

Exploring Content Virality in Facebook: A Semantic Based Approach

Reema Aswani, Arpan Kar, Shalabh Aggarwal, P. Vigneswara Ilavarsan

► **To cite this version:**

Reema Aswani, Arpan Kar, Shalabh Aggarwal, P. Vigneswara Ilavarsan. Exploring Content Virality in Facebook: A Semantic Based Approach. 16th Conference on e-Business, e-Services and e-Society (I3E), Nov 2017, Delhi, India. pp.209-220, 10.1007/978-3-319-68557-1_19. hal-01768500

HAL Id: hal-01768500

<https://hal.inria.fr/hal-01768500>

Submitted on 17 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Exploring content virality in Facebook: A semantic based approach

Reema Aswani*, Arpan Kumar Kar, Shalabh Aggarwal and P. Vigneswara Ilavarsan

¹Indian Institute of Technology, Delhi

*reemaswani@gmail.com

Abstract. In the current era of digitization specifically with the advent of Web 2.0, social media has become an imposing force in shaping up the way people perceive and react to the information around them. Social media platforms have empowered people to share almost instant feedback on the content posted by individuals and organizations facilitating two way interaction and better engagement among them. This continuous interaction among individuals and organizations creates huge amount of user-generated content (UGC) and associated tokens. This study attempts to understand various semantics that might affect the virality of Facebook posts. Several pages have been identified and shortlisted from domains including e-commerce, manufacturing, services and media. A total of 53,340 Facebook posts comprising of 37, 38, 168 words have been extracted using Facebook Graph API from each of the mentioned domains and subsequently analyzed using NOSQL databases. Further, the derived tokens are semantically grouped and used to gather insights by mapping to existing virality frameworks for identifying and ranking the ones that might be affecting the virality of a post. Findings indicate the virality of content shared has positive correlation with direct brand engagement, promotional offers, freebies and direct user mentions.

Keywords: social media, information propagation, content virality, semantic analysis, facebook analytics.

1 Introduction

Today, in the digital world circling around social media with recent advancements in web 2.0, the ways and means of marketing and promoting content are concentrated completely around how certain content becomes popular and subsequently “viral” [1, 2]. In other words, making sure that the shared content reaches out to as many people as possible and subsequently attracts user attention in the form of likes, comments and shares. Social media comprises of wide variety of platforms to enable users to engage with each other, dominant platforms include Facebook, Twitter, LinkedIn, Google, YouTube, etc. In terms of traffic, Facebook is the most popular website, as given by its Alexa ranking in April 2017. With the widespread use of social media specifically Facebook, the generic verb “Facebooking” has become prevalent and is often used to describe the process of browsing profiles of oneself and others [3, 4]. Within a span of

10 years, Facebook's active user base grew to more than 1 Billion users which is around 1/7th of the planet's population. This vast user base makes Facebook a very attractive marketing platform for marketers across the globe. Facebook is used by people across all age groups in most of the geographies and this makes the platform much more important than other social platforms. With time, it has grown to provide a lot more features and services that enable different media type; be it text, images, audio or video to be shared and promoted amongst current and potential consumers [5]. This rampant growth of interactive social media platforms like Facebook has not only benefited individuals for better communication and engagement but has also attracted the attention of organizations, public figures, news portals and e-commerce portals [6, 7, 8]. With the huge user base Facebook commands, it has become an important and cost effective tool for various organizations to introduce, promote, market, collect feedback about, the products and services, while engaging with the consumers in the virtual world [9].

These organizations and web portals have started leveraging the immense power of these platforms to maximize their visibility and customer outreach [10, 11, 12]. This has led to a widely used concept of 'viral marketing' that was first coined in 1997 to describe Hotmail's idea of promotion by inserting advertisements about its free email service in the end of users' outgoing emails [13]. In last two decades, virality of online content has gone through multiple transformations and different scholars have different views about the same but the basic premise remains the same where "virality" means that certain content has reached a good number of people and is being talked about. There are several factors that affect popularity of individuals and organizations on these social media platforms [14, 15].

