

Detection of pedophilia content online: a case study using Twitter

Raphael Reis Coelho¹, Jonice Oliveira¹

Universidade Federal do Rio de Janeiro (UFRJ)¹

Caixa Postal 21941-916 – Rio de Janeiro – RJ – Brasil

raphael.reis.coelho@ufrj.br, jonice@dcc.ufrj.br

***Abstract.** Social media allows users to consume a wide range of topics and content, such as pornography. The consumption of pornography, as well as the problems associated with this type of material, has grown over the years, especially among adolescents. Now, another type of pornographic content has been widely consumed: child pornography. Are social media helping to spread this content? To understand this issue, a survey was conducted based on data regarding the consumption of child pornography extracted from Twitter in 2020, followed by a series of analyzes that show a possible impact of twitter on the dissemination and consumption of this type of content.*

1. Introduction

The interactivity seen in online social media already affects and will continue to change the social landscape and the population consumption experience (KNIGHT, 2009). Today, we have a range of content that can be consumed through networks as news, entertainment, podcasts and pornography.

The pornography spread in the past two decades, especially through the Internet, allowing people of all ages to find, consume, create and distribute sexually explicit content. A growing amount of data reveals these phenomena are increasingly common worldwide (FLOOD, 2007).

The pornography consumption has increased with the internet. Several papers described the spread and trafficking of child pornography in P2P networks (WOLAK, 2014), (BISSIAS, 2016). However, there are still few studies relating the consumption of child pornography and social media. In this context, this research aims to contributing by developing a study that approaches how social media is used to share pedophilia. For this, a data sample was collected on Twitter, followed by an analysis of what was found.

2. Theoretical framework

Today, there are 4.3 billion people connected to the Internet around the world. This is equivalent to approximately 57% of the world population. Even greater is the number of people who own their own cell phone, estimated at 5.1 billion. Of this total, 3.4 billion users use social networks, with 3.2 billion doing so actively. On Twitter, for example, 10% of profiles are registered between 13 and 17 years old, while 23%, between 18 and 24. (HOOTSUITE, 2019).

Obviously, this increasingly ubiquitous access to the Internet can be positive when, for example, people of all ages turn to this medium to obtain information on sexual education and health (BARAK & FISHER, 2001). On the other hand, unmonitored access

for young people raises genuine concerns in light of the potential risks (Federal Bureau of Investigation, 2011).

Although child pornography has been a recognized problem for decades, until recently it was considered to be a “rather small and essentially specialist problem related to a much broader and significant problem” (TAYLOR & QUAYLE, 2003).

In the mid-1990s, there was a significant change in the nature of child pornography in terms of the generation, distribution and possession of images and videos. Before, child pornography was practiced mainly offline and therefore had less impact and was more easily traceable. However, in recent years, it has evolved into a much more difficult problem with the advent of social networks, transforming the internet not only into a transmission channel, but probably into a factor that makes child pornography possible in itself (TAYLOR & QUAYLE, 2003).

Today, on Twitter, it is possible to observe through the search terms, several users, mostly Latin Americans and South Americans, asking for the exchange of images and videos of children being sexually abused. Pedophiles now not only share contents online but also organize themselves in their own social networks through instant messaging applications sharing interests and experiences, blurring the edge between the virtual and real worlds (VITORINO, 2017).

3. Materials and methods

In view of the phenomenon reported in the previous section, data collection was conducted on Twitter. For this, was used a *crawler* developed in Python that used the *Twint* library (TWINT PROJECT, 2018) to collect tweets from January 1 to December 31, 2020.

The search term “cp telegram” was used for the collection. CP is the acronym for child pornography. This is the term most used by users to search for this type of content and go “unnoticed”, since many people do not know what it means. In addition, many other terms are used to refer to child pornograph. Expressions like “Caldo de Pollo”, “Códigos Postais”, “Centro Pokemon” are just some examples of codes used by these people.

