

# Automating the Sci-tech Papers' Coverage and Citation Report Based on Cloud Platform Via Mobile Phone or Tablet

**Chen Yuecong**

Library of North China Electric Power University,  
Beijing 102206, China

**Fang Yanhong**

Library of North China Electric Power University,  
Beijing 102206, China

**Wu Guiqin**

Library of North China Electric Power University,  
Beijing 102206, China

**He Qiong**

Library of North China Electric Power University,  
Beijing 102206, China

**Tian Huiyun**

Library of North China Electric Power University,  
Beijing 102206, China

**Abstract:** This paper introduces the mobile application of automatic database coverage and citation certification automation system based on cloud platform, which services for university Sci-tech Novelty Retrieval Station. It creates the university science and technology (sci-tech ) paper database and its citation database, which relate with the university teachers and students' information. Users can complete the certification on mobile phone or tablet computer. It can generate the certification report and QR code automatically, and users can print the certification report at Sci-tech Novelty Retrieval Station with the QR code. This can give great convenience to users and reduce the burden of novelty station.

**Keywords:** Sci-tech Novelty, Coverage and Citation Report, QR Code, Mobile APP

## 1. INTRODUCTION

Sci-tech papers' coverage and citation certification(STPCCC) is to search a paper in the citation database such as SCI(Science Citation Index), SSCI(Social Science Citation Index), EI(Engineering Index), etc. and give a certification report for its coverage and citations in this citation database. Sci-tech Novelty Search Station (STNSS) is a department in university, which undertakes not only the task of sci-tech novelty research but also the task

of STPCCC. Sci-tech papers' coverage and citation have been increasingly applied to evaluate scientific research in academic and research institutions:

- For teachers: the applications of various funds, the appraisal and awards for scientific research and the professional title evaluation, etc. all need the certifications.
- For students: scholarship of engagement and other movements such as seek a higher education, job application, etc. all are their competitive standings as a part of the application process.

With so much teachers and students in the universities, the number of sci-tech papers to be certificated is very large, so they are heavy tasks for STNSS. And these certification reports are usually for single use, when they are needed the next time the teachers and students will come to the STNSS for another new certification report, so there are many repeating works for novelty searchers.

At present, the sci-tech papers' certification in STNSS is accomplished by hand. Because the users(teachers and students in universities) usually don't know their papers are in which citation database and the spellings of authors vary from one citation database to another, the certification of sci-tech papers is very complex, and the novelty searchers must repeatedly search in various citation database

and communicate with users before finish the certification. There will be hundreds of users to STNSS for certifications in peak period, so this work takes up a lot of time and energy of STNSS and the sci-tech papers' certification has become a major job of STNSS [1-5].

To better understand the status of sci-tech papers' certification in STNSS this paper investigated three dozen university libraries and public libraries in China and searched the online web-based databases and search engines such as Baidu, Google, etc. There are the several ways to certificate a paper in STNSS of Chinese universities:

- Face-to-face: the users come to the university's STNSS to apply for the certification themselves, and they usually are teachers and students in the same university.
- E-mail: the application form with paper list is linked on the website of STNSS; the users fill the application form and send it by e-mail to STNSS. The application form usually includes: the name of the application institution, the author(s) and database to be certificated, which year(s) to search, citation and self-citation involved or not, etc. Most STNSS accept the application by e-mail way.
- Online submission: similar to the e-mail way, the STNSS website provides an online submission of application form. Some STNSS provide the way of website application submission, such as Tsinghua University STNSS, The National Library, etc.
- You yourself complete way: the users fill the certification report themselves and take the report to be verified, signed or sealed by novelty searchers of STNSS. Zhongshan University and North China Electric Power University permit this "You yourself complete" way.

The first three ways vary from the way how the

application submission is given to the STNSS, but the retrieval process of novelty searchers is all by hand. As for the "You yourself complete" way, although the novelty searchers do not need to fill the certification report, they also need to carefully check the paper, database, authors, etc. in this report. The work to check the papers one by one does not necessarily small than fill a new report [6-7].

Besides the above four ways, there are two articles about automation system of certification report:

[2] is a "cited reference retrieve service automatic generation system", users upload the paper list in WORD or EXCEL and input the author(s) online, the system can automatically transfer this data to the retrieval part. The java programs of the system can retrieve the data through the Application Programming Interfaces (API) of the citation database, such as SCI, SSCI, EI, etc.

