Construction of SCI publications information system for statistic

Xie Wu[1], Huimin Zhang[2], Jingbo Jiang[1]

[1] School of Computer Science and Engineering, Guilin University of Electronic Technology, Guilin, 541004, China

[2] School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, 541004, China

iip2012_xiewu@126.com

Abstract. There are over 8000 SCI (Science Citation Index) publications in the ISI (Institute for Scientific Information) Web of Knowledge database system. However, the publications are too many and it is difficult for new authors to choose the most suitable journals or periodicals to submit their research fruits of high level. So, some valuable information about SCI publications is collected, and the corresponding database is established. The records from this database are classified and counted. The statistical results show that the SCI publications information system is helpful to authors to issue papers.

Keywords. Database system, Science citation index, Publications, Statistics.

1   Introduction

ISI (Institute for Scientific Information) issues or updates periodically the information of SCI (Science Citation Index) publications and JCR (Journal Citation Reports) every year. The records of SCI journals and periodicals are more than 8000, and it is a very interest project to feel for some rules about SCI publications. Jacsó and Péter researched the ISI Web of Science database from three configurations of the h-index, h-core citation rate and the bibliometric traits [1]. Zhou Ping and Leydesdorff Loet compared the Chinese scientific and technical papers and citations database and the SCI data by journal hierarchies and interjournal citation relations [2]. Meho Lokman I gave a comparison of scopus and web of science about citation counting, citation ranking, and h-index of human-computer interaction researchers [3]. Li Jinfeng and his colleagues analyzed the trends on global climate change based on SCIE (Science Citation Index Expanded) [4]. Ball Rafael made a bibliometric comparison between two citation databases of SCI and SCOPUS [5]. Some fruits about IF (Impact Factor), rank, citation and other aspects of SCI journal and periodicals are gained, and this information may be useful to some new contributors.

However, these fruits are only fit for some special researchers, projects or domains. Because all the SCI periodicals are generally divided into thirteen classes with many subdistricts, and the data of publications vary year by year, it is difficult for diverse

authors, especially for some beginners, to find the most appropriate journal to issue high level research fruits when facing over 8000 SCI publications. The role of IF from ISI is also very limited. The SCI IF threshold values from13 JCR academic subjects are different. So, it is not reasonable to assess the level of SCI publications only in terms of IF by four JCR divisions without some comprehensive evaluation indicators. Moreover, some publications belong to the interdisciplinary scopes.

So, motivated by Ref. [1-8], the SCI periodicals and journals information including publication name, abbreviation, ISSN (International Standard Serial Number) number, organization, nationality, location, IF, Chinese name and discipline was collected, and a SCI publications information system is customized and constructed for statistical analysis. The remainder of this paper is organized as follows. Section 2 deals with the system requirement analysis for many new contributors. The relational tables of the SCI publications information system are designed in Sect 3, and the results of the statistical charts of this database system are shows in Sect 4. Finally, the discussion and conclusion are described in Sect 5.

2 Requirement analysis with modules

Many new authors are often not familiar with the existing ISI database system or its websites, and journal records are too massive for them, while the useful knowledge about their ideal publications is too little. So, the function requirement of this SCI publications information system mainly involves in the best selection from 8000 SCI records from ISI. This system is primarily divided into several modules as follows.

(1) Records selection module.

This module introduces the records from the SCI publications information system in detail. The fuzzy selection function is afforded to the new contributors to gain the publication name, abbreviation, ISSN number, organization, nationality, IF, Chinese name, etc.

(2) Data manipulation module.

It mainly covers the operations including inserting, updating and deleting records information of the SCI publications information system. These manipulations can be carried out through the back platform by DBA (database administration). The system DD (data dictionary) of important tables or attributes can be set up increasingly by the common administrator or super administrator.

(3) Statistical chart module.

It can implement the functions of producing intuitional statistical charts according to the chosen attribute or attribute sets. From these charts, decision-makers can learn the distribution rules by setting the attribute conditions of organization, nationality, IF and subjects, etc.

(4) Users management module.

This system can be visited by each super administrator, common administrator, DBA, registered user and guest, and their authority grades associating with this SCI publications information system decrease orderly. The former can grant or revoke the latter through modifying the permission of creating, modifying and dropping users.

3   System design of relational logistic model

After the detailed demand analysis of publications information system, the logical structure design needs to be followed. The relational logistic model is chosen to store publications information records, and there are several relational tables. These tables are designed and stored in SQL Sever DBMS (database management system). Here, a database named SCIData is created, and all the records of all tables are inserted and updated in SCIData with SQL Sever. These tables are shown as follows.

