

The Social Media Indicator 2: Towards a Software Tool for Measuring the Influence of Social Media

Robin J. Withaar, Gabriella F. Ribeiro, Robin Effing

Saxion University of Applied Sciences
P.O.Box 70000
7500 KB Enschede
robinwithaar@gmail.com
fonseca.gabriella@gmail.com
robin.effing@gmail.com

Abstract: The measurement of influence of social media has gained importance. This paper introduces a matrix to measure the influence of social media from individual users. This matrix comprises the parameters to measure personal influence for Facebook, Twitter, LinkedIn, YouTube and Google+ while taking into account current technology limitations for each of them. Furthermore, backgrounds are presented to sustain the matrix with future changes in application programming interfaces of social media channels. While most available solutions regarding social media measurement use a keyword approach, our solution presents a framework focused on the individual's use of social media. This matrix and its scoring instructions can be used in future software tools for measuring the influence of Social Media.

1 Introduction

Social Media is a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content [KH04]. Social media such as Facebook, Twitter, YouTube, Google+ and LinkedIn dominate the web landscape. While Facebook has currently more than a billion profiles and registers more than 175 million active users, which is in numbers almost the population of Brazil (190 million) and over twice the population of Germany (80 million), YouTube receives 10 hours of uploaded content every single minute [KH04]. Since organizations invest time and money in relationships through these social media channels, they are looking for analytical ways to evaluate their influence. Consequently, there is an increasing importance for analyzing their impact by measuring the influence levels of the various social media profiles. However, social media measurement is quite immature, just as web analytics was back in the mid 1990's [MC12] and most currently available tools focus on search-queries and keywords.

Social media is primarily about relationships between people [SO14] and therefore we argue that the influence of a single user depends on the size of one's network. Virtual followers and friends are clearly influencing signals: anyone who follows an individual is in some way acknowledging, if not directly influenced, the willingness to be reached and thereby the possibility of being influenced [SM13].

With the keyword approach, the retrieved statistics indicate the presence and awareness of the scanned keyword in the public mass. The approach we use in this research is based on data retrieved from a smaller group of individuals who are strictly connected to the organization. Our main objective is to measure the influence of someone in the community, not the entire mass. Figure 1 shows an overview example of how a message can reach other people. The marked section indicates the people who have a direct tie to the sender of a message. While the keyword tools primarily provide information about the mass, our approach focuses on the smaller marked group.

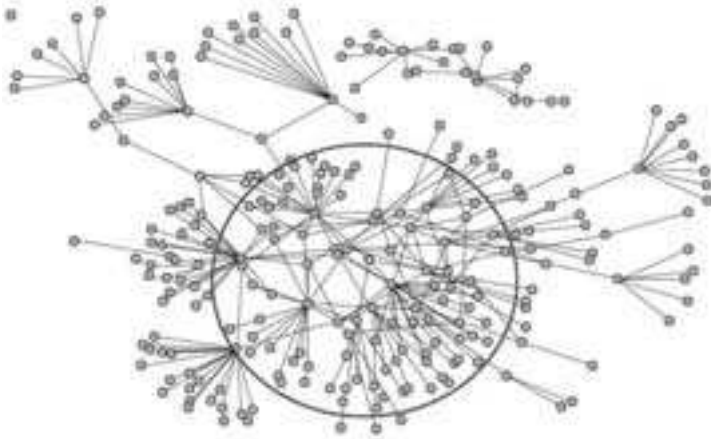


Figure 1: Social network diagram with selection of committed users

Manual observation of social media parameters can be time-consuming. The reliability of the observations can be questioned because when conducted manually, errors are easily made. With a software tool that is able to automatically analyze the social media profiles of all selected participants, this problem can be solved. Another advantage of a software tool is the capability to monitor all social media profiles of the participants in a more dynamic way with multiple measurement points in time.

The main aim of this paper is to introduce the new Social Media Indicator with an alternative way of measuring. With this matrix we expect to get a step further towards the development of software tools that can reliably indicate the influence that an organization and its individuals has on their social environment.

