How Impersonators Exploit Instagram to Generate Fake Engagement?

Conference Paper · January 2020

3 authors:

Koosha Zarei
Télécom SudParis - Institut Polytechnique de Paris
7 PUBLICATIONS   **8** CITATIONS

SEE PROFILE

Reza Farahbakhsh
Institut Mines-Télécom
71 PUBLICATIONS   **379** CITATIONS

SEE PROFILE

Noel Crespi
Institut Mines-Télécom
402 PUBLICATIONS   **2,861** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

IoT Applications Provisioning View project

FUSE-IT View project

How Impersonators Exploit Instagram to Generate Fake Engagement?

Koosha Zarei*, Reza Farahbakhsh*, Noël Crespi*

*CNRS Lab UMR5157, Télécom SudParis, Institut Polytechnique de Paris, Evry, France.

{koosha.zarei, reza.farahbakhsh, noel.crespi}@telecom-sudparis.eu

Abstract—Impersonators on Online Social Networks such as Instagram are playing an important role in the propagation of the content. These entities are the type of nefarious fake accounts that intend to disguise a legitimate account by making similar profiles. In addition to having impersonated profiles, we observed a considerable engagement from these entities to the published posts of verified accounts. Toward that end, we concentrate on the engagement of impersonators in terms of active and passive engagements which is studied in three major communities including "Politician", "News agency", and "Sports star" on Instagram. Inside each community, four verified accounts have been selected. Based on implemented approach in our previous studies [1], we have collected 4.8K comments, and 2.6K likes across 566 posts created from 3.8K impersonators during 7 months. Our study shed light into this interesting phenomena and provides a surprising observation that can help us to understand better how impersonators engaging themselves inside Instagram in terms of writing Comments and leaving Likes.

Index Terms—Impersonation; Instagram; Fake Engagement; Fake Profile; Fake News; Bot; Social Media.

I. Introduction

Impersonators are most commonly found on all major social media platforms including Facebook, Twitter, Instagram, YouTube and LinkedIn. Among these platforms, Instagram is widely used by celebrities and influencers with different level of popularity and visibility for their everyday activities and news propagation. On other hand, this is a golden opportunity for impersonators and recent studies show their considerable presence in Instagram. In this era, fake news and weaponized information is still a very hot topic and as it is presented in previous studies [2] [1], one of the most common ways of spreading fake news, disinformation, or false activities is using fake profiles, where malicious users create social media accounts impersonating a legitimate account and present themselves in profiles who are very similar to real persons in term of profile metrics. This activity named as "Impersonating" and impersonators are those accounts that are pretending to be someone well-known or representative of a known brands, company etc.

Furthermore, Fake Instagram profiles have a pretty clear plan—they make accounts appear more popular than they are and with bot services, they create fake engagement, too. Fake engagements strike social media and especially Instagram, which makes it considerably harder to understand which posts are genuinely getting the best reaction from legitimate accounts/followers.

From malicious activities in social media, a larger set of threats has been identified including brand abuse, fraud and follower farming. Therefore several lawsuits has been taken in place in United State (along with other countries), where criminal impersonation is a crime that is governed by states laws, which vary by state. It involved assuming a false identity with the intent to defraud another or pretending to be a representative of another person or organisation [3].

In this paper, we aim to understand the impact of impersonators on fake content production and propagation by analysing the tactics meant to lure the user attractions to the produced or fabricated content. Toward that end, we picked three distinct communities on Instagram including "Politician", "News agency", and "Sports star". Inside each category, we selected four top verified genuine accounts and we collected a great number of posts beside comments and likes (in a 7-month period). For each one, we detected and extracted the impersonator profiles based on our methodology presented in [1], and we ended up with 3.8K dataset. Next, we clustered them into three groups *C0-Fan-Pages*, *C1-Ordinary-Users*, and *C2-BotLike* based on profile and activity characteristics.

