Bing Xu *Editor*

2012 International Conference on Information Technology and Management Science (ICITMS 2012) Proceedings

Springer

2012 International Conference on Information Technology and Management Science (ICITMS 2012) Proceedings

Bing Xu
Editor

2012 International Conference on Information Technology and Management Science (ICITMS 2012) Proceedings

Editor
Bing Xu
Chongqing Three Gorges University
Wanzhou
Chongqing
China, People's Republic

Printed on acid-free paper

Preface

Welcome to 2012 International Conference on Information Technology and Management Science—ICITMS2012. The number of submissions has increased and more scholars such as the leading researchers, engineers and scientists in the domain of interest around the world participate actively in and start to enjoy this conference. We trust deeply that we will step on a higher floor of this conference this year in Chongqing.

As a global academic symposium on Information Technology and Management Science, the aim of ICITMS2012 is to facilitate the communication of academics between domestic and abroad, to construct an international communication platform, and also to exhibit the new accomplishments of Information Technology and Management Science.

We express our thankfulness to the authors who submitted their papers. We also thank the keynote speakers for their contributions to the conference. Our thanks also go to the Review Panel, to the members of the Program Committee, to the members of the Organization Committee, and to all others who supported the conference professionally. We are confident that this effort will make contributions to further scientific and industrial collaborations among the conference attendants and among all participating countries in the future. Moreover, what is specially appreciated is the great help and support of the sponsoring institutions, namely, Springer, Case Western Reserve University, and Cedarville University.

We hope and believe that everybody here will have an academic enjoyment during this conference in Chongqing.

ICITMS2012 Organization Committee

About This Book

Proceedings of the 2012 International Conference on Information Technology and Management Science compiles revised and extended articles written by prominent researchers participating in Information Technology and Management Science.

The text offers the latest information on advances in Information Technology and Management Science as an excellent reference work for researchers working in the area of Information Technology and Management Science.

Contents

Part III Theory and Practice of Modern Management

Part I
E-Business Engineering and Management

Research on China's Portal Websites' Competition Based on Social Network Methods

Ye Wang and Ting Wu

Abstract This paper analyzed and studied the competition of the portal websites in China through social network methods. Design the questionnaire survey, use the results of 113 persons to Sina, Sohu, Netease, Tencent and other big portals' channels as the sample data, apply the social network analysis software of Ucinet, analyze the competitive social network structure between any two of the big portals from the scale, centrality and utilize a clustering analysis of Concor for the samples, and finally make a conclusion on the competitive relationships of China's portal websites according to the results of the analysis of news channel.

Keywords Social network methods • Portal websites • Clustering • Ucinet

1 Overview of the Portal Competition by Social Network Methods

The content of social network is that each of the actors has more or less the relationship with other actors; social network analysis is to establish a model of these relationships, trying to describe the structure of group relations and study the influence of the structure to the group function or individuals within the group. The social network graph uses nodes to represent transmitters and the receivers of messages, uses the arrows to indicate the directions of information transmissions, uses to the thickness of the edges to mean the frequency of information transmission or the amount of information. The portal website is to provide certain types of comprehensive Internet information resources and provide the applications about information services. In China, the typical big portals are Tencent, Sina, Netease,

Y. Wang (✉) • T. Wu
School of Management, Tianjin Normal University, Tianjin, China
e-mail: wangyetjcn@hotmail.com; wutingwood@sina.com

Sohu, Phoenix, Chinanet, Xinhuanet, TOM and Yahoo. The article selected eight
popular channels of news, video, entertainment, finance, sports, fashion, military
and reading to conduct a questionnaire survey; got the most competitive portals for
each channel. Due to limited space in the article, only the data of news channel was
discussed in the paper, and other channels' data was similar to it and didn't need
too much discussion. Questionnaire data was used as the samples for analyzing,
transformed into the adjacency matrix. The paper applied social network methods
and the software of Ucinet, conducted the systematic analysis in terms of scale,
centrality, clustering and finally all summarized a detailed analysis from the results.

