Virtual Media Quality Index (W-Index) for Higher Institutions of Education

Kaufui Vincent Wong* and Samuel R. Ginkel

University of Miami, Coral Gables, FL 33146, USA

Abstract: The simplicity and ease of access of its website, has allowed YouTube to be embraced by the whole world and establish itself as the most prominent form of video sharing on earth. YouTube has become a widely used medium for individuals, corporations, and academic institutions alike. In the current work a W-index has been developed and applied to many academic institutions to evaluate the efficacy of the respective YouTube channel(s). This index was the creative product of the senior author who was inspired by the H-index developed by J.E. Hirsch (2005) to evaluate the productivity and impact of a researcher. The W-index will be used to evaluate the quantity and quality of a University's channel that is different from the traditional staples that define popularity and video effectiveness on YouTube, such as total views. The methodology used would be to see that if this W-index would correlate well to existing indices used for the evaluation of universities. Eventually, the W-index could serve as an indicator of whether or not a university needs to invest time or money into the development of better videos or more videos for their YouTube channel to maximize its impact on the academic community. A correlation between W-Index and three well-established and well-defined cybermetric rankings; US Rank, Impact Rank, and Excellence Rank, is established and used as the basis for the W-index's usefulness. The establishment of such a correlation indicates that the W-Index can also be used to evaluate the communication efficacy of individuals, despite the lack of any robust ranking system for individuals. The W-Index serves as a good indicator, based on consistent correlation coefficients among ranking systems analyzed, of a university or individual's success and communication efficacy. Rankings are useful as single numbers that contribute to decision making, simply because of their simplicity.

Keywords: Fame, popularity, universities, colleges, rank, social media.

1. INTRODUCTION

The Internet has revolutionized communication and has helped establish global interconnectedness on a scale that is unprecedented for humanity. Universities have grown with the internet in an effort to reach out to the world's youth and grab the attention of a broad spectrum of young individuals with different assets to offer. The information that universities link and display on various online media is constantly growing and evolving, and this information can be measured by cybermetric techniques in order to design university web rankings as a means of reference and further sub-analysis [1].

YouTube in particular is one of the most influential websites in operation, boasting more than a billion views per day, and approximately 6 billion hours of videos watched per month [2]. Universities and research organizations have been taking notice to YouTube's popularity in recent years and have been investing time and money into developing their own 'channels'. From these channels, academic institutions can connect to their student body and the general public with great ease. This is sought out by Universities because there has been an increase in demand for Universities to operate in the 'public sphere' over the last few decades [3]. This connectivity allows the development of videos that convey a wide variety of information, from broadcasting announcements about university life and operations, to outlining the importance and scientific details of specific research projects. The W-Index is a scale designed to measure the communication efficacy and activity of either an academic institution or individual's YouTube channel(s). There is a need for such a scale because normalized measures such as; total views, number of videos, and average views per video, do not give a good measure of an individual's or an institution's success (or lack thereof). The W-Index was inspired by the H-index created by Dr. Hirsch [4], used to measure both the research output and impact of an individual researcher. Instead of focusing primarily on researchers, the W-Index focuses on the output and impact of an entire University's YouTube channel, as well as an individual's YouTube channel. One popular video can distort a simple measure of an institution's effectiveness, as can the presence of many videos with few views. As a result, a more robust system must be used for interpreting the effectiveness of an institutions academic or informative output. The W-index is useful for determining the overall quality and impact of a University's YouTube channel.

Although the W-index can also be used to evaluate an individual's communication efficacy, the implications

*Address correspondence to this author at the University of Miami, Coral Gables, FL 33146, USA; E-mail: kwong@miami.edu

E-ISSN: 2409-9848/15

of such analysis are much better utilized by universities because of how actively they commit resources to intangibles. This information could be used to determine whether or not a university's communication is effective or ineffective, and persuade said university to either increase or reduce their budget for such endeavors. If a university has a low W-rating, or W-score then it may be beneficial for that institution to invest more time and money into the development of a proficient, more effective channel.

