A Content Based Study of Anonymous Online Social Networks & It's Contribution to Cyber –Bullying

Mohsina Zaman
CIS 533 Project Report

Abstract

After public online social networks, the user focus is shifting towards anonymous social networks. In anonymous social network, as the authors of posts are not identified, it has been observed that there is a possibility of harmful content. Therefore, in this study, a deeper look at the content of anonymous OSN is attempted in order to find if users of anonymous OSNs are indeed more prone to cyber – bullying. The results that I obtain by studying Whisper, does indeed support the theory of the risk of cyber-bullying being higher in anonymous OSNs compared to public OSNs.

Keywords: Anonymous social network, cyber-bullying, anonymous sensitivity, Whisper

Introduction

The advent of online social networks (OSN) has forever changed the way we communicate. It has made it easier to keep up to date with friends and family living close and far and also to make new friends all over the globe. Facebook and Twitter are one of the two most popular such OSN.

However, soon content posted on Facebook and Twitter began to be used to asses job candidates, support divorce litigation and to terminate employees [1]. Also, ever since the Snowden incident, the users have started to crave the use of OSN's which preserve their privacy. As a result, this has caused a rise in the use of anonymous OSNs.

Anonymous OSNs are OSNs which allows its user to post content anonymously or under pseudonym, unlike Facebook and some other OSNs which require the user's online identities to match their offline identity. Whispers, Yik yak, Ask.fm , Secret are some examples of such OSNs. Many may view anonymous OSNs as the saving grace of online social networking as users are able to express themselves without being held responsible or fear of their content being used against them in the future. However, there is grave disadvantage to such OSNs.

Anonymous social networks remove user accountability. As a result, a user is not afraid to post harmful content which may hurt an individual or a group as the content cannot be traced back. Critiques argue that anonymity undermine credibility and enables negative posts [1][3]. While supporters argue believe anonymous OSNs provide users a safe haven from prosecution and allow users to communicate without the fear of bullying or abuse.

While the supporters view are true, it is undeniable that anonymous OSNs have caused some grave incidents already. Several teenagers have faced extreme cyber bullying and has even led a few to commit suicide. For example, Ask.fm, which could be used via Facebook, has been linked to 5 teen suicides. A teenage after facing severe cyber bullying on Yik yak has opened a petition on change.org against Yik yak to close it down. The founders of Secret have taken the app off the market as the increase in cyber bullying on the site did not reflect the vision they had for the app.

While the incidents reported above have been corroborated by leading media in the world there has been no research work done to correlate cyber bullying with use of anonymous OSN. In this project, I have attempted to answer two questions:

Q1: Do anonymous OSNs have a larger number of anonymous sensitive posts and posts had high risk of cyber bullying?

Q2: Are users of anonymous OSNs at a higher risk of facing cyber bullying due to the anonymous sensitive posts?

Related Work

While extensive study has been carried out on public OSN very few work has been done with study of anonymous OSN. Out of 604 papers surveyed, 33 dealt with OSNs of which only 3 dealt with anonymous OSN. Also, of the papers on anonymous OSNs there is a distribution on the study of the structure of the OSN and the content of the OSNs.

The work done by Wang et.al [1], looks into the anonymous OSN Whisper and studies user interaction, content and also uncovers a security vulnerability. In [1] 24 million whispers written by more than 1 million users are collected over a period of 3 months. The data was collected by crawling the latest feed and the replies of each individual post. It was observed that at a time there was more number of replies than original posts. Regarding content, this paper has observed that most content on Whisper are highly personal content. Also by building Whisper interaction graphs, it has been observed that Whisper users interact with a higher sample of users than Facebook or Twitter. Also, connections between Whisper users are much less than compared to Facebook or Twitter (.033 against .059 or .048 respectively). Weak user communities have also been detected at Whisper, the explanation of which may be due to posts of same geographical areas being clustered together. Regarding user population growth, it has been observed that Whisper has a steady stream of new users and disengaged users (users who stop using the app) every day. The new users make a significant contribution to the overall whisper stream (>20%), however the overall content generation remains steady. Furthermore, the paper also studies content moderation in Whisper using keyword-based approach. They have observed Whisper deletes content which violate its user policy. Most deleted whispers are on sexually explicit and nudity related messages and are deleted within 24 hours of posting, on the other hand messages regarding personal expression, religion and politics are the least likely to be deleted. Lastly, security vulnerability has been identified in which a Whisper user is able to accurately track the location of another Whisper user by querying the Whisper server. The Whisper team had been informed and has taken steps to remove the vulnerability.

