Artificial Intelligence Research and Development
A. Cortés et al. (Eds.)

doi:10.3233/FAIA220317

Influence in Social Networks Through Visual Analysis of Image Memes

Carles ONIELFA [a,1], Carles CASACUBERTA [a] and Sergio ESCALERA [a,b]
[a] *Departament de Matemàtiques i Informàtica, Universitat de Barcelona, Spain*
[b] *Computer Vision Center, Spain*

Abstract. Memes evolve and mutate through their diffusion in social media. They have the potential to propagate ideas and, by extension, products. Many studies have focused on memes, but none so far, to our knowledge, on the users that post them, their relationships, and the reach of their influence. In this article, we define a meme influence graph together with suitable metrics to visualize and quantify influence between users who post memes, and we also describe a process to implement our definitions using a new approach to meme detection based on text-to-image area ratio and contrast. After applying our method to a set of users of the social media platform Instagram, we conclude that our metrics add information to already existing user characteristics.

Keywords. Memes, clustering, social media, social network, influence, culture, DBSCAN, CNN, graph

1. Introduction

A meme is usually defined as "an idea, behavior, phrase or usage that spreads within a culture" [1]. In the digital era, memes have adapted to new technologies and have become a phenomenon in contemporary web culture [2]. As a combination of humor, text, and a symbol, emoticons became one of the first types of Internet memes.

Even though Internet memes can exist as text, emojis, videos, or gifs, a common format is that of an image with superimposed text that conveys some type of merged message in an epigrammatic style. In the earlier days of the Internet, images with superimposed text began to propagate via e-mail and message boards. Later, social networks emerged, allowing memes to viralize [3]. Image memes have become an integral part of Internet culture. With the help of users they are born and reproduced, often mutated in the process. They are also used to spread political messages and ideologies. Compared to textual memes, image memes can condense their content and require less attention to be understood. Therefore, they are likely to be more effective [4].

Many studies have been carried out around memes, mainly focusing on their evolution [5], predicting their virality [4, 6], modeling their spread with mathematical models [7, 8], or devising algorithms for detecting them [3, 9]. But few, if any, have dug deeper behind the creators of memes.

[1]Corresponding Author: Carles Onielfa, Universitat de Barcelona, Gran Via de les Corts Catalanes 585, 08007 Barcelona, Spain; carlesonielfa@gmail.com.
This work has been partially supported by MICIN/AEI under projects PID2019-105093GB-I00 and PID2020-117971GB-C22, and by ICREA under the ICREA Academia programme.

Regarding human achievement, viral success is closely related to merit [10]. Therefore, it is natural that memes that were once uploaded anonymously are now being uploaded by users that are proud of their creations and sign their memes with their watermark. Some users who post popular memes have achieved massive followings, and this grants them enormous influence and reach. However, that would be true of any user on a social network with a big number of followers. What makes meme creators unique is that they not only have the power to reach their followers, but two factors greatly expand their scope. First, memes are meant to be spread and shared; hence, followers of meme creators, if they enjoy a meme, are likely to share it with their friends [11]. Second —and most importantly— from an original meme, other creators can mutate and alter the original to make their own, retaining core aspects of the meme such as the underlying image. If there was an idea or product within it, as the meme and its mutations viralize and are shared, the idea or product goes viral with it, achieving exposure orders of magnitude greater than the original reach of the creator of the meme.

In this study, we take over the task of providing tools to gain insight into creators and the relationship between them through a visual analysis of the content of their memes. Specifically, we provide a definition of a graph for visualizing relationships between users who post memes on social networks, together with metrics that evaluate and rank users (Section 3) and a process for experimentally building the corresponding graph (Section 4). We undertake the detection of image memes using a new approach, namely the extraction of features using a convolutional neural network (CNN) and the clustering of memes by their underlying images. As an example, we apply our definitions to a set of Spanish users of Instagram who publish memes, and comment on the results (Section 5).

We conclude that, thanks to our ranking, we are able to determine the users with the biggest potential for publishing viral memes, and that some of these would be overlooked by using standard metrics for determining influence.

2. Related Work

Recent studies in network analysis have analyzed how culture and behavior are spread via social network ties, yet without focusing on the phenomenon that revolutionized culture spread in social networks, namely memes. Likewise, studies in computer vision have analyzed image memes and have attempted to detect memes or cluster them together, without taking a look at the users who post them. This study bridges the gap.