2   Social Network Analysis of the Portal Competition

2.1   *Competitive Data of Portals*

Deal with the competitive relational data of portals; get the competitive relationship
matrix of news channel. It is a non-symmetric multi-matrix, according to the design
of the questionnaire; each value of the channel matrix represents the quantity of
(first choice, second choice). For example, the value "10" of the point (Sina, Sohu)
of news channel matrix in the first row second column indicates that the site of first
choice of 10 persons who read the news in the portal is Sina and second choice is
Sohu. Here, we only list the nine most important portals such as Sina, Sohu, Netease,
Tencent, TOM, Phoenix, Xinhuanet, Chinanet, Yahoo, and the rest of the small
portals or non-portals are summed up as "others", therefore there are 10 portals
for the survey.

2.2   *The Scale of the Competitive Network*

The scale of the network is the amount of all actors in the network, the size of the
portal competitive network is 10.

2.3   *The Competitive Network Graph*

The network graph is shown in Fig. 1. The nodes represent the portals; edges mean
the competitive relationships between the portals. The network graph is a directed
graph, the starting node of a edge is the site of first choice, the end node is the
second choice, and the edge is bidirectional or unidirectional, the bidirectional edge
indicates that the two nodes mutually are the first choice and second choice, and the
unidirectional edge indicates that one node is the first choice and the other is the
second choice.
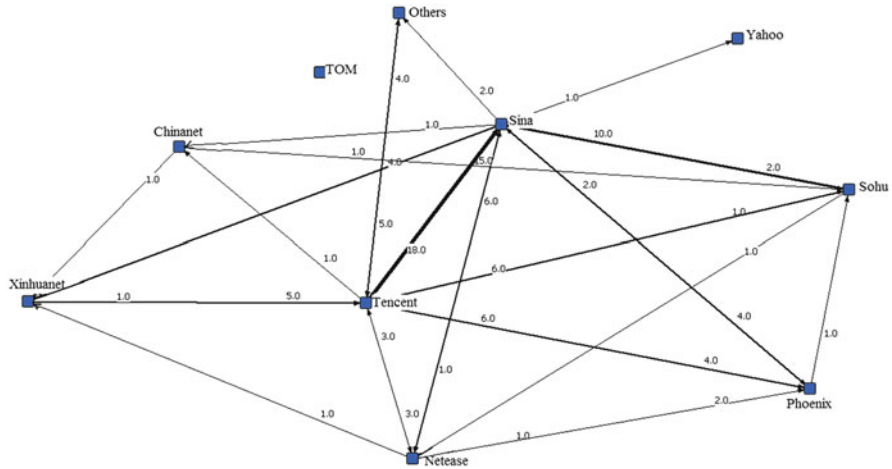
Fig. 1 Network of news channel

		1 Degree	2 NrmDegree	3 Share
1	Sina	46.000	28.395	0.291
4	Tencent	44.000	27.160	0.278
2	Sohu	19.000	11.728	0.120
3	Netease	13.000	8.025	0.082
6	Phoenix	13.000	8.025	0.082
7	Xinhuanet	11.000	6.790	0.070
10	Others	7.000	4.321	0.044
8	Chinanet	4.000	2.469	0.025
9	Yahoo	1.000	0.617	0.006
5	TOM	0.000	0.000	0.000

Fig. 2 Degree centrality of news channel

2.4   The Centrality Analysis of the Competitive Network

In the social network analysis, the centrality is divided into two categories, the absolute centrality (the points directly connected) and the relative centrality (standardized form) (Agarwal and Huan Liu 2008), the degree centrality is an indicator to evaluate the nodes important or not in the network, or how the power of the nodes in the network. Here, just list the degree centrality of news channel, as shown in Fig. 2. Seen from it, in the news channel competition relationship of 10 portals, the centrality degrees of Sina, Tencent, Sohu, Netease, Phoenix, and Xinhuanet are relatively high; they are in dominant positions and the most competitive of the news channel network.

