Proceedings of the Fourteenth International AAAI Conference on Web and Social Media (ICWSM 2020)

The Pushshift Telegram Dataset

Jason Baumgartner,[1] **Savvas Zannettou,**[2,*] **Megan Squire,**[3] **Jeremy Blackburn**[4,*]
[1]Pushshift.io, [2]Max-Planck-Institut für Informatik, [3]Elon University, [4]Binghamton University, [*]iDRAMA Lab
jason@pushshift.io, szannett@mpi-inf.mpg.de, msquire@elon.edu, blackburn@cs.binghamton.edu

Abstract

Messaging platforms, especially those with a mobile focus, have become increasingly ubiquitous in society. These mobile messaging platforms can have deceivingly large user bases, and in addition to being a way for people to stay in touch, are often used to organize social movements, as well as a place for extremists to congregate.

In this paper, we present a dataset from one such mobile messaging platform: Telegram. Our dataset is made up of over 27.8K channels and 317M messages from 2.2M unique users. To the best of our knowledge, our dataset comprises the largest and most complete of its kind. In addition to the raw data, we also provide the source code used to collect it, allowing researchers to run their own data collection instance. We believe the Pushshift Telegram dataset can help researchers from a variety of disciplines interested in studying online social movements, protests, political extremism, and disinformation.

1 Introduction

While the modern social media ecosystem is certainly dominated by a few major players, e.g., Facebook, Twitter, Reddit, etc., there are a variety of lesser known platforms with high active user bases. One such platform is Telegram, a mobile messaging app that has broader social media style features. Telegram is particularly interesting due to its use by social movements to disseminate information. For example, the Hong Kong protesters made use of Telegram to organize some of their activities. Unfortunately, not all uses of Telegram are generally positive: for example, Telegram is also home to a vast network of right wing extremist groups, who use it to organize as well as disseminate racist and violent ideology.

In this paper, we present the Pushshift Telegram Dataset. To the best of our knowledge, our dataset represents, by far, the largest collection of Telegram data made available to the public. While our dataset is available for download as static snapshots, it is also under periodic collection. The most current snapshot is available at https://zenodo.org/record/3607497. At the time of this writing, the Pushshift Telegram Dataset comprises 27,801 *channels* and 317,224,715 *messages* from 2,200,040 unique *users*.

Also, we make publicly our data collection source code at https://github.com/pushshift/telegram. The code can be reused by other researchers that want to take information from specific Telegram channels for their research.

In the remainder of this paper, we provide some background on Telegram, describe our dataset, and perform some general characterization of the dataset.

2 Background & Related Work

Telegram was started in August 2013 as an encrypted instant messaging platform. Telegram users provide a telephone number to access the service. Messages between users are stored in a centralized cloud-based storage and (with the exception of "secret chats") can be accessed by multiple devices that have been linked to a single user's account. In addition to the user-to-user secure messaging features, Telegram added broadcast *channels* for one-to-many communication in September 2015. Such broadcast channels can be created by any Telegram user, and other Telegram users can join or subscribe to the channel to read its content. Content in a channel mostly consists of messages, which take the form of text, still images, audio files, video files, and so on. Other messages sent in channels are *service messages*, for example status messages or errors, but these are largely invisible to regular users of the platform. These broadly-available and widely used Telegram channels are designed for information dissemination thus are the subject of this dataset project.

To illustrate the inner workings of the Telegram platform, we show some informative screenshots from Telegram's interface in Fig. 1. When a user opens the Telegram mobile app, she is presented with the landing page (see Fig. 1(a)) that includes all the channels and groups that the user participates. For instance, in this specific example, the user participates in the Telegram and The Washington Post channels, as well as "The Post's Public Chat" group. By clicking on a specific group or channel, the user is presented with the contents of the channel/group. For instance, in Fig. 1(b), we show the content when a user clicks on The Washington Post channel. In this screen, the user can get useful information like how many subscribers the channel has and all the messages posted along with useful information such as how many times each post was seen. Finally, by clicking on the sidebar button (top left corner in Fig. 1(a)), a user is

presented with the sidebar (see Fig. 1(c)) that allows her to perform various actions like create a new channel, create a new group, change settings, etc.

Previous work that used Telegram. A large body of previous work studies the Telegram ecosystem itself, or uses data from Telegram to study specific emerging research problems. Specifically, (Anglano, Canonico, and Guazzone 2017) and (Satrya, Daely, and Nugroho 2016) study the artifacts generated by the Telegram android application on the Android platform. They propose a methodology that enables the reconstruction of important information involved in the Telegram application like list of contacts, messages exchanged between users, and information about the channels and groups that the user is involved in.

