature of traffic on the Web [7, 34]. They have shown that studying traffic over a 14-day observation window, 7 days before and after the creation day, allows to consider for both short spikes in attention and weekly changes in traffic. The relative volume change is defined as:

$$\frac{\Delta V}{V} = \frac{V^{(b)} - V^{(a)}}{V^{(b)} + V^{(a)}} \tag{2}$$

where $V^{(b)}$ and $V^{(a)}$ are respectively the median traffic to neighbors in the 7 days before and after the creation of the article. According to eq. (2), $\Delta V/V > 0$ when the majority of traffic occurs before an article is created, i.e., attention toward the topic of the articles precedes its creation. When $\Delta V/V < 0$ attention tends to follow the creation of the hoax. Note that our traffic data covers a period spanning from December 2007 to August 2016. Since not all hoaxes in our dataset fell within that time frame, $\Delta V/V$ was calculated only for the 83 hoaxes (and their cohorts) whose creation dates fell within that period.

Having defined a way to quantify whether traffic to a given article preceded to followed its creation, our goal is thus determine whether it is the case that hoaxes tend to have a greater $\Delta V/V$ than legitimate articles. Unfortunately, we know very little about the distribution of $\Delta V/V$ over multiple pages, and how it has changed over the course of the history of Wikipedia. However, if hoaxes do



(a) Plain text length
(b) Wiki-link density
(c) Plain to markup text ratio
(d) External link density

Figure 4: Modified z-scores for all hoaxes in our sample relative to non-hoax articles in their cohorts for the four appearance features we considered. Hoaxes tend to have similar or slightly smaller count of plain text words (however with several higher-count outliers), lower external link density, higher plain to markup text ratio, and similar wiki-link density.

not differ from legitimate articles, then on average the difference the $\Delta V/V$ of a hoax and that of its cohorts should be zero. Therefore, we define:

$$D = \frac{\Delta V}{V} - \mathsf{E}\left[\frac{\Delta V'}{V'}\right] = \frac{\Delta V}{V} - \frac{1}{n}\sum_{i=1}^{n}\frac{\Delta V_i'}{V_i'} \tag{3}$$

where $\mathsf{E}\left[\frac{\Delta V'}{V'}\right]$ indicates the expected $\Delta V'/V'$ of legitimate articles. Thus, $D > 0$ indicates that, compared to its cohort, a hoax accumulates more attention preceding its creation; $D < 0$ indicates the opposite.

To estimate whether, prior to the creation of a hoax article, its topic receives more traffic than the topics of legitimate (i.e., non-hoax) articles in the same cohort, we computed the average difference D (see Equation (3)) and computed the confidence interval for the sample mean by resampling. We used in particular bootstrapping. Bootstrapping is a method in which the observations in a sample of size n are randomly resampled for a certain number of times to approximately simulate the sampling distribution of a parameter of interest. In our case, we are interested in the sampling distribution of the mean. Since each resampled vector, with size n, is randomly generated from the original sample, the mean will vary every time a new vector is created. By considering the means of all those vectors, we are able to sample the distribution of means, and compute the 95% CI of the mean of D. To perform bootstrapping, we resampled the original list of D values $10,000$ times with replacement. According to the Central Limit Theorem (CLT), the distribution of sample means approximates the normal distribution with the increase of sample size n [9]. The CLT holds true regardless of the original distribution of data.

Figure 5 shows the distribution of $\Delta V'/V'$ computed on its cohort, and its value of $\Delta V/V$, along with cohort mean, for a manually selected sample of hoaxes in our data. In general, we observe a trend in which hoaxes tend to have greater $\Delta V/V$ than their cohort. This means that the online attention toward hoaxes tends to more often precede their creation when compared to the attention their cohort members receive. Based on eq. (3), we can confirm this observation by calculating the difference between the mean $\Delta V'/V'$ of the cohort and the $\Delta V/V$ of the hoax. We measured a $D > 0$ in 75 out of 83 of the hoaxes in our data. The histogram in fig. 6 shows the distribution of the differences, and shows that the mean is approximately equal to 0.123, with a bootstrapped 95% confidence interval of $(0.1227, 0.1234)$.