(1) Publication basic information table.

In order to catch some useful information for new authors, these column attributes including publication name, ISSN, publication location, citation rate are collected together to process the raw information of journals or periodicals. The data item of every attributes can be updated easily every year.

Table 1. Publication basic information table

attribute name	description	data type	constraint
PubNameID	publication ID	Int	primary key
PubFullName	full name	Nvarchar(80)	not null
PubShortName	name for short	Nvarchar(80)	not null
Periodical	full name	Nvarchar(40)	
ISSN	ISSN number	Nvarchar(10)	not null
IF_ID	ID of ISI IF	Int	foreign key
AddressID	address number	Int	foreign key
CountryID	Country number	Int	foreign key
DisciplineID	Discipline ID	Int	foreign key

(2) Country of publication table.

This table is designed to search the distribution rules of SCI publications with many kinds or numbers. For example, the SCI journals from America or European countries are massive, and few SCI publications belong to those countries from Africa or Latin America. On the other hand, SCI publications in English-speaking countries are far more than that of non-English nations. So, the country is an important factor, and it is a foreign key from the publication basic information table.

Table 2. Country of publication table

attribute name	description	data type	constraint
CountryID	Country ID	Int	primary key
CountryName	Country name	Varchar(80)	check
Language	nation language	Varchar(20)	check

(3) Publication issue period table.

The issue periods of 8000 SCI publications can divided into several types, such as annals or yearbook, semiyearly, quarterly, bimonthly, monthly, semimonthly, weekly, tri-annual and daily. This table is helpful to classify the whole SCI journals or peri-

odicals, and it mainly reflects the speed of issuing publications. Authors tend to choose the journals with short periods, yet some important publications with high IF values issue papers for a long time. A new contributor has to make a choice by journal period.

Table 3. Publication issue period table

attribute name	description	data type	constraint
TypeID	TypeID	Int	primary key
Type	Type name	Nvarchar(30)	unique
IssuePeriod	Times in a year	Nvarchar(30)	enumeration
PubNameID	publicationID	Int	foreign key
Irregular	irregular cycle	Nvarchar(30)	
Issueyear	issue year	Datetime	not null

(4) IF table.

This table is one of the most important items to weigh the academic influence of SCI papers and publications. The IF values of every journal or periodical can be gained from ISI JCR every year. It is a very fair and common international rule for all SCI publications.

Table 4. IF table

attribute name	description	data type	constraint
IF_ID	impact factor ID	Int	primary key
IF_Year	IF value by year	Numeric(18, 2)	check
PubNameID	publicationID	Int	foreign key

4   System results with statistical charts

The SCI publications information system for statistic is developed with the tools of Microsoft Visual studio 2010 in Asp.net and C# languages, and all the records are stored in Microsoft SQL Sever 2005. All users can select their records by given attribute or attribute sets. Furthermore, one can carry out fuzzy union selections through several fields or attributes.

The statistical charts can be generated from Tab. 1 to Tab. 4 of the SCI publications information system. For example, according to the publication issue period table, one of the statistical sub-diagrams is shown as Fig. 1.
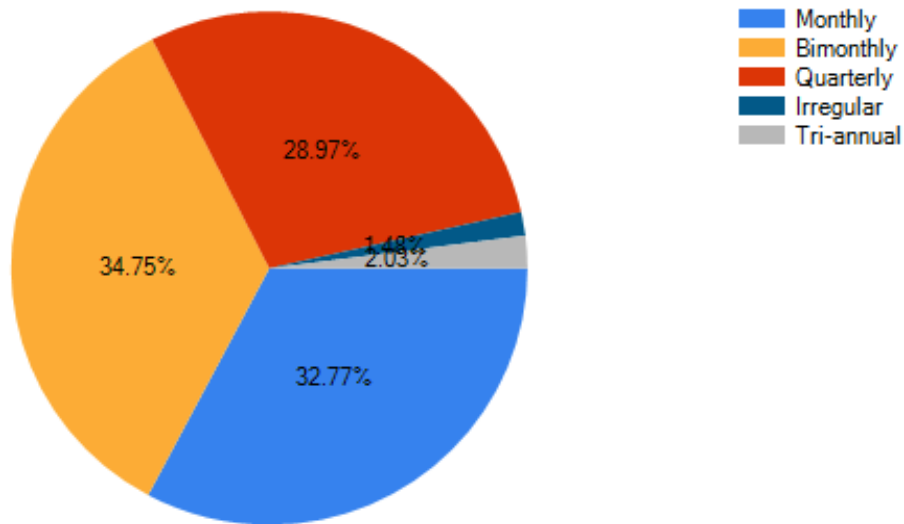
Fig. 1. One of the statistical results by issue period.

