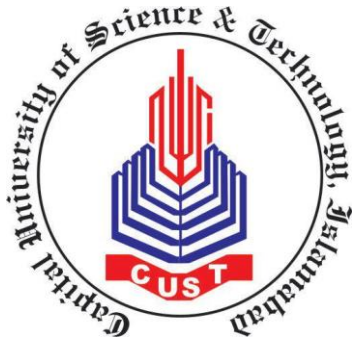
Evaluation of Bibliometrics and Altmetrics on large Scale Dataset



By

Bilal Ahmed
MS-131022

A thesis submitted to the
Department of Computer Science in partial fulfillment of the
requirements for the degree of
MASTER OF SCIENCE IN COMPUTER SCIENCE

Faculty of Computing
Capital University of Science and Technology
Islamabad
January, 2017

I dedicate my dissertation work to my family, teachers and friends. Special feeling of gratitude to my loving parents for their love, endless support and encouragement.

CAPITAL UNIVERSITY OF SCIENCE & TECHNOLOGY
ISLAMABAD

CERTIFICATE OF APPROVAL

Evaluation of Bibliometrics and Altmetrics on Large Scale Dataset

by

Bilal Ahmed

Reg No. Ms 131022

THESIS EXAMINING COMMITTEE

S No	Examiner	Name	Organization
(a)	External Examiner	Dr. Waqar Mehmood	COMSAT, Wah Cantt
(b)	Internal Examiner	Dr. Nayyer Masood	CUST, Islamabad
(c)	Supervisor	Dr. Muhammad Aleem	CUST, Islamabad
(d)	Advisor	Dr.Muhammad Tanvir Afzal	CUST, Islamabad

_____

Dr. Muhammad Aleem

Thesis Supervisors

April, 2017

_____

Dr. Nayyer Masood

Head

Department of Computer Science

Dated : April, 2017

_____

Dr. Muhammad Abdul Qadir

Dean

Faculty of Computing

Dated : April, 2017

Certificate

This is to certify that Mr. Bilal Ahmed has incorporated all observations, suggestions, and comments made by the external as well as the internal examiners and thesis supervisors. The title of his thesis is: **Evaluation of Bibliometrics and Altmetrics on the large scale dataset**.

Forwarded for necessary action

_____

Dr. Muhammad Aleem

(Thesis Supervisor)

ACKNOWLEDGMENT

First of all, I would like to thank my supervisor Professor **Dr. Muhammad Aleem**, and **Dr. Muhammad Tanvir Afzal** for their excellent supervision and great support during my study. Not only teaching me how to conduct research, they have also given me an insight into my future career. I have enjoyed a lot doing research because of my supervisor and a highly motivating research topic.

I would also like to thank my class fellow's students to help me and keep me motivating during this study. Although I have been through some difficult time during my research but they always insist not to give up and help me completing this research Thesis. I would like to thanks **Mrs.Imama Syed** also for making his dataset available for me to conduct my research.

Finally, I would like to thank my family for their love and moral support, and especially my brothers and sister for inspiration and always being there when I needed.

Bilal Ahmed

Ms 131022

DECLARATION

It is declared that this is an original piece of my own work, except where otherwise acknowledged in text and references. This work has not been submitted in any form for another degree or diploma at any university or other institution for tertiary education and shall not be submitted by me in future for obtaining any degree from this or any other University or Institution.

Bilal Ahmed

Reg. No, MS131022

January, 2017.

ABSTRACT

"Altmetrics", short form of alternative metrics, is a new technique of measuring the impact of an article on both scholarly and non-scholarly community based on some online activities. It acts as a complementary to the traditional bibliographic impact measurement techniques. This study explores the extent to which Altmetrics co-relate with the bibliographic impact measurement techniques: H-index and its variants. Traditional impact measurement approaches, i.e., Biblometrics seem unfit to compute impact of a scholarly document on general public, although they take into consideration the impact of an article on the scholarly community.

Only 0.2% of research article got indexed in CROSSREF and only such percentage got citation. It has to be investigated that whether those research work that got high citation in research community are equally shared and discussed on other web channels? The focus of this thesis is to evaluate Altmetrics with respect to traditional impact measurement techniques including H-Index, G-Index, M-Quotient, HC and HW and develop an argument that can altmetrics replace Biblometrics?

There are many techniques available that are used to reflect the popularity and quality work of an author in scholarly community. But due to the negligence of the impact created by an article on social media, these techniques, i.e., h-index and its variants are unable to fits in for calculating complete impact of an article both on scholarly and non-scholarly community. As Altmetrics claims to capture the impact of both scholarly and non-scholarly community, we would like to compare the efficiency of the correlation among the Altmetrics and Biblometrics.

Analysis is accomplished based on around 70,000 published papers of 45 categories from the field of mathematics. Number of tweets containing title, URL, or DOI of research papers has been compared with the H-index, G-index, HC-index, M-quotient and HW-index of the scholar.

Analysis has been done for around 57,155 authors and shows that with the very low twitter citation rate that is 1.47% almost no correlation exists between Twitter citation data and conventional impact measurement techniques. It implies that the social media metrics (Altmetrics) does not reflect the same kind of metrics as Bibliometrics indices. It further implies that Altmetrics should also be considered along with Biblometrics indices to access an author's work effectiveness.

LIST OF ACRONYMS

CUST Capital University of Science and Technology

FYP Final Year Project

MS Master of Science

MSC: Mathematics Subject Classification

Hw: Weighted h-index

Hc: Contemporary h-index

AMS: American Mathematical Society

IMU: International Mathematical Union

LMS: London Mathematical Society

POL: The Public Library of Science

Contents

LIST OF FIGURES

LIST OF TABLES

Chapter1. Introduction

This chapter gives introduction and background knowledge of Altmetrics, used to measure the quality of research, followed by the problem statement. Research question is defined and explained after problem statement, objectives of this research thesis in presented after research question. Finally the structure of this thesis reports is presented in the last section. Abbreviations and important definitions are also presented in the end of this chapter.

1.1 Background

Quality of research work has been measured by the impact factor of Journal or the number of citation it got from research community. Higher the research work is cited higher is the impact of a scholarly work. Current citation index is limited to only journals. The number of citation is dependent on the published work. Only the research community whether in education or corporations, is involved in citation and determining the quality of work. The main limitation of citation is it may take several years to cite the work because it must be used for future research.

Sometime research work that hasn't got any citation from the research community, might used in number of presentations, blogs and discussion elsewhere. Is there a need to define alternative metrics that can determine the quality of work or impact of research work? This direction gave birth to new idea known as Altmetrics presented in the following discussion.

1.2 Altmetrics

This is the era of social media where everything is shared and communicated through social media like Twitter, Facebook and Instagram etc. Traditional metrics only present one side of the picture. There is a need to take into account other conversions on the web but there is no system to give feedback to authors when their work is shared, discussed on other web channels.

Altmetrics can provide another mean to let the author know where his/her work is used on the web [3]. Since 2010 Altmetrics appears in post and website like Figshare[1], Impactstory[2], Altmetric[3] and Mendelely[4] take into account alterative metrics. The figure and table below present the current flaws in the citation system. We can see that on 0.2% of research article got indexed in CROSSREF and only such percentage got citation.

Article Views	PDF Downloads	CROSSREF
199913592	49450006	481771

Table 1: Comparison of Usage, Downloads and Citation of articles published on POL

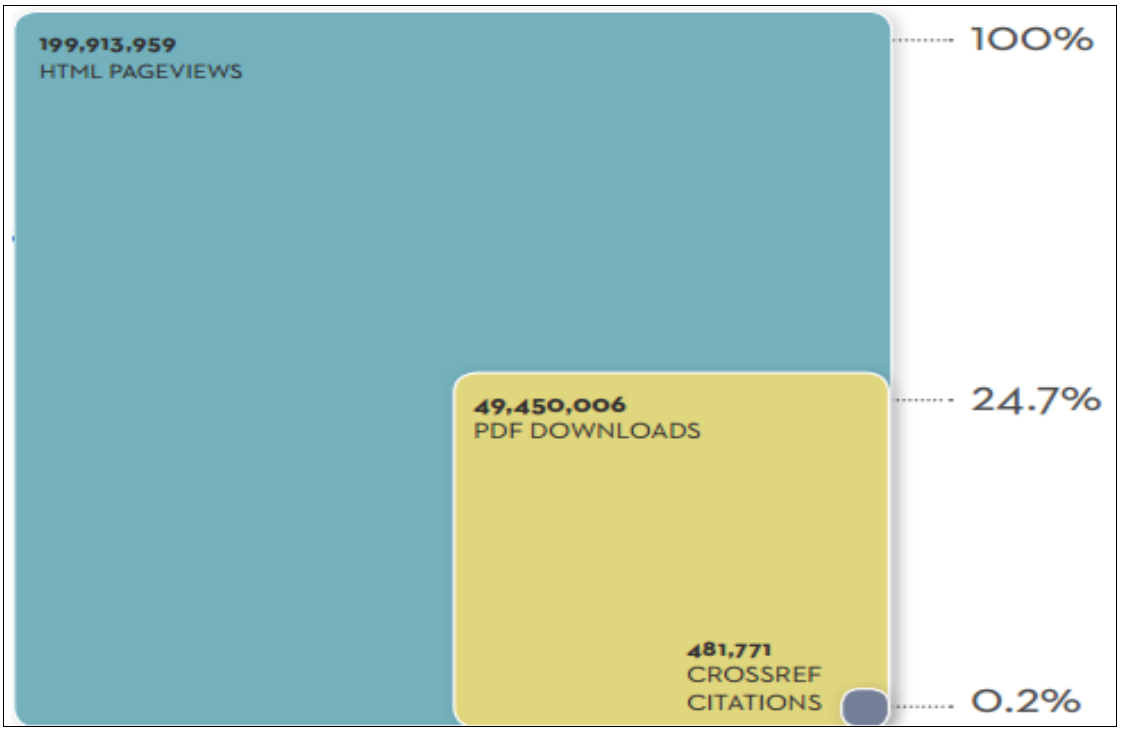


Figure 1: The Public Library of Science (PLOS) articles published until May, 2013[14]

[1] www.Figshare.com
[2] www.Impactstory.com
[3] www.Altmetrics.com
[4] www.Mendelely.com

Qualitative data and metrics other than traditional citation based metrics, which are used to measure the impact of scholarly work, are known as Altmetrics. Or it is also known as "Alternative Metrics". There are lot of metrics used like download the research articles (number of downloads) , citation in Wikipedia articles, discussed in blogs, discussed in news, used in reference mangers and mentions on blog and social networks like Facebook, Twitter or Instagram. Altmetrics keep tracking how the scholarly work is viewed, read, download, discussed and reused by the research community and public community and thus it determine the societal impact of research work. It also provide an opportunity to the scholar to track down who is following their work and what are the reviews about his/her research work which is difficult to determine in citation.

Altmetrics is an alternative way of determining the impact of research articles, improving the traditional citation based impact measures and give new insights to impact analysis[1]. There is no specific definition of Altmetrics but it is related to subfield of Infometrics and Webometrics[2]. Altmetrics is the study of non-traditional scholarly impact measurement techniques that are based on activity in web-based environments[3]. Altmetrics is an emerging field, which unlike the traditional citation impact metrics, such as impact factor and h-index, does not rely just on citation counts, but also takes into consideration other features of impact such as the number of knowledge bases that referred to the work, the number of times the work was viewed/downloaded, and/or mentions in social/news media [4] [5].

Due to its potential of measuring impact of scholarly work on both scholarly as well as non-scholarly community, measuring social media impact is gaining attention from researchers, reflecting the significance of Altmetrics. Indicators, e.g. research council arguing to use Altmetrics for evaluation of authors [6] and scholars wishing to include Altmetrics into

curriculum vitae[7], express the potential of these metrics. Activities on the social media platform like CiteULike, Facebook and Twitter can be monitored to predict the impact of scholarly article. Studies show that social media platforms like Twitter are used for dissemination [6] science popularization[8], and scholarly product promotion[9]. However, the authenticity and reliability of these metrics is controversial [10].

As compared to other social media platforms, Twitter is much more extensive [11] with 288 million active users; Twitter is one of the most popular platforms for dissemination of scholarly articles and a commanding tool to disseminate pointers (e.g. links) to information[12] with hash tag, @messages and re-tweet (Boyd et al., 2010). Terms like "Tweetations" [13] and "citation tweets" [3] are used for tweet count. Studies have also explored the involvement of research community for dissemination of links or documents via Twitter[11] although a detailed study has not yet been performed to support these claims. This research provides a comprehensive analysis to estimate the potential of Altmetrics in gauging scientific impact by analyzing mathematical publications. Bibliographic impact measuring techniques ignore the impact created by an article in non-scholar community. To measure the impact of a scholarly research in general community: scholar as well as non-scholar community, Altmetrics came into existence. Social media platforms are now the best way to share anything at any time. When it comes to scholarly research, it becomes critical to measure the impact of scholarly document and the way it is helping the non-scholarly community. Micro blogging platform Twitter is one of the best known mean to predict the impact of scholarly article in the near future [12].