Companies often create online ad campaigns and encourage UGC with discussions on social media anticipating that the content would be shared among others leading to popularity [16]. However, this does not happen always and some of these promotion efforts fail miserably and are not able to attract user attention. This makes us wonder whether virality is just random or are there specific characteristics that govern whether the content will be highly propagated and shared [17]. Facebook defines 'virality' as the percentage of people creating a story from a certain page post out of the total number of unique people who have actually seen the post. This becomes a good indicator of virality from the perspective of Facebook and can be very effectively explained in terms of likes, comments and shares a page post receives. The more the number of likes, comments and/or shares a post receives, the more popular or viral the post may be considered.

This study thus uses 53,340 publically available posts from Facebook (comprising of 37, 38, 168 words) from different industry domains with an aim to identify drivers of content virality [18]. Certain set of pages are selected from each domain where each page is exceptionally popular in its respective domain and has large number of followers. This study will thus prove to be beneficial in identifying various semantics that make a post in a certain domain popular or viral among the targeted audience. None of the existing studies explore content virality using a semantic token based approach with a focus on what tokens present in a post are likely to make the post go viral. The possible reasons of lack of such studies aiming at content virality semantics

may be limited application scope in traditional marketing from social media, limited dataset availability and scope of focus.

However, plethora of semantic analysis contributes towards causing the posts to be concealed from the target audience. Actual empirical data analysis will point to practical semantics and can be good pointer for marketers in planning their promotional content. Throughout the study, we also outline the data collection techniques which allow to get substantial data from an otherwise closed and restricted ecosystem of Facebook.

2 Literature Review

Literature highlights the emerging importance of social media in the current age of digitization [19]. Social media enables sharing and promotion of content on variety of topics including politics, technology, business, media and e-commerce to name a few [20, 21]. Studies also explore social media's business value and discussions surrounding selection process for selecting a platform for content promotion [22]. The impact of content shared amongst the users increases even more if the content becomes popular or goes viral. Several studies in literature highlight insights in social media marketing and content virality. Some of the studies focus on a single social media platform for the analysis of virality while others take into consideration multiple platforms and domains for the understanding of viral marketing [1]. Facebook has had its own large share of studies that have been conducted since its successful inception around a decade back. The growing sophistication of social media has opened new avenues for advertisers and marketers. This has resulted in a quest for reliable metrics to quantify the effectiveness of online messages [23]. Virality is thus considered as a combination of viral reach, effective evaluation, and message deliberation.

Social networks have become a great source of sharing opinions, ideas, information and beliefs [24]. With the availability of huge amount of data, the analysis of information diffusion has become an interesting area to explore [25]. Virality has thus become an indicator of online ad effectiveness [26]. Literature highlights several user behavior characteristics that greatly affect the popularity and subsequent virality of content [27, 28, 29]. The virality aspect includes posts [17, 30, 31], links and images/memes [32, 33, 34, 35, 36] that get popular.

Studies further focus on analysis of social posts and messages on popular social media platforms. However, only a couple of them focus on engagement metrics while others mostly deal with the communication aspects of the social media posts. Metrics surrounding virality on social platforms focus on analyzing buzz, appreciation, content traction and controversy [30]. Further, network structures and community metrics [37] also play a vital role in popularity and virality of content [38]. Studies demonstrate the importance of network retweets, follower networks and homophily when social contagions is spread [24]. Besides these, SPIN Framework [31] demonstrates the four metrics for viral content categorization as spreadability, propagativity, integration and nexus. Coursaris et al. [39] highlight how consumer engagement may be strategized using brand Facebook page messages. However, no study in literature

focuses on the semantics derived from the UGC that enables brands to identify drivers of content virality using Facebook brand pages. There are no discussions that explore the linkage between the motivation to propagate and the virality of the content. This study thus attempts to identify content semantics for four selected domains using Facebook brand pages under the listed categories. These semantics are known to attract user attention when present in the shared content.