Previous methods of analysis such as the g, f, z, and t index variants do an excellent job at evaluating the communication efficacy of individual's research. The g-index is a variation of the H-index that addresses the problem of a paper receiving further citations, by analyzing the highest number of papers that received g^2 citations. The f and t variants used a harmonic or geometric mean to determine the resultant f or t index values [5, 6]. The z variant considers an addition C factor in an attempt to combat the h-factors oversensitivity to a few highly cited papers, which can cause up to 75% of the authors to be neglected [7]. In addition to these variants, the tapered H-index also exists which takes into account the total number of citations received for a given researcher [8]. The W-index can also be used to evaluate an individual's YouTube channel communication efficacy. Analysis of an individual's YouTube channel's production and impact can also be performed using the W-index, but such an analysis is more difficult to legitimize as there is no existing 'ranking' system for how effectively and actively individuals communicate and display information to compare to the W-index to establish a strong correlation between W-index and some other measure of successful communication.

2. METHODS/THEORY

The W-index is the intellectual property of the senior author of the current work. W-index is computed by simultaneously considering the number of YouTube channel uploads, and the views for each video uploaded. If a channel has N videos, each video will be checked to see if the total views for that video exceeds n_i^2; if it does, n_i+1 of the videos (arranged in descending order of views, a good feature of YouTube) will be checked to see if it continues to satisfy this criteria; if it does not, the W-Index value will be n_i. In other words, the W-Index of a communicator is equal to n_i if there are at least n_i videos with n_i^2 views or more.

$$W - Index = n_i \text{ if no. of views is at least } n_i^2 \text{ in each of } n_i \text{ videos} \quad (1)$$

As an example, a communicator with 38 videos has 35 videos each with 1225 views or more, their W-index is given as 35. It does not matter if another communicator has 100 videos, but if only 35 of the videos have 1225 views or more, this second communicator also has a W-index of 35. It is evident that they must both have at least 35 videos. By taking into account the square of the views normalized across many uploads, the skew of a single popular video is eliminated, as is the skewing created by having many videos with very few views per video. With the skew from these two effects reduced this model can serve as an appropriate approximation of the effectiveness of an Academic Institution's YouTube channels. The relative W-scores of one field are only comparable to that of a similar field of work, because popularity differs between fields of study. In addition, the scores of individuals or corporations are not necessarily applicable or comparable to the scores of academic institutions.

The method of data collection involved the inspection of 75 of the top ranked University's YouTube channels, after sorting uploaded videos by popularity in descending order. A W score for each channel was determined by inspecting the total number of videos with views greater than n_i^2, and when the value of n_i^2 was less than expected the preceding numerical value was taken to be the given University's W-Index value.

The W-Index gives a larger number for the more popular university channel. Many channels have their athletic department's videos mixed in with their main channel, and as a result, athletic videos have been excluded from the calculation of the W-Index.

3. DATA ANALYSIS

To analyze the effectiveness of the W-Index, two types of rankings are used, a national or "US" rank, and an Impact Rank both based off the rankings outlined in [9-11].

By visual inspection, it is somewhat obvious there is a correlation between W-Index and the specified ranking systems. The correlation, best described by the R^2 value, is far from perfect but shows a somewhat strong relationship between W-Index and both the US Rank and the Impact Rank. The good correlation shows that a University's communication efficacy can be reasonably represented by the W-Index alone. In

general, the 'higher' a university's W-Index, the greater that respective university's ranking (where ranking is categorized in descending order).

