Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 122 (2017) 138–145

Procedia Computer Science

www.elsevier.com/locate/procedia

Information Technology and Quantitative Management (ITQM2017)

Multivariate Features Based Instagram Post Analysis to Enrich User Experience

Vatsala Mittal[a], Aastha Kaul[a], Santoshi Sen Gupta[a], Anuja Arora[a,*]

[a]*Jaypee Institute of Information Technology, Noida, India*

Abstract

In today's digital world, wherein user personalized content such as text, video, photos and much more have become an integral part of people's daily lives; photo intensive social media applications have acquired enhanced adoption in social media users through Instagram. By the time, Instagram- a photo sharing site continues to evolve and grow in popularity. Although Instagram has rapidly gained popularity among active social media users, analysis of the interaction and engagement among people on Instagram is missing and is almost an untouched area. In this paper, we inspect some prominent user interaction properties and photo properties to understand users' engagements towards posts on Instagram. The considered user interaction properties are hashtags, photo post time etc as users' posted photo context. On the other end, photo properties are user's posted photo features or image contexts such as image filters. We have performed these user interaction properties and photo properties analysis task on eight major cities' Instagram posts and further classified the posts of these eight cities in five categories using Non-negative matrix factorization and latent Dirichlet allocation algorithm. The four prime influencing analyses have been computed to get ecology of the users on Instagram photo posts, which are Time based analysis (TBA), Image Filter analysis (IFA), Image Hashtags analysis (IHA) and Image categorization analysis (ICA). Henceforth, this multivariate feature based Instagram analysis will help users to gain insight of popular content and make their respective content popular so as to reach out to a maximum number of people.

Peer-review under responsibility of the scientific committee of the 5th International Conference on Information Technology and Quantitative Management, ITQM 2017.

Keyword: Instagram Post Analysis; Instagram; Features Analysis; Non-negative Matrix Factorization; Latent Dirichlet Allocation; Image Filter Analysis.

1. Introduction and Related Literature

Instagram is a way to capture and share the moment application which allows its user to share pictures and videos either publicly or privately. Instagram is used by people as a platform to interact with each other, share personal photos, videos, views, and reviews on different topics of daily life, politics, sports, markets and much more. According to Pew research centre survey report, a number of brands marketing on Instagram have grown

* Corresponding author. Tel.: +91-9810982039
E-mail address: anuja.arora29@gmail.com.

1877-0509
Peer-review under responsibility of the scientific committee of the 5th International Conference on Information Technology and Quantitative Management, ITQM 2017.
10.1016/j.procs.2017.11.352

from 11.5% in 2013 to 32.3% in 2015. Moreover, it is one of the best platforms to advertise on – a Forrester report has shown that engagement with brands on Instagram is 10 times higher than Facebook, 54 $_{\text{times}}$ higher than Pinterest, and 84 times higher than Twitter. In fact, according to Pew research centre, 32% of online adults use Instagram of which female internet users are more likely to use Instagram than men. However, unlike many social media networks, Instagram doesn't offer its own analytics API to do analysis according to hashtags, location or other Instagram user properties/ image properties. A wide range of emerging problems have been evolved for researchers due to the rapid usage of social media and diversity of sharing content on these social media sites. Users post a variety of content on social media sites [13] such as text, image, audio, video etc., out of which photos and videos are considered to be complex content to analyze. As this research work depends on Instagram images as social media content, we first summarized the work done by various researchers in this direction which is primarily detailed in this section and depicted in figure 1.

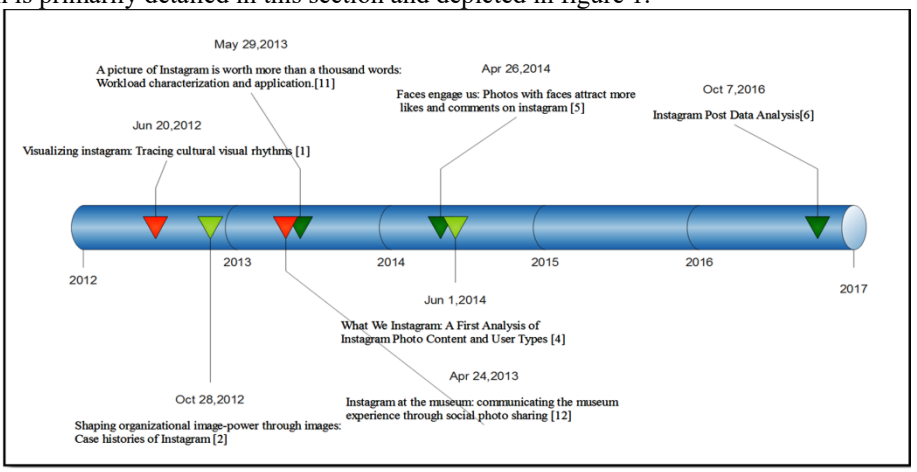

Fig. 1. Instagram Literature summary

The initial impetus to analyze Instagram was attempted by Hochman et. al. in 2012 [1]. In their work, they detailed about cultural analytics visualization techniques and experimented on 5, 50,000 images dataset which was collected by different Instagram users. Initially, analysis of 'New York' and 'Tokyo' photos was done and depicted a comparative visualization results that helped in differentiating among local colour usage, production rate, and hue's intensity. Basically, the dataset was explored in relation to various dimensions mentioned as Visual Rhythms such as brightness, saturation, color, texture etc.The paper used three fundamental techniques-

- ImagePlot*- Position Images according to their visual features.
- Montage† – a 2D grid of images and slice sequence of parts of images over specific time period.
- ImageJ‡ software was then used to perform visualization.

In the same year 2012, McNely [2] proposed and explored a coding scheme to understand organizational implementations of Instagram, within important news organization- a non-profit and a for- profit retailer. His work was primarily inspired by Spinuzzi [3] work. Spinuzzi mentioned in his work that Instagram Images were central to social software ecologies and had helped Faber [7] company to build their organization's self conscious, self-reflective management of public perception. Instagram has been the front runner in this and its adoption is significant and it has become a vital way to shape up the organizational image-power. McNaley's

* http://lab.softwarestudies.com/p/imageplot.html

† http://rsbweb.nih.gov/ij/plugins/image-montage/index.html

‡ http://rsbweb.nih.gov/ij/