Content based analysis is carried out by Correa et. al [2]. They have attempted to answer questions regarding anonymous sensitive content posted on Whisper and compared it to content posted in Twitter. And also the types of content posted on anonymous OSNs and if there is a linguistic difference between content on anonymous OSN and non-anonymous OSNs. They have also correlated user demographic with anonymous sensitivity of a post. To carry out their work, Correa et. al has used a crawler to collected 20.7 million Whispers posts along with the author name, timestamp, number of likes and dislikes. To measure anonymity sensitivity 100 Amazon Mechanical Turk were employed to rate 500 whisper posts and 500 tweets on anonymity sensitivity level from a range of 0.0 to 1.0 and computed the average level of each post to get the final level. It was observed that approximately 16% of whisper posts are above anonymity sensitivity level of 0.8 as compared to Twitter which has only 1% of posts above this level which suggests Whisper has a larger portion of sensitive posts than Twitter. Also, probability distribution of sensitivity level on Whisper is uniform which suggests whispers have posts on all levels of sensitivity. By studying content, Whisper posts were categorized and it was observed that majority of the posts fell under the categories of Confessions, Relationships, Meet ups, and Q&A. It has been observed that posts on NSFW (Not Safe for Work) and LGBTQ are at high levels of sensitivity. By studying user demographic, it has been observed gender affects anonymity sensitivity on NSFW related posts, older generation are more anonymous sensitive than the younger generation and college education has an increasing effect on anonymous sensitivity despite income. Lastly, a clear linguistic difference has been observed in Whisper posts vs. Tweets, Whispers have a higher percentage of 1st person singular pronouns, Whisper has double the number of *sexual* content and negative emotion content and content regarding wants, needs and wishes compared to Twitter. This linguistic difference allows the design of classifier which can be trained to automatically classify posts to anonymous and non-anonymous OSN.

Methodology

To begin with, I compiled a list of anonymous OSNs to study. Majority of the compiled anonymous OSNs are mobile phone applications with the exception of Whisper, which is also a mobile application but also has web – browser based feed. As crawling mobile application is still not clearly understood I chose Whisper as the anonymous OSN to study for my project.

A High Level Analysis of Whisper

Whisper is mainly a mobile application on the android and iPhone platform, but also has a web – browser based feed with limited functionality. Users post content on Whisper by superimposing their text on an image. Other users can heart (like) or reply to contents by posting whispers on their own. Whisper requires its user to be age 17 or older. Whisper does not require its user to provide any identification and a user can change their pseudonym each time they post content. Whisper does collect the location of the user to show content in their general area on the users feed. Users can add a certain school or group or automatically assign the user a school based on their location. On the mobile application, a user can view 4 feeds: latest, popular, nearby and posts from any group they are included in. The primary means of communication between users are via means of replying to certain whispers by whispers of their own. Whisper does provide a chat option, however it is harder to hide identity from fellow user if the chat option is used. The web-browser based feed of Whisper displays popular stories, which are collections of whispers on a certain topic and a Popular feed which displays the most popular whispers of the day. Popular whispers are determined by the highest number of likes and replies. The web-browser based feed however does not allow posting of content on Whisper, it is simply for viewing purposes.

Data Collection

Since Whisper has a web-browser based feed and my concentration was on collecting popular posts I decided to crawl the popular feed of whisper. I implemented a crawler in python using the Scrapy crawler library. Xpaths were used to define which elements to extract from the page. I collected whispers for 3 days running the crawler twice every day. However, as Whisper used infinite scrolling, which my crawler was not designed to handle, I was only able to collect a small number of whispers.

I collected a total of 167 whispers. Each Whisper consisted of the image source, text content, number of hearts and number of replies. A sample of the data collected is shown in Figure 1.

Labeling

After data collection was completed, I labeled each whisper on levels of anonymous sensitivity and potential for cyber bullying. This process needed to be done manually as otherwise it falls in the domain of natural language processing and sentiment analysis, which is completely different research area.

