2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Characterising and Detecting Sponsored Influencer Posts on Instagram

Koosha Zarei*, Damilola Ibosiola†, Reza Farahbakhsh*, Zafar Gilani†, Kiran Garimella‡, Noël Crespi*, Gareth Tyson†
**Institut Polytechnique de Paris, Télécom SudParis France.* {koosha.zarei, reza.farahbakhsh, noel.crespi}@telecom-sudparis.eu
†*Queen Mary University of London, United Kingdom.* {d.i.ibosiola, z.gilani, gareth.tyson}@qmul.ac.uk
‡*Massachusetts Institute of Technology.* garimell@mit.edu

***Abstract*—Recent years have seen a new form of advertisement campaigns emerge: those involving so-called social media *influencers*. These influencers accept money in return for promoting products via their social media feeds. We gather a large-scale Instagram dataset covering thousands of accounts advertising products, and create a categorisation based on the number of users they reach. We then provide a detailed analysis of the types of products being advertised by these accounts, their potential reach, and the engagement they receive from their followers. Based on our findings, we train machine learning models to distinguish sponsored content from non-sponsored, and identify cases where people are generating sponsored posts without labelling them.**

***Index Terms*—Influencers, Instagram, Social Media, User Behaviour Analysis, Sponsored Content**

I. Introduction

Social media has become an active part of billions of people's lives. Initially intended as a method to interact amongst friends, it has since become a platform used heavily for marketing. This is dominated by players such as Instagram who allow third parties to micro-target adverts at users matching certain criteria. This newfound ability has turned social media into a multi-billion dollar industry. However, in recent years a new breed of social advertisement has emerged. Driven by the rise of social media celebrities, so-called *influencers*, we have witnessed various companies marketing directly through celebrity endorsements (rather than via official advertisement channels on the platform) [1] [2].

A particularly interesting trend is the rise of *nano* influencers [3]. These users may only have a small number of followers, but often have reach into highly targeted audiences. This has radically increased the number of accounts actively promoting products, thereby making the industry challenging to regulate. To address this, standards bodies such as the UK's Advertising Standards Authority (ASA) have created a set of rules asking that such post are explicitly tagged with relevant hashtags (*e.g.* #ad) [4]. There have been a number of recent studies looking at influencer marketing including interviews with marketers [5], [6] and customers [7].

This paper makes two contributions. *First*, we present a broad characterisation of influencers on Instagram, and formally quantify their behavioural attributes. To this end, we collect a large-scale Instagram dataset consisting of 12k influencers (§II). We gather both Instagram posts and stories[1] from users who attach advert-related hashtags to their material. With this dataset, we proceed to explore influencer characteristics (§III). We see a range of promotion activity taking place, ranging from "mega" influencers with over 1 million followers to "nano" influencers with as few as 500 followers. We find a range of products being produced with health & beauty taking the top spot, followed by services, clothing and food featuring prominently. Naturally, mega influencers garner the greatest attention, as measured via likes and comments. Yet we find that small "nano" influencers sustain attention more effectively. Then, in our *second* contribution (§IV), we strive to build automated tools to detect sponsored posts that fail to explicitly tag themselves. We propose a Contextual LSTM Neural Network classifier which considers the text content and other metadata to identify whether a post is sponsored or not. Our findings have important implications for social media companies, marketers, researchers, and regulators wishing to better understand the behaviour of influencers.

II. Data Collection Methodology

A. Data Collection

Phase 1: Hashtags. It is first necessary to obtain a large list of influencer accounts. We compile the list by crawling all posts attached to a set of influencer-related hashtags. We turn to the UK's Advertising Standards Agency [4], which states that influencers should use the #ad, #advert or #sponsored hashtags in *any* posts that have been paid for. We expand this list with #advertising, #giveaway, #spon and #sponsor [8].