Generally, Articles are generally cited more frequently on social networks than on published papers[3]. Being cited shows the quality of the work produced by the specific author; there are many methods available to perform citation analysis on the basis of citations of the paper. H-

index, G-Index, M-Quintet, HC-index and HW-index are common ranking approaches for authors which reflect the productivity of an author with citation count. Articles are also cited online on social networks, this reveals the quality of that article due to which it has been cited. So authors should also be given appreciation on their work getting mentioned on social platforms. Keeping in view the importance of social network citation, we want to perform an analysis of both traditional bibliographic citation indexes (H-index, G-index, M-quintet, HC-index and HW-index) and Twitter citation (Altmetrics). In this work, our focal point is to identify authors that have higher indices value and their articles are trended on the Twitter as well. We try to figure out if the author with higher index value gets higher citation on Twitter? That would be estimated with the help of co-relation between bibliographic impact measurement techniques and Altmetrics. We also try to answer: how closely related they both are and, the other way around i.e. how different they are?

1.3 Research Gap

Many researchers evaluated the potential of Altmetrics on different domains as summarized in (Table a). They have used citation count as a Bibliometrics impact factor parameter only and compared it with the Altmetrics data. Although by analyzing co-relation between citation count and Altmetrics data they found very low co-relation between Bibliometrics and Altmetrics but no one yet explore other Bibliometrics reflective measures i.e. H-index and its variants on a large scale data.

Although many researchers attempt to find out the correlation between traditional impact measurement techniques with altmetrics but the only difference between previous work and ours is that not any of these researchers tries to conduct a comprehensive study on large scale dataset

with these parameters which we are using. Furthermore, so far no such study is conducted for the Mathematics publications as well.

Author name	Title	Data Set	Biblometrics parameters	Altmetrics parameters	Data collection platform(Bibliometrics)	Data collection platform(Altmetrics)	Co-relation
José Luis Ortega(2015)	Relationship between Altmetric and Bibliometrics indicators across academic social sites: The case of CSIC's members	10,000 author	Citation	Tweets and Facebook posts	Microsoft Academic Search and Google Scholar Citations	ResearchGate, Academia.edu and Mendeley	Less co-relation at authors level
Feng et. al. (2016)	Bibliographic Analysis of Nature Based on Twitter and Facebook Altmetrics Data	4276 *Nature* articles between January 2010 and June 2015	Citation	Twitter	online literature database	nature.altmetric.com Web of Science	0.52
Rodrigo et. al.(2014)	Do 'altmetrics' correlate with citations? Extensive comparison of Altmetric indicators with citations from a multidisciplinary perspective	1,380,143 distinct publications	citations	*Facebook walls* *Blogs* *Twitter* *Google+* *News outlets*	Not mention	Altmetric.com	Very less but platform dependent
Lutz Bornmann(2015)	Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics	Not mention	citation counts	Tweets	Not mention	Twitter Mendeley and CiteULike	CiteULike pooled r = 0.23 Mendeley pooled r = 0.51 Tweeter pooled r = 0.003
Stefanie Haustei et.al (2013)	Tweeting Biomedicine: An Analysis of Citations in the Biomedical Literature	1.4 million documents between 2010 and 2012	citations	tweets	Not mention	PubMed Web of Science (WoS)	Very low

Table 2: Co-relation between Altmetrics and Bibliometrics (Summary)

To cover this dimension we are going to perform a comprehensive study by using large scale data of Mathematic domain to explore the potential of Altmetrics

1.4 Problem Statement

This thesis will work on the following problems:

As it has been identified from critical analysis and research gap, evident from the table in section 1.3, no one evaluated Altmetrics for a particular domain on large scale dataset. In this work we are going to perform comprehensive analysis of Altmetrics for mathematics domain

Secondly, all of the previous work, as evident from the table 2, used citation as Bibliometrics reflective measure. We are going to use other Bibliometrics reflective measures such as H-index and its variants (Index, G-Index, M-Quotient, Hc-Index and Hw-Index) to find out the correlation of Bibliometrics with Altmetrics.

1.5 Objectives

The Objective of this thesis is not only to evaluates the potential of the Altmetrics to capture the impact of an article form social media platforms but also develop an argument about the Altmetrics that it can either replace the Bibliometrics or it can act as an complementary to the traditional impact factor measurement techniques.

In this research thesis we will try to analyze the capabilities of the altmetrics and try to develop an argument about the altmetrics as discussed in earlier section through a comprehensive study. By finding co-relation we will then able to get the appreciation of non-scholarly community along with scholarly community by declaring the altmetrics the alternative of biblometrics, if possible.

This thesis will provide a roadmap to get the appreciation of non-scholarly community along with the scholarly community. If the analysis shows some promise then we will be able to say

that the altmetrics can be used to calculate the impact of an author along with the biblometrics and the efficiency of an author then can be calculated by looking his altmetrics score.

1.6 Research Question

There are many techniques available that are used to reflect the popularity and quality work of an author in scholarly community. But due to the negligence of the impact created by an article on social media, these techniques i.e. h-index and its variants are unable to fits in for calculating complete impact of an article both on scholarly and non-scholarly community. As Altmetrics claims to capture the impact of both scholarly and non-scholarly community, we would like to compare the efficiency of the correlation among the Altmetrics and Bibliometrics. To do that Our Primary Research question that should be answered is as follow:

RQ1. Is there any correlation exists between Altmetrics and Bibliometrics impact factor measurement techniques i.e. (H-Index, G-Index, M-Quotient, HC and HW)?

RQ2. Author with higher Bibliometrics impact does get higher tweet count?

1.7 Dissertation Organization

The following sections will explain the structure and content of each chapter of this dissertation document.

Chapter 2: This chapter emphases on the techniques used for providing alternative metrics to the traditional metrics of citation.

Chapter 3: defines the proposed methodology adopted in this research. It includes data collection techniques, limitation of the dataset and other technique in order to conduct this research.

Chapter 4: present all the tables of experiment results and related discussions.

Chapter 5: this chapter concludes this dissertation and presents potential future research areas based on this research outcome. Further possibilities and approaches to further investigate this problem will also be discussed and recommendations will be presented at the end of the chapter.

1.8 Definitions and Abbreviations

MSC: Mathematics Subject Classification

hw: Weighted h-index

hc: Contemporary h-index

AMS: American Mathematical Society

IMU: International Mathematical Union

LMS: London Mathematical Society

1.9 Conclusion

This chapter introduced this dissertation topic, the research problem, and the key objective and research questions to be answered by this research. The structure of this dissertation was also presented and the content of each chapter were explained briefly.

Chapter 2 Literature Review

2.1 Introduction

In this chapter, light is shed on the research work based on existing research and the extent relevant literature for the altmetrics and its uses.

Altmetrics is not a single class of evaluation metrics which tell the importance of any scholarly article on web rather that it cover wide range of activities which can be grouped into following categories [14].

Viewed: *Accessing the research article online*

Saved: *Saving the online article*

Discussion: *Discussed the article on social networking website or talking about the article in news*

Recommendation: *Recommendation of article to other user on Answer question website or other recommender website*

Citation: *Formal citation of article in another article.*

According to Altmetric.com [15] in combination of different metrics provide a mean to know how many people are engaged to scholarly work which is the main drawback of traditional citation based impact. This helps us to understand why and where piece of research is used and they are quicker to accumulate than traditional citation metrics. Some of example of alternative metrics [16]are presented in the table below

Type	ALTMETRICS
Usage	Views, Clicks, Download, Play in video, View the presentation, Downloading the dataset or articles, software, apps and others.
Captures	Reading, Bookmarks in browser, Adding to favorites list, and code forks
Mentions	Wikipedia citation, discussed in blogs, comments and reviews.
Social media	Like, shares and comments on social media like twitter and Facebook etc.
Citation	ISI web of science, PubMed central etc.

Table 3: List of some Altmetrics

The impact of scholarly work is the key part of scholarly communication lifecycle, where understanding the impacts helps to understand the quality of research work[17]. Traditionally author's research quality has been accessed by the number of times it has been mentioned by other authors in their own research articles called "citation" which was considered as one of the most reliable mean of reflecting the impact of any article and also indicating the popularity of any author. Now with advancement of the web, there is a huge number of authors creating their contents and mentioning others on the digital web. Numbers in the citation database i.e. "WebOfScience" and "Scopus" show how well an author perceived in scholarly community but in the meantime, somehow, it ignore the impact of that scholarly product on non-authors [2]. To cater this issue, terms like "Webometrics"[18], "Scholarometer"[19] and "Altmetrics"[20] come into play. These terms are used to reflect the impact of an article not only on scholarly but also on the non-scholarly community as well. Jason Priem first used term "Altmetrics" with hash tag in his tweet on 29 September 2010, which is a short form of Alternative Metrics [21].

Authors	Paper title	Summary
Thelwall et al., 2013	Do altmetrics work? Twitter and ten other social web services. PloS one	detailed study has not yet been performed to support these claims of altmetrics
Liu & Adie, 2013	Realizing the potential of altmetrics within institutions.	Altmetrics claim to capture impact from a broader public but still it cannot replace the bibliographic scholarly impact
Zahedi, Costas & Wouters, 2014	How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. Scientometrics	it acts as a compliment to the traditional citation system
Thelwall et al., 2013	Do altmetrics work? Twitter and ten other social web services. PloS one	detailed study has not yet been performed to support these claims of altmetrics
Liu & Adie, 2013	Realizing the potential of altmetrics within institutions.	Altmetrics claim to capture impact from a broader public but still it cannot replace the bibliographic scholarly impact
Zahedi, Costas & Wouters, 2014	How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. Scientometrics	it acts as a compliment to the traditional citation system
Stefanie Haustein , Isabella Peters, Cassidy R. Sugimoto, Mike Thelwall, Vincent Larivière (2013)	Tweeting Biomedicine: An Analysis of Citations in the Biomedical Literature	▪ Data generated from social media activities can be used to reflect broad types of impact. ▪ The analysis is based on 1.4 million documents covered by both PubMed and Web of ▪ Science and published between 2010 and 2012. ▪ less than 10% of PubMed articles mentioned on Twitter ▪ Correlations between tweets and citations are low,

Table 4: Literature Review Summary

Although Altmetrics claim to capture impact from a broader public but still it cannot replace the bibliographic scholarly impact[22], nonetheless it acts as a compliment to the traditional citation system [23]. More ever some scholars do not encourage correlation coefficient for comparing Altmetrics and different citation indicators for papers published in different period due to time constraints. For example, a paper should poses more citation if it is older other than that paper which is comparatively resent because an article should buy some time to create an impact in traditional biblometrics impact factor measurement technique, but in case of online citation i.e. likes, download, online mention, it might not take so long to create an impact i.e. a paper can be discussed on the time of presentation in any workshop, on internet by any person who is present in the workshop at the time of presentation and he or she get impressed by the idea. Social bookmarking systems [24] are used to tag, share and bookmark an article over the social web and Twitter is one of the more frequently used platform which is use to share anything with shorter time span[25].

There are many sources available that provides altmetrics data about an article now a day. "Altmetrics.com" and "Impactstory.org" are two main sources that provide social impact data from different social platforms, including how many data and knowledge bases have referred to a work, it's article views, number of its downloads, and its mentions in social and news media. As for as biblometrics concern, multiple indices are available which can be used to calculate biblometrics like H-Index and its variants. To accommodate this research, we are considering some of the variants of H-Index i.e. H-index, G-index, HC-index, M-quotient and HW-index. These indexes are used to calculate the biblometrics data by their standard formulas.

Measuring the impact of an article is not a one-step shopping. Traditionally, as discussed earlier in this document that techniques like impact factor and h-index were used to assess the quality of

once work but those metrics were used at the age of printed world. Now as authors used to share their contents on line, the importance's of those metrics tend to be more limited. Today, SCOPUS and PubMed Citations, Mendeley saves, CiteULike saves, tweets, blog posts, media mentions, article views, appealing bookmarks, Wikipedia citations, downloads, how many data and knowledge libraries refer to it, or mentions in social media and news media etc., reflects the quality of not only the scholarly articles but people, journals, data sets, videos, source code repositories, web pages, books, book chapters, computer code, presentation slides, posters, blog posts, digital humanities projects etc.