In this study, We first, investigate the portion of the comment, like, and post that are distributed by impersonators across communities. Next, to understand what is being shared, we analyse the comments. In this regard, we use natural language processing (NLP) techniques to understand the context of the written text and analyse the semantic and sentiment aspects. The contribution of this study can be summarised as follow:

- We assemble a precious dataset of the content and activities of impersonators in three leading communities.
- provides a comprehensive analysis of the behaviour of impersonators in the shape of active and passive engagement.
- provides the first analysis of how impersonators create fake engagements across leading communities on Instagram.
- presents an investigation of the content that is produced by impersonators across communities which potentially lead us to type of fake contents.

The remaining of this study is as follows. Section II gives the related studies. The process of data crawling, the description of communities, validation and the dataset are described in section III. The concept of detection of impersonators is

specified in IV. Next, we investigate the behaviour of impersonators and the communities they target in section V. Next, we analyse the content that is distributed by impersonators in section VI. Finally, section VII shows future directions and concludes the study.

II. Related work

Fake account: Recent research has worked on related research problems and dedicated a fair amount of work to study a different aspect of OSNs. In this era, looking to behavioural aspect of users and understand the different pattern of activities is still a hot topic of research. Several studies tried to shed light on this direction by profiling users based on their activities and reactions. This work [4] presents a novel technique to discriminate real accounts on social networks from fake ones. The writers from this [5] study provide a review of existing and state-of-the-art Sybil detection methods with an introductory approach and present some of the emerging open issues for Sybil detection in Online Social Networks.

Bot: On the other hand, the huge existence of Bots can alter the perception of social media influence, artificially enlarging the audience of some people, or they can impact the reputation of a company. The problem of rising social bots are discussed in [6]. There are various strategies to tackle the problem of bot detection. [7] suggested a profile-based approach and [8] proposed a novel framework on detecting spam content. Also, [9] presented a machine learning pipeline for detecting fake accounts and authors in [10], [11] present a method to classify bots and understand their behaviour in scale.

Fake Engagement: . From this viewpoint, Authors in [12], focus on the social site of YouTube and the problem of identifying bad actors posting inorganic contents and inflating the count of social engagement metrics. They propose an effective method and show how fake engagement activities on YouTube can be tracked over time. Likewise, another study, [13], enumerate the potential factors which contribute towards a genuine like on Instagram. Based on analysis of liking behaviour, they build an automated mechanism to detect fake likes on Instagram which achieves a high precision of 83.5

User Behaviour: On another line of research, the authors in [14] [15] look at the profile and behavioural patterns of a user and discussed existing challenges on different OSNs. By integrating semantic similarity and existing relationships between users, it is possible to match profiles across various OSNs [16] [17]. Also, [18] conducted a detailed investigation of user profiles and proposed a matching scheme. On Instagram, for the sake of mitigating impersonation attack, [13] explored fake behaviours and built an automated mechanism to detect fake activities.

As far as our best knowledge, the problem of spotting and analysing the fake engagement is not studied in the literature and this is the first study that analyzed this phenomenon through the lens of impersonators on Instagram.

III. Data Collection

Considering the Instagram API policies, we implemented an exclusive crawler in Python to receive data and store in a MongoDB server in the form of JSON files. We use the official Instagram API [19] which is based on the Facebook Platform to gather all posts, comments, and likes. This will return posts concerning Instagram rules. Note that we only gather public data excluding any potential sensitive data. The whole data collection process is designed exclusively for research purposes and the data is stored in an anonymized format.

A. Communities and Case Studies

To investigate and understand the behaviour of impersonators, it is essential to have a dataset that consists of data from a variety of categories. Toward that end, we examined impersonators in three influential communities including politician, news agencies and sports stars. As a result, we are dealing with a wide range of profile characteristics and user behaviours. In such a scenario, we have targeted the top famous figures inside each community. All genuine accounts are official pages, have *Verified* Badge and are confirmed by Instagram [20]. Next we explain briefly each category and the target users inside each category in this study.

- **Politician** community is of high interest. Having a large number of followers, fan pages, oppositions and supporters are the main reasons for selecting this community. Additionally, Political Bot is a new phenomenon in this area. Donald J. Trump (@*realdonaldtrump*) the president of the United States, Barack Obama (@*barackobama*) the previous president of the United States, Emmanuel Macron (@*emmanuelmacron*) the president of France, and Theresa May (@*theresamay*) the Prime Minister of the United Kingdom (all at the time of writing this paper) are included as target users in our dataset.