User Influence in Social Networks. Many studies have reported that behaviors or preferences of people can spread via social ties in social networks, mainly getting their knowledge through surveys [12, 13, 14]. However, to our knowledge, only [15] has derived the users' characteristics from what the users post online. In [15] a CNN was used to classify the images that a user posted online. Then, the categories of the images of the users were compared among friends and random users to find that socially tied individuals are more likely to post images showing similar cultural lifestyles.

Meme Detection and Clustering. There have been studies that use memes and phrases extracted from news and blogs to track and study the dynamics of the news cycle [16] and research into clustering text-based memes on Twitter [17]. More in line with our study, the research in [18] was able to cluster image streams using perceptual image hashes (pHash). They identify memetic clusters using meme annotation from sites such as "Know Your Meme". One recent approach to meme detection is the Memesequencer model developed in [9]. However, research in [9] is limited to memes that have identifiable templates, previously documented on sites like Memegenerator or Quickmeme. Another approach to meme detection is the Meme-Hunter model from [3], which uses multi-modal deep learning. Their model combines image features, text features, and facial detection. However, they only consider memes as pictures with superimposed text in impact font or text placed in white space over a picture.

In comparison with these works, our approach to meme detection fits a much broader definition of meme and is more in line with the ever-changing landscape of memes, as it does not require the template to be previously cataloged.

3. Formalization of the Problem

In this section, we detail concepts about memes and formalize the context of our problem.

Definition 1. A *meme* is a virally transmitted image embellished with text, usually sharing pointed commentary on cultural symbols, social ideas, or current events.

This definition of meme could be expanded to contain videos, text or simply cultural references. However, within this study we only consider image memes.

Examples of memes can be seen in Fig. 1. Given a meme, we refer to its *meme format* as the underlying image of the composition. A meme format can often be used to create more than one meme by adding or changing the existing embellishments. An example of a meme format is shown in Fig. 2.

Figure 1. Three memes.

Figure 2. A meme format, known as “Galaxy Brain”.

For any meme format, there exists a meme that used it first. Given a set of users U and a meme format F, the *pioneer* of F within U is the user $u \in U$ who published the oldest meme with the format F. If the set of users is the set U_{tot} of all users on all social media platforms, then we refer to the pioneer as an *absolute pioneer*.

Definition 2. Given a set of users $U = \{u_0, \dots, u_n\}$, not necessarily belonging to the same social network platform, we define the *meme influence graph* of U as a directed weighted graph (U, E, w) with the following properties:

1. A pair (u_i, u_j) with $i, j \in \{0, \dots, n\}$ and $i \neq j$ is in the set E of edges if the user u_j has posted a meme whose format was pioneered by u_i.
2. $w(u_i, u_j)$ is the number of memes posted by u_j whose format was pioneered by u_i.

A meme influence graph $M = (U, E, w)$ is called *maximal* if $U = U_{\text{tot}}$, that is, if every user in every social media platform is in the set U.

Definition 3. Let (U, E, w) be a meme influence graph for users $U = \{u_0, \ldots, u_n\}$.

1. The *out-degree* of a user u_i is the number of outgoing edges $(u_i, u_j) \in E$ from u_i, that is, the number of other users who have used a meme format pioneered by u_i.
2. The *in-degree* of u_i is the number of incoming edges $(u_j, u_i) \in E$ to u_i, counting how many other users have pioneered meme formats that u_i has used.
3. The *weighted out-degree* of a user u_i is the sum $\sum_{j \neq i} w(u_i, u_j)$ of the weights of the outgoing edges from u_i. It is the number of memes published by other users who have used a meme format pioneered by u_i.
4. The *weighted in-degree* of a user u_i is the sum $\sum_{j \neq i} w(u_j, u_i)$ of the weights of the incoming edges to u_i, indicating how many memes have been published by u_i with a format pioneered by some other user in U.

The PageRank algorithm [19] applied to a graph measures the importance of each of its nodes taking into account the number of incoming edges and the importance of the source nodes of these edges. In short, a node will be important if other important nodes link to it. If A is the adjacency matrix for a graph (U, E, w), the *reverse PageRank* of the node u_i is the value that the PageRank algorithm for the graph with adjacency matrix A^t (transpose of A) assigns to u_i. By computing the PageRank in this manner, one gives importance to the outgoing edges instead of the incoming edges.

Definition 4. The *score* of a user u_i is the value that the reverse PageRank algorithm assigns to u_i.

