IECS: an Integrated E-Community System for Management, Decision and Service

Yu Bo, Hongding Wang, Zhang Peng, Yunhai Tong, Shiwei Tang, Dongqing Yang*

Contents

Abstract

The paper presents an Integrated E-Community System (IECS) for management, decision and service, designed for the e-government project of Haishu District of Ningbo, Zhejiang, China. The project need is to promote the integration of management information and service information of communities, providing a unified platform on which different departments of the district government can share and exchange community information, government officers can analyze information and make decisions, and the outside users can access and request services. To meet the project need, the IECS consists of five parts: 1) The Central DataBase (CDB) that stores all information related with management, decision and service of communities; 2) Information Extracting Subsystem (IES) that provides functions of extracting data from data sources, transforming and loading them into the CDB for system administrators; 3) Information Management Subsystem (IMS) that provides functions of querying and sharing of information for government users, and functions of information maintenance, rights and log management for system administrators; 4) Intelligent Analysis Subsystem (IAS) that provides functions of extracting analysis-related data from the CDB and loading them into the DW, and functions of multi-dimensional analysis and decision-making based on the DW and OLAP for government users; 5) Information Service Website (ISW) that provides functions of promulgating and collecting of information for government users and system administrators, and functions of browsing, querying and requesting of service information for outside users. The IECS supports management, decision and service of a government based on a unified data platform--the CDB, and ensures data security by providing different workplaces and rights for different users. In the real application, the system works well.

* School of Electronics Engineering and Computer Science, Peking University, Beijing, China
{yubo, hdwang, zhp, yhtong, tsw, dqyang}@db.pku.edu.cn

I. Introduction

A community is a basic unit of society in China, such as Zhongguanyuan Community, Haidian Strict of Beijing, China. Because the development of a community may make a great influence on management, decision and service of a city or district government, the e-community is an important part of building the e-government.

E-government refers to government' s use of technology, particularly web-based Internet applications to enhance the access to and delivery of government information and service to citizens, business partners, employees, other agencies, and government entities [1]. Nowadays, governments are going-on line and using the Internet to provide public services to its citizens. But current e-government projects pay too much attention to the Internet applications, neglect what information and service citizens actually want e-government to provide [2] and what the value of egovernment is [3]. How to find the demands of citizens and provide better and more useful services and information? How to exert the management function of government in the e-government environment? Karen Layne and Jungwoo Leeb in [4] gave a development model, namely the four-stage model for fully functional egovernment: 1) cataloguing, 2) transaction, 3) vertical integration, and 4) horizontal integration. Therefore, the future e-government could be an integrated information platform that covers all functions of e-government including not only the access and delivery of information and service on the Internet but also gathering, sharing and analyzing of information and service based on database or data warehouse. We develop an Integrated E-Community System (IECS) for management, decision and service. It has been applied to the e-government project of Haishu District of Ningbo, Zhejiang, China, with encouraging results. The IECS has four objectives:

- Building a central database that stores all information related with management, decision and service of communities, which is extracted from systems of different administrations (e.g. police station, health office, and family planning office) and external data sources.
- Developing a unified platform on which different departments of the district government can share and exchange information, government officers can handle official businesses, and outside users can access and request services.
- Intelligent analysis and decision supporting will be provided for government officials. By making use of the technologies of DW and OLAP, the system can find valuable information about the development of a community,

such as the number of poor or handicapped and the current status of public security, etc. Based on this information, governments can make more appropriate work planning and service goal.

- Enhancing the communication between governments and citizens. The system provides a website on which people can advance their requirements and the government can respond to them in time.

The remainder of this paper is organized as follows. Section 2 describes related works. Section 3 describes design and implementation of the IECS, particularly the architecture of the IECS and functions of four subsystems. Section 4 summarizes the features of system and suggests future research directions.

Ⅱ. Related works

A detailed analysis on the lessons, challenges and future directions of e-government is given in [3]. [4] describes different stages of e-government development and proposes a 'stages of growth' model for fully functional e-government, particularly the viewpoint of integration. [5] analyzes data using a structurational model, to identify issues in developing this initiative, and construct a framework to analyze future egovernment initiatives. [6] utilizes RDF and RDF schema to develop a semantic ecommunity. [7] discusses secret strategies for successful online communities. [8,9] present a knowledge management system designed for the Italian Department of Technology and Innovation, which supports work and knowledge sharing for all actors involved. Our works not only support work and knowledge sharing, but also stress on knowledge acquisition and the integration of management, decision and service.

