Research and Design on Auxiliary System for Citation Certificate of Academic Papers

Bin Lu[1], Changyu Liu[2,*], Weifeng Gao[1] and Zhenbang Rong[1]

[1]*School of Computer Science, Wuyi University, Jiangmen 529020, China*
[2]*College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China*
lbscut@gmail.com, [*]yezhich@gmail.com, gwf10@126.com, bondrong@163.com

Abstract

In this paper, considering the special situation of great demand for citation of academic papers in China, we design and implement a special auxiliary system of citation certificate of academic papers. At the beginning, we analyze the primary system requirements and divide the system into four functional modules, which are the EI retrieving and certificating module, the ISI retrieving and certificating module, the query as well as report module, and the original settings restoring module. Next, for the system design, we present details about these modules. Furthermore, we also illustrate the implement of the final system. The system speeds up the relevant staffs' work and improves their efficiency.

Keywords: Auxiliary System, Citation, SCI, EI, C#

1. Introduction

In the world all sorts of academic writings are published every year, and many of them are indexed by all sorts of document retrieving systems, such as Engineering Index (EI), Science Citation Index (SCI). EI is now a world-leading engineering information platform belonging to Elsevier Group Incorporation [1-2]. SCI is founded by Institute for Scientific Information Incorporation (ISI) [3-5], which has other retrieving systems, like ISI Proceedings and Social Sciences Citation Index (SSCI). ISI Proceedings has two important components, including Index to Scientific and Technical Proceedings (ISTP) and Index to Social Sciences & Humanities Proceedings (ISSHP).

At present, research institutions count and analyze the including and citations of their researchers' papers in China, in order to evaluate their research capabilities [6-9]. So, the researchers must provide the included proofs to show their research ability. Most of this work are finished by hand, and there is not an effective way of lighten the workloads of staffs. For that, we bring forward an auxiliary system of citation certificate of academic papers firstly, which is developed by the programming language C#. The auxiliary system speeds up the relevant staffs' work and improves their efficiency. Furthermore, the auxiliary system enhances office automation, and it can do most of work of citation certificate of academic papers. Besides, it also reduces the error rate on a complete manual operation basis.

2. Research Background

In China, SCI, EI and ISTP are the three most famous document retrieving systems [10-11]. So, scientific research personnel will provide included proofs of document retrieving systems, especially SCI, EI and ISTP, when they want to promote to a higher

---

[*] Corresponding Author

rank of special technology, declare or accomplish scientific research projects, and apply for some scholar roles. In addition, a good many master and doctor candidates have to do so when they apply for a scholarship and defend their dissertation. Most of universities and other scientific research institutions also show their scientific research strength by the methods. These are special phenomena only in China, which hardly appear in other countries [10-12]. For example, there is statistical data of published papers of 5 universities of Guangdong province included by the famous databases in the last six years, shown in Table 1. From Table 1, the quantity of the included papers of the 5 universities is huge and increases year by year on the whole. So, citation certificate of academic papers is a heavy workload by hand.

Table 1. Number of Published Papers in Major Databases for five Universities in Guangdong Province during Recent 6 Years

Database	EI Compendex	SCI	ISTP	SSCI
2009	7848	5917	1865	383
2010	9143	6811	1378	489
2011	11559	8352	1612	591
2012	11082	9883	1947	656
2013	12341	10982	1543	770
2014	13206	11554	838	814

In addition, it is very expensive to purchase electronic resources of these document retrieving systems, so there are a few research institutions that can provide the included proofs in China. Furthermore, not all research institutions buy enough resources. The institutions are usually libraries of many universities. As such, citation certificates only from the libraries are convincing. People are inclined to get the included proofs from the libraries of those prestigious comprehensive universities, for the universities have relatively complete resources. In this way relevant staff members are increased pressure and workloads. Moreover, the job of issuing the included proofs is very trivial and tiresome, which is also repetitive work. As far as these staffs are concerned, the heaviest work is that copying the bibliographical descriptions of papers to MS Word templates and setting the formats of bibliographical descriptions. So, decreasing these workloads is the highest requirement.

3. Requirement Analysis

As a key process of software engineering, the requirement analysis is also the basis of software development which faces users directly. The main task of requirement analysis includes describing the functions and performance of a new product deeply, setting the constraints of software design and defining other effective demand. Using requirement analysis technology, a systems analyst can determine users' needs. Then the functions of software systems are determined. For the required auxiliary system, major functions involve searching and extracting related information of papers and an effective statistical report.