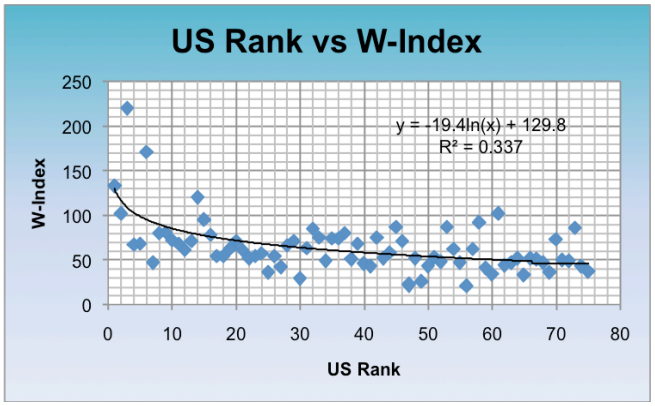

Figure 1: US Rank vs W-Index.

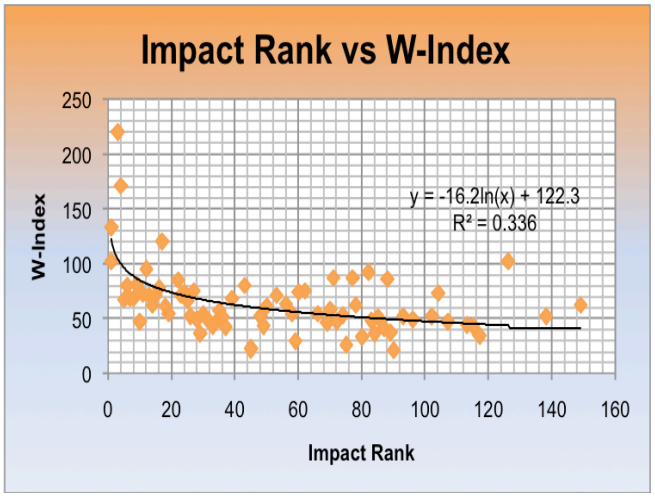

Figure 2: Impact Rank vs W-Index.

A third ranking is analyzed and compared to the preceding two to establish a more robust correlation between W-Index and a University's overall communication efficacy. This is the Excellence Rank [11].

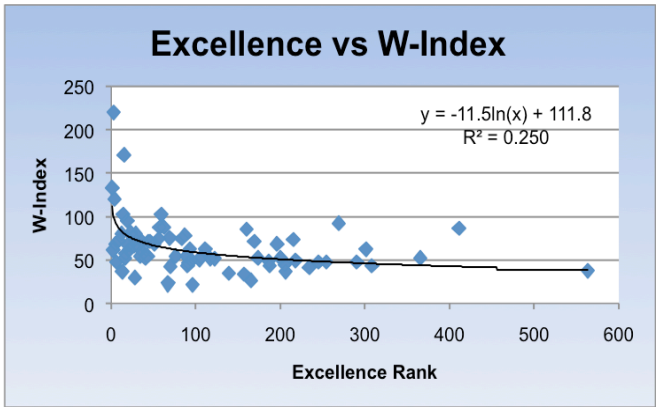

Figure 3: Excellence Rank vs W-Index.

4. RESULTS

The W-Index has relevance in most cases, however there are some factors that may taint the relevance of the W-index and reduce its validity for certain organizations. Some of these factors include the presence of multiple channels per university, channels that may have been abandoned and been sitting un-maintained and channels that have been intentionally increasing their own view count through some form of view manipulation. However, in the case of the seventy-five universities examined, no channels were found that might taint the relevance of the analysis performed, although some universities would separate the athletic related videos from the academic uploads allowing for an easier analysis to be performed. The presence of athletic or non-academic videos is taken into consideration when W-Index is determined, and these videos are excluded from the final calculation. One must not assume that the W-index is a perfect measure of a superior university, because a university's academic relevance cannot always be properly quantified with an analysis about one single form of media, in this case the efficacy of their YouTube channel(s).