(Sutikno et al. 2016) study the features that are available in three popular messaging applications: WhatsApp, Viber, and Telegram. They conclude that Telegram is the best messaging application in terms of security, WhatsApp is best in terms of ease-of-use, while Viber is another good option with many integrated features like in-app voice calls.

(Abu-Salma et al. 2017) undertake a user study to understand whether end-users understand the security features that Telegram offers. They find that most of the users tend to use less secure features of the Telegram application, and they overall feel secure because they "are using a secure tool." Also, the authors analyze Telegram's user interface (UI) and find that it includes a lot of technical jargon and inconsistencies, and that some of the security features offered by the platform are not explained clearly in the UI.

(Nikkah, Miller, and Young 2018) study the use of Telegram by Iranian immigrants during their immigration procedure by observing 30 Iranian immigration-related groups. They show several examples of how Telegram bots are used enforce specific policies, how groups are moderated, and how pinned messages are posted by administrators.

(Hashemi and Chahooki 2019) study the group features of 900K Iranian channels and 300K Iranian groups on Telegram with the goal to identify high-quality groups (e.g., professional and business groups) over low-quality groups (e.g., dating groups). They find that high-quality groups tend to have more phone numbers in their messages, have longer messages, and have more user engagement when compared to low-quality groups.

(Asnafi et al. 2017) examine the use of the Telegram platform in various Iranian academic libraries. They collect data from channels posted on the websites of the libraries and find that users mostly talked about news and information, book introductions, and various files. They also note that most of the messages contained images.

(Akbari and Gabdulhakov 2019) study the ban of Telegram by Russia and Iran following Telegram's refusal to give access to encrypted data on the platform.

(Dargahi Nobari, Reshadatmand, and Neshati 2017) perform a structural and topical analysis of content posted on the Telegram platform. By collecting data from 2.6K groups/channels and 219K messages, they build a graph based on the mentions, concluding that the mentions graph is extremely sparse and includes several separated connected components that indicates that users have a low tendency to mention other users.

(Naseri and Zamani 2019) examine the dissemination of news within the Telegram platform by collecting news articles posted in five news outlets' official Telegram channels. Furthermore, they make their dataset publicly available to the research community.

Other previous research focuses on studying how terrorist organizations like ISIS use Telegram for various purposes like dissemination of content and ideology, as well as recruiting fighters and terrorists (Prucha 2016; Yayla and Speckhard 2017; Shehabat, Mitew, and Alzoubi 2017).

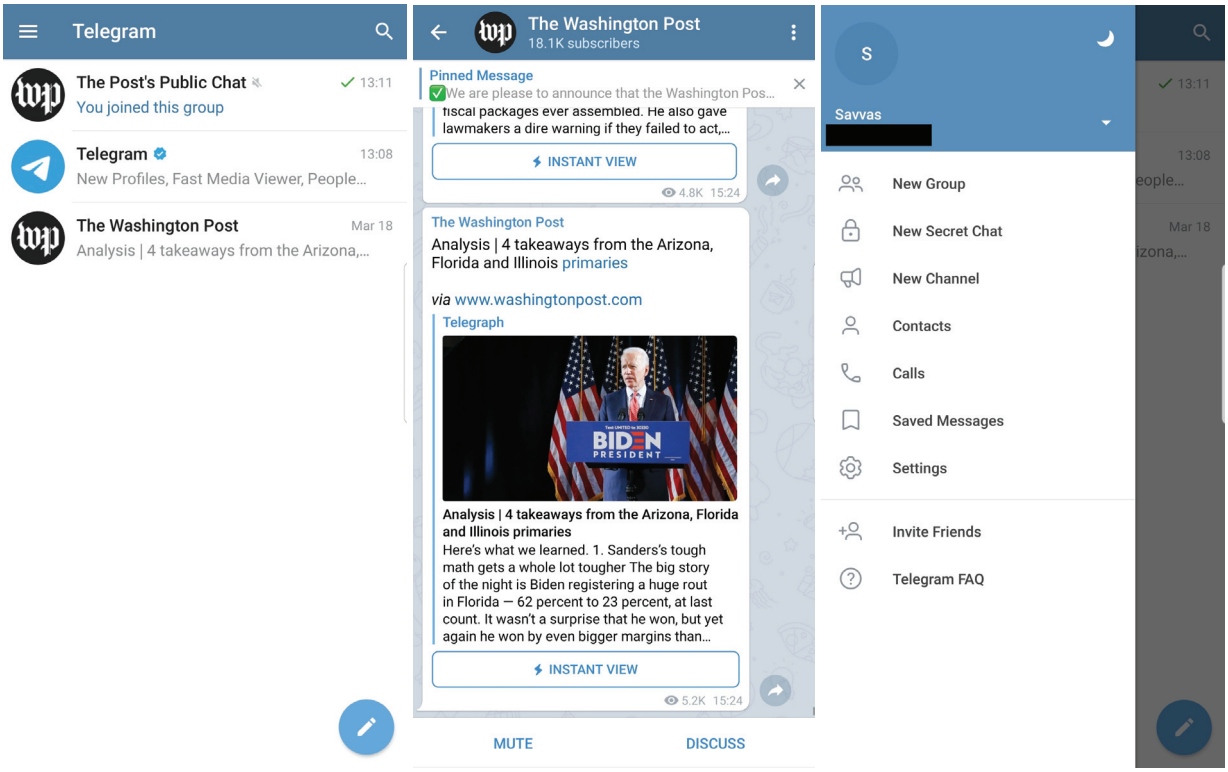
Other dataset papers. Since the main contribution of this paper is the rich dataset we release, here we provide a non-exhaustive overview of previous work that focuses on releasing datasets from various Web communities. (Garimella and Tyson 2018) study the WhatsApp messaging platform and they share their tools for obtaining public WhatsApp groups data, as well as a dataset from 178 groups that includes data for 454K messages posted from 45K users.