Accordingly, the required auxiliary system is divided into four functional modules, which are the EI retrieving and certificating module, the ISI retrieving and certificating module, the workload calculation module and the original settings restoring module. Specifically, for the EI retrieving and certificating module, the system should provide

functions for visiting the website of EI, retrieving the related papers and certificating citations of them. Analogously, for the ISI retrieving and certificating module, the system should provide functions for visiting the website of ISI, retrieving the related papers and certificating their citations including SCI, ISTP, ISSHP and SSCI. For the workload calculation module, the system should provide functions for statistical report of certifications by the operating user, manual backup, manual recovery of backup and setting the path of backup. For the original settings restoring module, as the name implies, the system should provide functions for restoring the original settings due to fault in operation.

4. System Design

After we understand system requirements, we can then model the corresponding business process. Subsequently we can concentrate on the detailed system design, which are customization browser design, the design of classes for papers, the design of Word template and the design of statistical report in this paper.

4.1. Customization Browser Design

On the basis of the process of manual operation, some messages obtained from related web page are imported into the corresponding position of Word files, in which the contents of the text are details of citation certificates. On the other hand, because we cannot access these academic databases of EI and ISI directly, we get related information only by reading webpage contents. There are three methods that can acquire information from remote appointed web page in C#, including the class WebClient, HttpWebRequest/HttpWebResponse and WebBrowser.

WebClient Class provides common methods for sending data to and receiving data from a resource identified by a URI [13]. By WebClient Class, people can invite resources of internet without using browser. But WebClient Class do not support cookies/session, which results in user not controlling automatic redirection of a URI.

HttpWebRequest Class provides an HTTP-specific implementation of the WebRequest class [14]. HttpWebResponse Class provides an HTTP-specific implementation of the WebResponse class [15]. HttpWebRequest Class and HttpWebResponse Class are mostly used to interact with HTTP servers directly, dealing with certificates, coding and cookies/session.

WebBrowser Class enables the user to navigate Web pages inside your form. The WebBrowser control lets you host Web pages and other browser-enabled documents in your Windows Forms applications [16].

In the proposed auxiliary system, we used WebBrowser control in order that users' operative habits are not changed. However, WebBrower control loads IE 7 rendering mode by default, and IE 7 is not enough in supporting more CSS and HTML technologies. In order to be compatible with the newest CSS and HTML pages, we have to call the newest IE rendering engine by some methods, such as setting a registry value of window system. In this way the auxiliary system can reduce the existing workload and not increase new workload additionally.

4.2. XML Classes Design

Because one important function of the proposed system is that it can retrieve the bibliographical description of papers and import them into relevant files of Word template and then print them, some basic classes for papers must be designed. Generally a published paper has some properties, such as title, authors, Author affiliation, journal title, issue number, volume number, page number, publication date, language type, document type, accession number. These properties must appear in certification. For ease of follow-

up operation, these properties are imported into some specific files of XML. An instance of related XML file is shown in Figure 1.
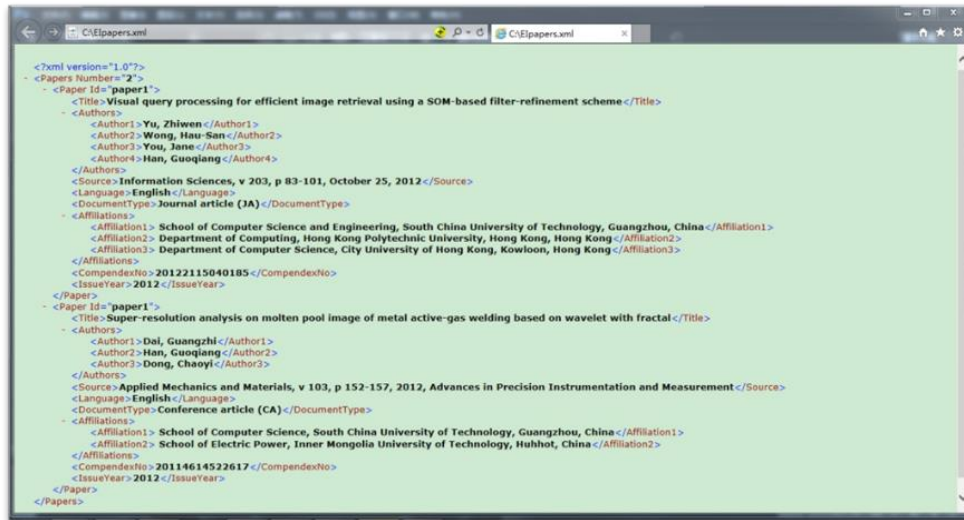