Phase 2: Post & Stories Collection. We then use the official Instagram API [9] to gather all posts and stories that include any of the above hashtags. Note that stories are similar to normal posts, yet they are automatically deleted after 24 hours (akin to Snapchat posts). Hashtag Engine is used [10] with a maximum of 30 unique hashtags. Our crawl for posts and stories ran between Sep 2018 and April 2019. This process identifies 12K accounts.

Phase 3: Account Collection. We next extract all accounts identified from the Phase 2 dataset and begin dedicated monitoring for all posts and stories generated by those users. (*i.e.,*

[1]Stories are time limited posts that automatically delete after 24 hours.

IEEE/ACM ASONAM 2020, December 7-10, 2020

Fig. 6: A non-declared sponsored post: 1) A verified Mega influencer is promoting. 2) The product is tagged and mentioned. 3) There is no sponsored metadata in the caption. 4) The post received reasonable amount of reactions.

that Mega influencers tend to dominate: they are the major publishers for 77% of product types we identify. This is in sharp contrast to Macro (14%) and for Micro (9%). These findings confirm the intuition that Instagram is dominated by promotions surrounding consumables such as food, retail and beauty. These cover 43% of all adverts.

IV. Predicting Sponsored Content

A. Classifier Design

We take a supervised deep learning approach, as we have a substantial ground-truth (annotated) dataset of sponsored *vs.* non-sponsored posts.

Dataset. For classification, we use the post dataset described in Section II-A. In total, the dataset consists of nearly 7K sponsored and 27K non-sponsored posts from all groups of influencers. As we are dealing with an imbalanced dataset, we use Random Under-Sampling to reduce the size of the non-sponsored class. Accordingly, we randomly select posts from non-sponsored class and remove them from the training dataset. While selecting the posts, we try to keep the diversity of influencer's profiles. Finally, we produce a dataset with 14k observations with an equal number of sponsored and non-sponsored labelled posts (Section II-A). Note that all these posts are generated by users who have posted at least one sponsored post. Hence, our classification task is limited to differentiating between sponsored and non-sponsored posts from influencers (rather than a more general audience). Based on our observations, the posts in the above dataset can be separated into three sup-populations:

(1) **Sponsored Post.** In a sponsored post, influencers normally try to directly or indirectly advertise a product. By adding sponsored metadata (Section II-A), the post is explicitly declared as an advertisement. We note that the main company is often tagged in the post.

(2) **Non-Sponsored Post.** In contrast to a Sponsored Post, this is a routine post by an influencer which does not include any direct or indirect advertisements.

(3) **Hidden Sponsored Post.** This sub-population is similar to the first (Sponsored Post) except that sponsored metadata (*i.e.,* hashtags) are removed. We remind the reader that The Advertising Standards Authority (ASA) [14] advises the use of such hashtags. Despite this, we posit that there is an incentive for influencers to avoid disclosure, as there is greater value in "personal" endorsements. "*Hidden Sponsored Post*", "*Hidden Advertisement*" and "*Non-declared Sponsored Post*" terms address the same meaning in this paper.

Feature Engineering & Pre-processing. Table III summarizes the features that are employed. We split the feature list into two main categories: (*i*) Post features comprise all features that are obtained from the content of the post, *e.g.* number of likes, the caption. And (*ii*) Publisher features cover features extracted from the profile of the publisher. For all text-based features such as "post caption" and "post hashtag", we remove all punctuation marks, stopwords and convert them to lowercase characters. Words are stemmed to reduce to their root forms.

TABLE III: Feature Set Used for Classification.