(Fair and Wesslen 2019) focus on Gab by releasing a dataset that includes 37M posts and 24M comments posted between August 2016 and December 2018. (Zignani et al. 2019) focus on Mastodon, a decentralized social network, by releasing a dataset of 5M posts: each post is associated with a label indicating whether the post or its content is inappropriate (according to the users that made the post). (Founta et al. 2018) provide a large scale dataset of tweets that are annotated on whether they are hateful, abusive, spam, or a normal tweet. To do this, they leverage crowd workers and annotate each tweet according to the majority agreement between all the crowd workers. (Brena et al. 2019) release a data collection pipeline and a large scale dataset related to the dissemination of news articles on Twitter. The data collection relies on a list of news sources and generates a large dataset of articles from these sources that are posted on Twitter. (Salem et al. 2019) focus on the Syrian War and release a carefully curated dataset of 804 news articles that are also labeled as real or fake. (Nørregaard, Horne, and Adalı 2019) release a set of 713K news articles collected between February and November, 2018, from 194 news sources. Also, for each news source in their dataset, they include ratings from eight different assessment sites that include, among others, scores related to the reliability, trust, bias, and journalistic standards of each news source.

3 Description of the Pushshift Telegram Dataset

Data Collection

Our data collection on Telegram is channel-based. Our goal is to collect data and metadata for publicly-viewable channels and public "chats". Typically each Telegram channel is set up by its owner to allow broadcast, or one-way, communication from a small set senders to a broader set of general channel users. However, each Telegram channel can also optionally have an associated "chat" that allows communication among all channel participants. Finally, some non-chat channels are not broadcast-only, but have themselves



(a) (b) (c)

Figure 1: Examples of the Telegram's inteface: a) the landing page with all the groups and channels of the user; b) example of a channel and the content posted within it; and c) the sidebar where a user can perform various actions.

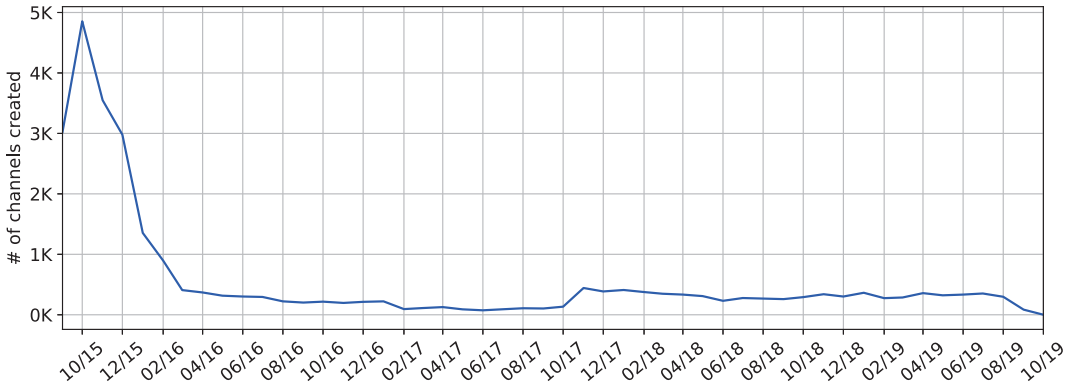


Figure 2: Number of channel creations per month in our dataset.

been configured to allow two-way communication between all participants on the channel.