The majority of users do not post explicit content directly to their profiles. They use other platforms to exchange this content. The main one is the Telegram instant messaging application. For this reason, “cp telegram” was the term used for the development of the research.

Tweets, comments, retweets, and likes that mentioned the keyword were collected without any filter of language or region. They were classified and flagged if they contained the content treated in this work or not by searching in tweet text some terms such as “material”, “intercambio”, “aggregame”, “link”, “DM”, “group”.

As this research aimed to identify if there is a continuous effect of the dissemination of illegal pornographic content, more precisely, pedophilia in social media, two research questions (RQ) were elaborated followed by null hypotheses.

RQ1: Has Twitter been used to share child pornography over time?

RQ2: Does this type of content create positive engagement? That is, people engaging in favor of this content?

Each question was evaluated using a metric:

Constancy: how constant is the volume of tweets with the term cp telegram?. Thus, was evaluated the absolute value of tweets as time t ($Va(t)$), the absolute value of tweets as time t-n, where n represents the number of the month ($Va(t-n)$). The evaluation criteria used was the division of each $Va(t+n)$ for $Va(t)$ is greater than 1, according to formula 1:

$$\frac{Va(t+n)}{Va(t)} \geq 1$$

Formula 1: Constancy metric evaluation criteria

Engagement: how attractive these contents are to other users to generate engagement (likes, comments and retweets)? Thus, was evaluated the absolute value of comments (Vco), likes (Vli) and retweets (Vre). Where the evaluation criteria used was the sum of $Vco(t)$, $Vli(t)$ and $Vre(t)$ is greater than $Va(t)$ according to formula 2:

$$\frac{Vco(t) + Vcu(t) + Vre(t)}{Va(t)} > 1$$

Formula 2: Engagement metric evaluation criteria

4. Analysis and results

A total of 5,616 tweets were collected, which were classified as adhering to the object of this research, comprised in the period from January 1 to December 31, 2020, as shown in figure 1.

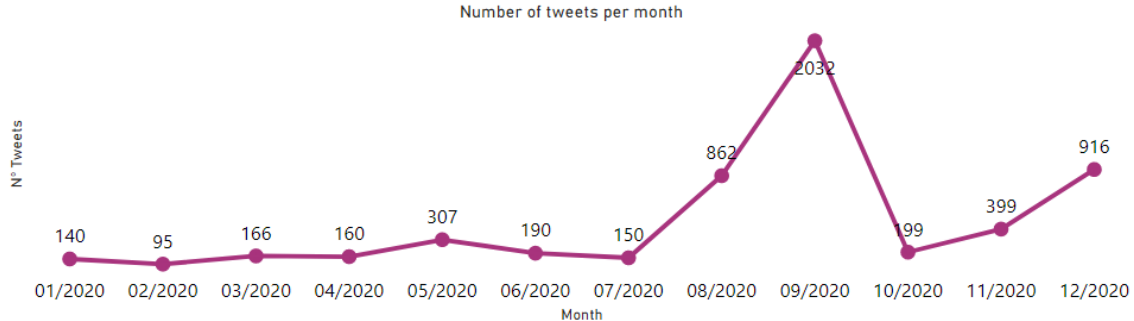


Figure 1: Tweets containing the term “CP telegram” in 2020.

In addition to the number of Tweets, the engagement of this type of content on the network was collected through comments, retweets and likes (Figure 2).