Ⅲ. Design and implementation of the IECS

3.1 Architecture

The IECS adopts a three-tier architecture (Fig.1):

- Data tier: It consists of a central database (CDB) and a data warehouse (DW). The CDB is an information-sharing platform that stores all information related with the management, decision and service of communities, such as demographics, population, sanitation, culture, sports, and requests, etc. Data from autonomous systems and external sources are extracted into the CDB using the special information extracting

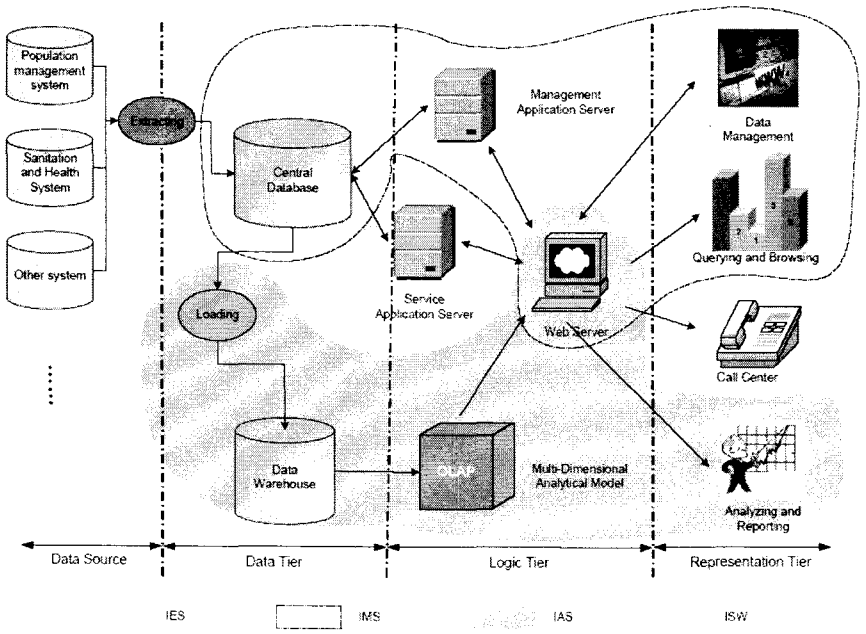
system (IES). The DW is an information-analyzing platform that stores analysis-related data extracted from the CDB according to the DW subject classification.

- Business logic tier: It consists of a management application server that implements information management logic, a service application server that implements information service logic, an OLAP server that implements information analysis logic, and a Web server that servers web pages to the representation tier across the Internet or an Intranet and handles HTTP requests.
- Representation tier: It contains data management tools, query and browsing tools, analysis tools and call center.

From the function point of view, the IECS consists of three sub-systems and one website (Fig.1): 1) Information Extracting Subsystem (IES) that provides functions of extracting data from data sources, transforming and loading them into the CDB for system administrators; 2) Information Management Subsystem (IMS) that provides functions of querying and sharing of information for government users, and functions of information maintenance, rights and log management for system administrators; 3)

Intelligent Analysis Subsystem (IAS) that provides functions of extracting analysisrelated data from the CDB and loading them into the DW, and functions of multidimensional analysis and decision-making based on the DW and OLAP for government users; 4) Information Service Website (ISW) that provides functions of promulgating and collecting of information for government users and system administrators, and functions of browsing, querying and requesting of service information for outside users. Therefore, the IECS is an integrated e-community system that supports management, decision and service of a government based on a unified data platform--the CDB, and ensures data security by providing different workplaces and rights for different users.

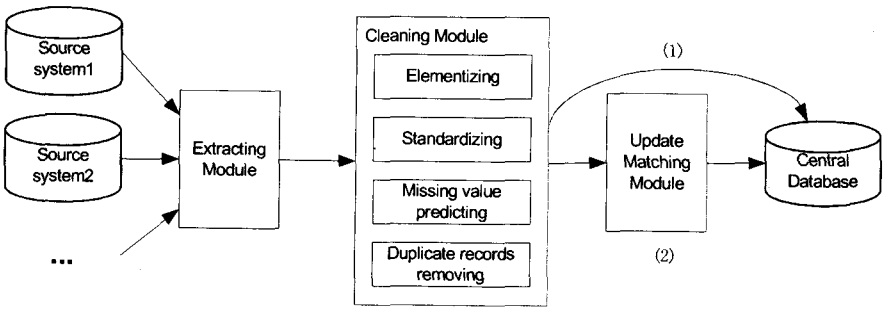
Fig. 1. Architecture of the IECS



3.2 Information extracting subsystem (IES)

The IES extracts data from data sources, transforms and loads them into the CDB. It is the workspace of system administrators or data entry executives. For the government that has no management system, the IES provides an on-line datagathering tool with which data entry executives in different level governments can directly input related data into the CDB. For the government that has some management systems (e.g. population management system, culture & sports facility management system, Resident's Committee management system), the IES provides a data-extracting tool for system administrators to extract data from data sources and load them into the CDB. Because data sources usually are autonomous and heterogeneous, we have to solve data quality problems such as missing value, data conflicts, approximate duplicate records before loading. In the latter scenario, The IES has the following major modules (Fig.2):

Fig. 2. Function modules of the IES



(1) initial extracting; (2) incremental updating

1) Extracting Module

We assume that data sources have been chosen. This module extracts related data from these data sources and loads them into a buffer.