To collect content and metadata from all of these types of channels and chats, we use Telethon[1], a Python interface to the Telegram API. We began with a seed list of approximately 250 primarily English-language broadcast channels and chat channels on Telegram. This initial list consists of 124 channels focusing on right-wing extremist politics and 137 channels cryptocurrency-related channels. To compile this list, we manually found mentions in news article mentioning Telegram channels related to right-wing extremist politics and cryptocurrencies, as well as from manual inspections within these channels. We elect to focus on these topics due to anecdotal evidence that emerged and indicated the exploitation of the Telegram platform by right-wing extremists (Anti-Defamation League 2019) and by users that coordinate scams related to the cryptocurrency mar-

[1]https://docs.telethon.dev/en/latest/

ket (Williams-Grut 2017). To grow the list of channels, we rely on the "forwarding" feature within Telegram wherein content can be forwarded between channels. Each time we discover content forwarded from a channel that is not already on our list, we add it, collect its data, and follow all of its channels. This snowball "crawling" approach has so far resulted in a list of 27,801 channels.

Channel data exposed by the Telegram API includes metadata such as the unique identification number, title, creation date, and various channel settings (e.g. usage configurations, administrator restrictions, and whether the channel is a bot), as well as the actual messages sent in the channel. Some calculated fields are also included, such as a current count of users, the identification number of the current "pinned" post, and a count of how many messages are unread.

Users communicate on Telegram channels by sending "messages" to the channel. Messages can either be original content posted to a channel, or can be forwarded content from another channel or from another user. The access rights to read a message can be restricted by the channel settings, for example allowing anyone to read the messages, or requiring users to "join" the channel before being able to view the messages. Message data exposed by the Telegram API includes metadata such as the datetime that the message was sent, whether it included media (e.g. images or video), and the identification number of the user who sent the message. For each message sender, Telegram provides details such as their username, whether they are a bot, whether they are a verified user, and so on.

We stored data and metadata for each channel/chat, message, and user in a PostgreSQL relational database management system. We currently do not collect media that accompanies the messages. This will change in the future as we find a more robust storage space solution.

Dataset Structure

Our static snapshot consists of three files:

- *Accounts metadata:* A newline delimited JSON file that includes the metadata for all the accounts that posted on any of the channels in our dataset (2.2M). Documentation for the fields included in this file are provided by the Telethon library and is available at https://tl.telethon.dev/constructors/user.html.
- *Channels metadata:* A newline delimited JSON file that includes the metadata for all the channels in our dataset (27.8K). The documentation for the fields included in the channels metadata file are available at https://tl.telethon.dev/types/chat_full.html and https://tl.telethon.dev/constructors/channel_full.html.
- *Messages:* A newline delimited JSON file including all the messages posted in these channels (317M). A documentation of the fields included in this file are provided by the Telethon library and is publicly available at https://tl.telethon.dev/constructors/message.html

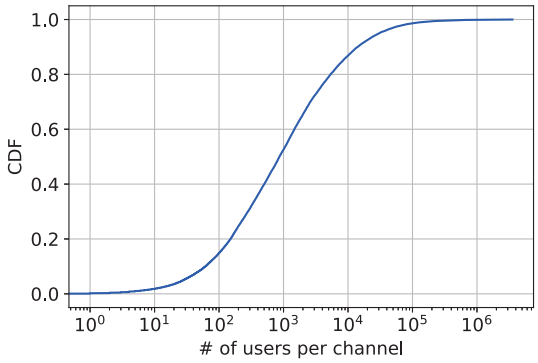


Figure 3: CDF of the number of registered users per channel.

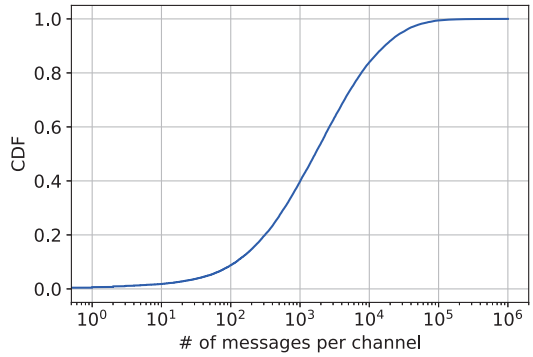


Figure 4: CDF of the number of messages per channel.

FAIR principles

We emphasize that our dataset fully conforms with the FAIR principles.[2] Specifically, our dataset is *Findable* since it is publicly available via the Zenodo service[3], which assigns a digital object identifier (DOI): 10.5281/zenodo.3607497.[4] Our dataset is also *Accessible* since it can be accessed by anyone in the world, while at the same time the format of the dataset is JSON, which is a widely used standard for data format. Due to the use of the widely known JSON standard, our dataset is *Interoperable* as almost every programming language has a library to work with data in JSON format. Finally, we release the full dataset and we provide a description of the dataset and the pointers to the Telethon API documentation that allow the interested researchers to understand the data and work with it. We ask that researchers cite our work if they use the dataset. Thus, our dataset is *Reusable*.