The remainder of this paper is structured as follows. First, we will elaborate on our method. Second, we will introduce the Social Media Indicator. Third, the scoring instructions and our algorithm for implementation in future software solutions are presented. Finally, there is a section of discussion with future research and limitations.

2 Method

In this section we will describe the current situation of social media monitoring and explain how we have designed a new version of the Social Media Indicator for measuring Social Media influence of individuals. The method consists of practical and theoretical

steps. We have based our approach on literature study, earlier field-studies and a design-science approach for developing a measurement matrix and a software tool.

Measuring influence is not about some arbitrary score; it's about following rules and best practices for how the measuring of influence has been approached for over 100 years [SM13]. In order to find a software tool that gives insights regarding the influence-level of Social Media on people, we have compared existing applications that were available to provide social media statistics. There are various online tools with this purpose. Some examples are Radian6 [RA10], Teezir [TE11], Coosto [CO07] and SocialMediaCheck [SO09]. However, these type of software tools take a keyword oriented approach. SocialMediaCheck, for instance, measures the number of times a company's name was mentioned on social media. The risk of these statistics is they can be irrelevant and unreliable because a higher number of posts does not necessarily relate to a higher level of influence. In fact, more than 90% of online mentions are spam [SM13]. Another example of a broadly used tool is Klout [KL08]. However, it still serves mainly as entertainment and is unable to provide serious influencer lists that business would be willing to pay for because the degree of certainty behind the data is too low [SM13]. Therefore it is necessary to investigate alternative approaches for retrieving data.

Most of the available tools to analyze the influence of a Social Media user are designed for the extraction of information [RE15], receiving less or none focus in research, while a big range of publications in this field present well-structured arguments and definitions of parameters but not consider limitations of technology. For instance, the Social Media Indicator proposed by Effing, van Hillegersberg and Huibers [EHH02] uses the number of posts and comments in blogs as parameters to come up to a final score. However, as blogs do not necessarily follow a strict pattern and can be implemented in different Content Management Systems, it is complicated to automate a process to retrieve data from them.

Another approach is to look at the behavior of the participants themselves on social media; the most common-known tool that uses this approach is also Klout [KL08]. After measuring an individual Social Media account, this tool provides a score between 0 and 100 that indicates how great one's influence is. This number is primarily based on how many friends an individual has on Social Media and how big the response of one's network is to one's posts based on replies and shares. The number this tool provides gives an insight of the individual's personal influence to his network compared to all other users on social media. Klout has been a subject for critique, since their method of calculation is not completely transparent and is sometimes considered biased [ES01].

By working with an earlier version of the Social Media Indicator matrix, as introduced by Effing, van Hillegersberg and Huibers [EHH02], we experienced some shortcomings in its first version. Although it provides well-structured questions for indicating contribution and interaction level in social media, the matrix does not cover the limitations and special characteristics of each social media channel.

We propose a conceptual measurement matrix based on two main concepts: contribution and participation. This division is based on the e-Participation ladder of Macintosh [MS05], the classification of parameters to measure ROI described by Hoffman and Fodor [FH03] and on the Social Media Participation Model defined by Effing, van Hillegersberg and Huibers [EHH02]. We consider contribution and participation to be the key-factors to measure influence. We argue that these two aspects of participation are complementary but not dependent. Consequently, the Social Media Indicator evaluates the use of social media by a user, providing separate scores for each of them.

In this research it is our goal to find a way of calculation that addresses every social media channel from the same level and delivers an equivalent value for each channel, regardless the number of available parameters. We base our new matrix on the theoretical backgrounds of the Macintosh participation ladder [MS05] and the ROI model described by Hoffman and Fodor [FH03] and their way of interpreting data. In the model of Hoffman and Fodor [FH03] we find a matrix where every possible action and reaction in social media is described. This matrix is divided by the concepts of brand awareness, brand engagement and word of mouth. The theory of Macintosh makes use of different terminology. She made a distinction in participation levels described as e-enabling, e-engaging and e-empowering. This theory is not specifically made for social media, but a general matrix to evaluate online participation practices.

The Social Media Indicator-2 is introduced in table 1. In that matrix, contribution is related to how a person contributes to a social network site and participation is related to the interaction the community performs with one's personal actions. To this extent, and compared with the parameters set by Hoffman and Fodor [FH03], contribution is a matter of exposure while participation is associated with engagement.