Altmetrics not only covers the impact of publication on the scholarly and non-scholarly public but it is more diverse and immediate as well. The best example of Altmetrics is PLoS Medicine (http://dx.doi.org/10.1371/journal.pmed.1001603) and Nature (http://dx.doi.org/10.1038/nature12961) is available. To acquire Altmetrics data for different disciplines you might require different means i.e. Web of Science, Scopus, Google Scholar or Other databases, such as PsycINFO etc. on the bases of the nature of data you required. Today both institutional and individual level tools are available online which provides Altmetrics data at one place. As for as institutional level tools are concerns, ORCID researcher profiles (http://orcid.org) and My Research Dashboard - Elsevier (https://www.myresearchdashboard.com) are the most common one. On the other hand tools for individual researcher are "ImpactStory" (http://impactstory.org), "ResearchGate", "Academia.edu", "Mendeley" (https://www.mendeley.com/) and "Altmetrics" are the most popular one.

Altmetrics is complementary to the traditional impact measurement analysis, helps to cover other types of impact beside the scholarly impact like societal, cultural and educational etc. Social

media attention some time called “Altmetrics donut” or “Altmetrics score” which reflect both quantitative attention in terms of “High attention-High score” value and qualitative attention in terms of “Weight according to different sources”. Beside the fact, quality of the Altmetrics yet remains disputed so care and modesty are needed when elaborating the results. Altmetrics indicators as for as twitter is a concern is: “Number of user who tweet or re-tweet a scholarly article”. As for as correlation is concerns many approaches can be used depends upon the nature of data, but Spearman correlation approach is considered one of the most suitable one when the analysis of Altmetrics is perform and compared with the traditional impact measurement approaches.

Traditional citation system although helps to measure the impact but failed to provide relevant paper to researchers. As apposite the Altmetrics not only measure the impact of an article but one can find relevant paper on the bases of the Altmetrics score which is more immediate and easy to propagate. As Altmetrics was described “The creation and study of new metrics based on social web for analyzing and informing scholarship” (http://altmetrics.org/about/), it still doing its job at best. Meanwhile Altmetrics is considered as a compliment to existing approaches but still it is relevant to see up to what degree Altmetrics is correlating with citation. That is why in manifesto the authors appealed and noted that: “Work should correlate between altmetrics and existing measures, predict citations from altmetrics, and compare altmetrics with expert evaluation.[26]”.

2.2 H-Index and its Variants:

H-index was first proposed by Jorge E. Hirsch back in 2005, in short Hirsch index called as h-index used to measure research output of any single scholar. Let me put it that way that h-index can be used as an indicator for lifetime achievement of an author, measured by the number of received citations score. The Hirsch index can be calculated for topics, journal citation, library

loans per category and predating its actual introduction here we will only use terms (i.e. citations and publications). The definition of h-index can be as follow: "A scientist has index h if h of his/her papers have at least h citations each and the other papers have less than h citations each [27]" . year later 2006 Glänzel [28]pointed out that this definition is not quite precise.

Publication	Number of citations
1	6
2	5
3	4
4	4
5	4
6	2

Table 5: Table for determining the h-index

According to original definition of Hirsch the h-index for Table 1 is not 5 because it is not true that there are 5 articles with at least 5 citations. Either the h-index cannot be 4 because it is not true that there are 4 articles with at least 4 citations (the 1st part of definition is logically correct) and the other ones have fewer than 4 citations (this part is now false). The h-index can precisely define as follows.

Consider the list of publications, ranked according to the number of citations received by each of these publications. Still an ambiguous question is that what happened with the same number of citations? In this case we know that publications are ranked between 1 to h from the Hirsch core, so for several publications with the same number of citation the Hirsch core can be determined by two techniques. One technique is to include all publications with h citations (in this case the Hirsch core may contain more than h elements) then in above table the value of h=3 with 4 elements, further it is noted that the last article from the list occupied R rank with citation C such

that C > R. The second technique is to give different rankings; the exact order does not matter to give different rankings, but anti-chronologically might be a good choice. The advantage of this technique is that the most recent articles have larger probability to belong to the Hirsch core than older articles. Hirsch core can said as a group of high performance publications. Ronald Rousseau proposed to call Hirsch core the set of the first h articles [[29]].

The h-index is no longer used as measure scientific achievement only for single researcher, but it is widely used to assess the research output of research group as well. When single researchers work in research groups, or countries, in addition to calculate h index values at the higher aggregate level it also possible to calculate h-index values[30] The h-index for research group can be defined as "The institute has an index h2 if h2 of its N researchers have an h1 index of at least h2 each, and the other (N−h2) researchers have h1indices lower than h2 each. The succession can then be continued, e.g., for networks of institutions or countries or other higher levels of aggregation".

There are different variants proposed for the h-index. Each proposed variant try to overcome the shortcomings h-index or its other variant. The g-index is the variant of h-index which is widely used. The g-index is introduced by Leo Egghe in 2006 [31]which overcomes weakness of h-index by capturing more citations that was not covered by h-index. Egghe defines the g-index "as the highest number g of papers that together received g2 or more citations. From this definition it is already clear that g = h". The g-index can found by comparing the cumulative citations against square of rank. When cumulative total of citation is equal/exceeds to rank square, then the g-index has been found. Another easy method to determined g-index is to first calculate the h-index and then apply following calculation

$$g = h(a^x) \qquad 1$$

Where a^x act as a vector of average citation.

Sometime we need to differentiate the authors on the basis of their career length for some purposes, the h-index only evaluate author on citation and publications base. Due to this reason, already in 2005 Hirsch had proposed to divide the h-index by number of years of an individual's research activity [28].

Burrell write in his paper that the h-index is approximately proportion to career length [32]. To compare the author with scientific careers we will divide the h-index by number of year from first publishing the paper. Following is definition of m quotient:

$$m\ quotient = \frac{h}{y} \qquad 2$$

Where '*h*' is value of h-index and *'y'* represent the number of years since publishing the first paper

Another technique proposed by Egghe & Rousseau in 2008 [33]which was sensitive to performance changes. This techniques known as hw-index (similar to ar-index) is variation of h-index, the h-index is weighted by citation impact and then it refers as hw-index. The hw-index defined as:

$$h_w index = \sqrt{\sum_{j=1}^{f_o} cit_j} \qquad 3$$

In h-index and its many variant the citation count is used to quantify the research output, another approach is proposed by Ajiferuke et al in 2010 [34]known as ch-index. In this technique instead

of citations, they proposed to use the number of citers for researcher production[34]. This technique minimizes the limited circle of authors citing the other's work.

For junior authors or recent authors the contemporary h-index is normally similar to their h-index. For more recognized authors there can be a difference between the two indices, representing that most of the papers included in their h-index have been published some time ago. While such the contemporary h-index continually provide relationship between junior authors and senior authors as compare to the h-index. Contemporary h-index is defined as "An Author has hc if hc of its number of publications get a score of Sc(j) $\geq$ hc each, and the rest publications get a score of Sc(j) $\leq$ hc [26] .

$$\mathbf{Sc(j) = \alpha(y(now) - y(j) + 1) * cit(j)} \qquad 4$$

Publication	Citations	Years	Cit/year	c4*4
1	8	2007	8/9=.88	3.52
2	6	2002	6/14=.42	1.68
3	4	2014	4/2=2	8
4	2	2009	2/7=1.12	1.12

h_c	Sc(j)≥h(c)
1	4
2	2
3	2
4	1

Table 6: Publications, citations and pub-year

To examining the ranking parameters, Author conducted experiments on an online bibliographic database containing data from conference and journal publications. For the case of conference and journal ranking, the indices contemporary h-index gives a more reasonable view for the ranking.

Chapter 3 METHODOLOGY

3.1 Introduction

This chapter explains the methodology to evaluate h-index and its extensions against the social citation or known as "Altmetrics". The data that we are using in this research thesis is related to data of Mathematics field.

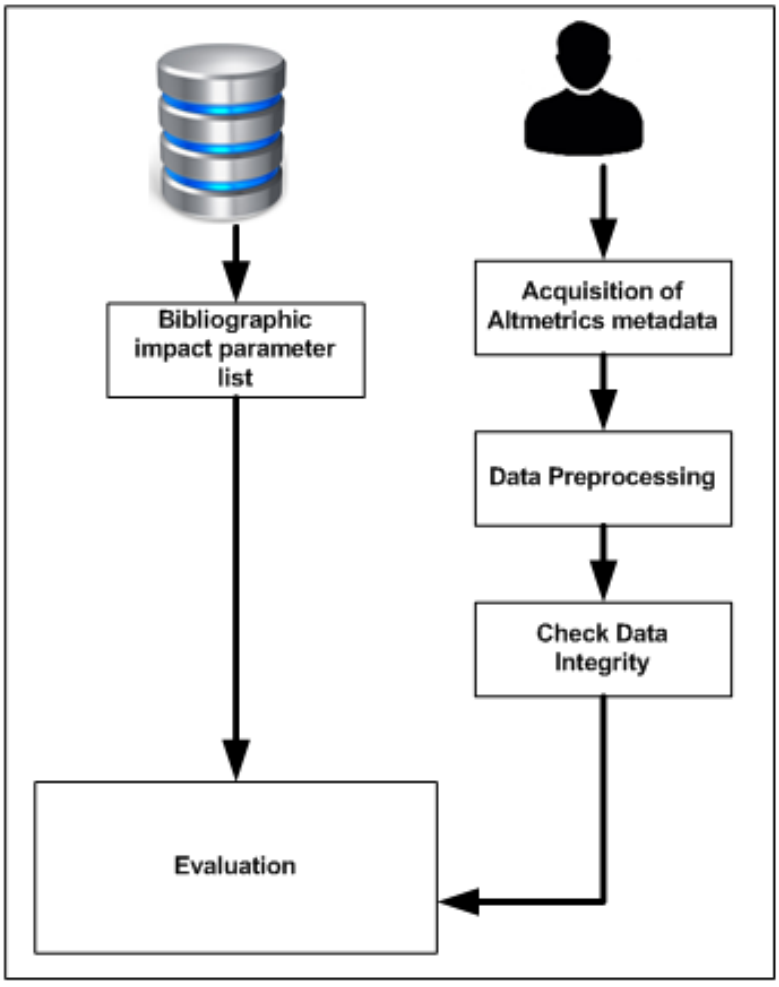


Figure 2: Proposed Methodology (abstract level)

Figure 2 presents the step in order to implement the proposed methodology at abstract level. Detail proposed methodology processes has been presented ahead in figure 3 on page 38. Each

step is discussed in detail in the following discussion. The comprehensive dataset of Mathematics domain has been taken to determine the potential of the altmetrics. Subsequently, h-index and its five variants (i.e. H-index, g-index, m-quotient, hc-index and hw-index) have implemented to acquire the correlation between the biblometrics and altmetrics. Further, Then the correlation is evaluated comprehensively against the altmetrics data as explained in the previous chapter.

Before the detail discussion we shed some light on the domain selection. As mentioned earlier that we are going to use authors dataset related to mathematics in order to make comparison between traditional biblometrics and altmetrics.

3.2 Domain Selection

We are examining the potential of a comparatively new approach to measure the impact of an article known as altmetrics by evaluating it against the well-known impact factor measurement approach known as biblometrics. To perform such task, we have selected the domain of Mathematics. The justification to select this domain is as follows:

i. Mathematics covers diverse area of knowledge

ii. Application of Mathematics are found in almost every fields of science

iii. None of the previous research has evaluated altmetrics data specifically in this domain.

iv. There was already work done in this domain for biblometrics prospective and we will like to see its results on the altmetrics (publications, citations, and co-author network), and we are motivated from already found results.

3.2.1 Taxonomy Building

Mathematics has applications in almost every field of sciences. Mathematics covers vast and diverse area of knowledge. We have taken data from Imama (Imama, 2015) to conduct a study

using different ranking parameter, (i.e., H-Index and its variants) to evaluate the competent authors in Mathematics domain. There are many subcategories in Mathematics domain. The certified classification of Mathematics field is known as Mathematics Subject Classifications (MSC). The latest version of this classification is MSC2010. This classification exhibits a hierarchical structure with three levels of categories. There are total 64 top level categories in this classification. 45 of these categories are related to pure Mathematics.