3 Research Methodology

This study uses a tokenization-based approach for mining semantics from the public content available on Facebook pages under four broad domains of e-commerce, manufacturing, media and services. A set of popular pages have been identified for each domain for the purpose of data extraction. Further, a semantic token-based approach is adopted to capture top 10 topics that apparently gain traction when it comes to popularity of content on the platform. We detail these activities in the subsequent subsections:

3.1 Domain and Page Selection

In order to understand virality semantics on a broader level, it would be unjust to single out a domain resulting in nullification of generalizability of the findings. The study thus considers four broad domains for the analysis are e-commerce, manufacturing, media and services. The domains selected are strictly business domains that deal with customer engagement on a regular basis. From each of the selected domains, Facebook pages of a list of dominant firms are identified and subsequently analyzed. The minimum criteria considered for page selection for the relevant domains is a minimum of 100,000 fans per followers. Considering multiple pages provides a deeper look at the content that is promoted and shared helping in an in depth analysis and comparison. A total of 53,340 posts have been considered for the analysis. Table 1 lists all the pages considered for analysis of virality per domain with number of followers/fans and number of pages per post.

Table 1. Data Description of selected pages in each domain.

Domain	Page	Followers/Fans	Number of Posts
E-Commerce	Amazon	26,653,543	6122
	AmazonIN	5,213,009	2127
	Flipkart	5,429,434	2083
	Myntra	3,166,625	1758
	Shopclues	2,708,682	2738
	Snapdeal	3,964,189	1978
	Total		
Manufacturing	Honda	4,061,973	2905
	Toyota	2,814,599	3474
	Ford	3,432,091	3625

	HeroMotoCorp	1,756,770	935
	TVS Motor Company	242,078	1168
	Total		12,107
Media	channelvindia	5,644,539	3383
	MTV	48,397,370	4248
	MTV India	11,600,934	2345
	TEDxEvents	1,415,214	1120
	Yahoo	13,726,619	1723
	ScoopWhoop	3,355,158	2378
	Total		15,197
Services	AmericanExpressIndia	6,243,134	934
	amazonwebservices	140,340	1156
	FacebookIndia	170,811,531	1357
	Google	19,656,574	625
	LinkedIn	1,121,976	924
	oyorooms	330,918	2395
	TataConsultancyServices	537,169	1839
	Total		9230

The number of followers of each page is very dynamic and changes at a very significant rate. The count provided above is recorded at the time of data collection. The data collection for the pages selected from every domain is conducted using Facebook's Graph API which is the primary way to extract data out of Facebook's platform. This HTTP-based API can be used for querying data, posting stories, managing advertisements, uploading photos and a variety of other tasks that are an essential part of any application. For the scope of this study, only the data query and read aspects of the API are used. Certain pre-requisites and conditions are considered while extracting the relevant data including a unique follower/fans threshold of 1,00,000, a minimum of 140 characters of textual content and availability of posts for more than a year at least. A python script is used to collect the data. Table 2 describes the attributes collected for each post including the domain name, page name, the actual content, likes, shares and comments. Post analysis an inference column is appended depending on the semantic analysis for popularity of the content.

Table 2. Description of attributes collected Facebook posts.

Field Name	Description
Domain	Predefined value out of E-Commerce, Media, Manufacturing and Services
Page Name/ID	Name/ID of the Facebook page for which the data is being collected, example: Amazon, MTV, Honda etc.
Post Message	Actual complete content of concerned post
Number of Likes	Total number of likes for the concerned post
Number of Shares	The total number of shares for the concerned post
Number of Comments	The total number of comments on the concerned post
Inference	A computed value that specifies whether the post is "vi-