Figure 7 shows the distribution of the means for each of the $10,000$ resampled vectors. It is worth noting that all of the means returned were positive, implying a greater $\Delta V/V$ for the hoax. In summary, our measured D indicates that the generation of hoaxes in Wikipedia is associated with prior consumption of information, in the form of online attention.

4 DISCUSSION AND FUTURE WORK

In this study we analyzed the role of attention within Wikipedia. More specifically, we focused on the relationship between online attention, as measured by traffic volume to Wikipedia articles, and the creation of intentionally false or misleading hoax articles. In particular, we focused on the change in traffic to neighboring articles around the creation of each hoax, using a list of known hoaxes compiled by Wikipedia contributors. We collected data about a set of known hoaxes from this list, along with their cohort formed by all non-hoax articles created on the same day as each hoax. These cohorts were used as the baseline for comparison purposes. A positive association was found between preceding traffic and the creation of hoaxes. To do so, we leveraged a large dataset of Wikipedia traffic logs. For each day of creation of hoax d, we collected the view count for $d \pm 7$ days for the out-link neighbors each hoax and of its cohort.

Following prior work [7], to assess the allocation of attention during that period, we calculated the relative traffic volume change, which accounts for potential confounding factors due to traffic fluctuations. We observe that almost 75% of hoaxes have a higher $\Delta V/V$ than their respective cohort. This indicates that, on average, hoaxes tend to have more traffic accumulated before their creation than after. We used resampling to determine the confidence interval of the average difference in relative traffic change between hoaxes and non-hoax articles.

This study has some limitations that need to be acknowledged. First of all, our results are based on a list of only 83 hoaxes. Even though we originally collected a dataset that was more than double in size, we were limited by the fact that not all hoaxes were covered in our traffic dataset. Prior work relied on a broader definition of