4 Dataset General Characterization

Channels. Overall, our dataset includes information from 27,801 channels. Fig. 2 shows how these channels are created over time: we plot the number of channels that are created per month. Telegram introduced channels as a feature

[2]https://www.go-fair.org/fair-principles/
[3]https://zenodo.org/
[4]https://zenodo.org/record/3607497

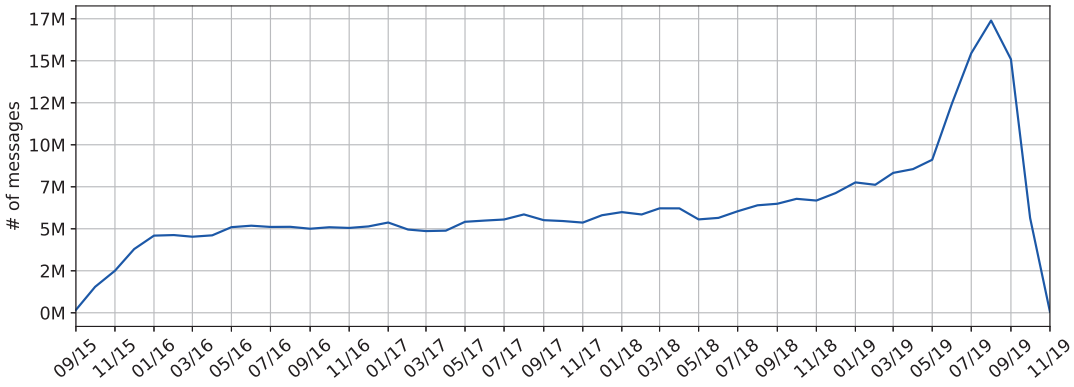


Figure 5: Temporal overview of the messages that are included in our dataset. We show the number of messages per month.

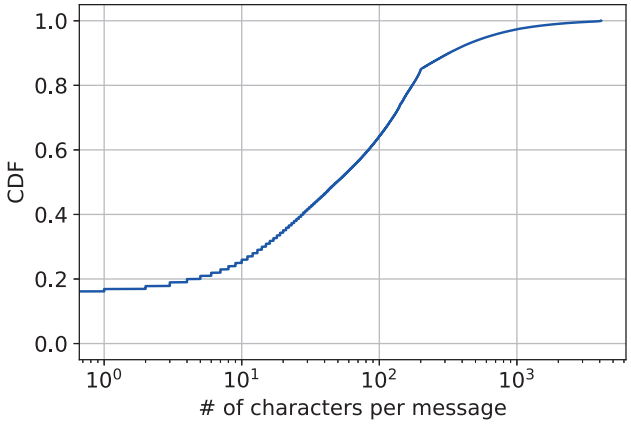


Figure 6: CDF of the number of characters per message.

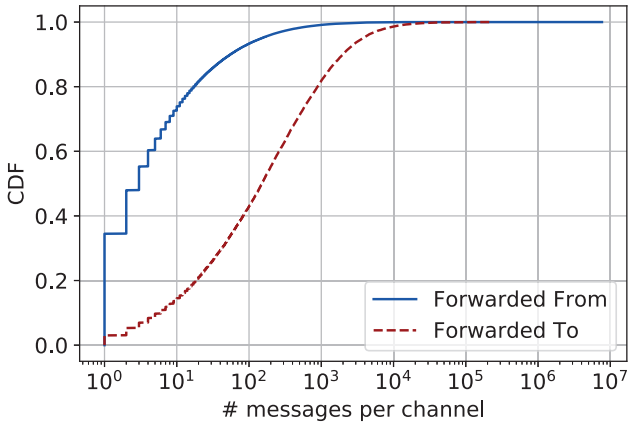


Figure 7: Distribution of number of messages per channel that were forwarded.

in September 2015, and our data includes 3,024 channels created in that very first month. We find that the channel creations spike on October 2015 with 4,854 channels created. New channel creations drop in subsequent months. Between March 2016-October 2017, and November 2017-August 2019, we observe a steady rate of channel creation, with the latter period having slightly more channels created.

Next, we look into the number of registered users per channel in our dataset. We find a mean number of registered users of 9823.3, while the median is 864. Fig. 3 shows the Cumulative Distribution Function (CDF) of the number of registered accounts per channel. We observe that the majority of the channels in our dataset (85.2%) have at least 100 registered users, while 47.4% of the channels have at least 1000 registered users.

Fig. 4 shows the number of messages per channel. We find, a mean of 6937.5 messages per channel, while the median is 1644. Also, we observe that 60% of the channels have at least 1K messages.