This categorization is combined effort of the two major mathematical reviewing databases, Mathematical Reviews[35] and Zentralblatt MATH [36]. These categories then used by different mathematics journals like Conformal Geometry and Dynamics, Journal of American Mathematical Society and Mathematics of Computation. There are many sources to acquire data of an author publication, Citations, Co-author networks such as: MathSciNet, Zbmath and WoS and Google Scholar. Their access (except Google Scholar) is not openly available, and WoS has very limited coverage of Mathematics domain[37] Author constructed dataset from Google Scholar (Imama, 2015). Google Scholar is an open access online resource which also does citation indexing contains comprehensive data on almost each field of science. Many authors compared the coverage of WoS with Google Scholar and one of the recent study revealed that Google Scholar growth is 13% more than that of WoS [38]. Furthermore the number of Google Scholar citation over the one year period increased by approximately 1.5% per month[38].

The main limitations that are pointed out are: double counting of citation and less frequent updating of dataset. To gather data from Google Scholar manually is a time consuming task. Author built crawler to crawl a data for the topics in the Mathematics Classification system (Imama, 2015). Imama gathered 69527 publications from Google Scholar. These publications were gathered by top rank relevant and highly cited results. There was two test applied to ensure

their relevancy with Mathematics domain. It was identified manually that some of publication were not actually the publications. Author verified data in two dimensions. First they checked the relevancy with Mathematics domain, and second was to check the replications of authors in the gathered data (author disambiguation). After preprocessing of data, finally 69198 publications were left after discarding 169 publications. After data gathering process the total number of authors was 57533 and among those 29263 have last same name, after verifying these same name authors7744 unique authors were identified. After preprocessing step, total authors were 57533 with 8,821,251 citations.

3.2.2 Dataset Limitations

First and very obvious Limitation of Altmetrics is that Altmetrics data providers may change or suspend their presence with the passage of time. Secondly, it is nearly impossible to collect complete data from the internet sources; there is always a chance to miss some of information while collecting data from internet sources. Although, we make sure that all the data collected regarding Altmetrics must be complete and up-to-date. However, still the chances of an error cannot be denied. For tweet count we relied on the Twitter API for tweet count (which itself caries some technical limitations), which are obtained through document's title, DOI and URL. Therefore, evaluations of results associated with Twitter are based on its internal criteria on which it considers a tweet related to a specific paper. Another basic question asked about social media platforms is reliability. Reliability itself comes with many questions about completeness, authenticity, comprehensiveness etc. In addition, we did not know how download speed, time zone restrictions, and server downtime etc. would affect available data [22]. On the bases of above mentioned limitations, we have ensured data integrity by crosschecking collected data with other Altmetrics data sources like Impactstory.org to minimize the effect of limitations.

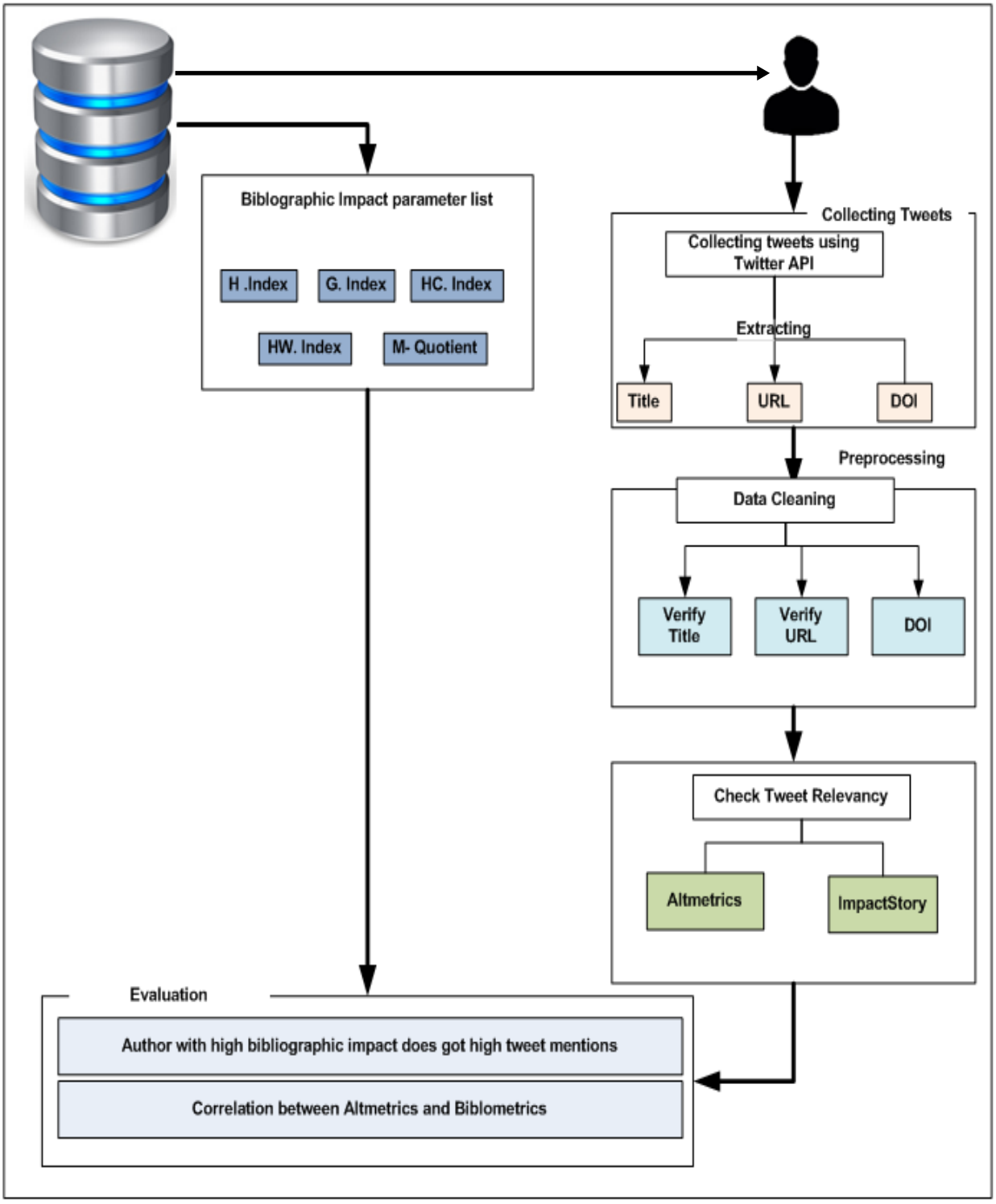


Figure 3: Detail proposed Research Methodology

3.3 Twitter Dataset

Twitter is a very popular social networking website founded by Jack Doresy and in 2006, used by millions of users. Twitter is used as platform where user can share their feeling about any events or news etc. A user can share a single tweet usually comprise of 140 words. We collect tweet related to the authors and their research work. Twitter API is used to collect these tweets. We are interested in the URL of author or their research work, the Document ID and title of the research article. The dataset of tweet comprised of those tweets related to the authors of mathematics which we consider for this research presented in figure 4 below.

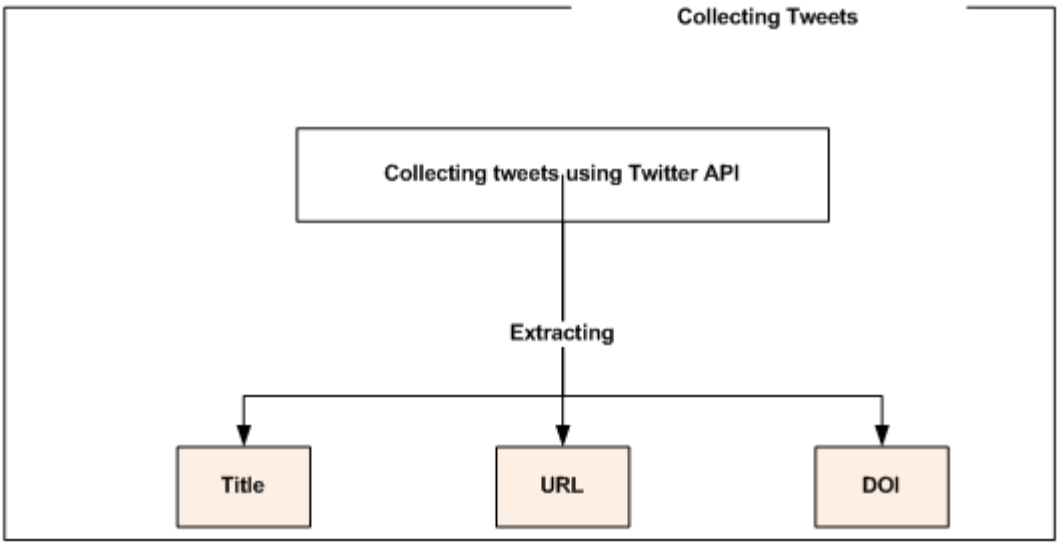


Figure 4: Tweets related to authors and their research articles are collected and Title, URL and DOI is extracted

3.4 Data Preprocessing

The collected dataset of tweet is treated with text preprocessing steps. Text preprocessing processes like strop words removal, deleting impurities and deleting ambiguous tweets. After treating the tweet dataset with text preprocessing processes next the Document object identifier (DOI), Title and URL (Universal Resource locator) has been verified shown in figure 5.

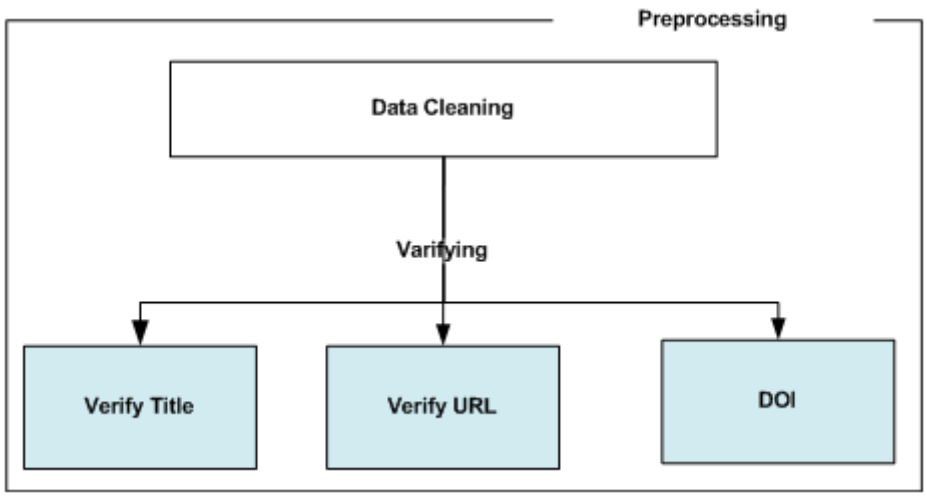


Figure 5: Data cleaning and verifying the Title, URL and DOI of the authors and articles.

3.5 Check Tweet Relevancy

To check that we got a valid Altmetrics score against any particular paper, we checked the relevancy of tweets by using online Altmetrics data providers like Altmetrics and ImpactStory for double check.

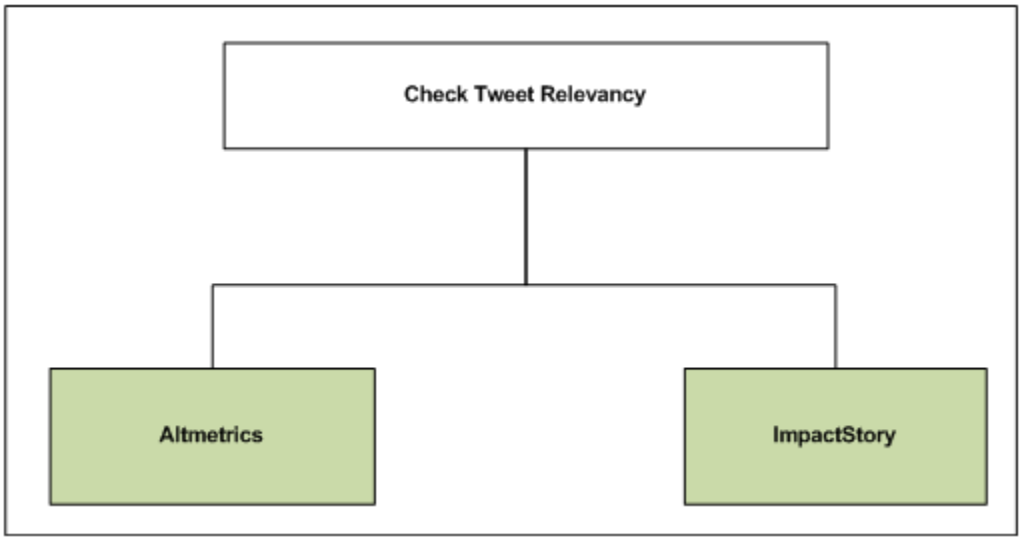


Figure 6: Check the tweet relevancy with two Altmetrics website a) Altmetriec.com, b) ImpactStroty.com

3.6 Biblometrics impact factor parameters

This module is dedicated to the creation of the ranked expert lists Hw-index, Hl-index, Hm-index, Hc-index, M-quotient, Fraction-count-on-citation and fraction-count-on-papers. In following discussion each of the above mentioned parameters are explained and depicted in figure 7 below.