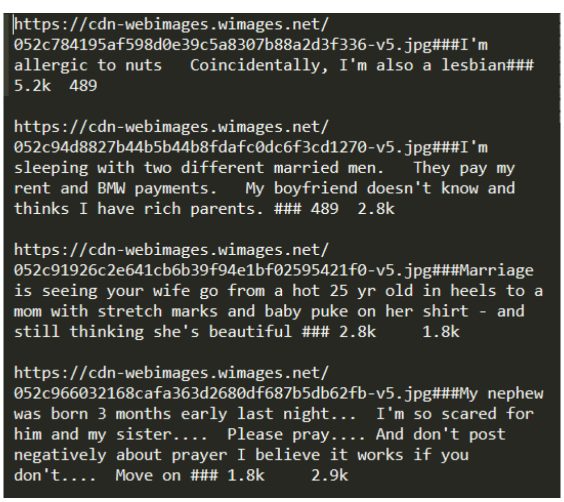
Each post was labeled on anonymous sensitivity based on my perception of how anonymous sensitive the whisper must be. The judging criteria were:

1) If I were the author, would I be comfortable with posting such content with my identity attached to it in a non-anonymous network?
2) From the point of view of the viewer, would their perception of me as a person be affected greatly by content if my identity were attached?

Based on the criteria's mentioned above, each post was labeled from 0 to 10 with increments of 1, with 0 being not anonymous sensitive at all (ex: content related to weather, sport) to 10 being the highest (ex: sexual, intimate details).

On the case of labeling for potential of cyber-bullying, the labeling was done based on the following criteria

1) Is the whisper generally negative or designed to harm the sentiment of a person or a group?
2) Does the whisper put the author at a risk of getting bullied by her/his fellow whisperers?



```
https://cdn-webimages.wimages.net/
052c784195af598d0e39c5a8307b88a2d3f336-v5.jpg###I'm
allergic to nuts   Coincidentally, I'm also a lesbian###
5.2k  489

https://cdn-webimages.wimages.net/
052c94d8827b44b5b44b8fdafc0dc6f3cd1270-v5.jpg###I'm
sleeping with two different married men.   They pay my
rent and BMW payments.   My boyfriend doesn't know and
thinks I have rich parents. ### 489  2.8k

https://cdn-webimages.wimages.net/
052c91926c2e641cb6b39f94e1bf02595421f0-v5.jpg###Marriage
is seeing your wife go from a hot 25 yr old in heels to a
mom with stretch marks and baby puke on her shirt - and
still thinking she's beautiful ### 2.8k     1.8k

https://cdn-webimages.wimages.net/
052c966032168cafa363d2680df687b5db62fb-v5.jpg###My nephew
was born 3 months early last night...  I'm so scared for
him and my sister....  Please pray.... And don't post
negatively about prayer I believe it works if you
don't....  Move on ### 1.8k     2.9k
```

Figure 1: Collected Data

```
Text likes replies ann bully
I had a hole in my leggings. I didn't know how to fix it
and I was in a rush so... I spray painted my leg the
colour of the leggings....### 2.9k### 414###2###0

So my friend face swapped the adele CD...###
414###1.5k###0###-5

Happy Pokemon Day!  I can't believe it's been 20 years.
Who was the first Pokemon you chose?  I chose
Squirtle.### 1.5k### 55###0###1

I draw on my eyebrows and my daughter walked up to me
like this.###55###1.2k###0###-5

I was Facebook stalking my crush and accidentally liked
a really old post. So I deleted my
Facebook.###1.2k###574###3###0

Both my parents died last year, but in Google maps,
they're still sitting on their front porch. I dread the
day that Google updates the pictures. ### 574###
1.1k###0###0

 Damn. Santa Monica Police have zero chill.### 1.1k### 27
```

Figure 2: Labeled Data

The posts were labeled on a scale from -5 to 5, so as not to confuse with the anonymous sensitivity. However, in the code the scale was adjusted to be from 0 to 10. Therefore, a post was assigned a level -5 if the post was related to passive items and a scale 5 if the content included intimate personal details which can be used to attack the author or designed to attack. A sample of labeled data is shown in Figure 2.

Metrics

Three main metrics were calculated in this project.

1) Percentage of posts at each level of anonymous sensitivity and cyber bullying.
2) Correlation between number of likes and levels of the two criteria.
3) Correlation between number of replies and levels of the two criteria.

1: The number of posts at each level were calculated and divided against the total number of posts.

2 & 3: Firstly, the average number of likes and replies over the entire dataset was calculated. This allowed the dataset to be divided into four groups, posts with number of likes above average, posts with number of likes below average, posts with number of replies above average and posts with number of replies below average. For each group then. The average number of likes and replies were calculated to gather an understanding on any possible correlation between popularity and average anonymity sensitivity and cyber-bullying level.

Results & Analysis

Percentage of posts

The following Table 1 shows the percentage of posts at each level of anonymous sensitivity & cyber – bullying potential.