Table 1 Unoverlapping clustering result

Number	Clustering members
1	TOM
2	Sina, Tencent, Others, Phoenix
3	Netease, Sohu, Xinhuanet, Chinanet, Yahoo

Table 2 Density matrix of news channel

	Clustering 1	Clustering 2	Clustering 3
Clustering 1		0.000	0.000
Clustering 2	0.000	5.000	2.000
Clustering 3	0.000	0.450	0.200

2.5 The Clustering Analysis of the Competitive Network

Next, the Concor analysis method (Shen Juan-hua 2007) is applied to non-overlapping clustering on the news channel. The specific result of clustering is shown in Table 1. Analyze the non-overlapping clustering result to the news channel. Its density matrix is shown in Table 2.

Derived from Tables 1 and 2, clustering 1 only has one member of TOM, its density is empty; members of clustering 2 are Sina, Tencent, Others and Phoenix, its density is 5.000, the internal structure is very closely linked; clustering 3 has Netease, Sohu, Xinhuanet, Chinanet and Yahoo, its density is 0.200, the internal structure is relatively loose. In the news channel, clustering 1 only has TOM separated into one class. TOM focuses on the young and trendy to provide services, including wireless Internet business and online advertising. Internet users have low awareness of TOM; it has not many direct effects on other large-scale comprehensive portals. The four members of the clustering 2 are closely linked, indicating a strong competition between them. For the news channel, the propaganda and marketing strategies of Sina, Tencent and Phoenix bring a high CTR for themselves, but there are still a part of Internet users who will watch news by other ways, for example, many people are accustomed to watch news on TV. Therefore, Sina, Tencent, Phoenix and Others can become a class of highly competitive portals. The mutual influence of members of clustering 3 is weak, mainly because they are similar awareness in the hearts of Internet users. In order to better development, these portals must enhance their visibility and take some unique marketing ways so as to create their own unique advantages, in order to gain more page views and users, and their competitiveness will be greatly enhanced.

3 Conclusion

This article analyzed the competition of 10 portals by social network analysis method, conducting the analysis of their scale, density, centrality, clustering, showing some structural features of the portals' channels from the perspective of

social networks, and finally obtained the relationships between the clusterings of the portals. However, due to the time limit, this paper selected only 10 portals which were the most commonly used by Internet users as questionnaire data, there were quite differences with the all portal data on the Internet, the relation matrices were limited to 1-mode. There were many inadequacies in the process of the research; only one analysis software was selected to analyze the data. The study depth of the social network analysis of the portal competitive relationship is also yet to be developed.

References

Agarwal N, Huan Liu (2008) Blogosphere: research issues, tools, and applications. SIGKDD Explor 10(1): 33–41

Shen Juan-hua (2007) Clustering analysis based on the correlation matrix and the mixed exponential distribution. J Suzhou Univ 8:3–8

Synergy Field and Its Application in Economic Space of Electronic Commerce

Dihui Lai

Abstract From the perspective of complexity theory and gauge theory, the paper puts forward the conceptual framework of "synergy field" and introduces "field" as the measure to depict the complexity of economic space. It studies the dynamic response characteristics of synergy field under the catastrophic state, and provides positive analysis. It also reveals the deep operating mechanism in the catastrophic process of economic space system of electronic commerce after the catastrophe happened, which contributes to grasping the E-commerce system's evolution condition and lays the foundation for further theoretical research and practical application.