2) Cleaning Module

This module solves data quality problems and transforms data in the buffer according to the standard format of the CDB. It includes:

- Elementizing: The IES parses address and name into elements. For example,
  Before elementizing: Jinfeng Road 13, Haishu District, Ningbo, Zhengjiang, 315002
  After elementizing:
  Street Address Number: 13
  Street Name: Jinfeng Road
  District: Haishu
  City: Ningbo
  Province: Zhengjiang
  Zipcode: 315002
  Community: null
- Standardizing: The IES transforms name, ID number, address, and phone number into standard format or type. For example, "8657482757756"→ "86-574-8275- 7756".
- Missing value predicting: The IES uses the most probable value to fill in the missing value. It can be determined with regression, lookup table, inference-based methods using decision tree induction, or a Bayesian formalism. For example, if there is a mapping rule *"Jinfeng Road 13 belongs to Wuhuang Community"* in the lookup table, the system can predict the above null *community* value is *Wuhuang Commnunity.*
- Duplicate records removing: The IES uses the domain independent Priority-Queue algorithm [10] to detect duplicate records, and then, the IES submits detected records to domain experts or users who decide whether remove duplications or not.

3) Update Matching Module

If it is the initial extracting, after cleaning, the IES directly loads cleaned

results into the CDB. Otherwise, extracting is an incremental update operation and the IES will call Update Matching Module to find records in the CDB matching with extracted data and update them.

3.3 Information management subsystem (IMS)

The IMS is an information and office management platform (Fig.3) that provides different level rights of information querying and information sharing on the CDB for government users based on user roles, and functions of information maintenance, rights and log management for system administrators. It is the workspace of government users and system administrators.

The IMS adopts the three-tier J2EE architecture (Fig.4): representation tier, business logic tier and data tier. Representation tier accepts queries of government users and represents query answers with HTML sheets and JSP. Business logic tier is the center of transaction processing which responds to queries of government users, processing data according to business logics, and returns query answers using EJB, JDBC and Servlet. Data tier, namely the CDB, is the core component of the IMS, which gathers, stores, and maintains e-government information.

The IMS contains the following functions:

1) Information querying and browsing

Government users can query and browse management information in the CDB based on user roles. They can set query conditions, browse detailed information satisfying query conditions, and print query results.

2) User right management

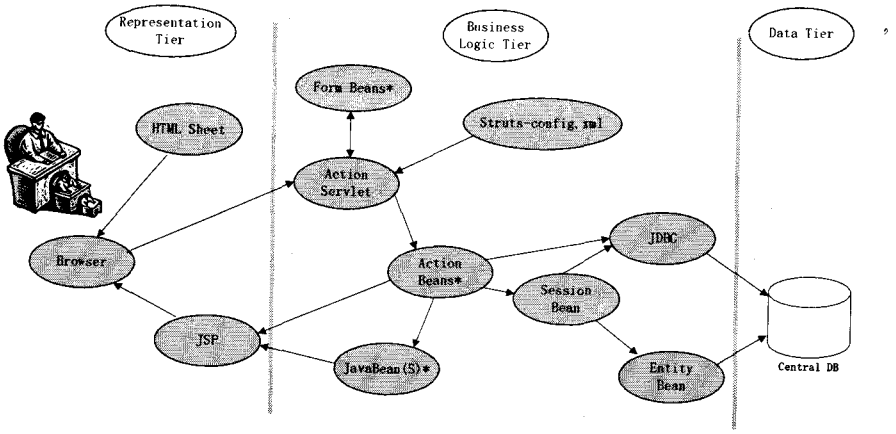
The IMS provides different-level access rights for different-level government users and ensures each user only can access information within his right scope. User rights are divided into four levels: the city government user, the district government user, the street government user, and the community user. Users in the same level have different roles, such as community official, community worker, data entry executive, etc. Different roles have different operation rights, such as querying, browsing, and modifying, etc. Which operation right a user has to some kind of information is decided by user role. System administrator sets user rights and user roles.

