Online Information Review

Research impact of general and funded papers: A citation analysis of two ACM international conference proceeding series

Cheng-Che Shen Ya-Han Hu Wei-Chao Lin Chih-Fong Tsai Shih-Wen Ke

Received 2 August 2015
Revised 11 January 2016
Accepted 15 January 2016

Research impact of general and funded papers

A citation analysis of two ACM international conference proceeding series

Cheng-Che Shen
Department of Information Management, National Chung Cheng University, Chiayi, Taiwan and
Department of Psychiatry, Taichung Veterans General Hospital, Chiayi, Taiwan
Ya-Han Hu
Department of Information Management, National Chung Cheng University, Chiayi, Taiwan
Wei-Chao Lin
Department of Computer Science and Information Engineering, Hwa Hsia University of Technology, New Taipei City, Taiwan
Chih-Fong Tsai
Department of Information Management, National Central University, Jhongli, Taiwan, and
Shih-Wen Ke
Department of Information and Computer Engineering, Chung Yuan Christian University, Zhongli, Taiwan

Abstract

Purpose – The purpose of this paper is to focus on examining the research impact of papers written with and without funding. Specifically, the citation analysis method is used to compare the general and funded papers published in two leading international conferences, which are ACM SIGIR and ACM SIGKDD.
Design/methodology/approach – The authors investigate the number of general and funded papers to see whether the number of funded papers is larger than the number of general papers. In addition, the total citations and the number of highly cited papers with and without funding are also compared.
Findings – The analysis results of ACM SIGIR papers show that in most cases the number of funded papers is larger than the number of general papers. Moreover, the total captions, the average number of citations per paper, and the number of highly cited papers all reveal the superiority of funded papers over general papers. However, the findings are somewhat different for the ACM SIGKDD papers. This may be because ACM SIGIR began much earlier than ACM SIGKDD, which relates to the maturity of the research problems addressed in these two conferences.
Originality/value – The value of this paper is the first attempt at examining the research impact of general and funded research papers by the citation analysis method. The research impact of other research areas can be further investigated by other analysis methods.

Keywords Citation analysis, Highly cited papers, Research impact, Research funding
Paper type Research paper

1. Introduction

Generally, publication of the research outcome in a suitable outlet is one of the final stages of conducting a research project. A research project may obtain financial support in a specific country depending on some relevant criteria, such as domain

Online Information Review
Vol. 40 No. 4, 2016
pp. 472-480

1468-4527
DOI 10.1108/OIR-08-2015-0249

importance, potential contribution, and so on. Different countries provide specific research funding sources for applications, e.g., the National Science Foundation in the USA and the European Research Council in Europe.

Since research funding is always limited, it may be decided, after a review process by relevant reviewers, that some research projects will not be granted support. Alternatively, some researchers that conduct research projects and get their research results published may not have even applied for research grants. Thus, published research papers can be broadly classified into the papers with and without financial support.

As the research papers published in the most prestigious international journals and presented at the most important conferences can be assumed to have high impact, which heavily affects future research, this raises the main question addressed this paper: is there a difference in the research impact of general and funded research papers? This is an interesting question because from the viewpoint of financial sponsors, understanding the return on investment (ROI) of the funded research projects is important and affects how they provide grants for future projects. Note that the word "return" here can be regarded as equivalent to the "research impact." In other words, one of the criteria for determining the success of financial support for a research project could be whether the research results published in leading journals and/or conferences have a higher impact than general papers without funding.

On the other hand, the authors of non-funded research projects (including ones that have been rejected) might like to know whether there is any difference in impact between their research projects and the funded projects. The results of the analysis might help them to review their current research and decide on future research directions.

The citation analysis method is used to analyze the research impact of published papers (Katerattanakul *et al.*, 2003). The research impact is calculated based on the number of citations papers have received in other publications. Generally speaking, the papers having a larger number of citations have a higher research impact than the papers with few citations in the same subject areas.

In citation analysis, the journal impact factor (Garfield, 2006) and h-index (Hirsch, 2005) are two widely used methods. The impact factor is a metric used to measure the influence or impact of journals in various subject areas, and it is calculated based on the average number of citations to recent articles published in that journal. On the other hand, the h-index is used as a measure of both the productivity and impact of the published work of a scientist or scholar. It is calculated based on the set of the individual's most frequently cited papers and the number of citations that they have received in other publications.