Keywords Synergy field • Economic space • Electronic commerce • Catastrophe

1 Introduction

Past researches have revealed the doublet features of the process of nature or social economic life (Scully and Svidzinsky 2010). Economic space system is the product of the interactive movement of natural science and social science, and the coupling of several complicated processes. The self-organizations of subsystems in economic space system take place all the time, which reflects the duality (humankinds, organizations, etc.) of the subjects (humankinds, organizations, etc.) and objects (trust, controls, etc.) The substantive character of both economic activities and quantum mechanics is the doublet. Only because there is the identity, does it become possible to set up the trans-subjective connection between the subject-object duality in economic activities and wave-particle dualism of quantum

D. Lai (✉)
School of Economics and Management, Tianjin Institute of Urban Construction,
Tianjin 300384, People's Republic of China
e-mail: ldhredarmy@vip.sina.com

mechanics. With regard to the nature of the field, there are many different ideas in academic circles. It is a special form of matters. The field theory of economic space is a new method and thinking for the research on the systems in economic space and the process of its self-organizing, and in fact it is a synergetics research with the instruction of synergetics methodology. Therefore, this paper raises the concept of "synergy field", which is a special type and space Synergy field theory is a dynamical model about various kinds of particles systems.

The fundamental unit of the process of field in economic space is particle, which is the basic element to maintain and promote the process and development in economic space of electronic commercial system. The model of synergy field, which is shown in Fig. 1.

There are four kinds of particles: technology, economy, market and institution. he interaction of the homogenous particles or heterogenous particles forms particle field; by means of the dynamic function, interrelated function and fluctuated function, depending on the crossed integration of certain material stream, energy stream and information stream, the basis kernel of self-organizational evolution, the foundation of quantum field, is formed. In the scope of quantum field, the particles and field quanta can transform mutually. The nature of particles can be gained from that of quantum field, and in this space-time condition, technology particles are the excitation state of technology filed, economy particles are the excitation state of economy field and system particles are the excitation state of system field. Part of the excitation state has been discussed in domestic documents and proved by demonstrations (Chen et al. 2005). The quantum field amplifies the particle force field to dimension of space complexity, which highlights the synergistic effect of the interior and exterior environment of E-commerce system, and the infinite degree of freedom that influences E-commerce is canonically quantized, thus the rudiment of synergy field is formed.

Synergy field stresses the synergy nature of field and it has the symmetry of space-time of the fundamental particles, that is, the particular combinations (the formations of certain structures) of the wave function density of the corresponding particles arise in the form of relativistic invariant in the field's Lagrangian function density (Yang 1987). The symmetry can be classified into global symmetry and partial symmetry according to the boundary it involves. Global symmetry is the symmetry that all the particles transform in the same state, and it can be found in the violent changes, such as economic crisis and natural disasters, of the interior and exterior environment of E-commerce. Local symmetry refers to the independent change of the particles in the field, such as the upgrade of the homepage developing tools and the spread of electronic authentication technology, etc. Global symmetry is slow relaxation variable, and it makes the field leave stable state for unstable state and exhibits undamped phenomenon when instability appears because of disturbance in the synergy field. While partial symmetry is fast relaxation variable and serves as the damping function that makes the field back to the stable state. With the rise and development of the ordered structure of the field, the two kinds of symmetries are interrelated and restricted and exhibits a synergy movement. The synergy movement manifests symmetry breaking of the field, and the existence of the breaking is the theoretical foundation of the existence of catastrophe in synergy field.

2 Application of the Dynamical Catastrophe of the Synergy Field

2.1 Formulation of the Field Equation Model of Synergy Field

Traditional field can only describe the object's undulatory property, but not its corpuscular property. To make field able to describe the microcosm, it must be quantized. What synergy field describes is a complicated system that is infinite degree of freedom. It divides the three-dimensional space of the synergy field into infinite volume elements ΔVi, and defines the coordinate No. i.

$$\varphi_i(t) \equiv \frac{1}{\Delta V_i} \int_{(\Delta V_i)} d^3x \varphi(\vec{x}, t) \tag{1}$$