- **News Agency** is another vital community in which top English language news broadcasters including BBC (@*bbc*), CNN (@*cnn*), FoxNews (@*foxnews*), and Reuters (@*reuters*) are considered. Use of Social Media is changing the relationship between the news agencies and the audience. This community has a large number of followers from various groups which make it very interesting category for the popuse of this study.

- **Sports Star** community represents top sports players in football and tennis. Nowadays, thanks to social media, we see sporting star's habits, milestones and personal lives every day on our phones. Fake news, Fake profiles, and Disinformation are considered as serious difficulties inside this community. Leo Messi (@*leomessi*), Cristiano Ronaldo (@*cristiano*), Rafael Nadal (@*rafaelnadal*), and Roger Federer (@*rogerfederer*) are selected.

B. Dataset

In this study, we use the dataset which is obtained from our previous studies [2] [1] and the primary target is to analyse the content that is generated by the impersonators and investigate the fake engagement. First, we target the previously mentioned

well-known figures on Instagram (see III-A) and collect their activity from October 2018 until April 2019. The activity includes posts, comments, likes, and user information. Based on our methodology, from the the pool of users who reacted in the shape of comment and like, we extract and identified 3.8K unique impersonators. Next, based on different metrics, we clustered impersonators into three main clusters (for more details please see IV). In total, our dataset includes 3.8K impersonators who generate 4.8K comments and 2.6K likes across 566 unique posts during the period of 7 months.

C. Validation

A natural risk is that a subset of the comment and likes that are given to posts may be generated by users who are not impersonators. So, we further perform manual annotation to validate the general correctness of our data. To validate our dataset, we manually looked at the profiles of the impersonators to verify if they were really impersonator. To validate profiles, accounts of three clusters are completely checked manually. We filter any incorrectly identified impersonators.

Ethics : In line with Instagram policies, user privacy and ethical consideration defined by the community, we only gather publicly available data that are obtainable from Instagram.

IV. WHO ARE IMPERSONATORS?

Phase 1: Impersonator Detection. An impersonator is someone who pretends or copies the behaviour or actions of another. Of course, there are many reasons for impersonating someone. In the first study [2] we answered questions like who are the impersonators? What is the rate of engagement in the shape of like and comment? How many impersonators exist? and What is the activity of this group? We studied politician community with 3 use cases (D. Trump, B. Obama, E. Macron) on Instagram and track their activity (with user reactions) for three months. We presented a methodology to detect impersonators based on the profile similarity and we discovered more than 200 fake accounts with different levels of similarity. Interestingly, While Trump held the most impersonators, but Macron contained the least (108 vs. 21).

TABLE I: Summary of Impersonators across Clusters

Clusters	Type	#of unique account	#of comment	#of like	#of post*
C0_Fan_Page	Fan Page	54%	52%	50%	36%
C1_Ordinary_User	Normal User	34%	37%	29%	24%
C2_Botlike	Bot	12%	11%	19.4%	40%
Total Number		**3.8K**	**4.8K**	**2.6K**	**566**

*the number of unique posts which impersonators reacted to.

Phase 2: Clustering. We, next in the second study [1] investigated that impersonators are more interested in which community? Among them, how many distinct hidden groups exist? what are their characteristics? and How impersonators are involved in terms of reactions? To answer these questions as we extended the dataset to 3 communities including 'politician', 'news agency', and 'sports star' communities with 12 famous verified use cases (see III-A), also we enhanced the detection methodology. So we ended up with 3.8k impersonators with various characteristics. We, next applied three

major clustering methods including K-means, Gaussian Mixture Model, and Spectral Clustering algorithms. We divided impersonators into 3 clusters based on 10 features such as *username similarity*, *name similarity*, *bio similarity*, *photo similarity*, *most common metrics (mcm)*, *number of followers*, *number of followees*, *number of media count*, *private status*, and *verified status*. Next, Based on their characteristics and behaviours we are calling them *C0-Fan-Pages*, *C1-Ordinary-Users*, and *C2-BotLike*. Table I summarises the dataset across clusters. Rows present clusters and columns present data types.