[3] designed a JSP-based online database retrieving and science technological novelty search system. The users login the system and fill the application form, the novelty searchers search through the APIs of citation database, and chat with users to confirm the retrieval result.

On the basis of the above investigation, this paper will study the automation system of certification report to reduce the workload of STNSS researchers.

## 2. System design

As the power and popularity of smartphones and mobile tablets, it is possible for us to complete the retrieval of citation databases via mobile APP technology, which is already familiar to users. The cloud computing is a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service using Internet technologies. The cloud computing is an exciting technology and a compelling discipline that has already exhibited profound implications on our life, work and share of knowledge. Remarkably, this happens regardless of the

geographical and temporal space in which knowledge seekers and knowledge providers physically exist [8]. This is just what we need to provide a mobile and automatic application of certification reports. The provision of cloud computing services can occur at the infrastructure level, platform level, and software level. Built on the cloud platform, the database of sci-tech papers together with the QR code technology will be used in this paper.

The sci-tech papers' certification is always accomplished in STNSS of universities, so if the teachers and students' information can be related with the database of sci-tech papers, it will be convenient for users to input the teachers or students' ID number to certify their papers on the mobile phone or tablet. The data will be processed by the server of cloud platform, and the users only need to confirm the certification report returned by the server. When confirmed, a QR code will be automatically generated; the users can download and save this code, which can be used to identify the certification report in STNSS. North China Electric Power University STNSS is taken as an example to better understand this research.

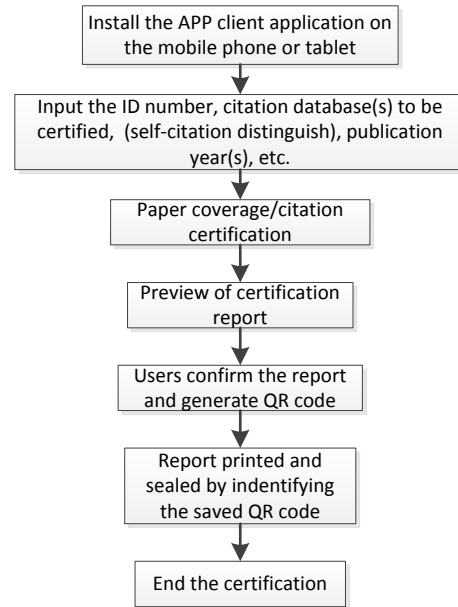


Fig.1. The design of system flow is as

### 3 System function module design

#### 3.1 Create the student and teacher information database (STID)

The data from common card system of North China Electric Power University include name, number ID, identifiable information (teaching & administrative staff, student, various temporary workers, etc.), department, telephone number, etc. For user convenient, the student and information database (STID) was created with the key information related with the coverage and citation certification system and derived from the STID.

Tab.1 Student and teacher information database

ID	ID number	Name	Identifiable Information	College/department	College/department identifier	English standard name	English standard college/department name
User unique identification	Common card ID number field	Common card name field	Common card identifiable information field	Common card department field	To identify college/department	From the common card name field	From the College/department field

The identifiable information field is to mark the user to be student or teacher; the college/department identifier field is to identify the user's college or department. These two fields are designed for the

data query and statistics. Because the citation database are all English database, so the pinyin4j-2.5.0.zip-JAVA Chinese Pinyin open source package<sup>[9]</sup> was used to converse the user name and college/department name to Pinyin, and then form the

English standard user name and English standard college/department name. According to the author common spelling method and the change/merge of college/department, the extended knowledge database of user name and college/department name was created for the matching of STID and university sci-tech papers database. The self-modify interface is provided for the user whose name has polyphone character, rare character or four-character name.

### 3.2 Download the sci-tech papers and its' citations and create the institute sci-tech papers database (ISTPD)

The ISTPD is created with all the sci-tech papers of North China Electric Power University in the citation databases(EI, SCI, SSCI, CPCI-S, etc.) of Engineering

Village and Web Of Science, these papers are retrieved by the author affiliation field as “North China Electric Power University”, “NCEPU”, “102206”, “071003”, “Beijing”, or “Baoding”. All the papers detail records are downloaded from the citation database and analyzed as follows. The citation databases are divided into two kinds: EI and Web Of Science databases (WOS). WOS include the citation databases such as SCI, SSCI, CPCI-S, etc. which are on the Web Of Science platform.