Post Features		Publisher Features	
Feature	**Type**	**Feature**	**Type**
post caption	*text*	number of follower	*numeric*
post hashtag	*text*	number of followee	*numeric*
number of likes	*numeric*	length of biography	*numeric*
number of comments	*numeric*	profile biography	*text*
length of caption	*numeric*	is verified	*numeric*
number of hashtags	*numeric*	external URL exist	*numeric*
number of mentions	*numeric*		
number of tagged users	*numeric*		

Model Architecture & Performance. Next, we propose a Contextual LSTM Neural Network architecture and we compare the result with a Random Forest classifier. Post metadata can enrich the information available for classification. So, to build the model, we combine the text with other metadata features (Table III). First, we tokenize text metrics (*e.g.* caption and hashtag) using the Keras Tokenizer Class [15] and then the result is fed to the LSTM layer which outputs a 64-dimension vector. Next, we attach numerical metadata (post and profile features) to this vector and pass it through 2 ReLU activated layers of size 128 and 64. Finally, it connects to an output layer that predicts the label. We use TensorFlow, Keras, and Scikit-learn libraries, and we run them on Google CoLab. Also, we use a random split of 80% (training set) and 20% (test set) and to avoid overfitting, we use 10-fold Cross-validation. For our labelled data, we observe high performance, with positive results across both classifiers. Our Contextual LSTM Classifier enhances the performance by nearly 5% (89% of accuracy).

TABLE IV: Classifier Performance

Model	Accuracy	Precision	Recall	F1
Random Forest	0.84	0.83	0.84	0.83
Contextual LSTM	0.89	0.88	0.87	0.89

B. Results

The results are shown in Table IV. We believe some influencers do not declare sponsored posts with defined metadata (§II-A), so, we use the trained model on the non-sponsored posts (which are not used in the training). This part includes nearly 20K posts from all 4 groups of influencers.

Results Summary. Our model detects that 17.7% sponsored posts are published without metadata (not declared as sponsored) which is approximately 3.8K posts across all influencers. We see a noticeable set of hidden sponsored posts, particularly for Micro and Nano influencers. Nano with 18%, micro with 21%, and macro with 11% of shared posts, advertise almost 4 times more than Mega. In detail: (1) **Mega:** Nearly 29% of Mega influencers publish one non-declared post. Out of 3K non-sponsored posts, 4.3% are detected as hidden advertisement entities. (2) **Macro:** 57% of Macro influencers have one hidden sponsored post. 13% of 3.6K non-sponsored posts are recognised as hidden advertisements. (3) **Micro:** is the largest group in terms of publishing non-declared advertisements. 78% of Micro influencers publish one sponsored post. 24% of the 5K non-sponsored content is detected as hidden advertisements. (4) **Nano:** More than 70% of nano influencers publish at least one non-declared advertisement. From 9.7K non-sponsored posts, 20% are detected as hidden advertisements.

Validation. To validate the above results, we manually validate 50% of the detected posts to confirm they are hidden sponsored posts. Across the identified content, we witness in 81% of posts, influencers are confirmed as promoting products, 11% are incorrect, and in 8% we are not able to verify they are commercial or not. This confirms that our model *is* effective at identifying influencer posts.

C. Text Feature Analysis

Finally, we manually explore which textual features (Table III) are most prominent in our prediction task. We do this to gain an understanding of what characteristics are most common specifically for (hidden) sponsored posts.

Post Caption holds the most valuable information of the post in the shape of the text (Fig. 6). Often the text includes details of products, promotions, names, *etc.* and encourages users to perform an action. An action could be to buy something, follow a page, join a competition, install an application, use a specified hashtag, *etc.* We also notice a peculiar way of writing the text or using keywords as follows. If X represents a *product* or *producer*: (*i*) the caption includes X as raw text in 94% of the sponsored posts. (*ii*) Also X is mentioned (followed by @ sign and becomes blue) in 91% of the posts. (*iii*) The caption could contain particular sentences such as "*thank you X*", "*many thanks to X*", "*X from this page*", "*my top choice is X*", "*go and follow X*", *etc.* . This happens in 78% of the sponsored posts. (*iv*) We observe call-to-action keywords such as "*link in bio*", "*download it*", "*watch my story*", "*Use discount code*" and "*comment/like to win*" in 53% of the posts.