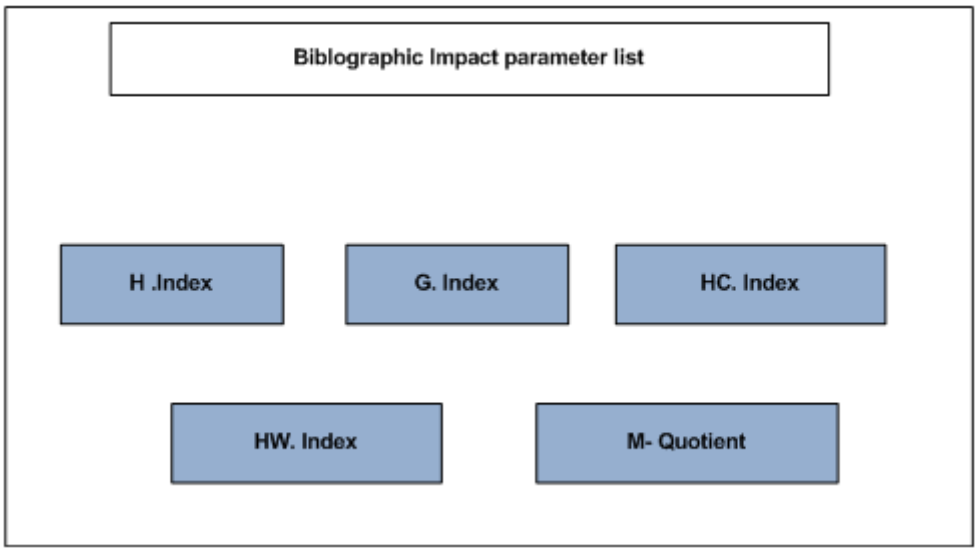


Figure 7: Traditional Bibliographic impact parameters

3.6.1 hw-Index:

The hw-index was calculated the hw-Index is defined as:

$$\sqrt{\sum_{j=1}^{r.} cit(j)} \quad where\ r(j) = \sum_{j=1}^{i} cit(j)/h\ r(j) \leq cit(j) \qquad 5$$

(10) The above formula calculates the hw-index of an author. By sum of those authors whose values should be less than or equal to that citation.

3.6.2 M-quotient:

Taking career length problem in mind, Hirsch himself proposed Hirsch's m or m-quotient in his original paper in which he proposed h-index. He computed m-quotient, by dividing h-index with the number of years since publishing the first paper. m quotient is an important ranking

parameter tool when any person needs to compare 40 authors with different lengths of scientific career. The main advantage of m-quotient is when a scholar gets a break from publications; his or her *h*-index can continue to decrease. The m-quotient is defined as:

$$m - quotient = \frac{h}{y} \qquad (6)$$

3.7 Evaluation

We have attained both biblometrics and altmetrics data of researchers according to each of these measurements, then a complete evaluation is performed.

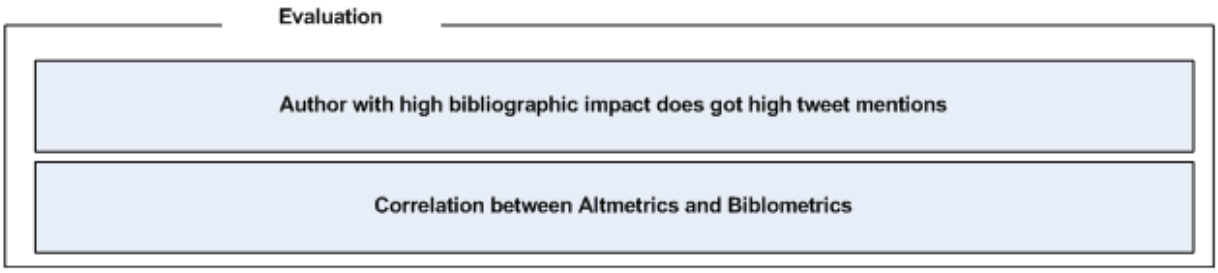


Figure 8: Research questions being answered in this research

3.7.1 Measure of Correlation between indexes and altmetrics

In evaluation, the first step is to find correlation between indexes and altmetrics data. Our first research question is to find the correlation between Biblometrics and altmetrics. The correlation between altmetrics and indexes is evaluated. The main purpose to find the correlation between these indexes and altmetrics is to discover to which extend these indexes are related to altmetrics date. Considering the ranked nature of the data the best suitable correlation measure is spearman. To measure the co-relation between Bibliometrics and Altmetrics, the standard formula of spearman co-relation will be used.

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}. \qquad (7)$$

We implement the spearman co-relation in excel to find out the co-relation by inputting both Bibliometrics and Altmetrics data. The more detail outcome of this research question is discussed in the chapter 4.

3.7.2 Author with higher indexes value dose get higher altmetrics value

The top ranked authors by all rankings are evaluated to answer this question. To achieve this task the top 20 authors have been picked which have higher h-index value and checked against the corresponding altmetrics data. The more detail outcome of this research question is discussed in the chapter 4.

3.7.3 Does altmetrics replace the biblometrics

This third Research question gives the contributing parameter to bring the awardees in top ranked list. To achieve the percentage the occurrence of awardees in top 10% to 100% of the ranked list of authors is checked. The detail outcome of this research question is discussed in chapter 4.

Chapter: 4 RESULTS & EVALUATION

4.1 Analysis

In different set of analysis, we first show the extent to which scholarly documents from the field of mathematics are found on the Twitter. We also explore the degree to which these citations vary over time. To perform this analysis data set of 69196 papers were used. Citation rate over Twitter, simply calculated by finding mean of tweeted papers, is calculated and distribution of tweets per article is mentioned. We also identify top twenty papers which have been cited frequently over Twitter.

In second phase of analysis our focus is to distinguish and identify number of papers that have been tweeted at least once. Percentage of tweeted documents or Twitter courage P% tweeted and the mean that is Twitter Citation Rate T/P tweeted (Priem 2010) were calculated. For this calculation, only those articles have been considered which were tweeted at least once. Exclusion of the articles which are never tweeted leaves us with 1618 papers. Spearman correlation has been calculated between Twitter citation of papers and traditional bibliographic impact factor techniques H-Index and its variants. Finally analysis of correlation between Twitter citation count and Index value of all 57533 authors are mentioned that is calculated with statistical correlation (Spearman's) that is a common approach to validate new matrices by examining the correlation between them. Journals that are most frequently tweeted have also been listed in the results.

4.2 Altmetrics Data-Set Limitations

Replication is considered as one of the big hurdles while dealing with the Altmetrics data. First and very obvious one is Altmetrics data providers may change or suspend their presence with the passage of time. Secondly, it is nearly impossible to collect complete data from the internet

sources; there is always a chance to miss some of information while collecting data from internet sources. Although, we make sure that all the data collected regarding Altmetrics must be complete and up-to-date. However, still the chances of an error cannot be denied. For tweet count we relied on the Twitter API for tweet count (which itself caries some technical limitations), which are obtained through document's title, DOI and URL. Therefore, evaluations of results associated with Twitter are based on its internal criteria on which it considers a tweet related to a specific paper. Another basic question asked about social media platforms is reliability. Reliability itself comes with many questions about completeness, authenticity, comprehensiveness etc. In addition, we did not know how download speed, time zone restrictions, and server downtime etc. would affect available data. On the bases of above mentioned limitations, we have ensured data integrity by crosschecking collected data with other Altmetrics data sources like Impactstory.org to minimize the effect of limitations.

4.3 RESULTS AND DISCUSSION

Analysis shows that 1021 out of 69196 scholarly documents get citation on the Twitter at least once which is very low Tweetation rate indeed. From these 1021 documents most of the tweeted papers, that is 63, were from year 2013, 45 were from year 2012 and more than 33 were from 2003, 2007, 1998, 2002, 2008 and 2006 equally. Figure10 shows the remaining papers tweet score by a document based on its publication year. The trend shows that latest papers got more citations on the tweeter than the older papers.

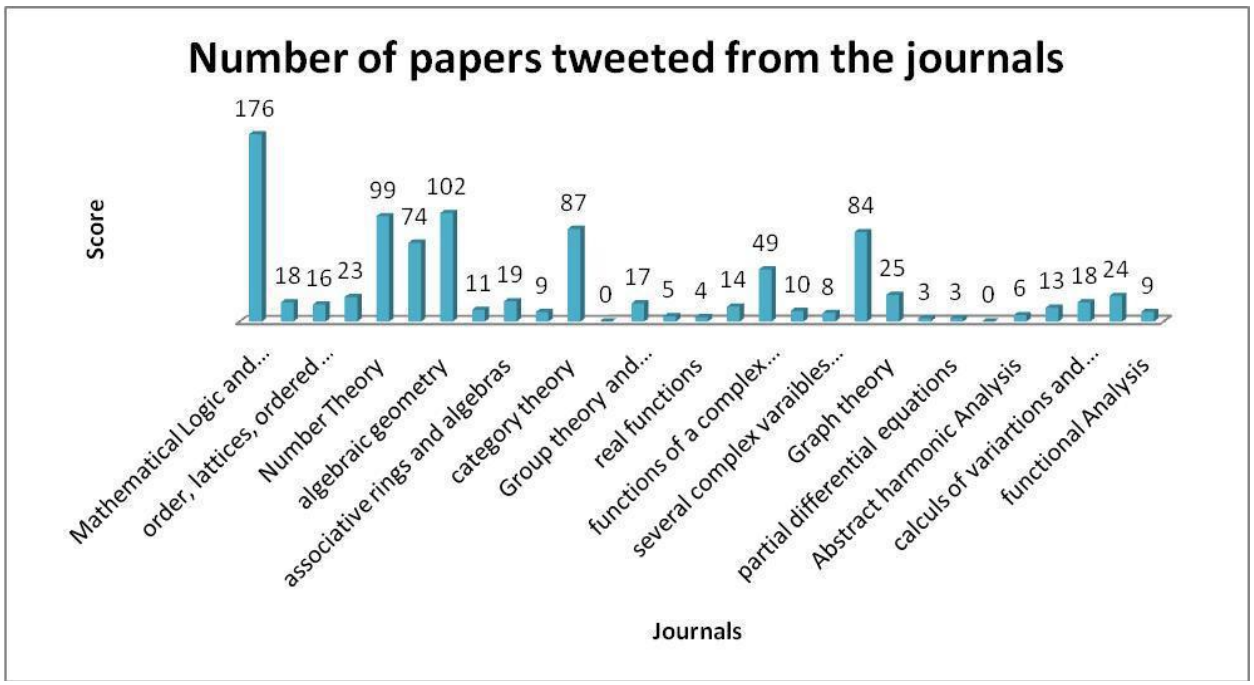


Figure 9: Shows the score by different categories

From 1021 tweeted papers only four papers score more than 10,000 citations over Twitter which covers 0.391% of the total tweeted documents, 1017 documents got at least one citation which covers 1.47% of the total dataset and Remaining 98.5% (68175 out of 69196) do not get any citation on Twitter. From the 45 categories of the mathematical domain only 29 categories were mention on Twitter that is 64.44% of the total categories. Remaining 16 categories (i.e., Round 34%) did not get any citation over Twitter. Figure 9 shows the results of categories mentioned frequently on Twitter. From the total tweeted Categories, "Mathematical Logic and Foundations" were the most frequent tweeted category; scored 176 tweets. "Algebra geometry" remains second with 102 tweet score. Likewise, form dataset of 57534 authors only 16% of them were tweeted at least once on the Twitter.

With respect to our research question no 1 that was, ***Is there any correlation exists between Altmetrics and Bibliometrics impact factor measurement techniques i.e. (H-Index, G-Index, M-Quotient, Hc-Index and Hw-Index)?***

As far as correlation is concerned, there is very low correlation exists in Bibliometrics and Altmetrics. As for as the H-index is concerns, co-relation between the Bibliometrics and Altmetrics was very low, that is 0.023091 but (Figure 13) the trend is positive (i.e., as H-index grows, tweets grow as well). On the other hand, with the co-relation of 0.025936, there is once again very less co-relation between G-Index and Altmetric but (Figure 14) the trend is positive (i.e., as G-index grows, tweets grow as well).