The similar fields of EI and WOS are as follows: Title/TI, Source/SO, Corresponding author/RP, DOI/DI, Conference name/CT, Conference date/CY, Conference location/CL, Document type/DT, Abstract/AB, Number of references/NR. The different fields are as follows:

Tab.2 The comparison of different fields of EI and WOS

	Article Number	Author	Author affiliation	Citations	Database
<b>EI</b>	Accession Number	Author	Author affiliation	No citation field	Database
<b>WOS</b>	UT(WOS)	AU, AF,CA	C1	TC, Z9	No database field

- Article Number (AN): The AN of EI is the Accession Number which is a 14 digital coding begins with years -20150900579594, for example. The article unique identifier of WOS is a 15 digital coding begin with “WOS” -WOS: 000329173200041, for example.
- Author and Author Institution(AI): These two fields are different as follows:

Tab.3 The comparison of Author and Author affiliation fields

	Author	Author affiliation
<b>EI</b>	Gong, Xue(1,2); Kong, Wei-Zhi(1); Yuan, Yi-Zhao(2)	(1) State Key Lab. of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China; (2) North China Electric Power University, Beijing 102206, China
<b>WOS</b>	Wang, XQ; Nordlander, E; Thorin, E (Wang, Xiaoqiang; Nordlander, Eva; Thorin, Eva)	[Wang, Xiaoqiang; Nordlander, Eva; Thorin, Eva; Yan, Jinyue] Malardalen Univ, Sch Sustainable Dev Soc & Technol, S-72123 Vasteras, Sweden; [Wang, Xiaoqiang] North China Elect Power Univ, Sch Renewable Energy, Natl Engn Lab Biomass Power Generat Equipment NEL, Beijing 102206, Peoples R China

STPCCC is for institution individual users, so it is needed to extract all the authors of the institution:

EI: As in Tab.3, the identification number (IN) in brackets first needs to find in the Author affiliation field. And use this number to find the authors with this IN in the Author field. For example, to North China Electric Power University in Tab.3, the IN is 2, and the authors with this IN are Gong, Xue and Yuan, Yi-Zhao.

WOS: The authors of an institution will be enclosed in square brackets[] before the institution name. For example, the author of North China Electric Power University is Wang, Xiaoqiang<sup>[10]</sup>.

- WOS databases have the Time Cited (TC) and Time Cited from Web Of Science Core Collection (Z9) fields, which is the citation counts in Web Of Knowledge Platform or

only in core collection. While EI has no time cited field.

On the basis of the above similarities and differences, the database will be combined with the similar data fields. The different data fields will be combined after a series of data processing. And then the ISTD and its citation database will be created.

### 3.3 The data update of ISTD and its' citation database

The institute papers and its' citations in Engineering Village and Web Of Science are updated frequently, so the ISTD and its' citation database

must update the databases according to the citation databases' updates. Considering the WOS databases have the time citation data of papers and are more complex than EI, so the data update of WOS database is discussed in this paper.

Web Of Science provides three kinds of APIs, as in Tab.4:

Tab.4 Comparison of APIs of Web Of Science

	Web Service Lite	AMR	Web Services Premium
<b>Free/Charge</b>	Free	Free	Charge
<b>Data fields can be accessed</b>	-Authors -Source -Keywords -UT -Title	-Time Cited -UT -DOI -PMID -Source URL -Citing Article URL -Related Article URL	Complete set of data available
<b>Retrieval way</b>	Batch retrieval	One by one retrieval	Batch retrieval
<b>Limitations</b>	Only specific data fields can be accessed, 2 requests per second, 100 records per request	Only provides the specific data fields of the paper, and cannot batch query.	Charge

- Web Service Lite (WSL): Follow the SOAP1.1 (Simple Object Access Protocol) to perform request / response message delivery. WSL can retrieve specific fields for a given search terms and limits the number of records to the response, batch retrieval within 100 records per request.
- Article Match Retrieval (AMR): According to the DOI, UT, Title, etc. of the record, it can only obtain relevant information of the corresponding record, such as UT, Authors, Source, Time Cited, etc. The

Information is limited to that corresponding record, and can be called "one by one" retrieval in comparison with WSL.

- Web Services Premium (WSP): WPS is written to follow SOAP 1.1 (Simple Object Access Protocol), JSR 224 (Java API for XML-based Web Services), and WSDL 1.1. All available data fields can be provided, and it is as enterprise-level support<sup>[11-16]</sup>.