(a) Active Engagement / community (b) Active Engagement / cluster

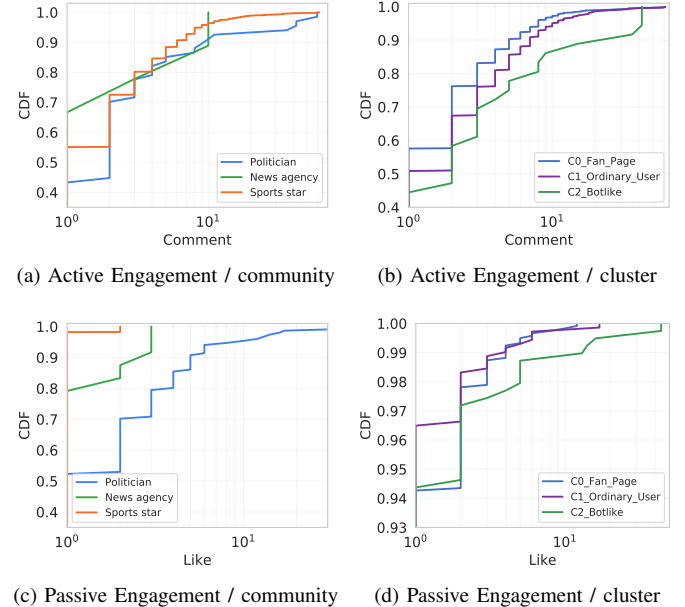(c) Passive Engagement / community (d) Passive Engagement / cluster

Fig. 1: Engagement per community and cluster: (a) CDF of number of comments issued by unique impersonator across communities (3 communities) and (b) across clusters (3 clusters). (c) CDF of number of likes issued by unique impersonator across communities and (d) across clusters.

V. WHAT IS THE FAKE ENGAGEMENT OF IMPERSONATORS?

In this part, we move through the activity of impersonators to analyse how impersonators are distributing engagement and in general what is the rate of fake engagement through different communities?

Active & Passive Engagement. First, let's look at the distribution of Active Engagement (comments) and Passive Engagement (likes) that are issued by impersonators across clusters and communities which is demonstrated in Figure 1. This figure displays the interest of impersonator amid communities/cluster. The first notable thing is that in Figure 1(a), while impersonators target all communities with a high number of comments, but politician and sports earn more (avg 4.89 vs. 2.81). This difference is even greater in passive engagement (Figure 1(c)) where sports star hosts the least number of likes compared to politicians (avg 1.01 vs. 2.75). Despite, the number of given comments in Sports star still high. Interestingly, across communities, impersonators mostly

prefer to engage in the shape of Active Engagement rather than Passive Engagement.

Next, let's look at the distribution over clusters in Figure 1(b)(d). Again these engagements are given by unique impersonators. Interestingly, in all clusters, we can see impersonators issue more engagement in the shape of active engagement rather than passive. This shows the importance of content that is trying to publish. Moreover, as we expected, botlike is the most active cluster in promoting both active and passive engagements: While 'C2_Botlike' distributes more comments and likes, but 'C0_Fan_Page' cluster issues fewer (avg. comments 5.9 vs. 2.29).

Post distribution. Next, we examine posts to see how Active and Passive Engagements are scattered among them. This distribution is exhibited in Figure 2. On average, 'C0_Fan_Page' issued 29.5, 'C1_Ordinary_Users' issued 33.2, and 'C2_Botlike' 2.04 comments per post and this numbers for like are 125.9, 260.8, and 19.3 per post respectively.