Post Hashtag is the next leading text feature which generally includes valuable information about the product and its producer. Influencers may use one or more hashtags for the corresponding product (Fig. 6). For example, if *#outfit* is the main hashtag representing the product, we also observe *#outfitday*, *#oufit*, *#bestoutfit*, *#trendyoutfit*, *#outfitstyle*, *etc.* In 97% of the sponsored posts, the name of the product or producer (or both) is listed as a hashtag. Hashtag count also helps differentiate sponsored content. Sponsored posts one to get hashtags trending, or (*ii*) to get more visibility in Instagram explorer, or (*iii*) to be easily found in search results, regularly have longer hashtag length:

Profile Biography is another feature that improves the accuracy of our classifier. In general, influencers usually put relevant information in their profile biography (which could be temporary): (*i*) In 63% of the biographies, information such as "sponsor info", "campaign details" or "promotion code" exists as raw text. (*ii*) In 54% of the profiles, influencers put sponsor "hashtag(s)" (which becomes blue) or mention "product/producer page" (followed by @ sign). (*iii*) In 34% of the profiles, there is a call-to-action phrase with special keywords such as "follow", "buy", "sale", "watch", "join", "check out", "promotion", "more info" *etc.* (*iv*) Profiles may include an External URL (in 21% of the profiles) or YouTube link (in 11% of the profiles), which redirects users to the main product webpage, producer website, or full youtube review video. Biography length is also helpful as influencers who do promotion usually have (*i*) promotion codes, (*ii*) sponsor contact detail, and (*iii*) product-related hashtags on their biography (longer biography).

V. Conclusion & Future Work

This paper has performed the first large-scale analysis of influencer behaviour on Instagram. Rather than solely discovering "celebrities", our methodology has exposed a large array of influencer types, including highly targeted nano influencers. We have further trained a DNN model to classify posts as sponsored, allowing us to identify seemingly sponsored posts that are not properly declared. As influencer income is taxable, we note that this may have financial ramifications that go beyond the issues of deceiving consumers.

References

[1] Thorsten Hennig-Thurau, Kevin P Gwinner, Gianfranco Walsh, and Dwayne D Gremler. Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of interactive marketing*, 18(1):38–52, 2004.

[2] Duncan Brown and Nick Hayes. *Influencer marketing*. Routledge, 2008.

[3] Luis V. Casaló, Carlos Flavián, and Sergio Ibáñez-Sánchez. Influencers on instagram: Antecedents and consequences of opinion leadership. *Journal of Business Research*, 2018.

[4] ASA. An influencer's guide to making clear that ads are ads, 2018.

[5] Sofie Biaudet. Influencer marketing as a marketing tool: The process of creating an influencer marketing campaign on instagram. 2017.

[6] NL Ewers. # sponsored–influencer marketing on instagram: An analysis of the effects of sponsorship disclosure, product placement, type of influencer. Master's thesis, UT, 2017.

[7] Vaibhavi Nandagiri and Leena Philip. Impact of Influencers from Instagram and YouTube on their Followers. *International Journal of Multidisciplinary Research and Modern Education*, 2018.

[8] US Federal Trade Commission. Us federal trade commission. https://www.ftc.gov/news-events/press-releases/2017/09/csgo-lotto-owners-settle-ftcs-first-ever-complaint-against.

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

[10] Instagram. Instagram hashtag search. https://developers.facebook.com/docs/instagram-api/guides/hashtag-search, September 2019.

[11] Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, and Krishna P Gummadi. Measuring user influence in twitter: The million follower fallacy. In *AAAI conference*, 2010.

[12] Instagram. What are the requirements to apply for a verified badge? https://help.instagram.com/312685272613322, September 2019.

[13] Json Hjh. Instagram stories vs feed ads. which is more effective at driving traffic? https://www.agorapulse.com/social-media-lab/instagram-stories-ads, Feb 2018.

[14] The Advertising Standards Authority Ltd. (trading as ASA). https://www.asa.org.uk.

[15] François Chollet et al. Keras. https://keras.io, 2015.