(Figures 15, 16 and 17 shows co-relation results of HW-Index, M-quotient and HC-Index respectively) With the co-relation figure of 0.067626 HW- Index shows very low correlation with the twitter citation but shows some promise as compare to other Bibliometrics parameters. Still the trend line shows the positive behaviour of the co-relation between HW-index and Altmetrics. In case with M- Quotient 0.008069 shows less correlation with the twitter data as compared with the H-Index value. Rank correlation with the HC-Index does not show positive intent with value -0.00411.

Author with higher Bibliometrics impact does get higher tweet?

Comparing the result of H-Index and G-Index of first twenty authors; there is a negative co-relation between Bibliometrics and Altmetrics. As for as H-index is concern, the co-relation is -0.378512 dictates no similarities among impact full authors in traditional system and authors famous on tweeter. Same is the case with the G-index, that is, with the co-relation of -0.379116, there is once again a negative co-relation between G-index and tweets of first twenty authors.

Figure 11 shows the correlation between Altmetrics and G-Index values and figure 12 show the co-relation between H-Index and Altmetrics.

Table 7 shows the top twenty authors that get higher citations on the Twitter. Average tweets of an author are calculated by simple mean that is 1.305. From 1618 tweeted authors 766 score 1 that is 47.34% of the total tweeted authors. While 13.53 get two tweets. 97.18% authors do not get any citation over Twitter. Of the top twenty authors who have highest H-Index value, only ten get mentioned on the micro blogging platform that is 0.979% of the total tweeted data set With around -0.4 correlation coefficient values, no correlation exists between the first twenty authors with higher H-Index value and the Altmetrics data.

Table 7: Top twenty authors who score highest on Twitter along with their H-Index and G-Index

S.NO	AUTHOR ID	AUTHER First Name	AUTHER's Last Name	TWEETS	G-Index	H-Index	CORRL-h	CORL-g
1	7866	L	**Zhongkui**	10000	1	1	-0.378512	-0.379116
2	20457	L	**Bo**	10000	2	1		
3	40238	GA	**Barnard**	10000	2	2		
4	57250	S	**Barzilai**	10000	1	1		
5	71	G	**Longo**	4948	10	7		
6	57697	EJ	**Horvitz**	4117	7	7		
7	21609	JH	**Palmieri**	4107	3	3		
8	59161	A	**Alves**	4088	1	1		
9	59162	Y	**Panis**	4088	1	1		
10	59163	D	**Trancart**	4088	1	1		
11	59164	JM	**Regimbeau**	4088	1	1		
12	2259	SO	**Hansson**	3547	2	2		
13	2260	T	**Grüne-Yanoff**	3547	1	1		
14	12899	AV	**Stepanov**	2815	3	2		
15	22238	C	**Petronio**	2762	4	3		
16	410	P	**Schuster**	2755	5	5		
17	228	T	**Coquand**	2754	14	8		
18	2261	H	**Lombardi**	2754	7	5		
19	54203	H	**Formby**	2471	2	1		
20	23400	R	**Steiner**	2353	13	13		

Same is the case with the G-Index value; from top twenty authors with higher G-Index value, only eight authors work was tweeted at least once on the Twitter that is 0.783% of the total tweeted dataset.

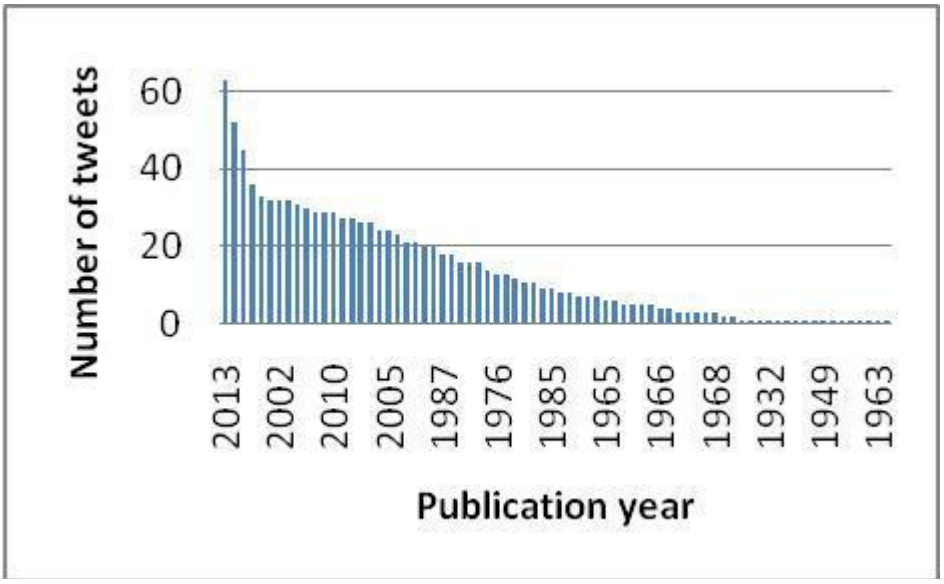


Figure 10: Latest papers get more attention on the Twitter as compare to the older one.

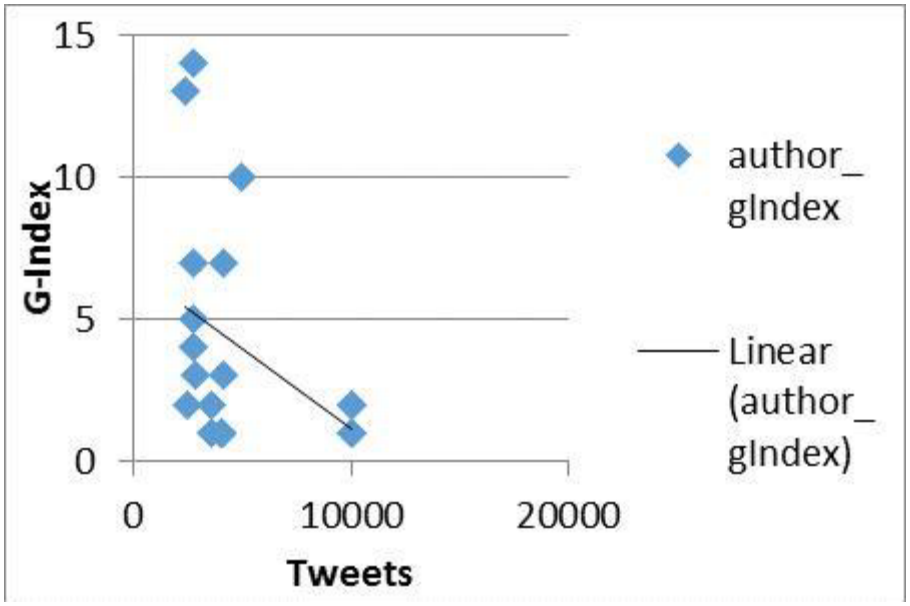


Figure 11: Co-relation Coefficient of G-index and Tweetation of first 20 authors [-0.379116].

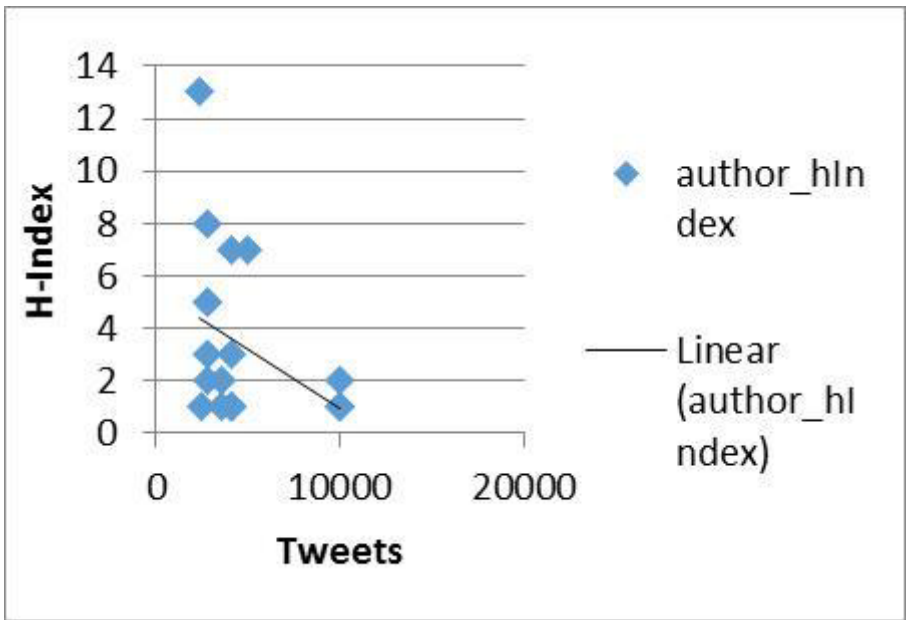


Figure 12: Correlation Coefficient of H-index and Tweetation of first 20 authors [-0.378512].

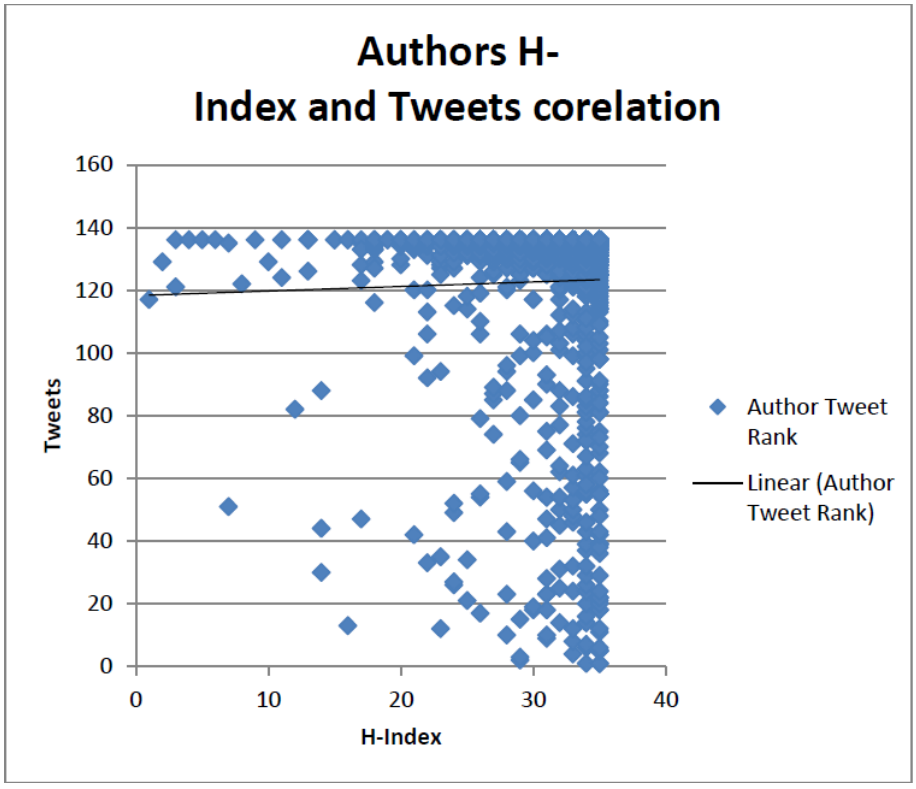


Figure 13: Rank correlation between H-Index and Twitter that is round 0.23091

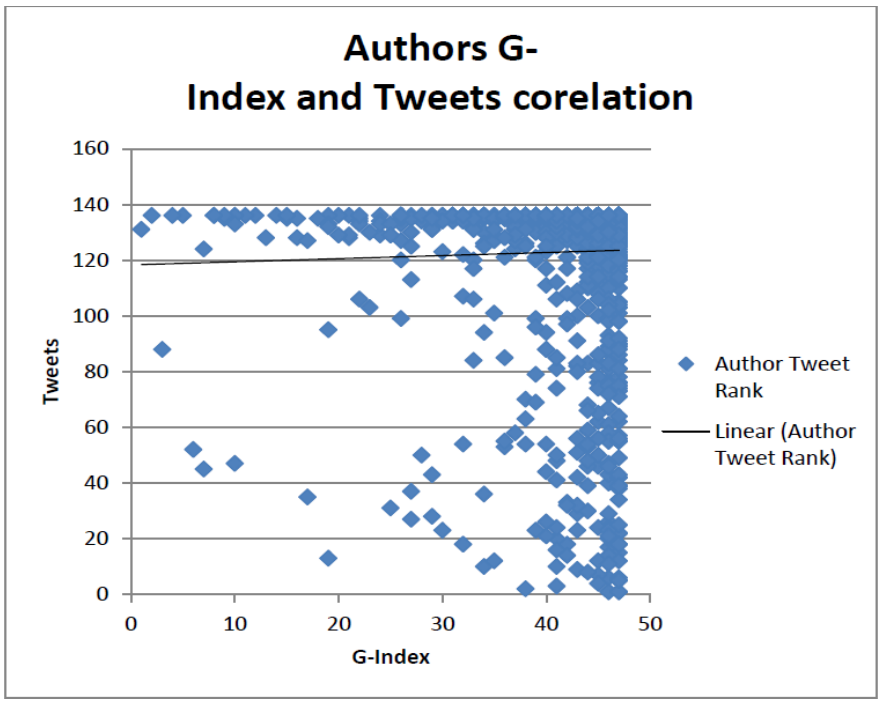


Figure 14: Rank co-relation between G-Index and Altmetrics that is round 0.025936.