For a maximal influence graph, degrees can be interpreted as follows. The out-degree of u_i is the number of users who have been inspired by memes of u_i, while the in-degree is the number of users who have influenced u_i when creating memes. The weighted out-degree of u_i is the number of memes that have been influenced by u_i, and the weighted in-degree of u_i is the number of memes from u_i that have been influenced by some other user. Since a user, when creating a meme, can be inspired by a meme from a user who is not the pioneer of the meme format, the influence from a pioneer on a user is assumed to be indirect. In the case of a non-maximal influence graph, we can also use the previous interpretations but with some nuances. Suppose that the pioneer is not the absolute pioneer of a meme format. In that case, there might not even be an indirect relationship of influence, since given a user u_j in a set of users U who published a meme with a format F with pioneer $u_i \in U$, there exists a possibility that u_j first saw the format F from another user $u_k \notin U$. Therefore, the relationship of influence on a meme influence graph that is not maximal has to be interpreted as potential influence.

4. Implementation

The process for building a meme influence graph (Definition 2) is shown in Fig. 3. The input is a set of users U and the output is the meme influence graph for those users. Even though the meme influence graph is defined for users of any social media platform, in this study we limited the scope to Instagram. Since Instagram only allows users to publish images and videos, it is likely to find users whose content is mainly image memes. Furthermore, Instagram is the third biggest social media platform [20] and, on this platform, it is common for brands to partner with influential users (influencers) and publish sponsored posts [21]. Thus, metrics for determining influential users are valuable.

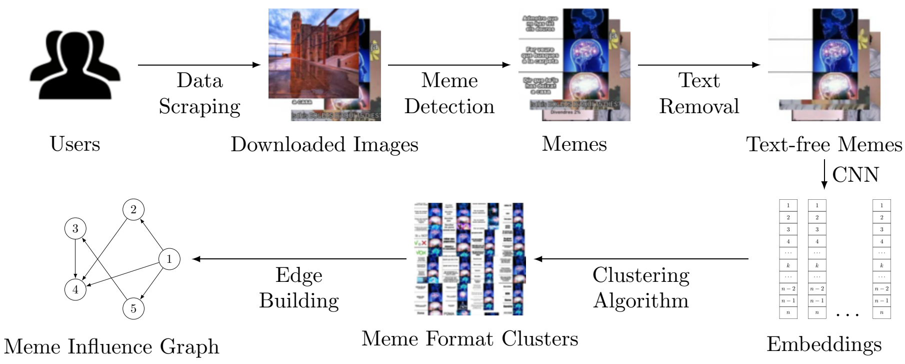

Figure 3. Flow diagram for the creation of a meme influence graph.

Data Scraping. Data for Instagram were extracted by storing the responses of Instagram's API to the calls that the browser made when browsing relevant data. The data accessed in this study are 100% public and accessible by anyone. Retrieving user data is not strictly necessary for building a meme influence graph, but having some user characteristics enables us to interpret the graph and compare the metrics from the graph to the existing features of the users. Relevant features are their usernames, amount of followers, number of posts, average comments and likes per post, and text-to-image area ratio (computed after processing all the images from the user).

Algorithm 1 Meme Detection Algorithm

Require: $I :=$ image to process of size (w,h)
$\alpha, \beta :=$ lower and upper bounds for the text-to-image area ratio
$\gamma :=$ minimum standard deviation threshold
Ensure: *true* if the image is detected as a meme, *false* otherwise
$p \leftarrow \text{text detection}(I)$ ▷ Detect areas containing text on the image using the CRAFT text detector [22]
$A_{\text{tot}} \leftarrow w \times h,\ A_{\text{text}} \leftarrow \text{area}(p)$ ▷ Compute total image area and text area
$r \leftarrow A_{\text{text}}/A_{\text{tot}}$ ▷ Compute text-to-image area ratio
if $r \notin (\alpha, \beta)$ **then** ▷ If the text-to-image area ratio is not within bounds
    return *false* ▷ The image has either only text or no text and it is not a meme
end if
$I_{\text{inpainted}} = \text{inpaint}(I, p)$ ▷ Inpaint I using the Navier–Stokes algorithm [23] with p as inpainting mask
$\sigma = \text{std}(I_{\text{inpainted}})$ ▷ Compute standard deviation of grayscale values of the inpainted image
if $\sigma \leq \gamma$ **then** ▷ If after text removal the image has high grayscale deviation
    return *false* ▷ There was no content of substance left after removing the text, so it is not a meme
else ▷ If after text removal the image has low grayscale deviation
    return *true* ▷ There was an underlying image after removing the text, so the image is a meme
end if