This paper is still in the research stage, so the free WSL and AMR are used to update the papers

and citations. Spring and Quartz are used to automatically execute the task<sup>[17]</sup>.

```
<bean id="sciTrigger"
class="org.springframework.scheduling.quartz.CronTriggerBean">
    <property name="jobDetail"
ref="sciJobDetail"/>
    <property name="cronExpression"
value="*0 0 1 ? * L"/>
</bean>

<bean id="citingTrigger"
class="org.springframework.scheduling.quartz.CronTriggerBean">
    <property name="jobDetail"
ref="citingJobDetail"/>
    <property name="cronExpression"
value="*0 0 3 ? * L"/>
</bean>
```

The symbolicTimeSpan of WSL is set as 1week to update data once a week. The citation data can be obtained by AMR. The URL (Uniform Resource Locator) of citations is written to a text file, and the detail information of citations, such as Title, Author, Source, Abstract, etc. can be obtained through wget, which is a computer program that retrieves content from web servers<sup>[18]</sup>. The detail information of citations is automatically saved in the citation database, also once a week. North China Electric Power University is as an example as follows<sup>[15-16]</sup>.

```
_search_queryParameters.setDatabaseID("WOS");
EditionDesc desc_sci = new EditionDesc();
desc_sci.setCollection("WOS");
```

```
desc_sci.setEdition("SCI");
EditionDesc desc_ssci = new EditionDesc();
desc_ssci.setCollection("WOS");
desc_ssci.setEdition("SSCI");
EditionDesc desc_istp = new EditionDesc();
desc_istp.setCollection("WOS");
desc_istp.setEdition("ISTP");
EditionDesc desc_isshp = new EditionDesc();
desc_isshp.setCollection("WOS");
desc_isshp.setEdition("ISSHP");
_search_queryParameters.getEditions().add(desc_sci);
_search_queryParameters.getEditions().add(desc_ssci);
_search_queryParameters.getEditions().add(desc_istp);
_search_queryParameters.getEditions().add(desc_isshp);
_search_queryParameters.setQueryLanguage("en");
;
_search_queryParameters.setUserQuery("AD=(n* china elec* power univ* same (102206 or 071003 or beijing or baoding or 215123))");
_search_queryParameters.setSymbolicTimeSpan(1week);
_search_retrieveParameters.setFirstRecord(1);
_search_retrieveParameters.setCount(100);
```

### 3.4 Build the cloud platform

For copyright protection of electronic data resources, most databases use IP addresses to

restrict access to external users. The coverage and citation retrieval must be completed in campus, which is inconvenient for users. If the data is distributed on cloud platform, the retrieval can be accomplished in any place with network not limited by the IP address. So the cloud platform is clearly needed for user convenient.

Google, Amazon, IBM, ORACLE Cloud, Aliyun, etc. are the well-known cloud computing vendors. In the research stage of this application system, Aliyun is considered to be rent to complete the system build. The cloud system includes: Server Load Balancer (SLB), Elastic Compute Service (ECS), Relational Database Service (RDS), Open Storage Service (OSS) and cloud shield. As in Fig.2, SLB and ECS are responsible for providing external HTTP services; SLB distributes traffic for multiple cloud servers; RDS stores the user information, sci-tech papers and its citations; OSS stores images, QR code, certification reports and other data. The cloud shield provides a total security services, including DDoS protection, host intrusion prevention, vulnerability detection, Trojan detection, and etc<sup>[19-20]</sup>.

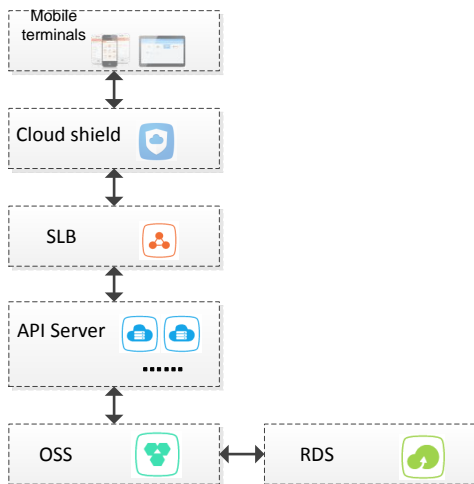


Fig.2 Cloud platform

### 3.5 Mobile applications

Three-tier architecture was used to the mobile APP application. Three-tier architecture is a client-server architecture in which the user interface (presentation), functional process logic ("business rules"), computer data storage and data access are developed and maintained as independent modules. The presentation tier displays information to users and is implemented in the client side. The business logic tier and data tier are on the database side.