Figure 1. An Instance of Related XML File

4.3. MS Word Template Design

As a part of office automation, the software of Word is widely applied, for its friendly user interface and powerful text editing functions. Using the technology of COM, application programs easily call a Word document and process it. The content of the related certificates is relatively constant, and the special bibliographical descriptions are dynamically changing. So, related Word templates are set in the designed format, in which some bookmarks are inserted at these locations of dynamic changes as markers to aid in modifying these dynamic contents automatically. A word template is shown in Figure 2. In Figure 2, the gray square brackets represent bookmarks of Word.
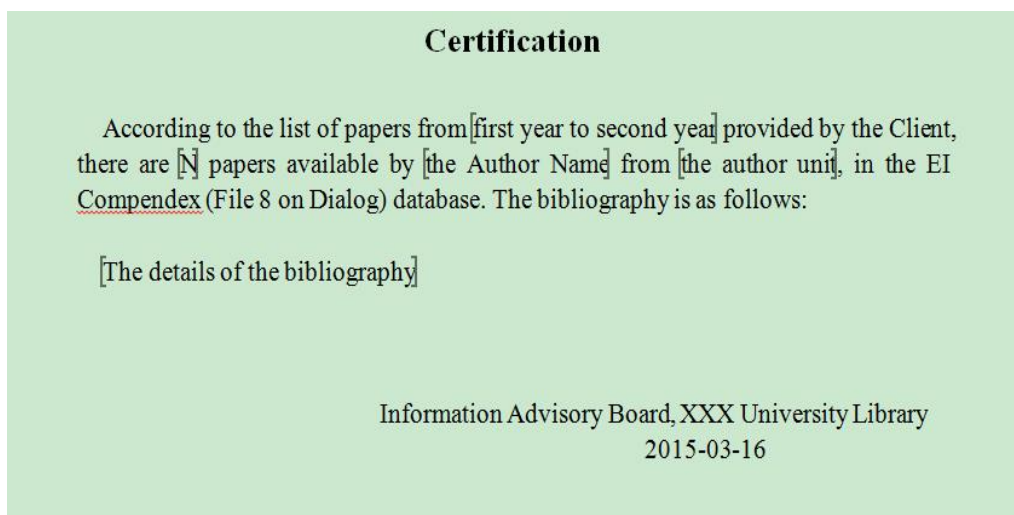
Figure 2. A MS Word Template

4.4. Statistical Report Design

Designing this function is to serve two purposes: First, aiding in calculating workload of users of the system, second, assessing academic level of the users' organization roughly. For the users, workload is probably linkage with their incomes. The design should record the details of certification operations. Because the people needing certificate are mostly in the same organization as users, the statistics could reflect academic level of the organization to a certain extent.

5. System Implementation

We implement the auxiliary system by Microsoft Visual Studio 2010, which is an excellent and powerful development toolkit for desktop applications, web applications, smart client, and office plug-ins and so on. The proposed system is a desktop application of the programming language C#, and it also used the office plug-in function of MS Visual Studio 2010. But the system only supports MS office 2007 and upward. In this section, we detail the implementation of main components of the system.

5.1. Main Interface

As mentioned previously, a WebBrowser control is used to get some necessary information of papers. Usually WebBrowsers are applied in combination with a TabControl. A TabControl contains tab pages, which are represented by TabPage objects that you add through the TabPages property. And a tab page contains a WebBrowser. In the auxiliary system, a specific custom control of TabPageis designed, which is the inheritance class of base class TabPage. By overriding the OnCreateControl method, a Tabpage is banded to a WebBrowser together at the initialization of the WebBrowser. The custom control is set an important attribute "Navigate", whose value is transformed from the Navigate method of WebBrowser. In addition, the auxiliary system can force new Tabpages to show the details of each paper, and the little of these new Tabpages are started with "EI" or "ISI" to show the difference. The main interface is shown in Figure 3.
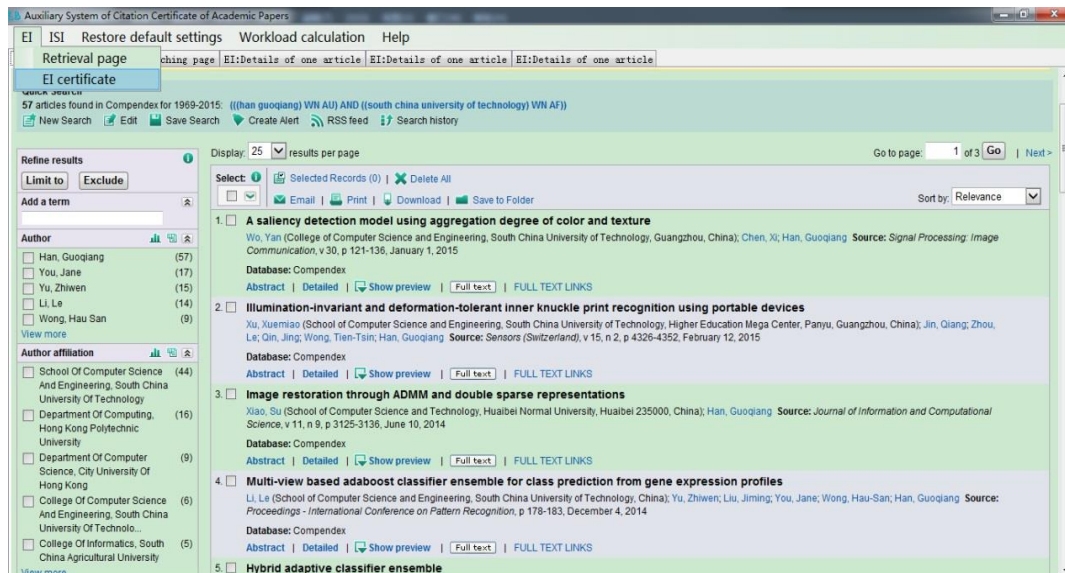