Fig. 3: BoxPlot of Comment Age: (a) comment issued by impersonator across three communities. (b) comment age issued by impersonator across three clusters.

interesting one is botlike cluster (median 203) which the range is between two fixed hours (1H to 10H). this behaviour reveals another peculiar characteristic of bots.
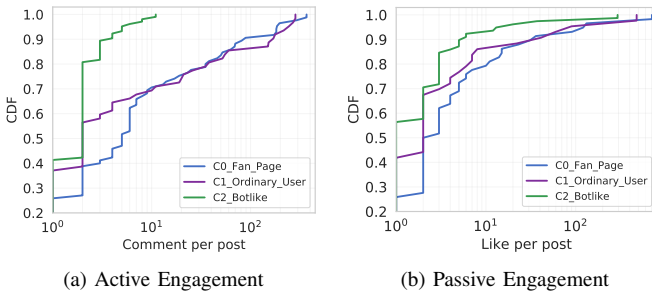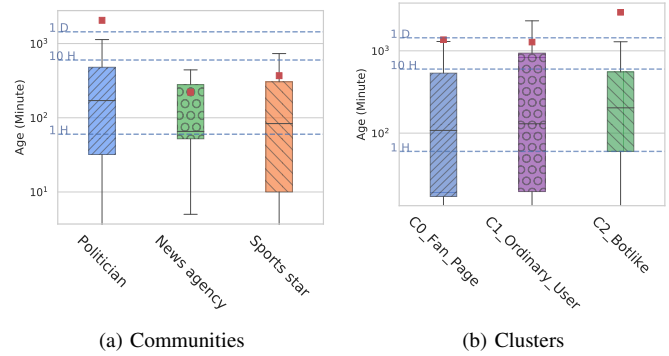
Fig. 2: Engagement per post: (a) CDF of number of comments issued by impersonator per post across clusters (b) CDF of number of likes issued by impersonator per post across clusters.

The first notable point is 'C1_Ordinary_Users' spot more posts compared to 'C2_Botlike'. This behaviour is the same in both comments and like engagements. Interestingly, in comments, $\leq 80\%$ of 'C2_Botlike' cluster aim mostly 3 posts while 'C0_Fan_Page' and 'C1_Ordinary_User' target 10 times more posts (110). This reveals bots are targeting some specific (and limited) posts over communities to issue active engagement (comment), but are delivering like to all posts.

Comment Age. Another important metric which needs to be studied is the time of publishing comment by impersonators. This can support that the comments are published at the same time or not. This metric can be measured across both communities and clusters. From the community viewpoint, in Figure 3(a), impersonators publish sooner in sports star than politician communities (median 83.5 vs. 170). Moreover, the news agency has the shortest range of age (median 83.2) and impersonators mostly comment from hour ≥ 1. This number for the sports star is from minute ≥ 10.

From cluster viewpoint, Figure 3(b), both fan page (median 108) and ordinary user (median 130) clusters begin to write immediately. Moreover, both have the largest range. The

VI. WHAT CONTENT DO IMPERSONATOR PUBLISH?

In this section, we aim to discuss what impersonators publish in form of comments? What is the content and topic of their comments? Is there any differences among clusters?

Analysis of Comments. First, let's look at the wordcloud of the most frequent word extracted from comments across communities (Figure 4). Words are colour-coded. Mostly, in Sports star community, we see 'king' and 'legend' keywords besides the name of the players which give positive support. Likewise, in politicians, we see the same trend and most dominant words are 'best', 'love', 'great', and 'thanks'. Moreover, some words such as 'follow' (refers to follow me), and 'story' (refers to check my story) are replicated which could be generated by bots that are mainly trying to attempt users to follow something.

Fig. 4: The most frequent words across communities.

Next, we aim to understand the general sentiment statistics across communities and clusters. This can give a good insight about the opinion expressed by impersonators. To do this, we use *Affin* library [21] which is one of the most popular lexicons that can be used. Figure 5 presents general sentiment and summary statistics for comments distributed by impersonators