Messages. Overall, our dataset includes information for 317,224,715 messages. Out of those, 3,069,829 are just service messages indicating an event that happened on a specific channel (e.g., user adds, errors, and other status messages). The remaining 314,154,886 messages are actual messages that include content shared on the channel by users. Fig. 5 shows the monthly number of non-status messages that are shared in our dataset. We observe that, during 2016 and 2017, message activity is somewhat stable with around 5M messages per month, while we find a peak in message activity during August 2019 with approximately 17M messages. This peak in August 2019 coincides with the addition of 19,000 new Telegram users during the Hong Kong protests (Iqbal 2019).

Fig. 6 shows the CDF of the number of characters per message, which gives an intuition of how lengthy Telegram messages are. We find that messages have a mean number of characters equal to 152.2, while the median is 49 characters per message. We also observe that a substantial percentage of messages (16.1%) have an empty message, likely indicating that users are sharing messages with multimedia and no textual message.

Forwarded Messages. We also assess how common it is to forward messages across channels on Telegram. Other popular messaging apps like WhatsApp limit the number of times a specific user can forward a specific message, in an attempt to limit the spread of misinformation (Kastrenakes 2019). Previous research suggest that this counter measure can offer substantial delays in the propagation of misinformation (de Freitas Melo et al. 2019). Telegram does not limit the number of times a message can be forwarded. In our dataset, we find that 25,601,073 (8.1%) of all the messages are actually forwards from previously posted messages on other channels, or forwards from messages other users posted to their personal user channels. This indicates that forwarding messages across Telegram channels and between users and broadcast channels is a common operation on the platform. Next we investigate the users and broadcast channels involved in the forwarding of messages. We find 346,937 user channels or broadcast channels where forwarded messages originate. At the same time we find 27,039 broadcast channels where messages were forwarded to. Fig. 7 shows the CDF of the number of messages per channel for the the channels that messages were forwarded from and to. We observe that the number of messages-per-channel in the channels forwarded *from* is substantially smaller than the number of messages-per-channel in the channels forwarded *to*. The *from* mean is 946.8 vs *to* mean of 73.9, while *from* median is 151 vs *to* median of 3. This difference is due to the number of channels that messages were forwarded from is substantially larger (346K vs 27K). Overall, these findings indicate that message forwarding across Telegram channels is a popular feature of the Telegram platform, and this feature can be studied by researchers to assess the effect of this feature in emerging phenomena like the spread of false information, hateful content, etc.

Media. Next we look into whether messages contain media attachments, and what the different types of media attachments are in our dataset. We observe that 50.1% of the messages include media attachments with 48.1% of all the messages containing exactly one attachment. Fig 8 shows the distribution of all the media attachments into types according to Telegram. We observe that in our dataset, 53.8% of the attachments are photos, 29.4% are documents, 16.5% are Web pages, while the remaining 0.3% of the attachments are Polls, Geo locations, Games, Contacts, Venues, or Invoices. Fig. 9 shows the CDF of the number of media per channel in our dataset. We observe that Telegram is a rich source of media: more than half of the channels have shared over 1K media attachments.

Mentions and Hashtags. Here we investigate whether Telegram users are using mentions and hashtags in their messages. Overall, we find 7,638,430 (2.3%) messages contain hashtags, and 87,029,573 (27.7%) messages contain mentions. This indicates that Telegram users are more frequently mentioning other users or channels when compared to using hashtags in their messages. Table 1 report the 20 most popular mentions and hashtags that we find in our dataset, as well as their respective percentage over the overall number of messages that include hashtags/mentions. In

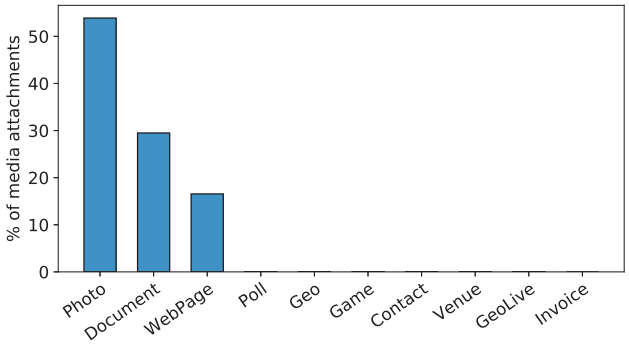


Figure 8: Distribution of the media attachments according to their type.

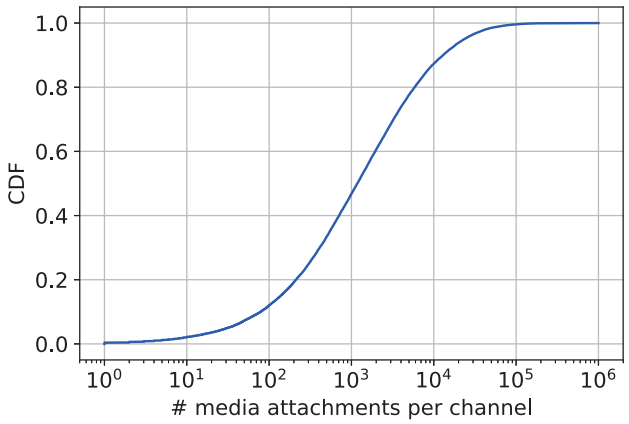


Figure 9: CDF of the number of media attachments posted per channel

terms of hashtags, we observe several international hashtags like #Trending, #Hot, #news. Interestingly, we also find divisive hashtags like #USAKillsYemeniPeople. In terms of mentions, we observe that most of them are related to Iran, hence highlighting the popularity of Iranian users on Telegram in general and in particular in our collected sample.