Therefore, synergy field φ is like numerable degree of freedom system, the time derivative of $\varphi_i(t)$ is:

$$\varphi_i(t) \equiv \frac{1}{\Delta V_i} \int_{(\Delta V_1)} d^3x \frac{\partial}{\partial t} \varphi(\vec{x}, t) \tag{2}$$

Lagrangian variable L can be rewritten:

$$L = \int d^3x \mathcal{L}(\varphi(x), \partial_u \varphi(x)) \to \sum_i \Delta V_i \overline{\mathcal{L}}_i(\varphi_i(t), \varphi_i(t), \varphi_{i\pm t}(t), \ldots) \tag{3}$$

Different $\varphi_i(t)$ are all independent degree of freedom representing the relative independence of the elements that influence electronic commercial activities. $\varphi_{i\pm t}(t)$ can adjacent to φ_i is origin of volume coordinates, they symbolize particle effect means of different describing stage in quanta, $\nabla\varphi$ is formed from it. Because $\dot{\varphi}_i(t)$ only appear in $\bar{\mathcal{L}}_l$, regular momentum is

$$M_i(t) = \frac{\partial L}{\partial \dot{\varphi}_i(t)} = \Delta V_i \frac{\partial \overline{\mathcal{L}}_i}{\partial \dot{\varphi}_i(t)} \tag{4}$$

Hamiltonian is obtained by Legendre transformation, and this provides conditions for searching for the possible energy to promote the evolution of the electronic commercial system.

$$H = \sum_i M_i \varphi_i - L = \sum_i \Delta V_i (\pi_i \dot{\varphi}_i - \bar{\mathcal{L}}_i) \tag{5}$$

In order to better characterize the status of synergy field, the particles in the field are set to be interdependent on time, and the quantum state not interdependent on time, thus Hamiltonian could be interchanged with the Heisenberg equation, which would meet the canonical commutation relation:

$$[\pi_i(t), \varphi_j(t)] = -i \frac{\delta_{ij}}{\Delta V_i} \tag{6}$$

Then φ_i and M_i would meet the Heisenberg equation.

$$\varphi_i(t) = i[H, \varphi_i(t)], M_i(t) = i[H, M_i(t)] \tag{7}$$

To adopt continuum and spread quantization method to the level of field, there is:

$\pi_i(t) \rightarrow \pi(\vec{x}, t)$, $\pi(\vec{x}, t)$ is a conjugated momentum to $\varphi(\vec{x}, t)$, $\pi(\vec{x}, t) \equiv \frac{\partial \mathcal{L}(\varphi, \partial_u \varphi)}{\partial \dot{\varphi}(\vec{x}, t)}$ its average value in volume element is:

$$\frac{1}{\Delta V_i} \int_{(\Delta V_i)} d^3 x\, \pi(\vec{x}, t) = \frac{1}{\Delta V_i} \int_{(\Delta V_i)} d^3 x \frac{\partial \mathcal{L} Y \varphi, \partial_u \varphi Y}{\partial \dot{\varphi}} = \frac{1}{\Delta V_i} \cdot \Delta V_i \frac{\partial \overline{\mathcal{L}}_i}{\partial \dot{\varphi}_i(t)} = \pi_i(t) \tag{8}$$

In continuous limit, Hamiltonian (5) of synergy field turn into:

$$H = \int d^3 x \mathcal{H}\left(\pi(\vec{x}, t), \varphi(\vec{x}, t)\right), \mathcal{H} = \pi \dot{\varphi} - \mathcal{L} \tag{9}$$

$\mathcal{H}$ is hamiltonian density, expression (5) is being isochronous commutative changed:

$$[\pi(\vec{x}, t), \varphi(\vec{x}', t)] = -i\delta^3\left(\vec{x} - \vec{x}'\right) \tag{10}$$

When transfer to continuous limit, $\delta_{ij} / \Delta V_i \rightarrow \delta^3(\vec{x} - \vec{x}')$, Heisenberg equation (7) turn into:

$$\dot{\varphi}(\vec{x}, t) = \mathrm{i} \left[H, \varphi_i(\vec{x}, t) \right], \dot{\pi}\left(\vec{x}, t\right) = \mathrm{i} \left[H, \pi \left(\vec{x}, t\right) \right] \tag{11}$$

The regularization of synergy field is built on the expressions (9), (10), and (11).