Level	Anonymous Sensitivity	Cyber-bullying Potential
0	19.76	12.57
1	13.77	11.37
2	12.57	11.38
3	7.19	2.4
4	7.78	1.8
5	7.19	23.35
6	7.19	5.39
7	6.59	5.99
8	5.39	11.96
9	7.19	5.39
10	5.39	7.19

Table 1: Percentage of posts at each level

The following Figure 4 shows a plot of the results displayed in the table above.

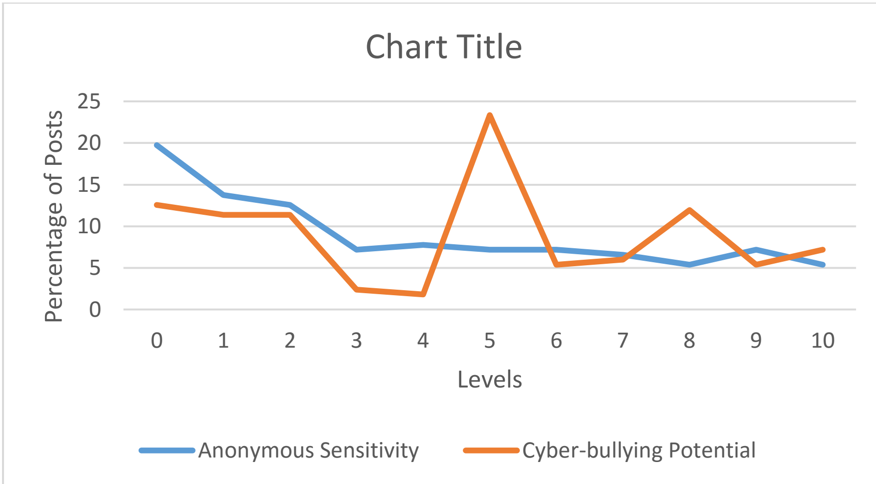
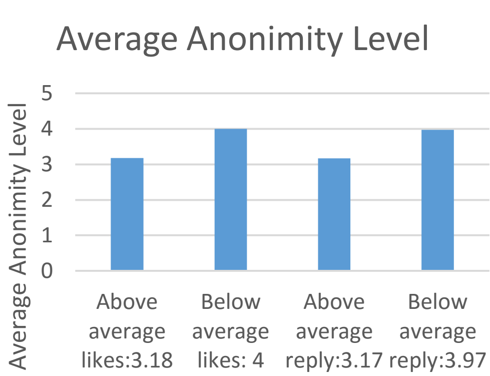


Figure 4: Plot of Percentage of post at each level

As it can be observed above, there are a higher percentage of posts at lower levels of each criterion. The percentage gradually decreases at higher levels of anonymous sensitivity. Cyber-bullying also follows the same trend with a sudden spike at level 5, which denotes posts with equal chances of getting and not getting bullied.

However, an interesting correlation can be observed. Both the percentage of posts at each level change in accordance to each other. There are similar percentage of posts at higher levels of anonymity sensitivity and cyber – bullying potential and similar percentage of posts at lower levels of the criterions. Therefore, as it has been studied that users of anonymous OSN tend to share more anonymous sensitive posts compared to non-anonymous OSN [1], it can be surmised that users of anonymous OSNs are at a higher risk of being related to cyber – bullying than users of non-anonymous OSNs.

Following in Figure 5 are the plots of the average levels at each of the 4 divisions of the dataset that was created.



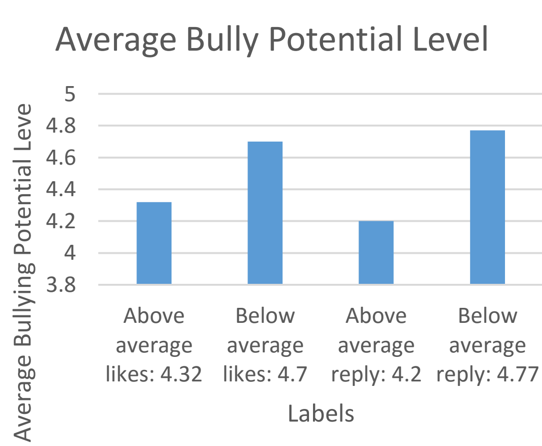


Figure 5: Average levels of likes and replies above and below average for each criterion

As it can be seen, for anonymous sensitivity, the levels are clustered between 3 and 4 and for average bully – potential the levels are clustered between 4 and 5. Both the average levels are at the lower range of the scale, which refutes the claim that in anonymous OSNs higher levels of anonymous sensitive and bully related posts obtain higher popularity. However, this observation is not conclusive as in the dataset a larger portion of the posts belong to the lower range of scales which give rise to the results that is displayed above. Therefore, in this study it is not possible to obtain a conclusive explanation of the correlation between popularity and anonymous sensitivity and cyber – bully potential.