We also plot the number of occurrences per hashtag/mention in Fig. 10. For hashtags we find a mean number of 14.3 occurrences per hashtag, while 55.4% of the hashtags occur only once in our dataset. For mentions we find a mean number of 82.4 occurrences per mention, while 49.3% of the mentions occur only once in our dataset.

5 Discussion & Conclusion

In this paper, we described the Pushshift Telegram Dataset, to the best of our knowledge, the largest and most comprehensive Telegram dataset available to date. Our dataset includes over 317M messages from 2.2M unique users across 27.8K channels. In addition to the data, we also release the source code we used to collect it. We argue that this implementation will be extremely useful to researchers that are interested in studying Telegram and its various aspects.

Our dataset can be used by researchers to advance the

Hashtag	% (out of 7.6M)	Mention	% (out of 87M)
Trending	3.23%	tahlilgarantala	2.59%
Hot	2.17%	tahlilgarantala_ons	1.15%
Syria	1.07%	tahlilgarantala_absh	0.89%
ULTIMORA	0.81%	noticiasul	0.59%
request	0.75%	tahlilgarantala_Seke	0.55%
Step_News	0.64%	Twitter_Farsi	0.48%
Nima	0.55%	ilnair	0.27%
habrahabr	0.55%	TEQNYEBOT_BOT	0.27%
Geral	0.51%	MyAsriran	0.23%
USAKillsYemeniPeople	0.48%	iran_times	0.21%
Mundo	0.44%	haberbulteni	0.20%
Venezuela	0.42%	khabaredagh	0.19%
Amazon	0.42%	BI20ST	0.19%
news	0.41%	Farsna	0.18%
Marvel	0.39%	K_BER1	0.18%
Sport	0.38%	KNWAT	0.15%
tw	0.35%	T5TTI	0.15%
soc	0.34%	caspiankhabar	0.14%
Economia	0.34%	alalamnewstv	0.14%
SouCurioso	0.33%	Khabar_Varzeshi	0.14%

Table 1: Top 20 hashtags and mentions that we find in our dataset

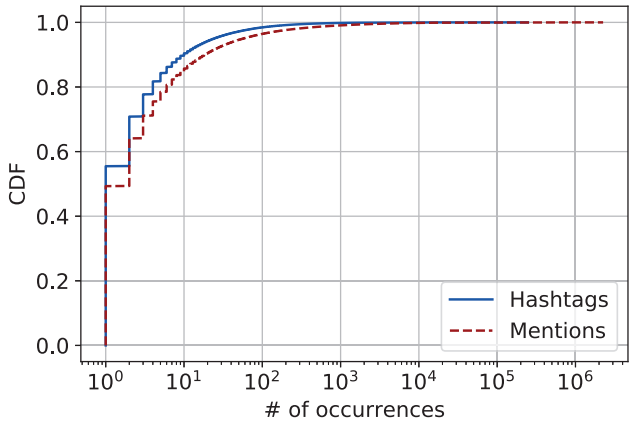


Figure 10: CDF of the number of occurrences per hashtag/mention.

frontier of knowledge around a variety of topics. For example, our dataset includes a large number of messages from right wing extremist groups, as well as more global movements like the Hong Kong protesters. In addition, the dataset includes a large number of channels related to finance and cryptocurrencies that might be useful for researchers studying finance scams and manipulation. Thus, researchers interested in understanding how computer mediated communication affects and is used by disinformation campaigns, scammers, violent organization, as well as more traditional political protests will find great value in the dataset presented herein. Finally, we believe that, despite the snowball sampling methodology, our dataset is quite comprehensive and can be used for general studies related to understanding the Telegram platform, as well as the overall messaging platform ecosystem. This is mainly because our snowball sampling was running for a long time period, hence substantially expanding the discovered Telegram channels.

References

Abu-Salma, R.; Krol, K.; Parkin, S.; Koh, V.; Kwan, K.; Mahboob, J.; Traboulsi, Z.; and Sasse, M. A. 2017. The Security Blanket of the Chat World: An Analytic Evaluation and a User Study of Telegram. Internet Society.