Figure 3. The Main Interface of the System

5.2. Printing the Certification

A great number of data from web pages are semi-structured, so they are not effectively used. Finally, these necessary data extracted from pages make up structured XML data,

which are imported into Word template for printing and database for making a count on workload. As shown in Figure 2, after hitting these second-level menus such as "EI certificate" or "SCI certificate", the system will call corresponding Word template and fill up them with these XML data. At last the printing function of Word will be called for papery certification. One certification of Word is shown in Figure 4.
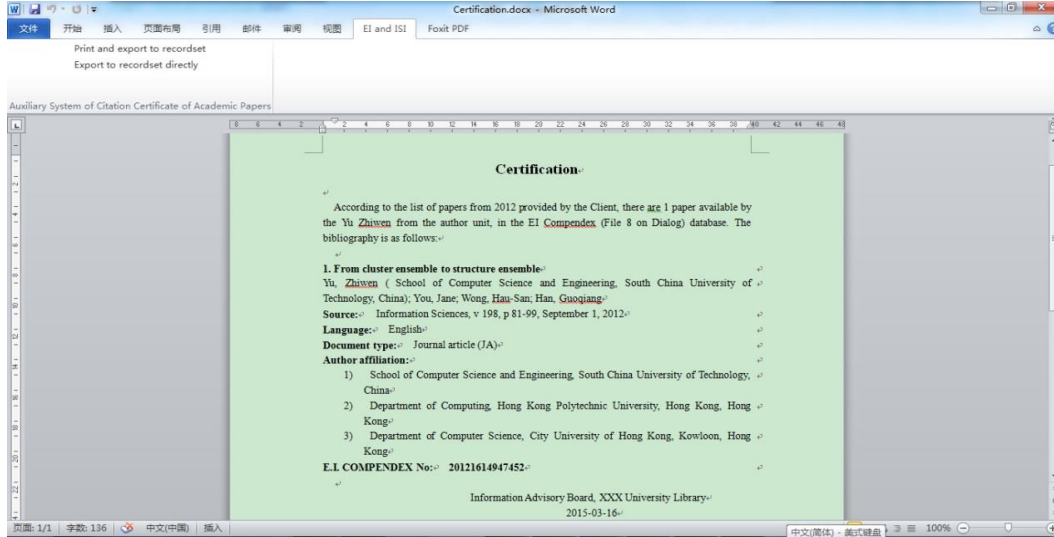

Figure 4. An Instance of Certification of MS Word

5.3. Statistical Report

We used VSTO technology to develop a Word plug-in for realizing the function. As is shown in Figure 4, the Word plug-in has two buttons, one is used to print papery certification as import something into database, and the other is used to import something into database directly in special cases. Then these imported data can be seen in the interface of statistical report, which is shown in Figure 5.

Figure 5. An Instance of Certification of Statistical Report

5.4. Restoring Original Settings

Because the system provide MS Word templates for printing papery certification, the original formats of these Word templates are very easy to be destroyed. Once those bookmarks of Word templates are damaged, the information of papers is not inserted into correct position, even cannot inserted into any location of Word templates. In order to ensure normal working, the system provide the function that restore Word templates. The function works through the original Word templates overlaying damaged word templates.

6. Conclusion

In this paper, we developed auxiliary system of citation certificate of academic papers, which aims at providing staffs of information department of Chinese academic institutions with effective and practical help. In addition, we described how we developed the system from technical aspects. At present, existing system applications have already demonstrated that the system significantly improved working efficiency and reduced labor effort. In the future, we want to expand more academic database supports such as CNKI (China National Knowledge Infrastructure) and provide more efficient services.