	ral”, “popular” or “ignored”
--	------------------------------

Computation of “inference” is done with the underlying consideration that the data collected for each post has associated count of likes, shares and comments for it. From a manual analysis of the sorted posts data, it is evident that the number of likes and shares and comments are mostly directly proportional to each other, i.e., as the number of likes increase, the number of shares increase and so the number of comment. It is hence safe to assume that we can rest our inference based on likes on the post. Exceptions occurring do not have an effect on the overall results due to minimal frequency. Further, for the purpose of this study if the number of likes on any post is greater than 1000 (which is 0.1% of our lower limit criteria for the number of likes any page should have), then the post is tagged as “viral”. Similarly, if the number of likes is more than 100 then, the post is considered as “popular” (it may become viral over a period). In case that does not happen, the post seems to have been “ignored” by the fans/followers of the page.

3.2 Data cleaning and token generation

The collected data (37, 38, 168 words from 53,340 posts) comprising of UGC from various pages grouped under different domains is further analyzed for identification tokens that primarily affect the virality of the post. The NOSQL database (here MongoDB) facilitates a quick read and write of the volume and structure of the data being handled. Python’s inbuilt NLTK (Natural Language Toolkit) library tokenizes the post messages based on words. The number of token words for each post depends on the length of the post message. An initial cleaning process eliminates all tokens with length less than 3 characters including common words like ‘is’, ‘if’, ‘in’ and ‘an’ to name a few. The frequency for final tokens is computed and a semantic token map is created. Individual frequencies of “viral”, “popular” and “ignored” posts are computed for the pages in every domain. Further, semantic ranking based on their relevance is captured by considering the following short listing criteria:

1. Remove the tokens (words) with relatively high “ignored” count. Their count in other two frequencies can also be high but a higher ignored count suggests that this is a commonly used word as the number of ignored posts was much higher than the other two combined.
2. Shortlist top 10 tokens to be rated if:
 - The viral count is more than 100; or
 - The popular count is more than 500.
3. The criteria for ranking the top 10 is a weight based approach:
 - For each count of viral frequency, a weight of 1 is given;
 - For each count of popular frequency, a weight of 0.25 is given;
 - For each count of ignored frequency, a weight of -0.1 is given.

Fig. 1. demonstrates the top 10 tokens for every domain with their respective weights have been shortlisted. The weights for similar tokens may vary across domains as the number of posts collected pertaining to it across domains also vary. On the other hand, Fig. 2. depicts a wordcloud of the popular discussions in each of the domain.

E-Commerce Token	Frequency Ignored	Frequency Popular	Frequency Viral	Weight
http	3625	4032	2230	2875.50
win	314	569	526	636.85
comment	97	187	479	516.05
see	792	362	450	461.30
ends	44	199	412	457.35
here	1579	854	375	430.60
rules	69	170	394	429.60
official	97	161	395	425.55
chance	238	375	352	421.95
deal	717	512	271	327.30

Media Tokens	Frequency Ignored	Frequency Popular	Frequency Viral	Weight
you	820	2080	349	787
for	797	1180	196	411.30
your	363	1031	138	359.45
with	331	776	151	311.90
http	601	827	158	304.65
this	321	651	158	288.65
channel	163	603	67	201.45
here	382	598	70	181.30
her	88	299	101	166.95
image	3	119	114	143.45

Manufacturing Token	Frequency Ignored	Frequency Popular	Frequency Viral	Weight
http	2722	3490	1277	1877.30
ford	978	266	423	391.70
fans	89	391	287	375.85
show	475	566	254	348
honda	1135	390	353	337
mustang	538	334	243	272.70
our	486	236	244	254.40
toyota	788	467	190	227.95
racing	660	546	127	197.50
see	773	326	165	169.20

Services Tokens	Frequency Ignored	Frequency Popular	Frequency Viral	Weight
http	3177	2677	1270	1621.55
your	1720	591	398	373.75
search	321	626	231	355.4
facebook	40	171	258	296.75
tcs	798	361	213	223.45
google	6	339	115	199.15
new	504	216	189	192.6
digital	260	354	127	189.5
linkedin	287	134	167	171.8
india	187	146	104	121.8

Fig. 1. Domain - wise token identification.