Social Medium	Contribution		Participation	
	Actions	Network	Interaction	Word of mouth
Facebook page	● # of posts	□ # of likes	□ # of talking about ¹ ● # of comments on post ● # of likes on post	● # of shares of all posts
Facebook profile	● list of posts ²	● # of friends ● # of subscribers	● # of comments on post ²	
Twitter	■ # of tweets	■ # of followers	■ # of reply's ■ # of favorites ● list of mentions ³	■ # of retweeted
YouTube	□ # of video's	□ # of subscribers	□ # of rates ² □ rating average ² □ # of likes ²	□ # of favourite ² □ # of views ²
LinkedIn	■ # of updates ⁴	■ # of 1st degree connections ⁴ ■ # of all connections ⁴		■ # of recommenders ⁴
Google+	□ list of activity's ⁵	□ list of activity's ⁵	□ # of comments per activity	□ # of +1 per activity □ # of re-shares per activity

1 Number based on activity of last 7 days

2 Can retrieve data of 25 posts per request

3 Can retrieve data of 20 tweets per request with a maximum of 800 tweets in total

4 Can retrieve data up to a maximum of 500kb per request

5 Can retrieve undefined activity's per request

□ Data is publicly available for developers

■ Data is available for developers by default but authentication of the user can be required depending on the privacy settings of the user

● Data is only available with authentication of the user

* Every post can individually be unreachable to measure depending on the privacy settings of the post

Table 1 – The Social Media Indicator-2

2.1 Social Media Indicator calculation

Every social media channel has its own type of posts and ways of replying and sharing. Aiming to address every social media channel on the same level and return a number that is independent and comparable to all social media, we have looked into the API's [API] of the selected social media channels and then made a list of all parameters that are relevant and possible to measure. We have included them in the matrix to get an overview based upon the type of parameter and what it is telling us. The classification was made based on the existing parameters of Hoffman and Fodor [FH03] and the Macintosh participation ladder [MS05]. From this point we can try to find a method to calculate the values of each parameter to a comprehensive and meaningful value that is comparable for each social network site. To retrieve data using the API of a social media channel you have to be an acknowledged developer of the social media channel. There are no additional costs to obtain the status of a developer in the included social media channels. The availability of each parameter in the matrix is indicated by a symbol. It is important to state that even the parameters that are available for developers by default may require user authentication depending on user's privacy settings [VI16]. Because a user always has the control in his own hands [BO18] regarding what he places online and makes visible for others to view, we can justify using all retrieved data from users without ethical issues [TI17].

In the matrix we decided to make a difference between the individual's contribution to a network and the participation of one's network to one's actions. The *contribution part* contains all data of one's actions and one's network. The *participation part* contains the data of all reactions to one's actions and all shares of one's actions by the network. We named the column with data of the reactions as *interaction* and the column with data of shares as *word of mouth*.

2.2 The Scoring Algorithm

In the first version of the Social Media Indicator calculation every possible parameter receives an equal value [EHH02]. All observed numbers are accumulated in that calculation. The shortcoming of that approach is that different social network sites present different parameters. For instance, Facebook offers more functions to communicate than Twitter does. This uneven number of parameters can bias the weight in the total score. When we determine the score of two people where one of them uses only Facebook and the other uses only Twitter, assuming the activity is equally for both of them, in the first version of the calculation the Facebook user gets a higher score than the Twitter user. When scores are all just added up, more parameters automatically results in a higher score [SO19]. To calculate scores that reflect the influence more reliably, we have to include the network as a leverage of the post. So we decide we need to multiply these two values by each other. This results in multiplying the number of posts to the size of the network instead of adding them up.