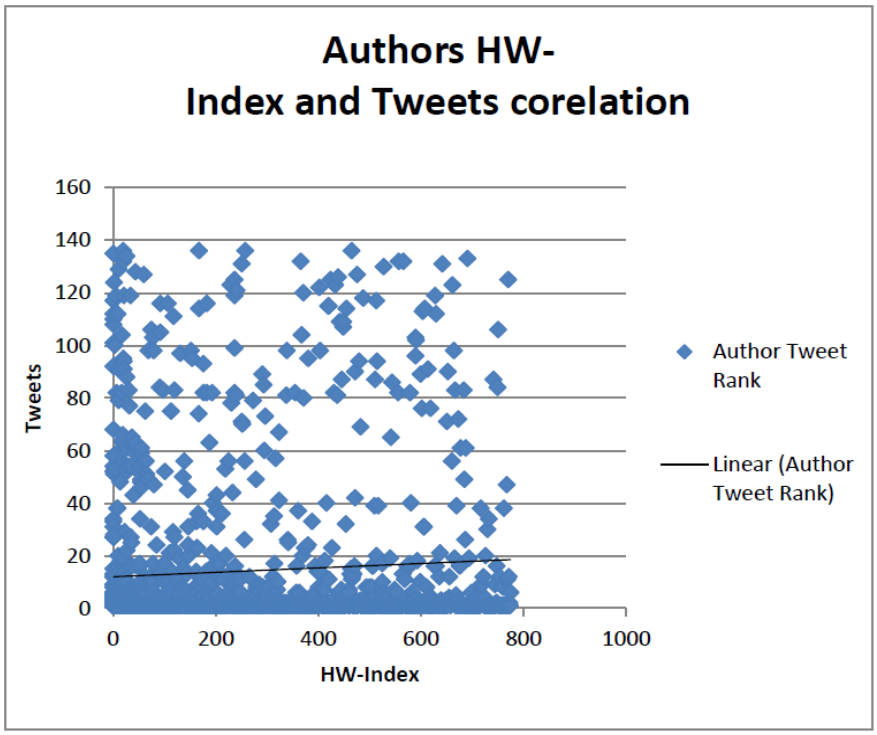


Figure 15: Co-relation between Twitter rank and HW-Index that is 0.067626

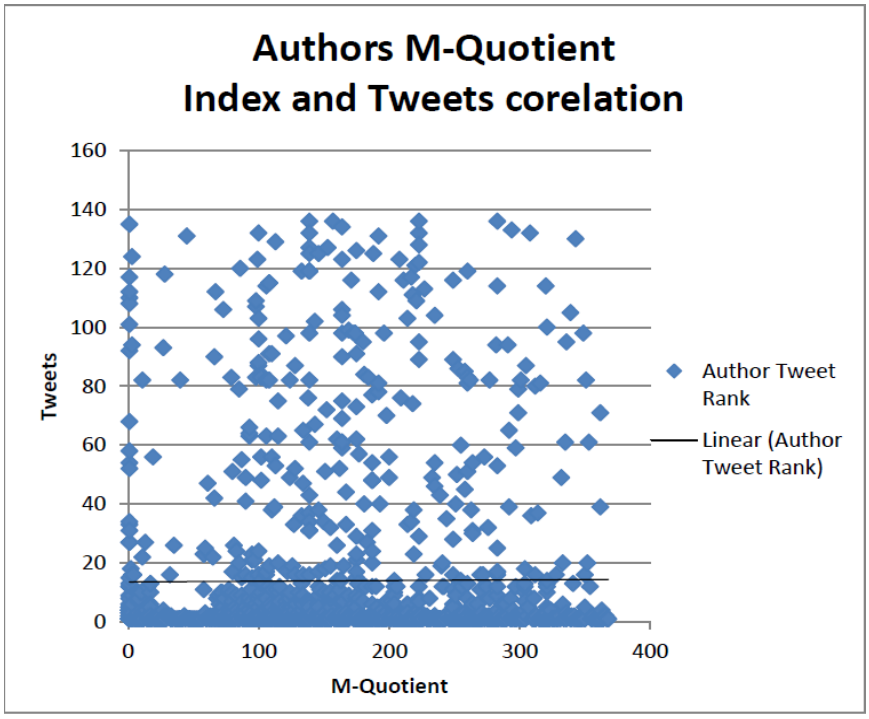


Figure 16: Co-relation between twitter and M-quotient that is 0.008069

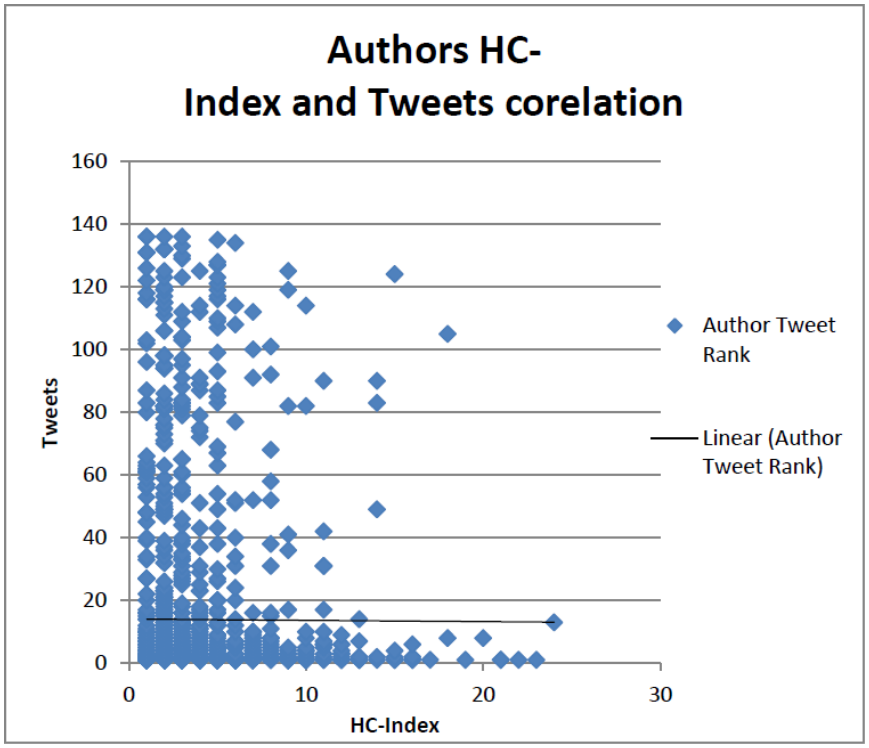


Figure 17: Co-relation between twitter and HC-Index that is -0.00411

Chapter 5 CONCLUSIONS

This study covers entire continuum of the mathematical domain with reasonable data for the evaluation of Altmetrics. We classify distinctive relationships between Altmetrics and bibliographic citations, which can be used as a roadmap to evaluate twitting behaviour of people in other fields of study. Furthermore from our result we found that people mention latest paper more frequently than the older ones. So, on the bases of that evidence we can predict that in future the need of Altmetrics will be fell more deeply.

5.1 Research Questions Answered

RQ1. Is there any correlation exists between Altmetrics and Biblometrics impact factor measurement techniques i.e. (H-Index, G-Index, M-Quotient, HC and HW)?

Over all there is a very less correlation between Bibliometrics and Altmetrics on the bases of H-index and its variants as a Bibliometrics impact factor parameters and tweets as Altmetrics parameters. With the co-relation figure of 0.067626 HW- Index shows some promise as compare to other Bibliometrics parameters (i.e., H-Index, G-Index, HC-Index and M-Quotient). On the other hand with -0.0041, HC-index shows the least but negative co-relation than others Bibliometrics parameters.

RQ2. Author with higher Biblometrics impact does get higher tweet count?

The outcome of less than 2% of documents mentioned on the Twitter shows a very low coverage of mathematical documents on Twitter, which most probably can be due to scholarly focus on the traditional sources of impact measurement techniques. However, we were able to demonstrate that some categories are more popular than others. Less correlation between indices and Altmetrics shows that Altmetrics and indices measurement techniques are far away from

each other which means; Altmetrics cover other type of impact that is not comparable with the traditional citation system and therefore Biblometrics indices should not be considered alone as being representative of an author's goodness, but altmetrics should also be taken into consideration to have a complete picture of the author's impact, that is his impact in author as well as non-author community.

5.2 Future Work

Although we discover many factors explaining the low and high correlations but we think that it need details investigation to discover many others. In future we are interested to unearth other factor behind this correlation. We are also interested to know the future of altmetrics. Altmetrics is the combination of some important metrics that consider some other web channels where authors work is cited and used. With fast growth of social media this becomes very important to use altmetrics. Lot of research work has been done on social media and altmetrics is one of them. We are interested to investigate that what is the future of altmetrics.

REFERENCES

[1] & D.-C. Galligan, F., "Altmetrics: Rethinking the Way We Measure," *j.serrev*, vol. 1, no. 39, pp. 56–61, 2013.

[2] J. Bar-Ilan, "JASIST@ Mendeley," in *Presented at ACM Web Science Conference Workshop on Altmetrics, Evanston, IL*, 2013.

[3] P. Priem, J., Taraborelli, D., Groth, "Altmetrics: a manifesto," 2010. [Online]. Available: http://altmetrics.org/manifesto/. [Accessed: 12-Dec-2015].

[4] P. McFedries, " 'Measuring the impact of altmetrics [Technically Speaking]," in *IEEE Spectrum*, 2012, p. 28.

[5] S. Galligan, F., & Dyas-Correia, "Altmetrics: Rethinking the way we measure'," 2013.

[6] Viney, " 'Altmetrics: research council responds," 2013.

[7] J. Piwowar, H. and Priem, "The power of altmetrics on a CV," in *ulletin of Association for Information Science and Technology*, 2013, pp. 10–13.

[8] Sugimoto, "Scholars on soap boxes: Science communication and dissemination in TED videos," *J. Am. Soc. Inf. Sci. Technol.*, vol. 64, pp. 663--674, 2013.

[9] Symes, " 'Decoupling hydrogen and oxygen evolution during electrolytic water splitting using an electron-coupled-proton buffer," *Nat. Chem.*, vol. 5, p. 403--409., 2013.

[10] R. Wouters, P., Costas, "Users, narcissism and control: Tracking the impact of scholarly publications in the 21st century. Utrecht: SURF foundation." [Online]. Available: http://www.surffoundation.nl/nl/publicaties/Documents/Users narcissism and control.pdf. [Accessed: 10-Nov-2016].

[11] C. R. Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, "Do altmetrics work? Twitter and ten other social web services," 2012.

[12] D. Boyd, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitte," in *43rd Hawaii International Conference on, IEEE*, 2010.

[13] G. Eysenbach, " 'Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact," *Med. Internet Res. JMIR Publ. Inc*, vol. 13, 2011.

[14] J. Lin and M. Fenner, "Altmetrics in Evolution : Defining and Redefining the Ontology of," vol. 2, no. 2, 2013.

[15] Altmetric.com, "Altmetrics can showcase the attention and influence of research," 2016. [Online]. Available: https://www.altmetric.com/about-altmetrics/what-are-altmetrics/.

[Accessed: 14-Jul-2015].

[16] H. Zhang, “ltmetrics: Measure Research Impact: Introduction to Altmetrics,” 2016. [Online]. Available: http://guides.library.oregonstate.edu/altmetrics.

[17] J. Priem, C. Parra, H. Piwowar, and P. Groth, “Uncovering impacts : a case study in using altmetrics tools.”

[18] M. Thelwall, “‘Webometrics’, Annual review of information science and technology,” *Wiley Online Libr.*, vol. 39, pp. 81--135, 2005.

[19] J. Kaur, “Scholarometer: A social framework for analyzing impact across disciplines,” *Public Libr. Sci.*, vol. 7, 2012.

[20] J. Priem, “‘Altmetrics: A manifesto,’” 2010.

[21] S. Haustein, “‘ altmetrics’: viewing acts on social media through the lens of citation and social theories’,” 2015.