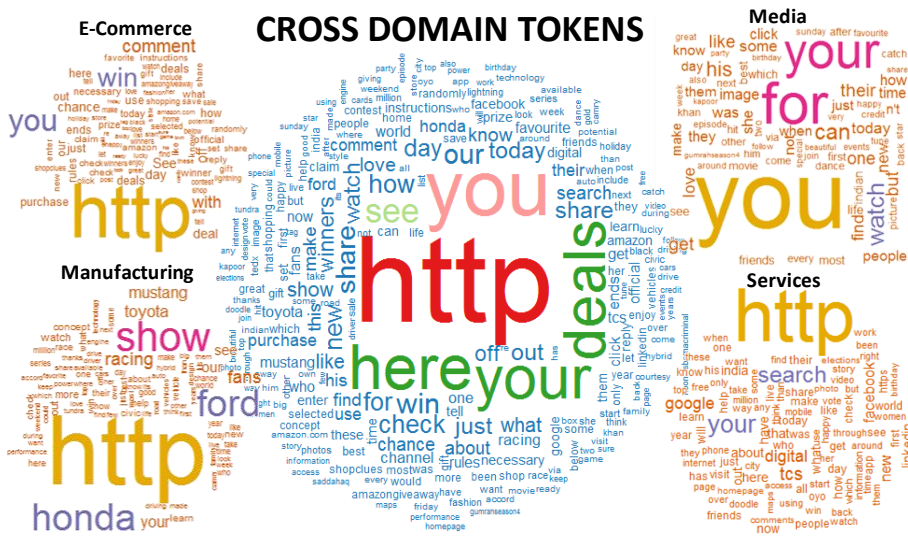


Fig. 2. Wordcloud for domain-wise and cross-domain tokens.

3.3 Cross-domain semantic analysis

Post semantic analysis for each domain, the identification of top tokens across all domains is done. The subsequent discussions analyze the relevance and impact of the top cross-domain semantics. Further, each token is analyzed using the SPIN Framework [31] to identify the impact of various parameters and how concepts of virality map to the same. The framework introduces four key factors for viral campaigns:

spreadability, propagativity, integration and nexus. The metrics relevant to this study are subsequently mapped to the identified tokens. In context with Facebook only spreadability comes into picture, since a single social media platform is considered (negating the relevance of integration) having a low cycle time (just a few clicks to share/like), high network size (limited by security features though), high level of content richness (allows multiple type of content amalgamation) and good content proximity (the options to share/like are visible very near to actual content that negates the need of considering propagativity. Further, the lack of enough data due to Facebook's security and privacy settings limits us for not considering nexus considers causal effect to future campaigns. Thus, spreadability is captured in the context of this study referring to the ability of the content of the message to appeal to the consumer in some way and prompt/motivate the consumer to take action on the same. Spreadability comprises of both likeability and shareability:

Likeability refers to the willingness of a consumer to use the content. This is often influenced by the degree to which the consumer finds the message stimulating and engaging.

Shareability is the willingness of the consumer to distribute the content. This is influenced the degree to which the consumer thinks that the content will impact others in their network as well.

From the analysis of the extracted data, it is evident that the shares of a post are directly proportional to the number of likes on a post. Further, Facebook specific findings in literature have emphasized consumers' expressing greater intentions to like a post and then subsequently sharing and commenting on it. This premise further strengthens the basis of the empirical data analysis that has been undertaken in this study of considering likeability as the primary metric of analysis [23].

Hence, an analysis on the basis of likeability may be a good indicator of shareability as well. Further, the identified tokens are semantically categorized into groups with their associated likeability. Table 3 demonstrates a list of semantics into which these token are grouped and captured across domains with their associated likeability ratings and a brief description as to why these semantics have an impact on virality.

Table 3. List of cross-domain semantics and associated likeability rating.