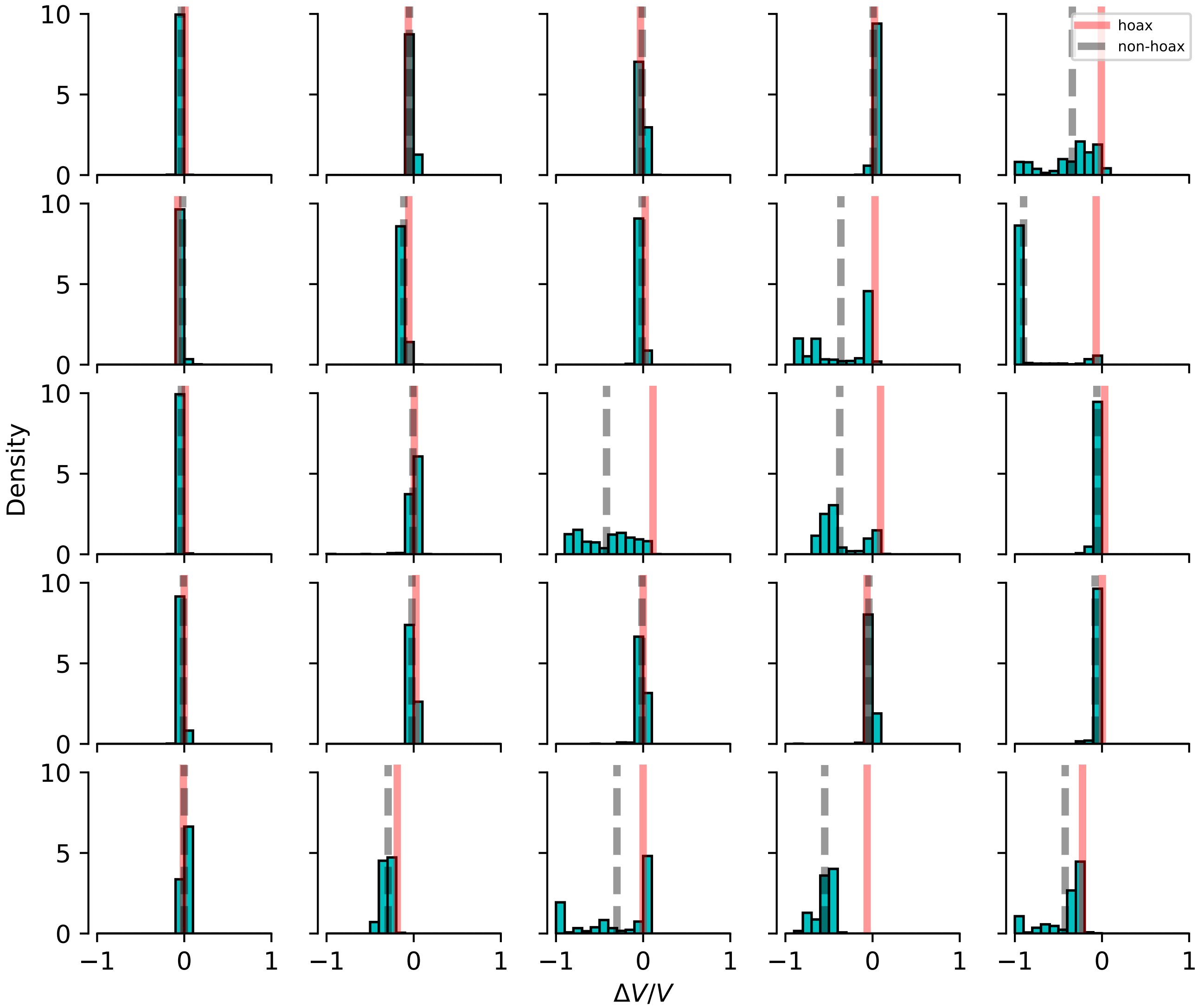


Figure 5: The distribution of the $\Delta V'/V'$ values for each cohort (turquoise blue histograms) in comparison to the $\Delta V/V$ of the respective hoax (black dashed line). The $\Delta V/V$ of hoaxes tend to, in general, be higher than the mean of their cohorts (red solid line).

'hoaxes' based on the 'New Page Patrol' process [22]. However, access to this list was not public due to the nature of the data, and thus our results were based on a smaller public list of known hoaxes [41]. Future work should extend our results to larger samples of hoaxes to ensure consistent results with prior work. Another limitation of our analysis is that the definition of topic was based only on the out-link neighbors. This was necessary since the data about in-links to hoaxes is not publicly accessible in the Wikipedia database, presumably due to the page deletion process. Instead, we parsed the body of articles (hoaxes and non-hoaxes) and extracted all wiki-links markup, which yield only out-links. In the future, we would like to extend our analysis to include in-links as well. Third, our traffic dataset is based on an older definition of pagecounts, which is affected by known limitations, including the presence of automated crawlers, which are known to generate large amounts of hits in a short periods of time. To mitigate for the presence of this type of outliers, in our definitions of traffic volume we rely on the median instead of the mean, which is more robust to outliers. However, in the future we would like to include a more recent traffic dataset that is not affected by this and other biases.

In conclusion, our study sheds light on an important factor affecting the supply of disinformation on the Web. Future work should extend our results to venues other than Wikipedia, for example social media platforms like Facebook or Twitter. In addition, other types

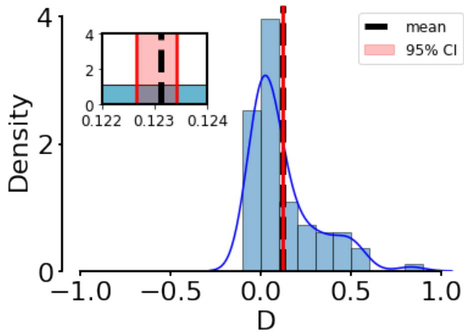


Figure 6: Histogram of the relative traffic change differences D (see Equation (3)). The black dashed line is the sample mean, and the red area the 95% bootstrapped CI. The blue solid line is a kernel density estimate. The inset shows the sample mean relative to the confidence interval.