in each cluster across communities. The output could be zero, positive, or negative number. In communities, we can see that the spread of sentiment polarity is much higher in sports star and politician as compared to news agency where a lot of the comments seem to be having a negative polarity. Practically, in the news agency, no negative comment exists. So, this community is a target for botlikes. However, we can observe a diverse range of comments in both politician and sports star communities. Form cluster viewpoint, the average sentiment for 'C2_Botlike' is 2, for 'C1_Ordinary_User' is 0.7, and for 'C0_Fan_Page' is 0.5. This claims bots distribute relatively positive text.
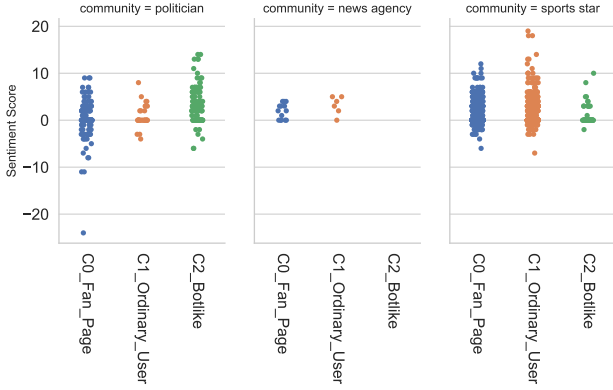

Fig. 5: Sentiment score for comments issued by impersonator across clusters and communities.

Next, in Figure 6 we visualize the frequency of sentiments across communities and clusters. Surprisingly, in Figure 6(a), while Fan pages published the most number of negative comments, but bots issued the least. Usually, fan pages are controlled by humans. We list some random comments of impersonators across clusters in Table II. *"I post trump memes every day! Check out my page?"*; this comment (row [1]) caught in D. Trump post attempting audience to follow fan pages. These kind of comments are repeated over different posts and clearly are published by bots.
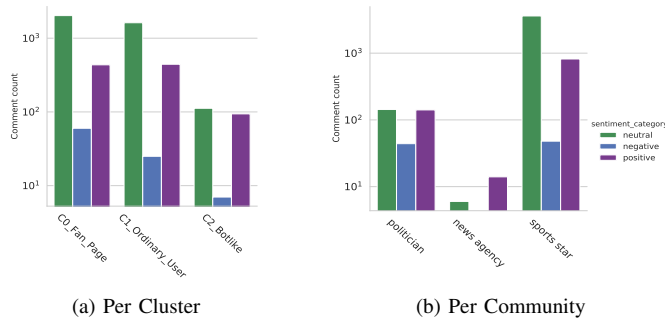

(a) Per Cluster   (b) Per Community

Fig. 6: Sentiment polarity: (a) Number of comment per cluster. (b) Number of comment per community.

In Figure 6(b), by considering communities, Sports star have more neutral comments due to the presence of comments

which are talking about sporting events without the presence of any emotion or feeling. Besides, both sports star and politician have the same rate of negative comments. This is an example of positive comment in politicians (Table II, row [3]): *"You are my president and I love you forever"*, and this is an example of negative in sports star (Table II, row [16]): *"kill ur self"*.

Duplication of comments. Next, we investigate how many duplicated comments are published by impersonators. This important metric confirms if clusters follow some particular patterns of publishing or hire automated bots to advertise something with a frequency. Toward that end, we implemented a similarity module and we were able to identify duplicated comments across communities which are demonstrated in Figure 7. To estimate the degree of similarity between comments, we use cosine similarity technique and we employ *NLTK* [22] library inside *scikit-learn* [23]. Comments that have a similarity of ≥ 0.95 (scale 0 to 1) are considered as duplicated text. Note that emojis are skipped from the measurement. As we observed, all clusters are utilising pre-defined text (with high rate of similarity) and common patterns to publish comments. It is crystal clear in Figure 7(a), *botlike* cluster distributes with a higher unusual rate (median 7) than ordinary users (median 3) and fan pages (median 2). To be sure that duplicated comments are distributed with pre-defined algorithms, we manually checked the comment time. $\geq 95\%$ of the duplicated comments have the same publishing age (from the post) which clarify our claim.

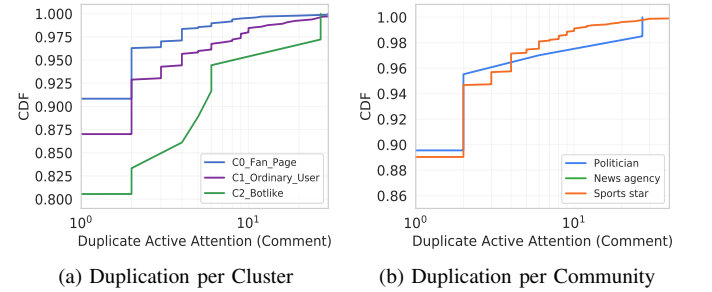(a) Duplication per Cluster   (b) Duplication per Community

Fig. 7: CDF of number of duplicated comments: (a) across clusters (b) across communities.