In the literature, many studies have used these methods to analyze researchers' performance (Clarke, 2009; Cronin and Meho, 2006; Hu and Chen, 2011) and journals' impact and quality (Haddow and Genoni, 2010; Mingers *et al.*, 2012; Vanclay, 2008). However, very few have used the citation analysis method to analyze the research impact of general and funded papers and examine the differences in number of citations between them.

Recently, Costas and van Leeuwen (2012) present some statistical analyses about the distributions of publications with funding acknowledgments among disciplines, countries, and document types. In addition, the research impacts of publications with and without funding acknowledgment are studied, but they only use the impact factors of the journal publications as the evaluation tool. On the other hand, Wang and Shapira (2011, 2015) studied whether the publications with funding acknowledgments have higher impacts than the ones without funding acknowledgments. However, they only

focus on the field of nanotechnology and the collected data are out of date, which is between August 2008 and July 2009.

The aim of this study is to collect information on papers published in two leading international conferences of computer science between 2011 and 2013, namely, ACM SIGIR and ACM SIGKDD, and to compare and analyze the number of papers with and without funding, the number of citations of general and funded papers, and the contribution distribution of the funded papers as well as their corresponding citation statistics.

The rest of this paper is organized as follows. Section 2 describes the research methodology and data collected for the citation analysis. Section 3 presents and discusses the analysis results. Finally, Section 4 concludes the paper.

2. Methodology and data

We collected the data from two prestigious international conferences held between 2011 and 2013, ACM SIGIR[1] and ACM SIGKDD[2]. These are the leading international conferences in the information retrieval and data mining fields and their publication acceptance rates in recent years have usually been lower than 20 percent. It should be noted that although these conferences publish regular and short (or poster) papers, only regular papers are considered in this study for citation analysis.

The Publish or Perish software[3] was used to determine each paper's citations. Moreover, the papers that are highly cited can also be identified based on the g-index proposed by Egghe (2006). In addition, whether each published paper has been prepared with or without financial support is also recorded, and the supporting countries are also identified.

After the data were collected, including citation statistics, funding information, and supporting countries, four hypotheses and one research question are formulated as follows:

H1. The number of funded papers published in ACM SIGIR and ACM SIGKDD is larger than the number of general papers.

Papers published in these two leading conferences should be very high quality and most of them should be financially supported because the importance and potential contribution of the research proposals has been reviewed by specific sponsors (cf. Section 3.1):

H2. The highly cited papers published in ACM SIGIR and ACM SIGKDD are mostly funded papers.

Similar to the proportions of the general and funded papers, the number of highly cited papers with funding should be larger than those of general papers without funding (cf. Section 3.2.1):

H3. The total citations of funded papers are larger than those of general papers in ACM SIGIR and ACM SIGKDD.

According to the first hypothesis, that the number of funded papers should be larger than the number of general papers, the total number of citations of funded papers should also be larger than those of general papers (cf. Section 3.2.2):

H4. The average number of citations per funded paper is larger than the average number of citations per general paper in ACM SIGIR and ACM SIGKDD.

Since the importance, potential contribution, etc., of funded papers are generally higher than general papers, the average number of citations of each funded paper should be larger than for each general paper (cf. Section 3.3.2).

What is the distribution by country for the number of funded papers, the number of highly cited papers, and the average number of citations per funded paper?

The answer to this question allows us to understand which supporting countries have a larger number of funded papers and highly cited papers as well as average citations per funded paper. This can show the ROI of the research projects funded by different countries (cf. Section 3.2.2).