Semantics	Likeability	Description
links	High	Posts which contain links get more attention
brand name	High	Posts which include the brand name give a feeling of inclusion to the consumer
see/show/ check/know /watch	High	Posts which prompt the consumer to go ahead and see something more, point to an interesting content ahead
your/you	High	Addressing the consumer directly tend to get more engagement.
win/chance/ deal/ends/	High	Engage the consumers by telling them that they can win something by liking this content/post.

claim/prize		
here/this	Medium	Such posts might be offering a good deal to the consumer on the current page itself.
call to action	Medium	Prompting the consumers explicitly to share or comment on the post might trigger the action.
image	Medium	Posts containing images get more attention in a limited time span as compared to textual content

It is evident from the likeability ratings of the cross-domain semantics that the posts containing URLs are most liked and shared by the users on social media indicating that http links have a “high” likeability rating. Further, the posts containing popular brand names like ford, honda, toyota, mustang, google, facebook, linkedin and tcs to name a few in our study also manage to get user attention and subsequent popularity. In addition to this, high content likeability is seen when the users are directly addressed and engaged by using words like “you/your” and offered promotional deals, offers, chances to win something. Besides these, posts offering something by using tokens like “here/this”, reference to current content, explicit mention of sharing, promoting and commenting and images in posts also gain some traction. Further, call to action words like “comment/share/like” also explicitly motivate users to propagate the content. The remaining content seems to have “low” likeability rating and thus is unlikely to be shared and propagated among the user social networks.

4 Conclusion

Social media platforms specifically have gained immense popularity not just for communication among individuals but also as a two-way interactive platform for online marketing and content promotion. In this era of digitization, the target of each marketer is now to produce content that gets as much attention as possible. These firms not only want their content to be consumed but also shared and propagated in the consumers’ respective social networks for greater popularity and outreach. The popularity of content not only depends on its inherent value or usefulness of the information but also several other factors as discussed in literature including network dynamics, nature of propagators, user behavior and lastly the content semantics as discussed in this study.

In this study, a total of around 53,340 Facebook posts’ comprising of 37,38,168 words from various domains is captured and analyzed to form tokens and subsequently rank relevant semantics associated with it that might affect the virality of the post. The domains under consideration comprise of e-commerce, media, manufacturing and services. Each domain has its own way of attracting consumers based on the content shared by them. Further, the consumers of each domain may not necessarily be interested in similar content and thus the analysis generates different set of tokens for each domain depending on the traction of the content among the consumers. A total of top 10 semantics (in the form of word tokens) are further shortlisted from each domain

with assigned weight metrics depending on whether the content is tagged as “popular”, “viral” or “ignored”. A cross-domain semantic analysis is subsequently done by mapping the metrics to the SPIN framework to explore whether the identified semantics belong to multiple domains. Facebook specific findings in literature have emphasized consumers’ expressing greater intentions to like a post, followed by sharing and commenting on it. This premise further strengthens the basis of the empirical data analysis that has been undertaken in this study of considering likeability as the primary metric of analysis.

The results indicate that posts that have more to offer to the consumer than just a simple text get more traction and have a higher likeability among the consumers. This “more” content may be in the form of web URL links or images or simple direct referencing to the content or the user in person. Also, it is seen that when the brand engages itself directly in the post via its name, then the probability of the content getting viral increases. This is indicative of the trust factor that the brand brings with itself and promotes engagement within the consumers. Findings also indicate that virality further depends on factors like explicitly telling people to like/comment/share a post or by providing lucrative offers to the consumers about discounts, deals, freebies etc., especially for a limited time. Hence, it is seen that virality over social media, specifically Facebook, is a broad area of study and one cannot be always sure whether the content will get the desired popularity and subsequent virality but a few semantic based incorporations in the content may attract significant user attention.