From community viewpoint, in Figure 7(b), impersonators hugely target politician (median 2) and sports star (median 2) communities by writing repeated comments and surprisingly there is no sign of duplication in the news agency. We, next check the text of duplicated comments by manual examination among all.

This part contains more experiments on the text and the correlation of the words that are used by clusters. As the matter of space, we include other plots in a separated report available online[1]. In general, 75% of C0_Fan_Page duplicated comments are emojis, and hugely they invite audience by predefined text such as "follow us" and "best of Ronaldo" to their pages to gain followers. In C1_Ordinary_Users, again

[1]https://sites.google.com/view/iengagement/home

we see $\leq 53\%$ of comments are emojis. Comments contain human-generated text and related hashtags are used to express the support. Most of the emojis are positive such as "heart", "like", and "thumbs up". C2_Botlike cluster contains $\leq \%70$ emojis in both positive and negative feeling. Comments hugely hold very short text, hashtags, and sometimes mentions (start with @ sign).

TABLE II: Some example of comments

#	Cluster	written comments by impersonators (randomly selected)
1		*I post trump memes every day! Check out my page?*
2		*king leo*
3	C0_Fan_Page	*You are my president and I love you forever*
4		*Thanks for supporting ALL Americans!*
5		*if u love messi like this comment*
6		*Follow @mtfoot for more!*
7		*Congratulations juventus team and @cristiano*
8		*King cristiano*
9	C1_Ordinary_User	*We are with you, if you not gone win the balondor, will be back stronger next year! forzaaa @cristiano Te Amo*
10		*You deserve all the best next time you want the hat-trick*
11		*Good luck legend, I hope you scored 3 goals tomorrow*
12		*Great win for @juventus, so happy @cristiano*
13		*More Americans are now employed than ever recorded before in our nation's history. President Donald Trump*
14		*President Trump miracle from God & for the country*
15	C2_Botlike	*Check my profile and my story*
16		*kill ur self*
17		*Beautiful pics Rafa! Thank you for sharing*
18		*thanks God that you are our president*

Note: Emojis are removed from the text.

VII. CONCLUSION AND FUTURE WORK

In conclusion, this paper has performed a first analysis of the content and engagements generated by impersonators on Instagram. Based on our previous studies, we did an investigation to discover the behaviour of impersonators and the generated content in three major communities. To the best of our knowledge, this is the first paper that conducts such analysis on Instagram. We used the dataset of nearly 4K impersonator which is extracted from our previous paper [1]. We analysed the distribution of issued Active and Passive engagement given by impersonator across three major communities. Next, we focused to the written comments by impersonators to perceive what kind of content do they publish. We obtained valuable knowledge by using various text analysis techniques which explains better the behaviours of impersonators.

As future work, This study could be extended from various angles: first, it is desirable to train a machine/deep learning classifier for comments. This can be done by considering some other important profile metrics alongside text features. As a result, this model could predict at first whether the content of the text is fake or not and second, evaluate whether the publisher of that comment is impersonator or not. Another perspective is to study other social media to understand is there any similar pattern across different platforms and can we correlate the identified profiles of impersonators.

REFERENCES

[1] Koosha Zarei, Reza Farahbakhsh, and Noel Crespi. Typification of impersonated accounts on instagram. In *2019 IEEE 38th International Performance Computing and Communications Conference (IPCCC) (IPCCC 2019)*, London, United Kingdom (Great Britain), October 2019.

[2] Koosha Zarei, Reza Farahbakhsh, and Noel Crespi. Deep dive on politician impersonating accounts in social media. In *2019 IEEE Symposium on Computers and Communications (ISCC) (IEEE ISCC 2019)*, Barcelona, Spain, June 2019.