We assume that, for measuring influence, the size of the network is more important than the number of posts. For instance, 10 posts to a network of 100 people shows more influence than 100 posts to a network of 10 people. The increment of influence is not linear to the number of posts. For example: the 11th post has a lot more value than the 101th post

in a certain time. Coming to these points, we decided that the logarithm function is an appropriate mathematical function to solve the issue of post increment. This problem also repeats itself when we look to the network size. The first attempt was to solve it the same way as we did with the posts. However, by using logarithms for the network too we would still be adding up the values because of the math rules.

$$\text{Log (a)} * \text{Log (b)} = \text{Log (a+b)}$$

Consequently, to create a value that returns a number that is increasing with every increment of the network a little less, we need to find another similar solution to a logarithm that solves this. To give more weight to the network size than to the number of publications it would be convenient that the increment of the network is a little larger than the increment of the publications. Therefore, we decided to look further for another function. As a result we created an exponential function with a divide of the network as paw so the increment is decreasing and the amount is controlled by the amount of the network itself.

$$\text{Log (publications)} * \text{network}^{\left(1 + \frac{1}{\text{network}}\right)}$$

In this formula, the amplification of one's network of followers is multiplied through one's actions. Considering reaching (network) has more value than action (publications), while the number of publications is parameterized by the $\text{Log}(\text{publications})$, we assume the value of the network increases in $\text{network}^{(1 + (1/\text{network}))}$. This way of scoring is expected to provide more representative scores.

3. Discussion and future research

Although the development of the Social Media Indicator is still in its infancy, the improved version 2, as presented in this paper, gives a theoretical ground for future software tools for measuring personal social media influence. However, we still need to validate the assumption that the network is a leverage of the posts and that one's actions, network and publications are related but do not share the same weight.

The next step for the improvement of the framework is formulating, an algorithm for each network that returns a number that represents the influence based on the available parameters, bringing all the social network channels to the same level.

With the presented matrix and scoring instructions we aim to help both scholars and business professionals to understand the impact of the individual use of social media.

Acknowledgements

Assistance provided by CNPq - Brazil.

References

- [ES01] Edwards, C; Spence, P; Gentile, C; Edwards, A.: How much Klout do you have - A test of system generated cues on source credibility; *Computers in Human Behaviour*, 2013.
- [EHH02] Effing, R.; Hillegersberg, J. Van; Huibers, T.W.C.; Social media and political participation: are Facebook, Twitter and Youtube democratizing our political systems? *Electronic Participation*, 2011. Pages 25-35.
- [FH03] Fodor, M.; Hoffman, D.: Can You Measure the ROI of Your Social Media Marketing?, 2010. MIT Sloan – Management Review. Vol. 52 No. 1.
- [KH04] Kaplan, A.; Haenlein, M.: Users of the world, unite! The challenges and opportunities of Social Media; *Business Horizons*, Elsevier, vol. 53(1), pages 59-68, January 2010.
- [MS05] Macintosh, A.; Smith, E.: Citizen participation in public affairs. *EGOV 2002*, pp. 256–263.
- [API-1] Facebook <https://developers.facebook.com/docs/reference/apis/>
- [API-2] Twitter <https://dev.twitter.com/docs/api/1.1>
- [API-3] Youtube <https://developers.google.com/youtube/>
- [API-4] LinkedIn <https://developer.linkedin.com/apis>
- [API-5] Google+ <https://developers.google.com/+api/>
- [CO07] <http://www.coosto.nl/>
- [KL08] <http://klout.com/>
- [SO09] <http://www.socialmediacheck.nl>
- [RA10] <http://www.salesforcemarketingcloud.com/>
- [TE11] <http://www.teezir.com/>
- [MC12] Murdough, Chris. Social Media Measurement: It is not impossible. *Journal of Interactive Advertising*, vol. 10, n 1, Fall 2009.
- [SM13] Sponder, Marshall. *Social Media Analytics: Effective Tools for Building, Interpreting and Using Metrics*. McGraw Hill Companies, USA, 2012.
- [RE15] Results, E. S. (2012). How social technologies drive business success, (May).
- [VI16] Vitak, J. (2012). The Impact of Context Collapse and Privacy on Social Network Site Disclosures. *Journal Of Broadcasting & Electronic Media*, 56(4), 451-470.
- [TI17] Think before you post: social media etiquette. (2013). *Surgical Technologist*, 45(4), 172-174.
- [BO18] Boyd, D. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210–230.
- [SO19] Social, H., & Works, M. (n.d.). *The Power of Like 2 How Social Marketing Works*.