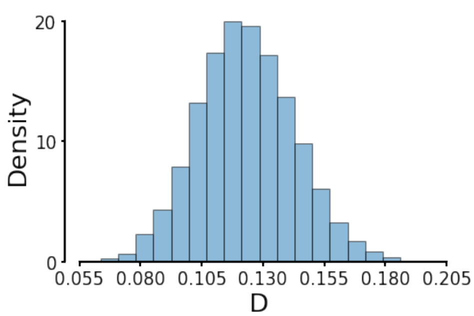


Figure 7: The sampling distribution of means obtained by bootstrapping $10,000$ samples with replacement.

of media (like video, audio, etc.) should be considered — hoaxes do not only come in the form of textual articles, and attention is an effective incentive for people to keep spreading more disinformation. Future work should also consider studying the role of attention in versions of Wikipedia other than English. We expect similar trends to ones observed here to apply to non-English language editions as well. However, the signal may be weaker owing to lower traffic volume of non-English language editions. A comparative analysis of the role of attention in the supply of disinformation across cultures could shed more light about these type of threats to the Web as a whole.

REFERENCES

[1] 5j9. 2022. Github Inc., – wikitextparser. https://github.com/5j9/wikitextparser. Last accessed: 09-March-2022.

[2] B. Thomas Adler, Luca de Alfaro, and Ian Pye. 2010. Detecting Wikipedia Vandalism using WikiTrust - Lab Report for PAN at CLEF 2010. In *CLEF 2010 LABs and Workshops, Notebook Papers, 22-23 September 2010, Padua, Italy (CEUR Workshop Proceedings, Vol. 1176)*, Martin Braschler, Donna Harman, and Emanuele Pianta (Eds.). CEUR-WS.org, Aachen, Germany, n.p. http://ceur-ws.org/Vol-1176/CLEF2010wn-PAN-AdlerEt2010.pdf

[3] Hunt Allcott and Matthew Gentzkow. 2017. Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives* 31, 2 (May 2017), 211–36. https://doi.org/10.1257/jep.31.2.211

[4] Marco Amoruso, Daniele Anello, Vincenzo Auletta, Raffaele Cerulli, Diodato Ferraioli, and Andrea Raiconi. 2020. Contrasting the Spread of Misinformation in Online Social Networks. *Journal of Artificial Intelligence Research* 69 (nov 2020), 847–879. https://doi.org/10.1613/jair.1.11509

[5] Tim Berners-Lee, Robert Cailliau, Ari Luotonen, Henrik Frystyk Nielsen, and Arthur Secret. 1994. The World-Wide Web. *Commun. ACM* 37, 8 (aug 1994), 76–82. https://doi.org/10.1145/179606.179671

[6] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information Credibility on Twitter. In *Proceedings of the 20th International Conference on World Wide Web* (Hyderabad, India) *(WWW '11)*. Association for Computing Machinery, New York, NY, USA, 675–684. https://doi.org/10.1145/1963405.1963500

[7] Giovanni Luca Ciampaglia, Alessandro Flammini, and Filippo Menczer. 2015. The production of information in the attention economy. *Scientific Reports* 5 (2015), 9452. https://doi.org/10.1038/srep09452

[8] Caitlin Dewey. 2015. The story behind Jar'Edo Wens, the longest-running hoax in Wikipedia. Last updated: 24-October-2018.

[9] William Feller. 1991. *An Introduction to Probability Theory and Its Applications* (2nd ed.). John Wiley & Sons, Inc., New York, USA.

[10] Ruth García-Gavilanes, Anders Mollgaard, Milena Tsvetkova, and Taha Yasseri. 2017. The memory remains: Understanding collective memory in the digital age. *Science Advances* 3, 4 (2017), e1602368. https://doi.org/10.1126/sciadv.1602368

[11] R. Stuart Geiger and Aaron Halfaker. 2013. When the Levee Breaks: Without Bots, What Happens to Wikipedia's Quality Control Processes?. In *Proceedings of the 9th International Symposium on Open Collaboration* (Hong Kong, China) *(WikiSym '13)*. Association for Computing Machinery, New York, NY, USA, Article 6, 6 pages. https://doi.org/10.1145/2491055.2491061