5 Implications and Future Scope

This study is an empirical study for identifying semantics that may have an effect on virality of Facebook posts. A number of semantics are identified and ranked as using the data collected from 53,340 posts comprising of 37,38,168 words across four prominent business domains. This study also brings out potential linkage between motivation and virality metrics/semantics. Previous literature also focuses largely psychographics primarily personality dimensions of users. However, there are no discussions that highlight the motivation behind users’ propagating the content resulting into popularity and subsequently virality on the social networks. Thus, motivation is an important and interesting topic of research in the domain of social media virality. This study thus highlights the motivation for content shareability from Facebook posts that became popular and subsequently viral using empirical and factual metrics.

Future studies may focus on combining user related metrics like age groups, user reputation, demographics, etc. along with their user behavior characteristics, usage and consumption patterns. Further this may subsequently be used for improvising the understanding of how virality can be impacted by fine tuning the content to get best results. Future work may also focus more on improvised semantic analysis to correlate motivation and virality aspects. This may be done by considering network specific metrics like centrality, betweenness, cliques, reciprocity and propinquity to name a few. Network analytics may be beneficial in providing useful insights surrounding causality of drivers for content shareability and virality.

References

- [1] Leskovec, J., Adamic, L. A., & Huberman, B. A.: The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)* 1(1), 5 (2007).
- [2] Mangold, W. G., & Faulds, D. J.: Social media: The new hybrid element of the promotion mix. *Business horizons* 52(4), 357-365 (2009).
- [3] Facebooking, the rage on college campuses. *The Seattle Times*. <http://www.seattletimes.com/business/facebooking-the-rage-on-college-campuses/>, last accessed 2017/03/31.
- [4] Junco, R.: The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement. *Computers & Education* 58(1) 162-171 (2012).
- [5] Lampe, C., Ellison, N. B., & Steinfield, C.: Changes in use and perception of Facebook. In: *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pp. 721-730. ACM, San Diego, CA, USA (2008).
- [6] Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D.: 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).
- [7] Waters, R. D., Burnett, E., Lamm, A., & Lucas, J.: Engaging stakeholders through social networking: How nonprofit organizations are using Facebook. *Public relations review* 35(2), 102-106 (2009).
- [8] Hanna, R., Rohm, A., & Crittenden, V. L.: We're all connected: The power of the social media ecosystem. *Business horizons* 54(3), 265-273 (2011).
- [9] Viswanath, B., Mislove, A., Cha, M., & Gummadi, K. P.: On the evolution of user interaction in facebook. In: *Proceedings of the 2nd ACM workshop on Online social networks*, pp. 37-42. ACM, Barcelona, Spain (2009).
- [10] Qualman, E.: *Socialnomics: How social media transforms the way we live and do business*. John Wiley & Sons (2010).
- [11] Handayani, P. W., & Lisdianingrum, W.: Impact analysis on free online marketing using social network Facebook: Case study SMEs in Indonesia. In: *2011 International Conference on Advanced Computer Science and Information System (ICACSIS)*, pp. 171-176. IEEE, Jakarta, Indonesia (2011).
- [12] Stelzner, M. A.: Social media marketing industry report. *Social Media Examiner* 41, 1-10 (2011).
- [13] Beeler, A.: Virus without a cure. *Advertising Age* 71(17), 54-55 (2000).
- [14] Quercia, D., Lambiotte, R., Stillwell, D., Kosinski, M., & Crowcroft, J.: The personality of popular facebook users. In: *Proceedings of the ACM 2012 conference on computer supported cooperative work*, pp. 955-964. ACM, Seattle, Washington, USA (2012).
- [15] Sabate, F., Berbegal-Mirabent, J., Cañabate, A., & Lebherz, P. R.: Factors influencing popularity of branded content in Facebook fan pages. *European Management Journal* 32(6), 1001-1011 (2014).
- [16] Berthon, P. R., Pitt, L. F., Plangger, K., & Shapiro, D.: Marketing meets Web 2.0, social media, and creative consumers: Implications for international marketing strategy. *Business horizons* 55(3), 261-271 (2012).
- [17] Berger, J., & Milkman, K.: Social transmission, emotion, and the virality of online content. *Wharton research paper* 106 (2010).
- [18] Berger, J., & Milkman, K. L.: What makes online content viral?. *Journal of marketing research* 49(2), 192-205 (2012).
- [19] Kaplan, A. M., & Haenlein, M.: Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons* 53(1), 59-68 (2010).
- [20] Shirky, C.: The political power of social media: Technology, the public sphere, and political change. *Foreign affairs* 28-41 (2011).