When browsing the user's publications (or posts), the only essential information are the images within them. As with the user data, features from publications were also valuable for later study, so they were extracted as well. Features extracted from posts were the user who posted it, the text, the amount of comments and likes, the date and time when the post was published, the hashtags added by the user, and whether the image had been detected as a meme or not (after processing it). In the case of Instagram, one publication can contain more than one media attachment (we call these publications *albums*) and the attachments can be images or videos. Videos have been treated as images using their first frame. The first frame of a video is a good representation of the media in this context, since meme videos using the same format have very similar first frames.

Meme Detection. In line with our broad definition of meme, the task that the meme detection algorithm had to perform was to discard images with no text or no underlying image. The process used to perform this task is described in Algorithm 1.

Embeddings. From a text-free meme, we have to extract features to have a lower-dimensional representation of the source image that enables us to determine differences in content between two images by comparing their features. Using text-free memes instead of the original memes with text makes the underlying meme format exposed. This diminishes the differences between memes using the same meme format and makes it easier to cluster them together in the next step.

We use the convolutional neural network VGG16 [24] pre-trained with weights from the ImageNet challenge. This neural network was chosen because it gave good results for characterizing memes in [6], which had a broader meme definition than [3] and [9]. To adapt the network to the task at hand, we set the output to the second-to-last fully connected layer, bypassing the classifier layers and giving an output of 4096 dimensions.

Deep Image Clustering. To cluster memes into groups sharing the same meme format, we input the embeddings into a clustering algorithm, namely Density-Based Spatial Clustering of Applications with Noise (DBSCAN [25]). The DBSCAN algorithm works well with a large number of samples and uneven cluster sizes. It includes outlier removal while only requiring tuning of one parameter. We apply principal component analysis (PCA) to reduce the dimensionality of the samples to 1024 for improving efficiency.

Meme Influence Graph. Finally, we build the meme influence graph (Definition 2). We add input users as nodes and then, for each cluster, we create edges from the pioneer of a cluster to the authors of the rest of the memes of that cluster. After building the graph, we compute the metrics defined in Section 3.

Scalability. Meme detection and feature extraction are parallelizable. Clustering can be scaled by using an incremental DBSCAN implementation [26] or a highly parallelizable one [27]. Scaling is limited by the speed at which data can be extracted from Instagram.

5. Results

This section contains the results of using our implementation to build the meme influence graph for a selected set of users. The implementation was coded using Python. A NoSQL database was used for locally storing the data generated at each step of the process.

Data Scraping. The study set includes 91 users and 457,101 media. Users were selected starting from two sets of 5 and 11 users who posted memes in Spanish and Catalan, respectively, and had a large amount of followers. We added users who appeared in the "Related Accounts" section of the profiles of the starting users and also posted memes. This criterion was established to obtain a set of users that we expect to be densely connected in their meme influence graph. The amount of users was limited by the speed at which the data could be extracted from Instagram. The time frame of the posts was limited to a period comprised between January 1st, 2017 and April 23rd, 2022. Within the study set, there are users who post general topic memes but also some who post topic-specific memes, such as football-themed, music-themed, or region-themed.

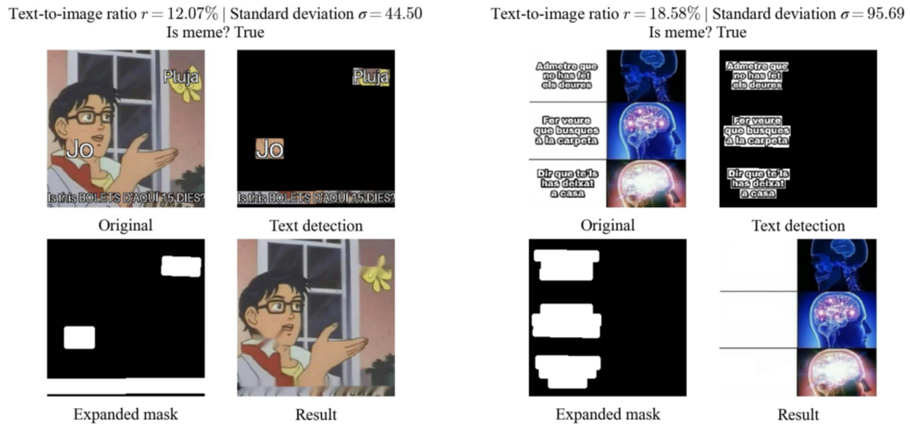

Figure 4. Images correctly detected as memes. Classification is based on text-to-image area ratio and standard deviation of gray values of the inpainted image converted to grayscale (Algorithm 1).