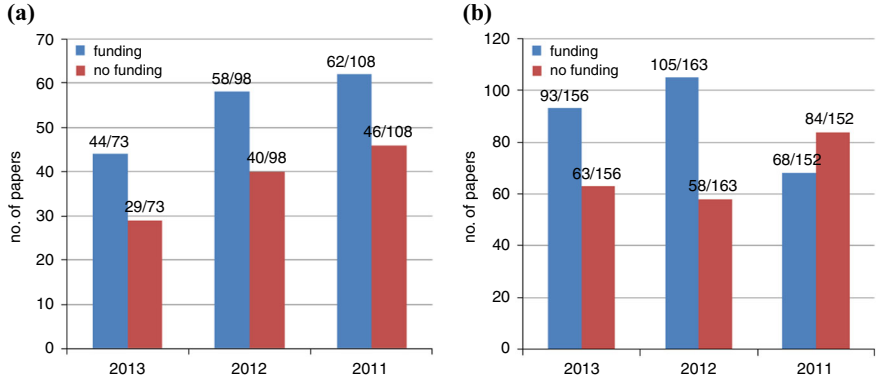
3. Results and discussion

3.1 Number of papers with and without funding

Figure 1 shows the number of papers with and without funding published in ACM SIGIR and ACM SIGKDD. As we can see, more funded papers are published in both conferences than general papers, except for ACM SIGKDD 2011. In other words, the number of funded papers published in ACM SIGIR and ACM SIGKDD is significantly larger than the number of general papers ($p < 0.01$), except for ACM SIGKDD 2011. Therefore, this result mostly proves the first hypothesis. In other words, the results of research projects having some financial support usually have a higher chance of being published in the leading conferences than do the results of general projects without funding.

Specifically, in ACM SIGIR, for the period from 2011 to 2013, the proportion of published papers with funding shows a gradual increase, with the numbers being 57.4 percent (2011), to 59.2 percent (2012), and 60.3 percent (2013). However, papers published in ACM SIGKDD do not show this trend, with the numbers being 44.7 percent (2011), 64.4 percent (2012), and 58.5 percent (2013).

One possible reason for this result may be because ACM SIGIR began in 1978 whereas ACM SIGIR began in 1998, which means that the research problems covered in ACM SIGIR have been studied for a longer time than those covered by ACM SIGKDD. In other words, ACM SIGIR focusses on more mature problems than ACM SIGKDD does. In this case, the research projects to be funded should show a certain level of potential contribution to the information retrieval field from the project reviewers' viewpoint. As a result, research results with some financial support are more likely to be recognized by the ACM SIGIR committee as productive of high-quality papers for publication. Therefore, the likelihood of the acceptance of research results for general research projects without funding to be accepted for publication in some conferences started more recently. However, further studies should be conducted to demonstrate this issue in the future.



Notes: (a) ACM SIGIR; (b) ACM SIGKDD

Figure 1. Yearly distribution of the numbers of papers with and without funding

3.2 Citation analysis results

3.2.1 The number of highly cited papers. Figure 2 shows the number of highly cited papers with and without funding from ACM SIGR and ACM SIGKDD. These numbers demonstrate that the research impact is generally higher for most funded projects than most general projects without funding. In other words, the number of funded papers that are highly cited and published in ACM SIGIR and ACM SIGKDD is significantly larger than the number of general papers ($p < 0.01$). Therefore, this finding proves the second hypothesis, that most highly cited papers are based on funded research projects. However, this does not necessarily mean that general research projects cannot provide high research impact in the future.

3.2.2 Total citations and average citations per paper. Figure 3 shows the citation numbers of general and funded papers published in ACM SIGIR and ACM SIGKDD. Since the number of funded papers is larger than the number of general ones and most highly cited papers are funded, there is no doubt that the citation number of funded papers is larger than that of general papers in both conferences, with a high level of significant difference ($p < 0.01$). This proves the third hypothesis, that the total number of citations of funded papers is larger than those of general papers in ACM SIGIR and ACM SIGKDD.

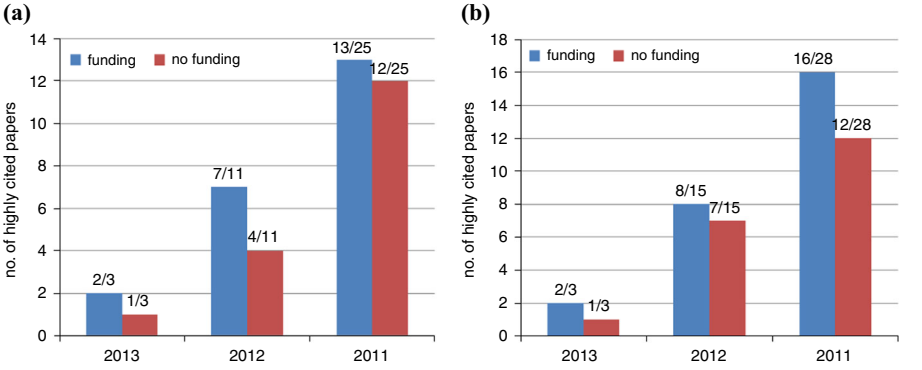


Figure 2.
Yearly distribution of the number of highly cited papers with and without funding