Table 1: Curve Fitting and Correlation Comparison between Ranking Systems

Ranking	Curve Fitting Eq.	R^2
US	y = -19.45ln(x) + 129.88	0.3373
Impact	y = -16.24ln(x) + 122.36	0.3363
Excellence	y = -11.58ln(x) + 111.89	0.2506

It can be seen that each of the curve fitting equations in the form of simple logarithmic expressions, have very similar y-intercepts and x-coefficients, suggesting a high degree of similarity amongst each of the comparisons. This similarity is a strong indication that the W-Index can properly represent the efficacy of a University simply through an analysis of its respective YouTube channels.

If there existed similar measures of an individual's communication efficacy, a similar analysis could be performed on a variety of individual educator's channels and solidify the accuracy and value of the W-index as a measure of communication efficacy and prowess. Because universally applicable index for individuals does not seem to exist, the correlation previously established between university success and

the W-index could prove the relevance of the W-index and serve as a framework for its use on future work about individuals.

The rankings chosen were taken from a reputable site/database, Webometrics, and are established *via* sophisticated analysis and cross referencing of many different criteria and generally agreed upon measures of university success and prowess. Cybermetric analyses serve as good indicators for describing university activity and success, and are useful for reflecting its institutional success [10]. The legitimacy and relevance of webometrics is outlined by several statements and facts from their website, "Webometrics is continuously developing ways to improve each ranking, changing or evolving the indicators and weighting models to provide a better classification. Few rankings are able to be stable between editions if correcting errors or tuning indicators are not used. Each ranking system is evaluated every six months by the Cybermetrics Lab (Spanish National Research Council, CSIC), and provides reliable, multidimensional, updated and useful information about the performance of universities from all over the world based on their web presence and impact," [11].

4.1. US Ranking

US ranking is determined by broadly analyzing the available open source data pertaining to University rankings in the United States, and then compiling and normalizing the results to achieve a final national ranking [11, 12]. A low numerical value corresponds to a high ranking, with 1 being the top ranking.

4.2. Impact Ranking

The method of determining 'impact ranking' involves an evaluation of all external in links that University websites receive from outside sources. These links are an indicator of how successful a University is at putting out valuable information that influences other research oriented groups or individuals. The links also are an indicator of an institution's academic success and prowess based on widely reviewed criteria. Two main backlink search providers, Majestic SEO, and Ahrefs provide the necessary data that is collected. Each use web search algorithms, popularly called crawlers, that are used to establish a wholesome data return free of errors or data gaps. Finally, the indicator is determined using the following formula [11]:

$$Impact\ ranking = \sqrt{(\#\ of\ backlinks)} * (\#\ of\ domains\ originating\ these\ backlinks) \quad (2)$$

Then the maximum from the normalized results is used as the respective University's impact ranking [11]. The higher ranked institution has a lower integer value.

4.3. Excellence Ranking

By analyzing academic papers that have been published in high impact international journals, an excellence ranking is subsequently determined. The proportion of a university's papers that are represented in the top tenth percentile of the given university's relevant scientific field are used for calculation, as simply using the total number of papers is not an appropriate measure. More than 5200 universities were analyzed based on their productive output from 2003-2010, Universities with zero papers were not included in the analysis [11]. The lower the integer value indicates a higher ranking.

5. DISCUSSION AND CONCLUSION

Performance indices such as the one proposed here, are popular because it encapsulates a lot of complex information into a single number. The W-index correlates with the US ranking and does simplify to a single number the complicated matrix used to arrive at the US rankings. The public make use of ranking because the relevant information (which matters to them in their decision making) is put into a nutshell and is simple to process.

One potential issue with the W-index is dynamicity, as new uploads are constantly being added and more views are accumulated with time. Further analysis can be performed that could potentially allow the prediction of future W-Index values, as is done in [13] to predict H-index values based on theoretical and experimental 'production rates'. Similar channel output rates could be calculated and allow an institution to determine if it needs to submit more uploads, or tweak the content of the uploads in an attempt to improve their channel's ranking and consequently impact on the online community. However, further analysis of time-dependent W-index can be performed after a significant period of time, like 6 months or a year, and it can be determined if W-index value variation with time is a significant noise factor in the analysis of communicator efficacy [14, 15]. Additionally, this variation with time is a characteristic of all quality indices, including the H-index.