[12] Michael H. Goldhaber. 1997. The attention economy and the Net. *First Monday* 2, 4 (Apr. 1997), n.p. https://doi.org/10.5210/fm.v2i4.519

[13] Yevgeniy Golovchenko, Cody Buntain, Gregory Eady, Megan A. Brown, and Joshua A. Tucker. 2020. Cross-Platform State Propaganda: Russian Trolls on Twitter and YouTube during the 2016 U.S. Presidential Election. *The International Journal of Press/Politics* 25, 3 (2020), 357–389. https://doi.org/10.1177/1940161220912682

[14] Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. Fake news on Twitter during the 2016 U.S. presidential election. *Science* 363, 6425 (2019), 374–378. https://doi.org/10.1126/science.aau2706

[15] Andrew Guess, Jonathan Nagler, and Joshua Tucker. 2019. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances* 5, 1 (2019), eaau4586. https://doi.org/10.1126/sciadv.aau4586

[16] Andrew M. Guess, Brendan Nyhan, and Jason Reifler. 2020. Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour* 4, 5 (01 May 2020), 472–480. https://doi.org/10.1038/s41562-020-0833-x

[17] Aaron Halfaker, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. 2013. The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist* 57, 5 (2013), 664–688. https://doi.org/10.1177/0002764212469365

[18] Aaron Halfaker and John Riedl. 2012. Bots and Cyborgs: Wikipedia's Immune System. *IEEE Computer* 45, 3 (2012), 79–82. https://doi.org/10.1109/MC.2012.82

[19] Manoj Harpalani, Michael Hart, Sandesh Singh, Rob Johnson, and Yejin Choi. 2011. Language of Vandalism: Improving Wikipedia Vandalism Detection via Stylometric Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Portland, Oregon, USA, 83–88.

[20] Boris Iglewicz and David C. Hoaglin. 1993. How to Detect and Handle Outliers. In *The ASQC Basic References in Quality Control: Statistical Techniques*, Edward F. Mykytka (Ed.). Vol. 16. ASQC, Milwaukee, WI, USA.

[21] Gary King, Jennifer Pan, and Margaret E. Roberts. 2017. How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument. *American Political Science Review* 111, 3 (2017), 484–501. https://doi.org/10.1017/S0003055417000144

[22] Srijan Kumar, Robert West, and Jure Leskovec. 2016. Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes. In *Proceedings of the 25th International Conference on World Wide Web* (Montréal, Québec, Canada) *(WWW '16)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 591–602. https://doi.org/10.1145/2872427.2883085

[23] David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, and Jonathan L. Zittrain. 2018. The science of fake news. *Science* 359, 6380 (2018), 1094–1096. https://doi.org/10.1126/science.aao2998

[24] MediaWiki contributors. 2021. Extension:TextExtracts — MediaWiki. https://www.mediawiki.org/w/index.php?title=Extension:TextExtracts&oldid=4940004. Last accessed: 9-March-2022.

[25] MediaWiki contributors. 2022. API:Main page — MediaWiki. https://www.mediawiki.org/w/index.php?title=API:Main_page&oldid=5019333. Last accessed: 9-March-2022.

[26] MediaWiki contributors. 2022. API:Revisions — MediaWiki. https://www.mediawiki.org/w/index.php?title=API:Revisions&oldid=5037632. Last accessed: 9-March-2022.

[27] Chitu Okoli, Mohamad Mehdi, Mostafa Mesgari, Finn Årup Nielsen, and Arto Lanamäki. 2014. Wikipedia in the eyes of its beholders: A systematic review of scholarly research on Wikipedia readers and readership. *Journal of the Association for Information Science and Technology* 65, 12 (2014), 2381–2403. https://doi.org/10.1002/asi.23162

[28] Martin Potthast, Benno Stein, and Robert Gerling. 2008. Automatic Vandalism Detection in Wikipedia. In *Advances in Information Retrieval*, Craig Macdonald, Iadh Ounis, Vassilis Plachouras, Ian Ruthven, and Ryen W. White (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 663–668.