Notes: (a) ACM SIGIR; (b) ACM SIGKDD

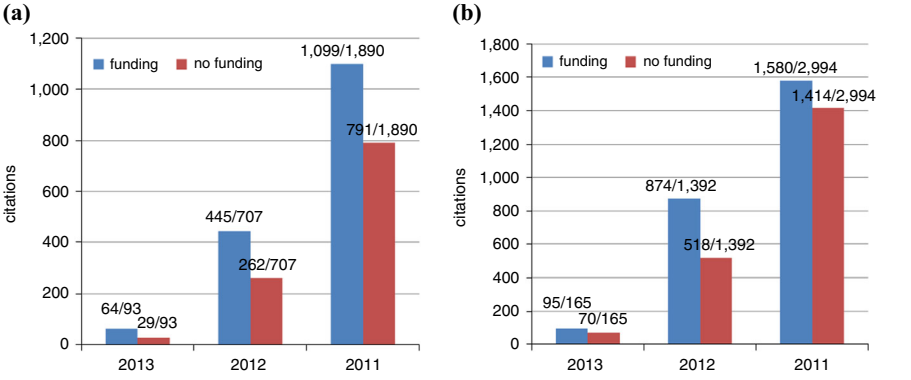


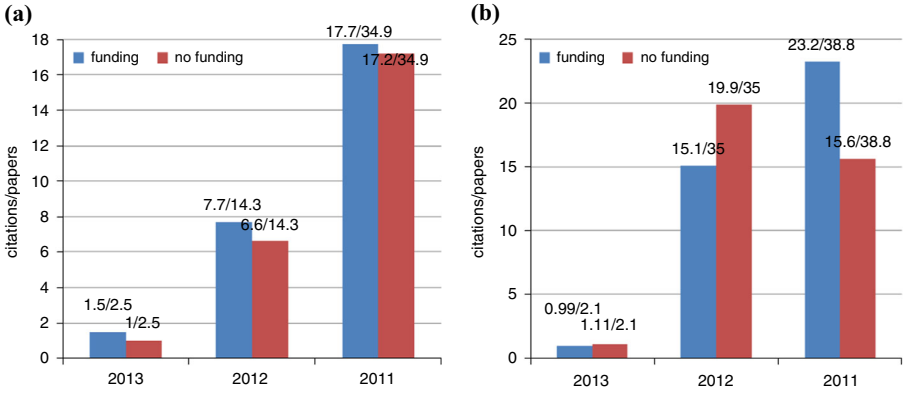
Figure 3.
Yearly distribution of the citation numbers of general and funded papers

Notes: (a) ACM SIGIR; (b) ACM SIGKDD

Furthermore, Figure 4 shows the average citation numbers of general and funded papers. In ACM SIGIR, the differences of average citations from 2011 to 2013 between general and funded papers (i.e. citations/paper) are significant ($p < 0.01$). However, the differences are only between 0.5 and 1.1 cites/paper. This means that even the general papers without funding still have some research impact in the information retrieval field.

For ACM SIGKDD, the results for 2011-2013 are different. Specifically, in the most recent two years (i.e. 2013 and 2012) the average citation numbers have been higher for general papers than funded papers. This result indicates that depending on the research field, funded papers do not always have a higher citation impact than general papers. Therefore, the fourth hypothesis, that the average number of citations of a funded paper is larger than those of a general paper is only proven for ACM SIGIR.

3.2.3 Distribution by country. Tables I and II list the distribution by country for the number of funded papers and number of highly cited papers published in ACM SIGIR and ACM SIGKDD, respectively. Note that we only report the top seven supporting



Notes: (a) ACM SIGIR; (b) ACM SIGKDD

Figure 4. Yearly distribution of the average number of citations for each general and funded paper

	China	Germany	The Netherlands	Singapore	Spain	UK	USA
2013 (73)	8 (0)	1 (0)	3 (0)	4 (0)	1 (1)	1 (0)	17 (0)
2012 (98)	11 (1)	3 (1)	3 (1)	2 (0)	7 (0)	3 (0)	17 (1)
2011 (108)	15 (4)	1 (0)	2 (1)	3 (0)	2 (0)	1 (0)	28 (9)

Note: Distribution by country for the number of funded ACM SIGIR papers (number of highly cited papers)

Table I. The number of funded ACM SIGIR papers (number of highly cited papers)