3) Log management

Logs are used to record the access and operation information of each user and help system administrators to maintain the system better. By utilizing the log function of DBMS, log management of the IECS can record the time, contents and objects of user operation, monitor the running of the system and ensure the security of the system.

Fig. 3. Information management subsystem

Fig. 4. Architecture of the IMS



3.4 Intelligent analysis subsystem (IAS)

The IAS is a decision-supporting platform based on the DW and the OLAP, developed for multi-dimensional analysis and the generation of statistics reports for government users. It takes data from the CDB, transforms and loads them into the data warehouse with ETL tool, and then performs multi-dimensional on-line analytical processing based on the user's analyzing requests. The IAS can help government officials make correct decisions about the development of a community and provide appropriate and timely services for citizens, enterprises or organizations. Therefore, it can improve the working effect of government.

1) ETL

ETL tool termly extracts data from the CDB, transforms and loads them into the DW in order to maintain data integrity and consistency. An ETL task can be broken into following four steps:

- Extracting data from the CDB and loading them into the buffer of data warehouse. ETL tool builds a materialized view log for tables from the CDB and extracts changed data incrementally.
- Transforming data structure. Data in the CDB and data in the DW are different in structure. Through building views over tables in the buffer, ETL tool transforms
- Updating the DW. ETL tool only updates changed data in the DW. For unchanged DW. data, it only modifies their time-stamps.
- Clearing data in the CDB. Because the CDB only stores current data, when data are out of date and extracted into the DW, they are cleared from the CDB immediately.

2) Subject of DW and multi-dimensional analytical model

In the Haishu District's e-government project, we divided community information into 10 subjects, as presented in Table 1.

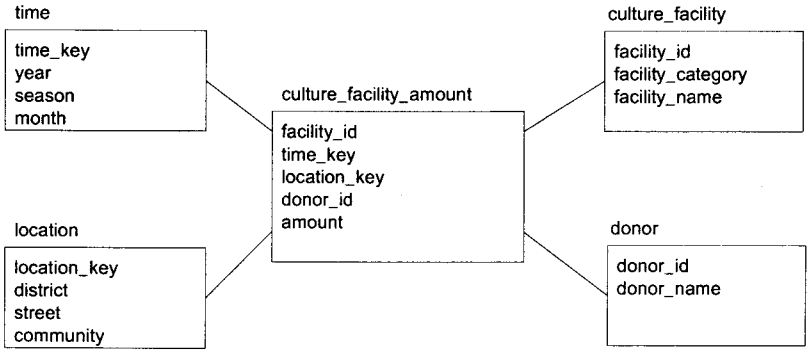
Table 1. Subjects of DW

Subject	Main information
Basic information	Brief, Memorabilia, Prize, Statute, Building, Meeting, Election, Finance, etc
Population and family	Permanent residents, temporary residents, children, etc
Employment statistics	Unemployment, re-employment, laid-off workers, re-employment of laid-off workers, hidden employment, etc
Social security	Poor, Handicapped, Old, Low income family, Donation, Funeral, Marriage, etc
Community security	Security guard, released man, arrested man, psychopath, house lease, etc
Organization and party	Political organization, women organization, sports party, etc
Environment and health	Health facility, Sanitation, infection prevention, etc
Civilization	Culture facility, School, Culture event, etc
Social service	Service item, Volunteer, Service facility (hotel, shop), etc
Family Planning	Childbearing age women, Married women, etc

Each subject has a number of information tables. Users can design one or several multi-dimensional analytical models based on these tables. The OLAP server provides the support of multi-dimensional analyzing. For example, if a user want to know the development information of culture facilities in a district, he can design the

following model (Fig .5) in *civilization* subject:

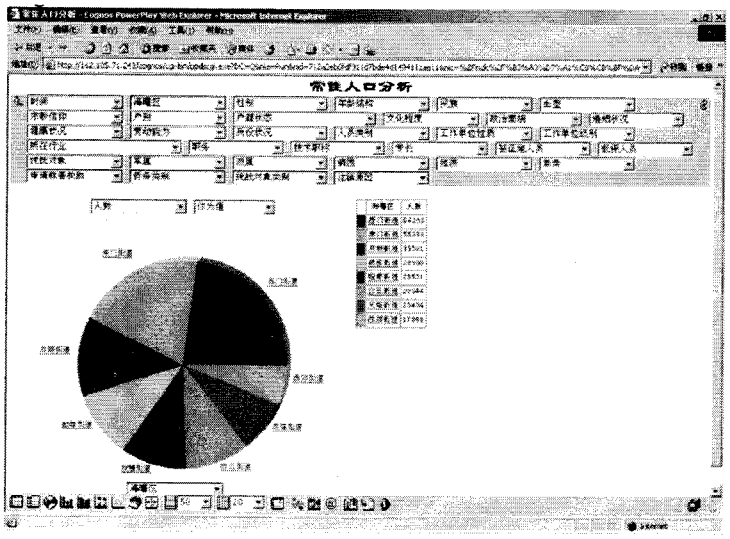
Fig. 5. A multi-dimensional analytical model on *civilization* subject