Akbari, A., and Gabdulhakov, R. 2019. Platform surveillance and resistance in iran and russia: The case of telegram. *Surveillance & Society* 17(1/2):223–231.

Anglano, C.; Canonico, M.; and Guazzone, M. 2017. Forensic analysis of telegram messenger on android smartphones. *Digital Investigation* 23:31–49.

Anti-Defamation League. 2019. Telegram: The Latest Safe Haven for White Supremacists. https://www.adl.org/blog/telegram-the-latest-safe-haven-for-white-supremacists.

Asnafi, A. R.; Moradi, S.; Dokhtesmati, M.; and Naeini, M. P. 2017. Using mobile-based social networks by iranian libraries: The case of telegram messenger. *Libr. Philos. Pract* 2017(1).

Brena, G.; Brambilla, M.; Ceri, S.; Di Giovanni, M.; Pierri, F.; and Ramponi, G. 2019. News Sharing User Behaviour on Twitter: A Comprehensive Data Collection of News Articles and Social Interactions. In *ICWSM*.

Dargahi Nobari, A.; Reshadatmand, N.; and Neshati, M. 2017. Analysis of Telegram, An Instant Messaging Service. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2035–2038. ACM.

de Freitas Melo, P.; Vieira, C. C.; Garimella, K.; de Melo, P. O. V.; and Benevenuto, F. 2019. Can WhatsApp Counter Misinformation by Limiting Message Forwarding? In *International Conference on Complex Networks and Their Applications*, 372–384. Springer.

Fair, G., and Wesslen, R. 2019. Shouting into the Void: A Database of the Alternative Social Media Platform Gab. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 13, 608–610.

Founta, A. M.; Djouvas, C.; Chatzakou, D.; Leontiadis, I.; Blackburn, J.; Stringhini, G.; Vakali, A.; Sirivianos, M.; and Kourtellis, N. 2018. Large scale crowdsourcing and characterization of twitter abusive behavior. In *ICWSM*.

Garimella, K., and Tyson, G. 2018. Whatapp Doc? A First Look at Whatsapp Public Group Data. In *Twelfth International AAAI Conference on Web and Social Media*.

Hashemi, A., and Chahooki, M. A. Z. 2019. Telegram group quality measurement by user behavior analysis. *Social Network Analysis and Mining* 9(1):33.

Iqbal, M. 2019. Telegram revenue and usage statistics (2019). https://www.businessofapps.com/data/telegram-statistics/.

Kastrenakes, J. 2019. WhatsApp limits message forwarding in fight against misinformation. https://www.theverge.com/2019/1/21/18191455/whatsapp-forwarding-limit-five-messages-misinformation-battle.

Naseri, M., and Zamani, H. 2019. Analyzing and Predicting News Popularity in an Instant Messaging Service.

In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1053–1056.

Nikkah, S.; Miller, A. D.; and Young, A. L. 2018. Telegram as An Immigration Management Tool.

Nørregaard, J.; Horne, B. D.; and Adalı, S. 2019. Nela-gt-2018: A large multi-labelled news dataset for the study of misinformation in news articles. In *ICWSM*.

Prucha, N. 2016. Is and the jihadist information highway–projecting influence and religious identity via telegram. *Perspectives on Terrorism* 10(6).

Salem, F. K. A.; Al Feel, R.; Elbassuoni, S.; Jaber, M.; and Farah, M. 2019. FA-KES: A Fake News Dataset around the Syrian War. In *ICWSM*.

Satrya, G. B.; Daely, P. T.; and Nugroho, M. A. 2016. Digital forensic analysis of telegram messenger on android devices. In *2016 International Conference on Information & Communication Technology and Systems (ICTS)*, 1–7. IEEE.

Shehabat, A.; Mitew, T.; and Alzoubi, Y. 2017. Encrypted jihad: Investigating the role of telegram app in lone wolf attacks in the west. *Journal of Strategic Security* 10(3):27–53.

Sutikno, T.; Handayani, L.; Stiawan, D.; Riyadi, M. A.; and Subroto, I. M. I. 2016. Whatsapp, viber and telegram: Which is the best for instant messaging? *International Journal of Electrical & Computer Engineering (2088-8708)* 6(3).

Williams-Grut, O. 2017. "Market manipulation 101": "Wolf of Wall Street"-style "pump and dump" scams plague cryptocurrency markets. https://amp.insider.com/ico-cryptocurrency-pump-and-dump-telegram-2017-11.

Yayla, A. S., and Speckhard, A. 2017. Telegram: The mighty application that isis loves. *International Center for the Study of Violent Extremism*.

Zignani, M.; Quadri, C.; Galdeman, A.; Gaito, S.; and Rossi, G. P. 2019. Mastodon Content Warnings: Inappropriate Contents in a Microblogging Platform. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 13, 639–645.