Meme Detection. The meme detection parameters (α, β, γ) in Algorithm 1 were found experimentally by selecting a random sample of the images on the dataset and splitting it into memes and non-memes using the meme detection algorithm. False positives and false negatives were manually identified and the thresholds were adjusted. This step was repeated several times until adjustments to the values were negligible. With this procedure, the following values were found: $\alpha = 0.018$, $\beta = 0.4$, $\gamma = 26$. Since standard deviation can vary depending on the size of the image, the images were resized to 224×224 pixels, matching the required dimensions for an input image to the neural network VGG16. In Figs. 4 and 5 we can see examples of how our meme detection algorithm processes the images. After applying the detection algorithm to all the images in our dataset, 342,984 out of 457,101 (75%) of images were detected as memes.

Embeddings and Clustering. The inputs to the VGG16 neural network are the text-free memes generated in the previous step, in a size of 224×224 pixels. The embeddings were reduced in dimensionality to 1024 components using PCA. The DBSCAN clustering algorithm was used with the cosine distance as metric, a minimum samples per cluster value of 3, and an epsilon value $\varepsilon = 0.12$. The epsilon parameter for the algorithm was tuned manually by selecting a small number of memes with popular meme formats and visualizing their clusters with an initially big epsilon value. The epsilon value was lowered in small increments until the only memes left in the selected meme's cluster were memes with the same meme format. The clustering was able to group memes using the same meme format (Fig. 6). On our dataset, the algorithm found 13,663 clusters containing 82,801 memes, and 260,183 memes were detected as noise.

Influence Graph and Metrics. We built the meme influence graph and computed the metrics defined in Section 3. The graph for our anonymized set of users is shown in Fig. 7, where high score nodes can be easily identified. By computing the Pearson correlation coefficient between each of the pre-existing characteristics of the users and each of our metrics, we found that the follower count had low positive correlations with score ($\rho = 0.29$), weighted in-degree ($\rho = 0.27$), and weighted out-degree ($\rho = 0.30$). Hence we conclude that the number of followers, which is frequently used for determining the importance of a user [28], was unable to tell the difference between incoming influ-

Figure 5. Images correctly classified as non-memes because of high text area and low standard deviation of grayscale values, respectively (Algorithm 1).

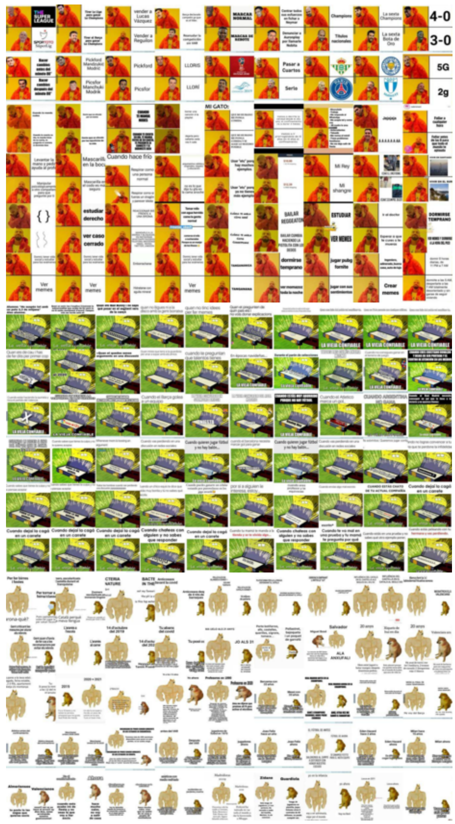

Figure 6. Some of the images in three clusters.

ence and outgoing influence in our set of users. The amount of posts published by these users had a high correlation with score ($\rho = 0.69$), weighted in-degree ($\rho = 0.85$), and weighted out-degree ($\rho = 0.73$). This matches the intuition that the more posts a user makes, the more opportunities for their memes to influence or be influenced. No correlation higher than 0.10 was found for average likes and comments per post with influence per post (weighted out-degree/media count), indicating that user engagement in posts does not correlate significantly with more influential memes.