[29] Diego Sáez-Trumper. 2019. Online Disinformation and the Role of Wikipedia. *CoRR* abs/1910.12596 (2019), 17 pages. arXiv:1910.12596

[30] Similarweb LTD. 2022. Top Websites Ranking. https://www.similarweb.com/top-websites/. Last accessed: Mar-09-2022.

[31] Herbert A Simon. 1971. Designing organizations for an information-rich world. In *Computers, communications, and the public interest*, Martin Greenberger (Ed.). Vol. 72. Johns Hopkins Press, Baltimore, 37–52.

[32] Koen Smets, Bart Goethals, and Brigitte Verdonk. 2008. Automatic Vandalism Detection in Wikipedia: Towards a Machine Learning Approach. In *Proceedings of the 2008 AAAI Workshop on Wikipedia and Artificial Intelligence: An Evolving Synergy*. AAAI, Palo Alto, CA, USA, 43–48.

[33] Kate Starbird. 2017. Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter. In *Proc. of the International AAAI Conference on Web and Social Media*. AAAI, Palo Alto, CA, USA, 230–239. https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15603

[34] K. Thompson, G.J. Miller, and R. Wilder. 1997. Wide-area Internet traffic patterns and characteristics. *IEEE Network* 11, 6 (1997), 10–23. https://doi.org/10.1109/65.642356

[35] Fernanda B. Viégas, Martin Wattenberg, and Kushal Dave. 2004. Studying Cooperation and Conflict between Authors with History Flow Visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vienna, Austria) *(CHI '04)*. Association for Computing Machinery, New York, NY, USA, 575–582. https://doi.org/10.1145/985692.985765

[36] Claire Wardle and Hossein Derakhshan. 2017. *Information disorder: toward an interdisciplinary frameworkfor research and policy making*. Technical Report DGI(2017)09. Council of Europe, Strasbourg, FR.

[37] Wikimedia Foundation, Inc. 2022. Help: Page name. https://meta.wikimedia.org/wiki/Help:Page_name. Last accesse: 09-March-2022.

[38] Wikimedia Foundation, Inc. 2022. Page view statistics for Wikimedia projects. https://dumps.wikimedia.org/other/pagecounts-raw/. Last accessed: 09-March-2022.

[39] Wikimedia Foundation, Inc. 2022. Wikimedia Statistics – English Wikipedia. https://stats.wikimedia.org/#/en.wikipedia.org. Last accessed 09-March-2022.

[40] Wikipedia contributors. 2021. Wikipedia:Vandalism does not matter. https://en.wikipedia.org/wiki/Wikipedia:Vandalism_does_not_matter. Last accessed: Mar-09-2022.

[41] Wikipedia contributors. 2022. Wikipedia:List of hoaxes on Wikipedia. https://en.wikipedia.org/wiki/Wikipedia:List_of_hoaxes_on_Wikipedia. Last accessed: Mar-09-2022.

[42] Wikipedia contributors. 2022. Wikipedia:New pages patrol. https://en.wikipedia.org/wiki/Wikipedia:New_pages_patrol. Last accessed: Mar-09-2022.

[43] Wikipedia contributors. 2022. Wikipedia:Size of Wikipedia. https://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia. Last accessed: Mar-09-2022.

[44] Wikitech contributors. 2021. Portal:Toolforge — Wikitech. https://wikitech.wikimedia.org/w/index.php?title=Portal:Toolforge&oldid=1932575. Last accessed: 9-March-2022.

[45] Taha Yasseri, Robert Sumi, and János Kertész. 2012. Circadian Patterns of Wikipedia Editorial Activity: A Demographic Analysis. *PLOS ONE* 7, 1 (Jan. 2012), 1–8. https://doi.org/10.1371/journal.pone.0030091

[46] Savvas Zannettou, Tristan Caulfield, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. 2019. Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web. In *Companion Proceedings of The 2019 World Wide Web Conference* (San Francisco, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 218–226. https://doi.org/10.1145/3308560.3316495

[47] Ahmad Zareie and Rizos Sakellariou. 2021. Minimizing the spread of misinformation in online social networks: A survey. *Journal of Network and Computer Applications* 186 (2021), 103094. https://doi.org/10.1016/j.jnca.2021.103094