The user interface is in the presentation tier. Users select the database and years to be retrieved, through SLB, the data is analyzed by the business logic tier. The data processed by the data tier will be returned to the presentation tier, and then presents to users. The automatic data acquisition from Engineering Village and Web Of Science platform, the creation and identification of certification reports, etc. are all accomplished by the data tier to reduce the burden of the client<sup>[21-23]</sup>.

The data access service gets the institute sci-tech papers and citations from Engineering Village and Web Of Science platform according to the fixed time. The data will be matched with the STID and stored in the database with database identification, such as EI, SCI, SSCI, etc. The data retrieval service needs to process the data of the client request, match the database to be retrieved and return the result to users.

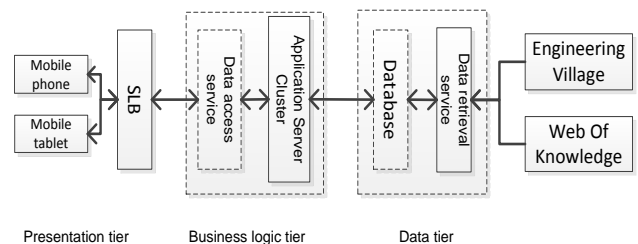


Fig.3 Mobile application frame

### 4 Use impact assessments

The retrieval procedure in this paper is greatly simplified than the traditional manual method, the

comparison is in Fig.4. In this paper, users need to install mobile application on mobile devices like mobile phones or tablets, and input the ID number and database to be certified. The system can automatically retrieve the papers covered or cited. After the users confirm the papers, the certification report will be automatically saved on cloud servers

and return a QR code to users. Users can print the certification report by identifying the QR code in STNSS and the retrieval procedure is end at that time. The novelty searchers do not need to do the retrieval process by hand as in the traditional retrieval method.

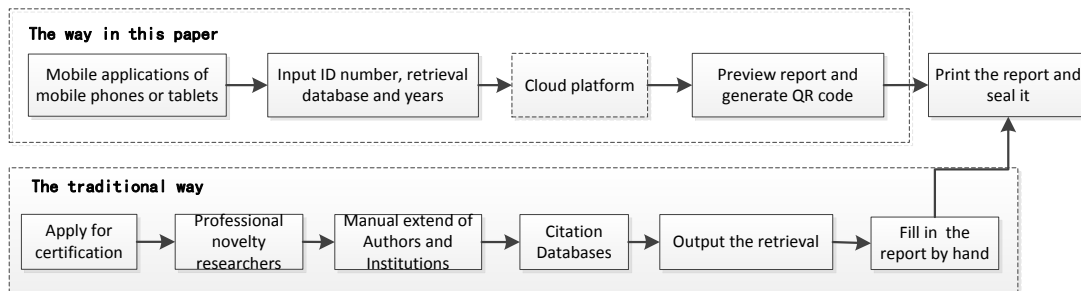


Fig.4 The procedure comparison of two ways

To the STNSS, information provided by users, retrieval method, the report generation, etc. are greatly improved in this paper, as in Tab.5 In the traditional way, after the users apply for the certification, the novelty researchers need to translate the author and institution, etc. The diversification of spelling and name repetition will confuse the searchers, and they must communicate with customers and try to retrieve repeatedly and check the results one by one. If encounter hundreds of papers included retrieval, the report needs 1-2 working days to complete.

After using the method in this paper, the data preparation is completed by computer servers, the certification report only needs one or two minutes to complete, the report body and attachment can be generated automatically according to the predefined standard format, without manual adjustment. The STNSS only need to provide a computer to identify the QR code and print the report. The novelty searchers are not involved in the retrieval and free from the heavy retrieval work.

Tab.5 The comparison of two ways

	User information	Retrieval method	Report generation	Reuse of report	Time needed
<b>The way in this paper</b>	Id number	Self-service	Automatic	QR code identification, can reuse	Several minutes
<b>The traditional way</b>	Provide paper list	Professional searchers	Fill by hand	Cannot reuse	1-2 working days

## 5 Conclusions

The automation system of certification report based on platform in this paper is still in research and testing phase, a period of time need to test and assess the system before it can be applied to practical work. There are many situations to be

considered in the system, such as the users not in this institute cannot use this method, polyphone character, rare character, etc. After the system is mature, this system can also provide the basis for the school scientific research management and the staff assessment.