Acknowledgments

References

[1] http://www.engineeringvillage.com
[2] N. Osorio, "The engineering index: The past and the present", 2010 ASEE Annual Conference and Exposition, **(2010)**, June 20-23, United states.
[3] http://www.webofknowledge.com
[4] M. Shirabe, "Identifying SCI covered publications within non-patent references in US utility patents", Scientometrics, vol. 101, no. 2, **(2014)**, pp. 999-1014.
[5] K. Y. Chuang and Y. S. Ho, "Bibliometric profile of top-cited single-author articles in the Science Citation Index Expanded", Journal of Informetrics, vol. 8, no. 4, **(2014)**, pp. 951-962.
[6] H. B. Chen, Y. Yang, Y. Yang, W. Jiang and J. C. Zhou, "A bibliometric investigation of life cycle assessment research in the web of science databases", International Journal of Life Cycle Assessment, vol. 19, no. 10, **(2014)**, pp. 1674-1685.
[7] X. J. Wan and F. Liu, "Are all literature citations equally important? Automatic citation strength estimation and its applications", Journal of the Association for Information Science and Technology, vol. 65, no. 9, **(2014)**, pp. 1929-1938.
[8] Y. Wang and C. Y. Tian, "Introduction of agricultural engineering journals embodied by databases of EI and SCI, etc", Transactions of the Chinese Society of Agricultural Engineering, vol. 28, no. S2, **(2012)**, pp. 316-322.
[9] X. J. Lin and R. Li, "Comparative literature study between investigations in foreign Science Citation Index journals and Chinese core domestic journals in the treatment of low back pain with acupuncture", Journal of Traditional Chinese Medicine, vol. 34, no. 3, **(2014)**, pp. 373-380.
[10] M. H. Biglu and M. Ghavami, "Authorship, institutional and citation metrics for publications on postmenopausal osteoporosis", Osteoporosis International, vol. 25, no. 4, **(2014)**, pp. 1337-1343.
[11] G. Parthasarathy and D. C. Tomar, "Trends in Citation Analysis", 1st International Conference on Intelligent Computing, Communication and Devices, **(2014)**, April 18-19, pp.813-821, India.
[12] H. Z. Zhao and X. H. Hu, "Language model document priors based on citation and co-citation analysis", 1st Workshop on Bibliometric-Enhanced Information Retrieval, Co-located with 36th European Conference on Information Retrieval,**(2014)**, April 13, pp.29-36, Netherlands.
[13] http://msdn.microsoft.com/zh-cn/library/system.net.webclient(VS.95).aspx
[14] http://msdn.microsoft.com/zh-cn/library/system.net.httpwebrequest(v=vs.110).aspx

[15] http://msdn.microsoft.com/zh-cn/library/system.net.httpwebresponse(v=vs.110).aspx
[16] http://msdn.microsoft.com/zh-cn/library/system.windows.forms.webbrowser(v=vs.110).aspx.

Authors

Bin Lu, he is currently a lecturer in the School of Computer Science at Wuyi University. He received his Ph.D. degree in 2013 from South China University of Technology. He is a reviewer of the Journal of Yangtze River Scientific Research Institute since 2010. His main research interests include complex network and machine learning.

Changyu Liu, he joined the Communication and Computer Network Lab of Guangdong as a PhD student in 2010 at South China University of Technology, advised by Prof. Shoubin Dong. He was a visiting scholar at the School of Computer Science, Carnegie Mellon University, from September 2012 to October 2013, advised by Dr. Alex Hauptmann. Then, he worked with Prof. Mohamed Abdel-Mottaleb and Prof. Mei-Ling Shyu at the Department of Electrical and Computer Engineering, University of Miami, from October 2013 to September 2014. His research interests include computer vision and machine learning.

Weifeng Gao, he is currently a lecturer in the School of Computer Science at Wuyi University. He received the M.S. degree in 2006 from Wuhan University of Technology. He is interested in network security, computer graphic and image.

Zhenbang Rong, he received B.S. Degree in computer science from Nanjing University of Science and Technology, Nanjing, China in 2001 and then he received the M.S. Degree in computer science from Sun Yat-Sen University, Guangzhou, China, in 2006. Since 2013, he has been working toward the Ph.D. Degree in the Department of Computer Science, Sun Yat-Sen University, Guangzhou, China. Since 2006, he has been a teacher of computer science at Wuyi University. He is interested in computer network, quantum computation and quantum information.