	Canada	China	Germany	Japan	Singapore	Spain	Taiwan	USA
2013 (156)	2 (0)	16 (0)	0 (0)	2 (0)	5 (0)	2 (0)	4 (0)	58 (1)
2012 (163)	1 (1)	19 (1)	6 (0)	1 (0)	2 (0)	2 (0)	2 (0)	64 (6)
2011 (152)	1 (1)	6 (2)	2 (0)	1 (0)	1 (0)	1 (0)	2 (1)	41 (8)

Notes: Distribution by country for the number of funded ACM SIGKDD papers (number of highly cited papers)

Table II. The number of funded ACM SIGKDD papers (number of highly cited papers)

countries that have relatively larger numbers of published papers. In addition, the numbers in brackets followed by the year indicate the total numbers of funded papers. For the numbers in other brackets, they indicate the total numbers of highly cited papers.

These results can reveal which supporting countries send large proportions of the funded papers published in leading conferences. Specifically, the USA and China are the top two supporting countries. In addition, Tables I and II indicate that the supporting countries with large numbers of funded papers published in leading conferences are likely to have more highly cited papers.

However, if we examine the average number of citations of each funded paper for each supporting country, some interesting results are found. Tables III and IV list the contributions for the average number of citations of funded papers published in ACM SIGIR and ACM SIGKDD, respectively. Note that the numbers in brackets followed by the year indicate the baseline average citations per funded paper (cf. Figure 4), and numbers larger than the baseline are italic.

As we can see, there is no direct relationship between the top supporting countries (such as the USA and China) and the higher research impact of their funded papers. In other words, although the top supporting countries have larger numbers of funded papers, including highly cited papers, the average number of citations of each funded paper from the top supporting countries is not always higher than the baseline average citations per funded paper. In particular, the Netherlands and Canada have the most effective ROI for ACM SIGIR and ACM SIGKDD, respectively, since although their number of funded papers is relatively small, the average number of citations per funded paper is relatively large.

4. Discussion

According to the results shown in Section 3, we can observe that there are generally more published papers, which were funded than the ones without funding in the two prestigious international computer science conferences, i.e. ACM SIGIR and ACM SIGKDD. In other words, the research projects having some financial support could make larger contributions than the ones without funding, which are likely to be published in these leading conferences.

Specifically, more funded papers are highly cited than general papers. This indicates that the funded papers are likely to have higher research impacts than general papers.

Table III.
The average number of citations of funded ACM SIGIR papers

	China	Germany	The Netherlands	Singapore	Spain	UK	USA
2013 (1.5)	*1.5*	1	*1.67*	1.3	*7*	0	1.3
2012 (7.7)	6.6	*10.3*	*11.7*	3	5.4	6	6.4
2011 (17.7)	*17.7*	15	*29.5*	8.3	15.5	*18*	*20.5*

Note: Distribution by country for the average number of citations of funded ACM SIGIR papers

Table IV.
The average number of citations of funded ACM SIGKDD papers

	Canada	China	Germany	Japan	Singapore	Spain	Taiwan	USA
2013 (0.99)	*1.5*	*1.6*		0.5	0.2	*1*	*1*	0.8
2012 (15.1)	*25*	7.7	3.5	7	6	10.5	8	9.1
2011 (23.2)	*62*	21	17	5	*27*	21	*24.5*	23

Note: Distribution by country for the average number of citations of funded ACM SIGKDD papers

However, for ACM SIGKDD 2012 and 2013, we found that the average citations of each general paper are larger than the one of each funded paper. Therefore, funded papers do not always have a higher citation impact than general papers depending on the research field.

Moreover, a large proportion of the funded papers published in these leading conferences are supported by the USA and China. We also can observe that the top supporting countries with large numbers of funded papers are likely to have more highly cited papers. However, if we look at the average citations per funded paper supported by the USA and China as the top two supporting countries, they are not always higher than the baseline average citations per funded paper. In other words, although the Netherlands and Canada have relatively smaller numbers of funded papers published in ACM SIGIR and ACM SIGKDD than the USA and China do, their average numbers of citations per funded paper are relatively large. This indicates that the Netherlands and Canada have the most effective ROI among all of the top supporting countries.

5. Conclusion

This study conducts a citation analysis of the papers with and without funding published in two leading international conferences from 2011 to 2013, which are ACM SIGIR and ACM SIGKDD. Several findings can be obtained from the analysis results. First, in most cases, the number of funded papers published for ACM SIGIR and ACM SIGKDD is larger than the number of general papers. Second, among the highly cited papers, the number of funded papers is larger than those of general papers, for both conferences. Third, related to the above findings, the research impact (i.e. total number of citations) of funded papers is larger than that of general papers. Fourth, the average number of citations of a funded paper is not always larger than those of a general paper, which is the case in ACM SIGKDD. Finally, the USA and China are the top two supporting courtiers with larger numbers of funded papers published at both conferences. However, the Netherlands and Canada are the top two supporting countries having a higher average number of citations per paper than the baseline.