[3] uslegal. https://definitions.uslegal.com/c/criminal-impersonation/, 2019.

[4] L. Caruccio, D. Desiato, and G. Polese. Fake account identification in social networks. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 5078–5085, Dec 2018.

[5] Devakunchari Ramalingam and Valliyammai Chinnaiah. Fake profile detection techniques in large-scale online social networks: A comprehensive review. *Computers Electrical Engineering*, 65:165 – 177, 2018.

[6] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Commun. ACM*, 59(7), 2016.

[7] White J. S. Hudson B. Voter B. R. Matthews J. N. Gurajala, S. Profile characteristics of fake twitter accounts. 2016.

[8] S. Shehnepoor, M. Salehi, R. Farahbakhsh, and N. Crespi. Netspam: A network-based spam detection framework for reviews in online social media. *IEEE Transactions on Information Forensics and Security*, 12(7):1585–1595, July 2017.

[9] Cao Xiao, David Mandell Freeman, and Theodore Hwa. Detecting clusters of fake accounts in online social networks. In *Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security*, AISec '15, pages 91–101. ACM, 2015.

[10] Zafar Gilani, Reza Farahbakhsh, Gareth Tyson, Liang Wang, and Jon Crowcroft. Of bots and humans (on twitter). In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, ASONAM '17, pages 349–354, New York, NY, USA, 2017. ACM.

[11] Zafar Gilani, Reza Farahbakhsh, Gareth Tyson, and Jon Crowcroft. A large-scale behavioural analysis of bots and humans on twitter. *ACM Trans. Web*, 13(1):7:1–7:23, February 2019.

[12] Yixuan Li, Oscar Martinez, Xing Chen, Yi Li, and John E. Hopcroft. In a world that counts: Clustering and detecting fake social engagement at scale. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pages 111–120, Republic and Canton of Geneva, Switzerland, 2016. International World Wide Web Conferences Steering Committee.

[13] Indira Sen, Anupama Aggarwal, Shiven Mian, Siddharth Singh, Ponnurangam Kumaraguru, and Anwitaman Datta. Worth its weight in likes: Towards detecting fake likes on instagram. In *Proceedings of the 10th ACM Conference on Web Science*, WebSci '18. ACM, 2018.

[14] Francesco Buccafurri, Gianluca Lax, Serena Nicolazzo, and Antonino Nocera. Comparing twitter and facebook user behavior. *Comput. Hum. Behav.*, 52(C):87–95, November 2015.

[15] Bang Hui Lim, Dongyuan Lu, Tao Chen, and Min-Yen Kan. #mytweet via instagram: Exploring user behaviour across multiple social networks. IEEE/ACM ASONAM '15, pages 113–120. ACM, 2015.

[16] Ali Choumane, Zein Al Abidin Ibrahim, and Bilal Chebaro. Profiles matching in social networks based on semantic similarities and common relationships. In *Proceedings of the International Conference on Compute and Data Analysis*, ICCDA '17, pages 14–18. ACM, 2017.

[17] Katharina Krombholz, Dieter Merkl, and Edgar Weippl. Fake identities in social media: A case study on the sustainability of the facebook business model. *Journal of Service Science Research*, 4(2), Dec 2012.

[18] Oana Goga, Patrick Loiseau, Robin Sommer, Renata Teixeira, and Krishna P. Gummadi. On the reliability of profile matching across large online social networks. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 1799–1808. ACM, 2015.

[19] Instagram. Official api graph instagram. https://developers.facebook.com/docs/instagram-api, September 2019.

[20] Instagram Verified Badges. Instagram verified badges. https://help.instagram.com/854227311295302, September 2019.

[21] Finn Årup Nielsen. A new anew: Evaluation of a word list for sentiment analysis in microblogs, 2011.

[22] Edward Loper and Steven Bird. Nltk: The natural language toolkit. In *In Proceedings of the ACL Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics. Philadelphia: Association for Computational Linguistics*, 2002.

[23] F. Pedregosa, G. Varoquaux, and A. et al. Gramfort. Scikit-learn: Machine Learning in Python . *Journal of Machine Learning Research*, 12:2825–2830, 2011.