## References

- [1] Development Center of Science and Technology, Ministry of Education [EB/OL]. (2014-11-27) [2015-01-09].  
<http://www.cutech.edu.cn/cn/zxgz/2014/11/1417029767876562.htm>.
- [2] Wang Xueqin; Hao Dan; Zheng Fei, etc. Practical Application Research of a Cited Reference Retrieve Service Automatic Generation System [J]. Library and Information Service, 2014,16:131-137.
- [3] Shi Xiaoqing; Wei Jiangxing. Design and Implementation of JSP-based Online Database Retrieving and Science Technological Novelty Search System ——A Case Study of Shandong University (Weihai) [J]. Journal of Modern Information, 2014, 03:131-134+138.
- [4] Liang Hongni, Hu Xiaofei. The Analysis of the Database Retrieving and Data Processing [J]. Information Studies: Theory & Application, 2009, 04:96-99.
- [5] Zhou Meihua. Investigation of database retrieving and data processing service under web service[J]. Library Work And Study, 2011,12: 106-109.
- [6] Library of Tsinghua University. Data retrieving and data processing [EB/OL]. [2014-12-20].  
<http://lib.tsinghua.edu.cn/service/retrieval.html>.
- [7] Library of Tsinghua University. Domestic and foreigner university libraries[EB/OL]. [2014-12-20].[http://lib.tsinghua.edu.cn/find/find\\_library.html](http://lib.tsinghua.edu.cn/find/find_library.html).
- [8]Gartner IT Glossary > Cloud Computing [EB/OL]. [2015-01-13].  
<http://www.gartner.com/it-glossary/cloud-computing>
- [9]Chinese to Pinyin [EB/OL]. (2006-06-21) [2015-01-12].  
<http://sourceforge.net/projects/pinyin4j/files/pinyin4j-2.5.0/pinyin4j-2.5.0/>.
- [10] Yu Jian; Wu Xia; Zhao Chunmei. Design and Application of Author Discrimination Software on Papers [J]. New Technology of Library and Information Service,2014,04: 78-84.
- [11] The Difference of AMR and WEB SERVICE LITE [EB/OL]. (2012-09-05) [2015-03-05].  
<http://mednoter.com/what's-the-difference-between-a-mr-and-web-service-lite.html>.
- [12] ISI Web of Knowledge<sup>SM</sup> Web Services Premium v.2.1.1 Guide [EB/OL]. (2010-11-23) [2015-03-06].  
[http://science.thomsonreuters.com/tutorials/wsp\\_docs/soap/Guide/](http://science.thomsonreuters.com/tutorials/wsp_docs/soap/Guide/).
- [13] Web Services FAQ[EB/OL]. [2015-03-06].  
[http://wokinfo.com/products\\_tools/products/related/web\\_services/ws\\_faq/](http://wokinfo.com/products_tools/products/related/web_services/ws_faq/).
- [14] Article Match Retrieval FAQ[EB/OL]. [2015-03-06].  
<http://wokinfo.com/directlinks/amrfaq/>.
- [15] Web of Knowledge Web Services Lite v. 3.0[EB/OL]. (2012-09-07) [2015-03-06].  
<https://www.accesowok.fecyt.es/wp-content/uploads/2014/10/WebServicesLiteguide3.0.pdf>.
- [16]Zhang Chunxiao. Institute Paper Statistical Analysis System Based on Web of Knowledge Web Services

- [J]. Journal of Modern Information, 2013, 01: 98-102.
- [17] Spring+Quartz to achieve timing task [EB/OL]. (2013-12-31) [2015-03-12]. <http://blog.csdn.net/u010397369/article/details/17465649>.
- [18] Wget [EB/OL]. [2015-03-12].<http://en.wikipedia.org/wiki/Wget>
- [19] Aliyun [EB/OL]. [2015-03-20]. <http://www.aliyun.com/?spm=5176.383846.2.1.dkPFJW&lang=en>
- [20] Aliyun Cases[EB/OL]. (2013-12-13) [2015-03-20]. <http://bbs.aliyun.com/read/149110.html?spm=5176.7189909.0.0.w2UMXi>.
- [21] Multitier architecture [EB/OL]. [2015-03-22]. [http://en.wikipedia.org/wiki/Multitier\\_architecture](http://en.wikipedia.org/wiki/Multitier_architecture)
- [22]Ye Li. Mobile Library Construction and Development Based on the APP Application——A Case Study of The Library of Wuhan [J]. Journal of Library Science, 2014, 03: 116-118.
- [23]Yang Dong-wu. The Research and Development of Mobile Library Inquire System Based on J2ME [J]. DATABASE AND INFORMATION MANAGEMENT, 2014, 08: 51-53+55.