In order to find the relationship between the attributions of countries and ISI SCI publications numbers, the above tables can be connected and selected. When the country table is connected to the publication basic information table by CountryID, the histogram is shown as Fig. 2.
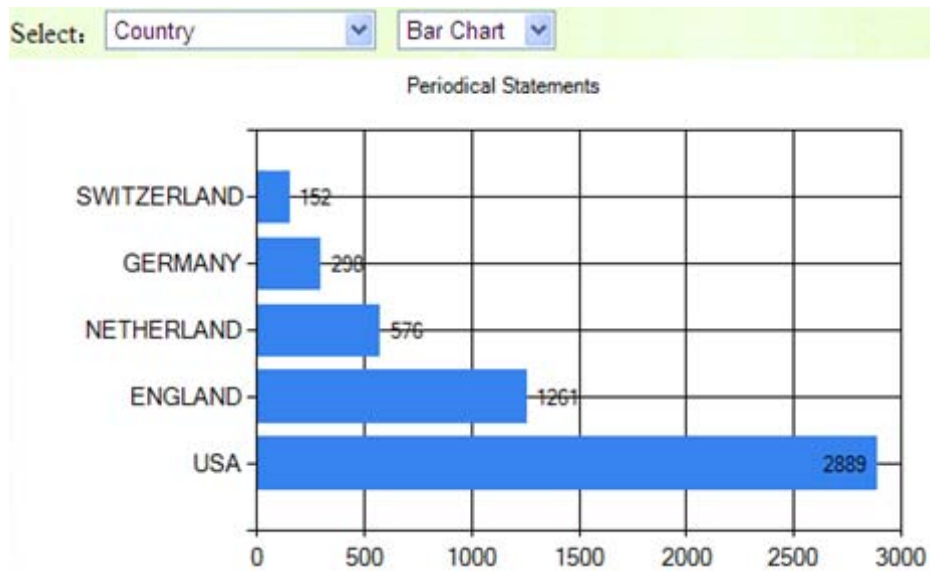

Fig. 2. One of statistical results by country

5 Summary

In this paper, we finish the work of system demand analysis and relational logical model design in detail, and a SCI publications information system is developed with the IDE (Integrated Development Environment) tools of Microsoft Visual Studio 2010 and SQL Server2005. It can be seen that from the results of this system software interfaces and the statistical charts, using this SCI periodical the information system, one can select some useful information by providing several important attributes of publication name, abbreviation, ISSN number, organization, nationality, location, etc. This software is easy to operate with simple and friend interfaces. It is very helpful for many new authors to choose the most appropriate SCI journals or periodicals to submit and issue their research fruits with high efficiency. These statistical charts from the selected data records of relational tables can provide some true and intuitive decision-making information to many new or initial contributors with little subjective assumption.

Acknowledgements

References

1. Jacsó, P.: The H-index, H-core Citation Rate and the Bibliometric Profile of the Web of Science Database in Three Configurations. Online Information Review 35(5), 821-833 (2011)
2. Zhou, P., Leydesdorff, L.: A comparison between the China scientific and technical papers and citations database and the science citation index in terms of journal hierarchies and interjournal citation relations. Journal of the American Society for Information Science and Technology 58(2), 223-236 (2007)
3. Meho, L. I., Rogers, Y.: Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. Journal of the American Society for Information Science and Technology 59(11), 1711-1726 (2008)
4. Li, J.F., Wang, M.H., Ho, Y.S.: Trends in research on global climate change: A Science Citation Index Expanded-based analysis. Global and Planetary Change 77(2), 13-20 (2011)
5. Ball, R., Tunger, D.: Science indicators revisited Science citation index versus SCOPUS: A bibliometric comparison of both citation databases. Information Services and Use 26(4), 293-301 (2006)
6. http://apps.webofknowledge.com
7. Fu, H. Z.: The most frequently cited adsorption research articles in the Science Citation Index (Expanded). Journal of Colloid and Interface Science 379(1), 148-156 (2012)
8. Hung, K. C., Lan S. J., Liu J.T.: Global trend in articles related to stereotactic published in science citation index-expanded. British Journal of Neurosurgery Colloid and Interface Science 26(2), 258-264 (2012)