Limitations

There are several limitations which may have had an effect on the study. Firstly, the dataset that was analyzed was not large enough to have captured the entire essence of an anonymous OSN. However, for the scope of the project, a larger dataset would have proved to be cumbersome as it required each post to be labeled manually to each criterion.

Secondly, as the posts were labeled only by me, there is a high possibility of personal bias on the labeling. It may be the case that posts that I deem as low anonymous sensitivity or bully potential may be deemed as at a higher level of anonymous sensitivity or bully potential by a different individual.

Conclusion

In this study, I have attempted to study an anonymous OSN, Whisper. I have attempted to answer questions regarding the number of anonymous sensitive posts in anonymous OSN, whether the users of anonymous OSNs are at a higher risk of cyber – bullying and to find a correlation between popularity and anonymous sensitivity and cyber – bullying.

Regarding the number of posts at each level and correlation of popularity with the two criterion, no substantial conclusion could be reached as a small dataset does not represent a large OSN properly.

However, it was observed that posts at a higher levels of anonymous sensitivity are at a higher risk being linked to cyber bullying. Therefore, if it can be assumed that users of anonymous OSN tend to share more anonymous sensitive posts compared to non-anonymous OSN [1], it can be concluded that users of anonymous OSN are at a higher risk of falling victim to cyber bullying.

Future Work

The results provided in this study maybe further used to carry out a study to obtain concrete results between anonymous OSNs and cyber bullying. This in turn can be used to help operators of such OSNs to combat harmful behavior. It may also be used to develop a plugin to alert concerned friends or family of a user if they are falling victim to traumatizing cyber – bullying so that secure steps can be taken.

References

[1] Wang, G., Wang, B., Wang, T., Nika, A., Zheng, H., & Zhao, B. Y. (2014, November). Whispers in the dark: analysis of an anonymous social network. In *Proceedings of the 2014 Conference on Internet Measurement Conference* (pp. 137-150). ACM.

[2] Correa, D., Silva, L. A., Mondal, M., Benevenuto, F., & Gummadi, K. P. (2015, April). The Many Shades of Anonymity: Characterizing Anonymous Social Media Content. In *Ninth International AAAI Conference on Web and Social Media*.

[3] Bernstein, M. S., Monroy-Hernández, A., Harry, D., André, P., Panovich, K., & Vargas, G. G. (2011, July). 4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community. In *ICWSM* (pp. 50-57).

[4] Lakkaraju, H., McAuley, J. J., & Leskovec, J. (2013). What's in a Name? Understanding the Interplay between Titles, Content, and Communities in Social Media. *ICWSM*, *1*(2), 3.

Appendix

The Crawler

The crawler was implemented using the Scrapy framework in Python. Scrapy is an open source web crawler library written in Python. The Scrapy framework is built around "spiders" which are self-contained crawlers given a set of instructions. The framework requires each Spider to have a unique name and extracts data by using "Xpaths" or "CSS" selectors. For my project I used Xpaths.

For example, if the required data is enclosed within the html body tag and in a div with an id of "unique", the corresponding Xpath would be "(//body//div[@id="unique"]).extract()".

Therefore, I first obtained the HTML source code of the page, inspected the page to design the required Xpath then incorporated that Xpath within my spider.

Firstly, a crawler was written to extract all the popular posts from the home page of Whisper. However, as Whisper incorporates infinite scrolling, which my spider was not designed to handle, I obtained the AJAX pagination up to depth 7. I chose depth as 7 as I observed after this depth, the whispers did not receive many likes or replies.

Using the unique AJAX pagination link of each depth, I wrote 8 crawlers for each page. Thus, this is how I incorporated infinite scrolling within my crawler. This is definitely not the most efficient method, however, due to time limitations I chose this path.

In the dataset the text, number of likes, replies, anonymity sensitivity level and bully potential level were separated by the string '###' which was later used as the delimiter.

Metrics

Once the dataset was properly labeled, several programs in C++ were written. Their functionalities were

1) Read dataset, count the number of posts, calculate the total, count the number of posts at each level of each criterion, and compute the percentage of posts at each level of each criterion.
2) Read the dataset; calculate the mean of total number of likes and total number of replies. Divide the dataset into 4 portions as explained above and write them in 4 different in text files.
3) For each of the 4 files, read the data and compute the average of each criterion and write the results in a text file.