Although these findings are interesting and provide some contribution to relevant knowledge, this study is only a first attempt at examining the research impact of general and funded research papers by the citation analysis method. More specific and detailed areas can be further examined and their research impact compared by other analysis methods.

Notes

1. www.sigir.org/
2. www.sigkdd.org/
3. www.harzing.com/pop.htm#download (calculation date: between May 12, 2015 and May 14, 2015).

References

Clarke, R. (2009), "A citation analysis of Australian information systems researchers: towards a new era?", *Australasian Journal of Information Systems*, Vol. 15 No. 2, pp. 23-44.

Costas, R. and van Leeuwen, T. (2012), "Approaching the 'reward triangle': general analysis of the presence of funding acknowledgements and 'peer interactive communication' in scientific publications", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 8, pp. 1647-1661.

Cronin, B. and Meho, L. (2006), "Using the h-index to rank influential information scientists", *Journal of the American Society for Information Science and Technology*, Vol. 57 No. 9, pp. 1275-1278.

Egghe, L. (2006), "Theory and practice of the g-index", *Scientometrics*, Vol. 69 No. 1, pp. 131-152.

Garfield, E. (2006), "The history and meaning of the journal impact factor", *Journal of the American Medical Association*, Vol. 295 No. 1, pp. 90-93.

Haddow, G. and Genoni, P. (2010), "Citation analysis and peer ranking of Australian social science journals", *Scientometrics*, Vol. 85 No. 2, pp. 471-487.

Hirsch, J.E. (2005), "An index to quantify an individual's scientific research output", *Proceedings of the National Academy of Sciences*, Vol. 102 No. 46, pp. 16569-16572.

Hu, P.J. and Chen, H. (2011), "Analyzing information systems researchers' productivity and impacts: a perspective on the h index", *ACM Transactions on Management Information Systems*, Vol. 2 No. 2, Article No. 7.

Katerattanakul, P., Han, B. and Hong, S. (2003), "Objective quality ranking of computing journals", *Communications of the ACM*, Vol. 46 No. 10, pp. 111-114.

Mingers, J., Macri, F. and Petrovici, D. (2012), "Using the h-index to measure the quality of journals in the field of business and management", *Information Processing & Management*, Vol. 48 No. 2, pp. 234-241.

Vanclay, J.K. (2008), "Ranking forestry journals using the h-index", *Journal of Informetrics*, Vol. 2 No. 4, pp. 326-334.

Wang, J. and Shapira, P. (2011), "Funding acknowledgment analysis – an enhanced tool to investigate research sponsorship impacts: the case of nanotechnology", *Scientometrics*, Vol. 87 No. 3, pp. 563-586.

Wang, J. and Shapira, P. (2015), "Is there a relationship between research sponsorship and publication impact? An analysis of funding acknowledgements in nanotechnology papers", *PLoS ONE*, Vol. 10 No. 2, pp. 1-19.

About the authors

Cheng-Che Shen is a PhD Student at the National Chung Cheng University and a Physician of Psychiatry Department, Chiayi branch, Taichung Veteran General Hospital, Taiwan. His research interest include, psychiatry, medical informatics, and medical data mining. His research has appeared in *Neurology*, *PLoS ONE*, and *Oncology*.

Ya-Han Hu is an Associate Professor at the Department of Information Management, National Chung Cheng University, Taiwan. His research interests cover data mining and its applications.

Wei-Chao Lin is an Assistant Professor at the Department of Computer Science and Information Engineering, Hwa Hsia University of Technology, Taiwan. His research interests are machine learning and artificial intelligence applications.

Chih-Fong Tsai is a Professor at the Department of Information Management, National Central University, Taiwan. His current research focusses on multimedia information retrieval and data mining. Chih-Fong Tsai is the corresponding author and can be contacted at: cftsai@mgt.ncu.edu.tw

Shih-Wen Ke is an Assistant Professor at the Department of Information and Computer Engineering, Chung Yuan Christian University, Taiwan. His research focusses on information retrieval, text mining, and e-learning.