From the three different ranking systems analyzed; US, Impact, and Excellence, it is evident that there is a definite correlation between W-Index and a University's ranking. It was shown that the higher the ranking (small rank number), the higher the W-Index. It can be reasoned that the W-Index can be used, at least a partial indicator of a University's success, or communication efficacy. The robustly defined and thoroughly refined ranking indicators used for correlation analysis serve as a basis for the establishment of the W-index as an effective measure of the general efficacy and impact of a given university. The fact that each of these correlation coefficients is rather close in magnitude further strengthens the argument that communication efficacy of virtual media, for both Universities and individuals can be accurately represented by the W-Index.

REFERENCES

[1] Orduna-Malea E and Ontalba-Ruiperez J. 25 October 2012. Proposal for a multilevel university cybermetric analysis model. An International Journal for all Quantitative Aspects of the Science of Science. Communication in Science and Science Policy 10.1007/s11192-012-0868-5.

[2] YouTube. YouTube n.d. Web. 12 Jan. 2015. <http://www.youtube.com/yt/press/statistics.html>.

[3] Vasileiadou E, van den Besselaar P 2006. Linking shallow, linking deep. How scientific intermediaries use the Web for their network of collaborators. ISSN 10(1): 1137-5019.

[4] Hirsch JE. 15 November 2005. An index to quantify an individual's scientific research output. PNAS 102 (46): 16569-16572. http://dx.doi.org/10.1073/pnas.0507655102

[5] Quesada A. April 2009. Monotonicity and the Hirsch index. Journal of Infometrics 3(2): 158-160. http://dx.doi.org/10.1016/j.joi.2009.01.002

[6] Schreiber M. October 2010. A new family of old Hirsch index variants. Journal of Infometrics 4(4): 647-651. http://dx.doi.org/10.1016/j.joi.2010.05.002

[7] Peterson AM. October 2013. The Z-index: A geometric representation of productivity and impact which accounts for information in the entire rank-citation profile. Journal of Infometrics 7(4): 823-832. http://dx.doi.org/10.1016/j.joi.2013.07.003

[8] Sangwal K. July 2012. On the relationship between citations of publication output and Hirsch index h of authors: conceptualization of tapered Hirsch index h_T, circular citation area radius R and citation acceleration. A Journal of Scientometrics 93(3): 987-1004.

[9] United States of America. Ranking Web of Universities. Webometrics. Retrieved January 12, 2015. <http://webometrics.info/en/North_america/United%20States%20of%20America>.

[10] Aguillo IF, Granadino B, Ortega JL, Priesto JA. August 2006. Scientific research activity and communication measured with cybermetrics indicators. Journal of Information Science 57(10): 1296-1296.

[11] Methodology. Ranking Web of Universities. Webometrics n.d. Web. Retrieved 21 Nov. 2014. <http://webometrics.info/en/Methodology>.

[12] Frequently Asked Questions. Webometrics N.p, Dec 2012. Web. Retrieved 12 Jan 2015. <http://webometrics.info/en/node/36>.

[13] Burell QL. January 2007. Hirsch's h-index: A stochastic model. Journal of Infometrics 1(1): 16-25. http://dx.doi.org/10.1016/j.joi.2006.07.001

[14] Kongo T. 1 January 2014.An alternative axiomatization of the Hirsch index. Journal of Infometrics 8(1): 252-258. http://dx.doi.org/10.1016/j.joi.2013.12.005

[15] Schreiber M. April 2013. How relevant is the predictive power of the h-index? A case study of the time-dependent Hirsch index. Journal of Infometrics 7 (2): 325-329. http://dx.doi.org/10.1016/j.joi.2013.01.001

Received on 23-02-2014 Accepted on 09-03-2015 Published on 30-06-2015

DOI: http://dx.doi.org/10.15377/2409-9848.2015.02.01.2