The analysis results can be represented by crosstabs, reports, histograms, and pie charts (Fig.6). In the Haishu District's e-government project, we design 74 multidimensional analytical models in 10 subjects in order to provide as many analyzing requests as possible for government users.

Just as the IMS does, the IAS also adopts three-tier architecture and provides user right management.

Fig. 6. Representation of analysis results in the IAS



3.5 Information service website (ISW)

The ISW is a citizen-friendly semantic network (Fig.7) that promulgates the management and service information and supports resource discovery [6]. We employ RDF as the main metadata description language and have developed many metadata specifications to make the e-community more useful, particularly for the citizens in the district who use the e-community. The ISW offers 16-category service information (e.g. medical care, shopping, hotel, finance, and training), a number of management information (e.g. government notice and public policy) and some analytical reports (e.g. the public security reports and the environmental health reports).

The ISW is an information service platform on which governments can promulgate service information and respond to outside users' s requests, and outside users can query service information and release new requests. Moreover, the ISW also loads valuable information into the CDB. The detailed information about the ISW can be seen in [6].

Fig. 7. Information service website – China 81890 Service Network

Ⅳ. Summary and future works

We have introduced the design and implementation of the IECS. Different with other e-government systems, the IECS has three distinct features as follows:

- The IECS is an integrated e-community system, which not only supports work and knowledge sharing, but also

stresses on knowledge acquisition and the integration of management, decision and service of a government.

- The IECS is an analytical e-community system, which can help government officials to make correct decisions about the development of a community by making use of DW and OLAP techniques, and provide appropriate and timely services for outside users.
- The IECS can facilitate the re-engineering of government process and service innovation. The integration of management, decision and service can change effectively the business process of a government, improve the effect of service, and finally accelerate process re-engineering and service innovation.

By taking advantage of IECS, the practical project designed for Haishu District of Ningbo, China works well, and now, the district government is enhancing the efficiency of work and inhabitants has got more benefits from the e-community. Of course, to realize fully functional IECS, there is still many works to do, future works include:

- How to add geographical information into the IECS and offer the query service of geographical information.
- How to further integrate management information and service information and improve the efficiency of extracting, loading, querying and browsing.
- How to implement data mining and provide more valuable management information for governments and more appropriate service information for outside users.

References

1. McClure, D. L. Statement of David L. McClure, U.S. General Accounting Office, before the Subcommittee on Government Management, Information and Technology, Committee on Government Reform, House of Representatives. http://www.gao.gov. 2000.
2. Duncan Aldrich, John Carlo Bertot, and Charles R. McClure. E-Government: initiatives, developments, and issues. Government Information Quarterly 19 (2002) 349-355.
3. Paul T. Jaeger, Kim M. Thompson. E-government around the world: Lessons, challenges, and future directions. Government Information Quarterly 20 (2003) 389-394
4. Karen Layne, Jungwoo Leeb. Developing fully functional E-government: A four stage model. Government Information Quarterly 18 (2001) 122-136.
5. Paul Raj Devadoss, Shan L. Pan, and Jimmy C. Huang. Structurational analysis of egovernment initiatives: a case study of SCO. Decision Support Systems 34 (2002) 253-269
6. Youlin Fang, Weihua Zhang, Dongqing Yang, and Shiwei Tang. Towards a Semantic ECommunity. EGOV 2003, LNCS 2739, pp. 476-479, 2003.
7. Amy Jo Kim, Community Building on the Web: Secret Strategies for Successful Online Communities, by, Paperback, (2000).
8. Albolino S., Morici R., Schael T., and Sciarra G. A KM solution for the empowerment of Local Authorities, international Workshop "Knowledge Management in e-government-KMGov2002 "(2002).
9. Roberta Morici, Eugenio Nunziata, and Gloria Sciarra. A Knowledge Management System for E-government Projects and Actors. KMGov 2003, LNAI 2645, pp. 304-309, 2003.
10. Monge Alvaro E., Elkan Charles P. An Efficient Domain-independent Algorithm for Detecting Approximately Duplicate Database Records. Department of Computer Science and Engineering, University of California, 1997.
11. http://www.81890.gov.cn