We found communities of users sharing memes that match a certain topic or geographic area using the Clauset–Newman–Moore community detection algorithm [29] on our graph. There were communities posting football-related memes and others related to territories. Most users were not included in any community with a relevant trait.

6. Conclusions

We presented a graph and metrics on it that serve as tools to visualize the influence and the relationships of meme creators, and provided a pipeline for constructing the graph and computing the metrics. This process was implemented using a novel approach to meme detection, deep features extraction, and DBSCAN clustering.

Our ranking method could be applied, for example, in order to select candidates from a set of users for a marketing campaign using memes. By basing our criteria on their

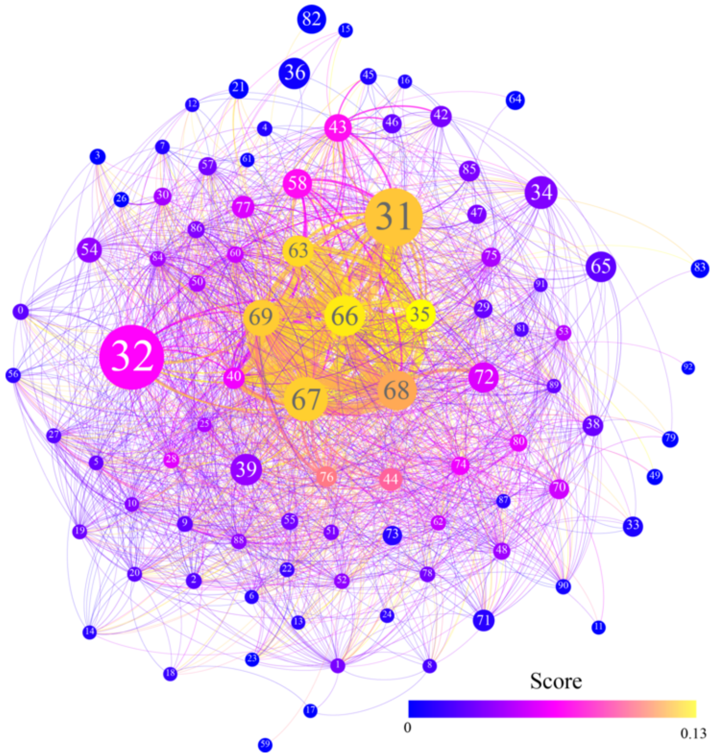

Figure 7. The meme influence graph for our set of 91 users. Nodes are labeled from 1 to 91; node size represents follower count; node color represents score from 0 (blue) to 0.13 (yellow) ; edge thickness represents edge weight; edge color matches the source node color; edge directions are represented clockwise.

scores, we ensure that memes generated by the selected users have the highest chance of viral spread through other users and reach an audience bigger than their group of initial followers. Using our graph, we can detect users who can be considered as "hidden gems", that is, users with a high score although they may not rank high with respect to their number of followers. For example, user #35 in Fig. 7 has the highest score but ranks 13th regarding the number of followers.

Limitations. This small-scale experiment does not attempt to characterize Instagram as a social network or extract information about general meme format use or virality, although observing such characteristics would be feasible if the experiment were scaled to encompass enough users. The set of users in this article does not represent a general population; therefore, the methodology can be extrapolated to other sets of users but not the results. The metrics and connections also need to be carefully interpreted according to their definition as explained in Section 3, since it is very likely that users are related with other users not represented in the graph.

References

[1] Blackmore S, Dugatkin LA, Boyd R, Richerson PJ, Plotkin H. The power of memes. Scientific American. 2000;283(4):64-73.