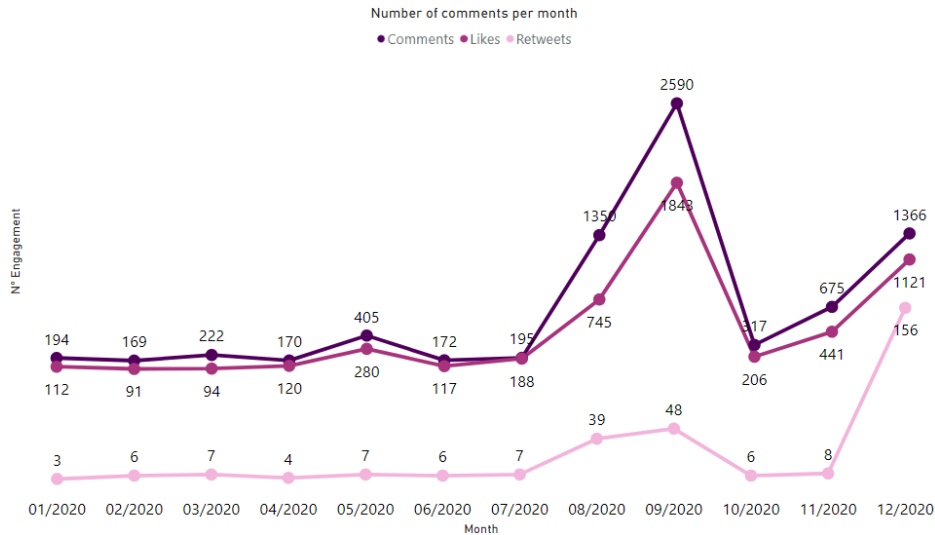


Figure 2: Comments, likes and retweets on tweets containing "CP telegram" in 2020.

Here another perception of the phenomenon is brought to light: the engagement of this term reached 7,825 comments, what means other people interacted in these tweets with the intention of viewing and/or exchanging child pornography, in addition to 5,358 likes in these publications. On the other hand, the retweet function was not something popular with these users, only 297 having been made in the period, with a considerable outlier in December.

5. Discussion and results evaluation

Constancy: According to figure 1, there is constancy over time, with some significant increases in the number of tweets made using the term "cp telegram". In Table 1, this relationship can be seen through the criterion used where, after one month t , the subsequent months reach a value greater than one in most cases, especially if we look at the initial months. However, it is necessary to mention that many tweets and accounts are deleted over time, mainly from these users, either due to a report made, or as an attempt to erase traces, decreasing the number of tweets.

It is also worth mentioning the collection took place in two moments: the first was made in September and the second in December. Both months were the biggest peaks and the months before these were the second in terms of tweets and engagement. This corroborates the comment made earlier that tweets are deleted over time, either by Twitter itself or by the user.

Even with the tweets being deleted, Through these data it is possible to answer positively to RQ1, where it is concluded that Twitter has been used constantly to share child pornography.

Table 1: Proportion of tweets in relation to previous months

t+n/t	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Jan	1	0,67	1,18	1,14	2,19	1,35	1,07	6,15	14,51	1,42	2,85	6,54
Feb	-	1	1,74	1,68	3,23	2	1,58	9,07	21,39	2,9	4,2	9,64
Mar	-	-	1	0,96	1,85	1,14	0,9	5,19	12,24	1,2	2,4	5,52
Apr	-	-	-	1	1,91	1,19	0,93	5,39	12,7	1,24	2,5	5,72
May	-	-	-	-	1	0,61	0,49	2,8	6,62	0,65	1,3	2,99

Jun	-	-	-	-	-	1	0,79	4,53	10,7	1,04	2,1	4,82
Jul	-	-	-	-	-	-	1	5,74	13,54	1,32	2,66	6,11
Aug	-	-	-	-	-	-	-	1	2,36	0,23	0,46	1,06
Sep	-	-	-	-	-	-	-	-	1	0,09	0,2	0,45
Oct	-	-	-	-	-	-	-	-	-	1	2	4,60
Nov	-	-	-	-	-	-	-	-	-	-	1	2,3
Dec	-	-	-	-	-	-	-	-	-	-	-	1

Engagement: To better understand how the dissemination of this type of content impacts the network, can be analyze the engagement it causes: comments, likes and retweets (Figure 2). Thus, it is clear that the actual number of users looking for child pornography is significantly greater than just the number of tweets, as shown in Table 2

For all months of the time series, the sum of the absolute numbers of comments, retweets and likes, was greater than the absolute number of tweets. With this data, it is possible answer affirmatively to RQ2 which concluded that this type of content makes other users engage positively about it.