[22] J. Liu, “‘Five challenges in altmetrics: A toolmaker’s perspective,” *Wiley Online Libr.*, pp. 31–34, 2013.

[23] R. Costas, “ ’Do ‘altmetrics’ correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective,” *J. Assoc. Inf. Sci. Technol.*, 2014.

[24] S. Reher, “Social Bookmarking in STM: Putting services to the acid test,” 2010.

[25] X. Shuai, “‘How the scientific community reacts to newly submitted preprints: Article downloads, twitter mentions, and citations,’” 2012.

[26] M. Doerfel, Stephan and Zoller, Daniel and Singer, Philipp and Niebler, Thomas and Hotho, Andreas and Strohmaier, “What Users Actually do in a Social Tagging System: A Study of User Behavior in BibSonomy,” *ACM Trans. Web*, vol. 10, no. 2, p. 14:1--14:32, 2016.

[27] H.-D. Bornmann, Lutz and Daniel, “Does the h-index for ranking of scientists really work?,” *Scientometrics*, vol. 65, no. 3, pp. 391--392, 2005.

[28] A. Braun, Tibor and Gl{"a}nzel, Wolfgang and Schubert, “A Hirsch-type index for journals,” *Scientometrics*, vol. 69, no. 1, pp. 169--173, 2006.

[29] R. Egghe, Leo and Rousseau, *Introduction to informetrics: Quantitative methods in library, documentation and information science*. Elsevier Science Publishers, 1990.

[30] J. Hu, Xiaojun and Rousseau, Ronald and Chen, “In those fields where multiple authorship is the rule, the h-index should be supplemented by role-based h-indices,” *J. Inf. Sci.*, vol. 36, no. 1, pp. 73--85, 2010.

[31] L. Egghe, “Theory and practise of the g-index,” *Scientometrics*, vol. 69, no. 1, pp. 131--152, 2006.

[32] Q. L. Burrell, “On the h-index, the size of the Hirsch core and Jin’s A-index,” *J. Informetr.*, vol. 1, no. 2, pp. 170--177, 2007.

[33] R. Egghe, Leo and Rousseau, “An h-index weighted by citation impact,” *Inf. Process. \& Manag.*, vol. 44, no. 2, pp. 770--780, 2008.

[34] D. Ajiferuke, Isola and Lu, Kun and Wolfram, “A comparison of citer and citation-based measure outcomes for multiple disciplines,” *J. Am. Soc. Inf. Sci. Technol.*, vol. 61, no. 10, pp. 2086--2096, 2010.

[35] ASM, “Mathematical Reviews,” 2016. [Online]. Available: http://www.ams.org/mr-database. [Accessed: 03-Mar-2016].

[36] Z. MATH, “Zentralblatt MATH - Serials Database,” 2015. [Online]. Available: http://www.zblmath.fiz-karlsruhe.de/serials/. [Accessed: 03-Mar-2016].

[37] H. Bosman, Jeroen and Mourik, Ineke van and Rasch, Menno and Sieverts, Eric and Verhoeff, “Scopus reviewed and compared: The coverage and functionality of the citation database Scopus, including comparisons with Web of Science and Google Scholar,” 2006.

[38] A.-W. Harzing, “A longitudinal study of Google Scholar coverage between 2012 and 2013,” *Scientometrics*, vol. 98, no. 1, pp. 565--575, 2014.

Appendix A

Modified List of Categories of Mathematics Field

1. **Abstract harmonic Analysis**
   - 1.1. Amenable groups
   - 1.2. Lp-spaces

2. **Algebraic geometry**
   - 2.1. Affine fibrations
   - 2.2. Elliptic surfaces
   - 2.3. Picard group
   - 2.4. Riemann-roch theorems
   - 2.5. Rigid analytic geometry

3. **Algebraic Topology**
   - 3.1. Elliptic cohomology
   - 3.2. Fiber spaces
   - 3.3. H-spaces and duals
   - 3.4. J-morphism
   - 3.5. K-theory
   - 3.6. Loop spaces
   - 3.7. Orbifold cohomology

4. **Approximations and Expansions**
   - 4.1. Chebyshev systems
   - 4.2. Pade approximation

5. **Associative rings and algebras**
   - 5.1. Hopf algebra
   - 5.2. Lattices over orders
   - 5.3. Nil and nilpotent radicals
   - 5.4. Quasi-frobenius rings

6. Calculs of variartions and optimal control optimization

6.1. Differential games

6.2. Duality theory

6.3. Frechet and gateaux differentiability

6.4. Hamilton-Jacobi theories

6.5. Inverse problems

6.6. Minimax problems

7. Category theory

7.1. Adjoint functors

7.2. Epimorphisms, monomorphisms

7.3. Functor categories

7.4. Monoidal categories

8. Combinatorics

8.1. Generalized remsey

8.2. Infinite Graphs

8.3. Matroids, geometric lattices

8.4. Matroids,geometric

8.5. Polyominoes

8.6. Q-calculus

9. Commutative algebra

9.1. Cluster algebras

9.2. Cohen-macaulay modules

9.3. Formal power series rings

9.4. Morphisms

9.5. Seminormal rings

9.6. Witt vectors

10. Convex and discrete geometry

- 10.1. Convex sets without dimension restrictions
- 10.2. Helly-type theorems
- 10.3. Isoperimetric problems
- 10.4. Lattice polytopes
- 10.5. Matroids
- 10.6. Spherical and hyperbolic convexity

11. Difference and functional equations

- 11.1. Stochastic difference equations

12. Differential geometry

- 12.1. Classical differential geometry
- 12.2. Differential line geometry
- 12.3. Euclidean space
- 12.4. G structures
- 12.5. Kinematics
- 12.6. Projective connections

13. Dynamical systems and ergodic theory

- 13.1. Cellular automata
- 13.2. Chaotic dynamics
- 13.3. Ergodic theorems
- 13.4. Homoclinic and heteroclinic orbits
- 13.5. Index Theory
- 13.6. Lattice dynamics
- 13.7. Monotone flows
- 13.8. Morse-smale systems
- 13.9. Nonholonomic dynamical systems
- 13.10. Notions of recurrence
- 13.11. Partially hyperbolic systems

13.12. Soliton theory

13.13. Symbolic dynamics

14. Field Theory and Polynomials

14.1. Hilbertian fields

14.2. Homological methods

14.3. Nonstandard arithmetic

14.4. P-adic differential equations

14.5. Skew fields

14.6. Topological semifields

15. Functional Analysis

15.1. Barrelled spaces, bornological spaces

15.2. Locally convex frechet spaces

15.3. Saks spaces

15.4. Sequence spaces

15.5. Sobolev spaces

16. Functions of a complex variable

16.1. Bergman spaces, fock spaces

16.2. Boundary value problems

16.3. Hardy spaces

16.4. Klein surfaces

16.5. Kleinian groups

16.6. Meromorphic functions

16.7. Quasiconformal mappings

17. General algebraic systems

17.1. Automorphisms, endomorphisms

17.2. Infinitary algebras

18. General Topology

18.1. Fuzzy topology

18.2. Hyperspaces

18.3. Moore spaces

18.4. P-minimal and P-closed spaces

18.5. Quotient spaces

18.6. Spectra

18.7. Syntopogenous structures

19. Geometry

19.1. Affine analytic geometry

19.2. Configuration theorems

19.3. Discrete geometry

19.4. Laguerre geometry

19.5. Linear incidence geometry

19.6. Metric geometry

19.7. Mobius geometry

19.8. Polar gemetry

19.9. Ring geometry

19.10. Steiner systems

20. Global analysis analysis on manifolds

20.1. Bifurcation theory

20.2. Critical metrics

20.3. De Rham theory

20.4. Hodge theory

20.5. Pfaffian systems

21. Graph theory

21.1. Hypergrpahs

21.2. Ramsey Theory

21.3. Random Graphs

22. Group theory and generalizations

22.1. Braid groups, Artin groups

22.2. Conjugacy classes

22.3. Fuzzy groups

22.4. Nilpotent groups

22.5. Orthodox semigroups

23. Harmonic analyisis on euclidean spaces

23.1. Conjugate functions

23.2. Convolution factorization

23.3. Fourier series

23.4. Harmonic analysis

24. Integral equations

24.1. Eigenvalue problems

24.2. Fredholm integral equations

24.3. Integro-ordinary differential equations

24.4. Integro-partial differential equations

24.5. Volterra integral equations

25. Integral transforms operational Calculus

25.1. Laplace transform

25.2. Radon transform

26. K-theory

26.1. Steinberg groups and k2

26.2. Whitehead groups and K1

27. Linear and multi linear algebra; matrix theory

27.1. Clifford algebras, spinors

27.2. Fuzzy Matrices

27.3. Hermitian, skew-hermitian

28. Manifolds and cell complexes

28.1. Cobordism and concordance

28.2. Diffeomorphisms

28.3. Differential topology

28.4. Flatness and tameness

28.5. Isotopy and pseudo-isotopy

28.6. PL-topology

29. Mathematical Logic and Foundations

29.1. Algebraic logic

29.2. Axiom of choice and related propositions

29.3. Computability and recursion theory

29.4. Fuzzy set theory

29.5. Godel numberings and issues of incompleteness

29.6. Lukasiewicz and post algebras

30. Measure and integration

30.1. Fractals

30.2. Fuzzy measure theory

31. Non associative rings and algebras

31.1. Color lie Algebra

31.2. Graded lie algebra

31.3. Leibniz algebra

31.4. Modular lie algebra

31.5. Vertex operators

32. Number Theory

32.1. Automorphism groups of lattices

32.2. Bell and Stirling numbers

32.3. Bernoulli and Euler number and polynomials

32.4. Bilinear and hermitian forms

32.5. Binomial coefficients; factorials; q-identities

32.6. Dedekind eta function dedeking sums

32.7. Diophantine inequalities

32.8. Fibonacci and Lucas number and polynomials and generalizations

32.9. Galois cohomology of linear algebraic groups

32.10. Galois Theory

32.11. Hecke-petersson operators

32.12. Jacobi forms

32.13. K-theory of quadratic and Hermitian forms

32.14. Non convex bodies

32.15. Nonholomorphic modular forms

32.16. The frobenius problem

32.17. Thue-Mahler equations

32.18. Weil representation

33. Numberical Analysis

33.1. Monte carlo methods

33.2. Numerical differentiation

33.3. Numerical integration

33.4. Numerical Linear Algebra

33.5. Smoothing, curve fitting

33.6. Splines

33.7. Stiff equations

34. Operator theory

34.1. C-semigroups

34.2. Difference operators
34.3. Functional calculus
34.4. Hermitian and normal operators
34.5. Ill-posed problems
34.6. Jacobi operators
34.7. Kernel operators
34.8. Markov semigroups
34.9. Perturbation theory
34.10. Random operators
34.11. Riesz operators
34.12. Spectral operators

35. Order, lattices, ordered algebraic structures

35.1. Fuzzy lattices
35.2. Modular Lattices, Complemented Lattices
35.3. Noether Lattices
35.4. Stein manifolds

36. Ordinary differential equations

36.1. Fuzzy differential equations
36.2. Lattice differential equations
36.3. Spectral theory
36.4. Weyl theory

37. Partial differential equations

37.1. Boltzmann equations
37.2. Close-to-elliptic equations and systems
37.3. Hamilton-jacobi equations
37.4. Nonlinear elliptic equations
37.5. Overdetermined systems
37.6. Schrodinger operator

37.7. Singular elliptic equations

37.8. Soliton solutions

37.9. Strong solutions

37.10. Weak solutions

38. Potential theory

38.1. Axiomatic potential theory

38.2. Dirichlet spaces

39. Probability theory

39.1. Combinatorial probability

39.2. Fuzzy probability

39.3. Geometric probability

39.4. Limit theorems

39.5. Markov processes

39.6. Stochastic analysis

39.7. Stochastic processes

40. Real functions

40.1. Fuzzy real analysis

40.2. Lipschitz classes

40.3. Quasi-analytic functions

41. Sequences series summability

41.1. Lacunary inversion theorems

41.2. Tauberian constants

42. Several complex varaibles and analytic spaces

42.1. Automorphic forms

42.2. Geometric convexity

42.3. Holomorphic convexity

42.4. Kahler manifolds

42.5. Lelong numbers

42.6. Milnor fibration

42.7. Pseudoholomorphic

42.8. Semi-Analytical sets

42.9. Twistor theory, double fibrations

43. Special functions

43.1. Airy functions

44. Statistics

44.1. Decision theory

44.2. Distribution theory

44.3. Linear regression

44.4. Multivariate analysis

44.5. Parametric inference

45. Topological groups, lie groups

45.1. Ergodic theory

45.2. Infinite-dimensional lie groups