An Analysis, Consideration and Development in Software Engineering

Shweta Ohri
Jagannath International Management School, Vasant Kunj, New Delhi-70 India

Abstract

This paper presents a module and automated method for assessment and trends analysis in software engineering compared with the prior studies. To achieve a more reasonable evaluation result, we take into consideration more high-quality publications, the rank of each publication analyzed, and the different roles of authors named on each paper in question. According to the few papers published during 2008 to 2013, the statistics of research subjects roughly follow power laws, implying the interesting effect. We believe that our results would provide a valuable insight for young scholars and graduate students to seek possible potential collaborators and grasp the popular research topics in software engineering.

Keywords: Systems and Software Engineering, Assessment, Trends Analysis, Research Publications, Power Law

I. INTRODUCTION

Scientific research is a primary mechanism by which a discipline (or a field) attempts to accomplish its advances. In order to better understand where the discipline (or field) in question has been, and to consider where it may be going, the analysis of research conducted within it has been widely recognized as a reasonable and feasible method [1], mainly including assessment and trend analysis.

For a specific discipline (or research field), such a method presents its history and current status and predicts future directions through the statistics on a large number of papers published in peer-reviewed journals, which provides various audiences with important reference for different purposes. For example, an assessment of scholars, institutions and countries (or regions) is valuable to evaluate the performance of research institutions and their scholars in a quantitative and comprehensive way [2], while the trend analysis for a certain research field is of importance to those newcomers who are seeking for future research directions and possible collaborative research opportunities [3].

Software engineering is a relatively new research field derived from computer science. Over six decades, from 1948 until today, its importance has been widely recognized by more and more scholars within the field of computing, and it becomes an active and promising subdivision of the computing field. Like other disciplines, such as cancer [4], agriculture [5] and geographic information system [6], the assessment and trend analysis have long been applied to software engineering [7], but there are several problems that remain unsolved [2]. These are:

- Since there are only seven journals selected as the result of a survey at most, the size of samples (i.e., the number of referred papers published in these journals) is small, implying that the results may be one-sided.
- Because the keywords analyzed were collected from the Top 15 scholars to best describe their research focus, they are likely to be subjective and biased, which may not be used to reasonably reflect the trends and hot topics in software engineering.
- The scoring schemes for leading scholars and institutions were designed using the evaluation rule proposed in [7], which overlooks the leadership role of few scholars among all authors of a multiple-authored paper.

To the best of our knowledge, the latest paper of the annual survey of publications in systems and software engineering from 1994 hasn’t been published till now, despite few of reports on the subdivisions of software engineering such as agile software development [8]. Thus, the main goal of this paper is twofold: on one hand, we will present a new assessment of scholars, institutions and countries (or regions) in software engineering from 2008 to 2013, as well as a survey of trend analysis of this field over the past six years; on the other hand, a more reasonable and general method for assessment and trend analysis, which overcomes the above-mentioned existing problems in prior studies, will be proposed to accomplish such a study with more publications than ever before.

Furthermore, it is worth to note that the study in this paper is actually based on empirical evidence, that is, the results may rely mainly on the data analyzed. To reduce data errors and ensure the repeatability of our results, we selected 24 prestigious journals and 12 famous international conferences (research track) in systems and software engineering, and obtained author list, institution list, keyword list and other information of each paper under discussion from the Elsevier EI (Engineering Village) Compendex database1.

In summary, the main contributions of this study are described as follows.

1) In addition to an assessment of software engineering scholars and institutions (2008-2013), this paper also presented two interesting results, namely, noteworthy scholars and institutions in a specific research field as well as popular research trends, based on a large sample of 7638 research papers published in 36 different publications during this period.
2) In consideration of the different roles of authors of a paper, this paper proposed a new evaluation rule for scholars, institutions and countries (or regions) based on the one [7] frequently used in prior studies; furthermore, our method for assessment and trend analysis was implemented by a software program, thus leading to automated data processing rather than manual operation.

3) We found that the distributions of scholars, institutions, countries (or regions) and keywords roughly followed power laws in terms of their corresponding scores, and theoretically proved that small data errors (e.g., few of papers are missing) of our method have hardly any impact on the Top 15 (or even 20) ranking results.

II. STATISTICS OF EXPERIMENTAL RESULTS

After performing the whole process of our method, we collected 7638 EI paper records, which contain more than 14 thousand authors, more than four thousand six hundred institutions, about 200 countries (or regions), and more than 6 thousand keywords. For the keywords, we further classified these keywords obtained into two types, namely, macro-keyword and micro-keyword. The macro-keywords represent the standard keywords defined by the EI Compendex database, e.g., classification terms, which reflect macro-level research subfields in software engineering. The micro-keywords denote those user-defined or uncontrolled keywords, which imply micro-level research topics.

In this paper, we utilized four frequently-used functions, namely, exponential function, polynomial function, logarithmic function and power function [32], to fit the curves of the scores of sorted scholars, institutions, countries/regions and keywords. As shown in Table 2, the distributions of the scores of experimental subjects except macro-keyword are best described by power laws, suggesting that only a few of leading scholars or institutions do receive much higher scores than those in the long tail. The finding implies that the Matthew effect also exists in software engineering research, and it highlights the assessment of top scholars and institutions as well as the trend analysis in this field.

III. CONCLUSIONS

As we know, the assessment of scientific research is not a simple job. It is very hard to reach a widely recognized evaluation method for such a study. Although software engineering is a young discipline, the prior studies on the assessment of scholars and institutions have been reported. This paper presents a software-aided method for assessment and trend analysis, which can be used in software engineering as well as other research fields in computer science (or other disciplines).

The method proposed in this paper is modular and automated compared with the method in prior studies [7, 10-22, 2]. Besides, it takes into consideration more publications (including conference proceedings), the rank of each publication analyzed, and the different roles of authors in accomplishing a paper. According to the unified data source of the EI Compendex database, this paper presents two levels of research trend changes and those noteworthy scholars and institutions in a given research field, in addition to the assessment of scholars, institutions and countries/regions. Hence, we believe that the results could provide useful guidance on the selection of appropriate potential advisors or collaborators and the popular research topics in software engineering for newcomers or young scholars.

Our future work will focus primarily on applying this method to other research fields in computer science or other disciplines. On the other hand, we will improve the method with the feedback from randomly selected scholars involved in questionnaire surveys.

REFERENCES