[2] Laineste L, Fiadotava A. Globalisation and ethnic jokes: A new look on an old tradition in Belarus and Estonia. The European Journal of Humour Research. 2017;5(4).
[3] Beskow DM, Kumar S, Carley KM. The evolution of political memes: Detecting and characterizing Internet memes with multi-modal deep learning. Information Processing & Management. 2020;57(2):102170.
[4] Ling C, AbuHilal I, Blackburn J, De Cristofaro E, Zannettou S, Stringhini G. Dissecting the meme magic: Understanding indicators of virality in image memes. In: Proceedings of the ACM on Human-Computer Interaction. vol. 5. ACM New York, NY, USA; 2021. p. 1-24.
[5] Bauckhage C. Insights into Internet memes. In: Proceedings of the International AAAI Conference on Web and Social Media. vol. 5; 2011. p. 42-9.
[6] Barnes K, Riesenmy T, Trinh MD, Lleshi E, Balogh N, Molontay R. Dank or not? Analyzing and predicting the popularity of memes on Reddit. Applied Network Science. 2021;6(1):1-24.
[7] Bauckhage C, Kersting K, Hadiji F. Mathematical models of fads explain the temporal dynamics of Internet memes. In: AAAI Conference on Web and Social Media. vol. 7; 2013. p. 22-30.
[8] Weng L, Flammini A, Vespignani A, Menczer F. Competition among memes in a world with limited attention. Scientific Reports. 2012;2(1):1-9.
[9] Dubey A, Moro E, Cebrian M, Rahwan I. Memesequencer: Sparse matching for embedding image macros. In: Proceedings of the 2018 World Wide Web Conference; 2018. p. 1225-35.
[10] Yucesoy B, Barabási AL. Untangling performance from success. EPJ Data Science. 2016;5(1):1-10.
[11] Wibowo T, et al. Usage of meme as information sharing media. Jurnal Ilmiah Betrik: Besemah Teknologi Informasi dan Komputer. 2020;11(3):165-71.
[12] Bond RM, Fariss CJ, Jones JJ, Kramer AD, Marlow C, Settle JE, et al. A 61-million-person experiment in social influence and political mobilization. Nature. 2012;489(7415):295-8.
[13] Lewis K, Kaufman J, Gonzalez M, Wimmer A, Christakis N. Tastes, ties, and time: A new social network dataset using Facebook.com. Social Networks. 2008;30(4):330-42.
[14] Christakis NA, Fowler JH. Social contagion theory: Examining dynamic social networks and human behavior. Statistics in Medicine. 2013;32(4):556-77.
[15] You Q, García-García D, Paluri M, Luo J, Joo J. Cultural diffusion and trends in Facebook photographs. In: Proceedings of the International AAAI Conference on Web and Social Media. vol. 11; 2017. p. 347-56.
[16] Leskovec J, Backstrom L, Kleinberg J. Meme-tracking and the dynamics of the news cycle. In: 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2009. p. 497-506.
[17] Ferrara E, JafariAsbagh M, Varol O, Qazvinian V, Menczer F, Flammini A. Clustering memes in social media. In: Advances in Social Networks Analysis and Mining. IEEE; 2013. p. 548-55.
[18] Zannettou S, Caulfield T, Blackburn J, De Cristofaro E, Sirivianos M, Stringhini G, et al. On the origins of memes by means of fringe web communities. In: Internet Measurement Conference; 2018. p. 188-202.
[19] Page L, Brin S, Motwani R, Winograd T. The PageRank citation ranking: Bringing order to the web. Stanford InfoLab; 1999.
[20] Keepios. Global social media statistics; 2022. Available from: https://web.archive.org/web/20220513111601/https://datareportal.com/social-media-users.
[21] Chen J. Instagram statistics you need to know for 2022; 2022. Available from: https://web.archive.org/web/20220513234913/https://sproutsocial.com/insights/instagram-stats/.
[22] Baek Y, Lee B, Han D, Yun S, Lee H. Character region awareness for text detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2019. p. 9365-74.
[23] Bertalmio M, Bertozzi AL, Sapiro G. Navier-Stokes, fluid dynamics, and image and video inpainting. In: 2001 IEEE Computer Society Conference on CVPR. vol. 1. IEEE; 2001. p. I-I.
[24] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.
[25] Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: KDD'96, Proceedings of the Second International Conference on Knowledge Discovery and Data Mining; 1996. p. 226-31.
[26] Chakraborty S, Nagwani NK. Analysis and study of incremental DBSCAN clustering algorithm. arXiv preprint arXiv:14064754. 2014.
[27] Götz M, Bodenstein C, Riedel M. HPDBSCAN: highly parallel DBSCAN. In: Proceedings of the Workshop on Machine Learning in High-Performance Computing Environments; 2015. p. 1-10.
[28] Influencer Marketing: A Research Guide; 2020. Available from: https://web.archive.org/web/20210627064440/https://guides.loc.gov/influencer-marketing/metrics-and-costs.
[29] Clauset A, Newman ME, Moore C. Finding community structure in very large networks. Physical Review E. 2004;70(6):066111.