- [21] Patino, A., Pitta, D. A., & Quinones, R.: Social media's emerging importance in market research. *Journal of Consumer Marketing* 29(3), 233-237 (2012).
- [22] Nagle, T., & Pope, A.: Understanding social media business value, a prerequisite for social media selection. *Journal of Decision Systems* 22(4), 283-297 (2013).
- [23] Alhabash, S., McAlister, A. R., Rifon, N. J., Quilliam, E. T., Sternadori, M., & Richards, J. I.: A different take on virality: the relationship among motivations, uses, and viral behavioral intentions on Facebook and Twitter. In: American Academy of Advertising. Conference. Proceedings (Online), p. 24. American Academy of Advertising. (2013).
- [24] Weng, L., Menczer, F., & Ahn, Y. Y. Virality prediction and community structure in social networks. *Scientific reports* 3 (2013).
- [25] Iribarren, J. L., & Moro, E.: Affinity paths and information diffusion in social networks. *Social networks* 33(2), 134-142 (2011).
- [26] Tucker, C.: Virality, Network Effects and Advertising. (2011).
- [27] Centola, D.: The spread of behavior in an online social network experiment. *science* 329(5996), 1194-1197 (2010).
- [28] Morales, A. J., Borondo, J., Losada, J. C., & Benito, R. M.: Efficiency of human activity on information spreading on Twitter. *Social Networks* 39, 1-11 (2014).
- [29] Roelens, I., Baecke, P., & Benoit, D. F.: Identifying influencers in a social network: The value of real referral data. *Decision Support Systems* 91, 25-36 (2016).
- [30] Guerini, M., Strapparava, C., & Özbal, G.: Exploring Text Virality in Social Networks. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011). pp. AAAI Publications, Barcelona, Spain (2011).
- [31] Mills, A. J.: Virality in social media: the SPIN framework. *Journal of Public Affairs* 12(2), 162-169 (2012).
- [32] Ienco, D., Bonchi, F., & Castillo, C.: The meme ranking problem: Maximizing microblogging virality. In: IEEE Int. Conference on Data Mining Workshops, pp. 328-335. IEEE, Sydney, TBD, Australia (2010).
- [33] Weng, L., Flammini, A., Vespignani, A., & Menczer, F.: Competition among memes in a world with limited attention. *Scientific reports* 2 (2012).
- [34] Guerini, M., Staiano, J., & Albanese, D.: Exploring image virality in google plus. In: International Conference on Social Computing (SocialCom), pp. 671-678. IEEE, Alexandria, VA, USA (2013).
- [35] Weng, L., Menczer, F., & Ahn, Y. Y.: Predicting successful memes using network and community structure. In: Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, pp.535-544. AAAI Publications (2014).
- [36] Deza, A., & Parikh, D. Understanding image virality. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1818-1826. IEEE, Boston, MA, USA (2015).
- [37] Zubeck, P. P., Chowdhury, I., & Katona, Z.: Information communities: the network structure of communication. *Social Networks* 38, 50-62. (2014).
- [38] Harrigan, N., Achananuparp, P., & Lim, E. P.: Influentials, novelty, and social contagion: The viral power of average friends, close communities, and old news. *Social Networks* 34(4), 470-480 (2012).
- [39] Coursaris, C. K., Van Osch, W., & Balogh, B. A.: A Social Media Marketing Typology: Classifying Brand Facebook Page Messages For Strategic Consumer Engagement. In: Proceedings of the 21st European Conference on Information Systems (ECIS), p. 46. Association for Information Systems, Utrecht, Netherlands (2013).