Table 2: Proportion of engagement in relation to the number of tweets.

Months	Total Comments	Total <i>retweets</i>	Total Likes	Total <i>Tweets</i>	Results
Jan	194	3	112	140	>1
Feb	169	6	91	95	>1
Mar	222	7	94	166	>1
Apr	170	4	120	160	>1
May	405	7	280	307	>1
Jun	172	6	117	190	>1
Jul	195	7	188	150	>1
Aug	1350	39	745	862	>1
Sep	2590	48	1843	2032	>1
Oct	317	6	206	199	>1
Nov	675	8	441	399	>1
Dec	1366	156	1121	916	>1

Although the policy for using Twitter explicitly states that the platform has zero tolerance for sexual exploitation of minors, whether through images, videos or posts (TWITTER, 2020), there is still no concrete measure to combat the phenomenon reported in this research.

As mentioned earlier, time is the biggest limitation on this research. The collection in previous periods does not provide a very broad view of the amount of child pornographic content that is posted daily on the network. For this reason, the collection has not been extended to years before 2020.

6. Conclusion

Despite the presented limitation, it is possible to conclude that a phenomenon is occurring in social media, which has been used as a vehicle for the distribution and consumption of child pornography, thus confirming the initial objective of this research. In this way, children and adolescents may come into contact with this type of material what may cause damage for these individuals and for society as a whole.

Action by those responsible for the functioning and moderation of social media on this type of content is necessary to guarantee that they are not made available on the network and that people who interact with these materials suffer the appropriate measures within the executable laws in the locality of these users.

References

- BARAK, A., & FISHER, W. A. (2001). Toward an Internet-driven, theoretically-based innovative approach to sex education. *Journal of Sex Research*, 38, 324–332.
- BISSIAS, George et al. Characterization of contact offenders and child exploitation material trafficking on five peer-to-peer networks. *Child abuse & neglect*, v. 52, p. 185-199, 2016.
- FLOOD, Michael. Exposure to pornography among youth in Australia. *Journal of sociology*, v. 43, n. 1, p. 45-60, 2007.
- FEDERAL BUREAU OF INVESTIGATION. (2011). A parents' guide to Internet safety. Disponível em: <http://www.fbi.gov/stats-services/publications/parent-guide>. Acesso em: 25 set 2020.
- HOOTSUITE. Essential insights into how people around the world use the internet, mobile devices, social media and e-commerce digital. Disponível em: <https://p.widencdn.net/kqy7ii/Digital2019-Report-en>. Acesso em: 25 set 2020
- KNIGHT, Kristina. NIELSEN: Social Networking 4th most popular online activity. Bizreport, 2009. Disponível em: http://www.bizreport.com/2009/03/nielsen_social_networking_4th_most_popular_online_activity.html. Acesso em: 04 de julho de 2020.
- TAYLOR, Maxwell; QUAYLE, Ethel. *Child pornography: An internet crime*. Psychology press, 2003.
- TWINT PROJECT. Twint – Twitter Integration Tool. 2018. Disponível em: <https://github.com/twintproject/twint>. Acesso em: 27 ago 2020.
- TWITTER. Política sobre a exploração sexual de menores. 2020. Disponível em: <https://help.twitter.com/pt/rules-and-policies/sexual-exploitation-policy>. Acesso em: 29 nov 2020.
- VITORINO, Paulo et al. Leveraging deep neural networks to fight child pornography in the age of social media. *Journal of Visual Communication and Image Representation*, v. 50, p. 303-313, 2018.
- WOLAK, Janis; LIBERATORE, Marc; LEVINE, Brian Neil. Measuring a year of child pornography trafficking by US computers on a peer-to-peer network. *Child Abuse & Neglect*, v. 38, n. 2, p. 347-356, 2014.