The establishment of the synergy field equation illustrates that the field has gauge invariance, that is the evolution of the E-commerce system is not supposed to change with the change of time-space scale on the time-space point, and the material's degree of freedom inside and outside the system does not influence the evolution of the system; the evolutional state of the system shows the integral cooperativity. The core of the equation is to reveals that there is symmetry breaking in economic space: The state of the inner energy of the E-commerce system is fully symmetrical, and the priority of the systematic ground state is the same, but actually only one of these systematic ground states is prominent; although the inner symmetry itself is not destroyed by outside elements, its expression mode is restrained. Understanding from the perspective of mathematics, spontaneous symmetry breaking indicates that there is asymmetric solution for the equation, and with the spontaneous breaking, numerous ground state with inner symmetry produce energy that may determine the direction of the system's evolution to be well-prepared for the transformation of the system.

2.2 Catastrophe Mechanism of Synergy Field in Economic Space

There is a events of catastrophe in the process of the self evolution of synergy field. The catastrophe of the synergy field is the process of the partially or completely fundamental catastrophe of the field's basic elements and structure, and this catastrophe is based on the change of the field itself or the mutual function between the field and its environment. Before the spontaneous symmetry breaking takes place, the slight rise of the synergy field is under the critical state: the field keeps a pre-established change based on the present stable interface–the stable state, and the initial points of the direction of the rise and fall are both the original stable state of the field. Once the rise and fall become intense, the range higher than the unstable critical point, it is not until the bifurcation arises and the symmetry breaking takes place, would there be qualitative changes, and the evolutional process of the field becomes transition from gradual catastrophe.

The transition of synergy field from stable state to tremendous rise and fall indicates that the field is a system of potential energy, and the state of bifurcation can be described by its equilibrium surface, the aggregate of all the points which meet the need that the first derivative (or two first partial derivatives) of the

Fig. 2 Catastrophe surface
of synergy field
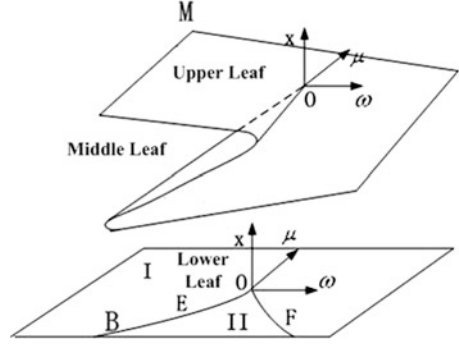
potential-energy function is zero. Suppose the potential function of the cusp
catastrophe model is:

$$\mathcal{V}(x) = x^4 + \mu x^2 + \omega x \tag{12}$$

Formula (12) is corresponding to the two one-dimensional successive dynamical
systems that control the parameters. The phase space is three-dimensional, and
the function of synergy field x is the state variable, the variation level of the
field's environmental elements μ and the level of the mutual function between the
environmental elements and the field ω is given by the control variable.

$$4x^3 + 2\mu x + \omega = 0 \tag{13}$$

From Fig. 2, the set of singular points is the subset of M to meet the equation
$6x^2 + \mu = 0$ and the conclusion that the set of bifurcation points B is $8\mu^3 + 27\omega^2 = 0$
can be gained by solving formulas (12) and (13) simultaneously.

The set of bifurcation points B is the projection of the crumple of profile of
equilibrium M on the controlling plane below, and it divides the controlling plane
into two ranges: I and II. When the controlling locus, shaped by the quantum field
in synergy field, it enters the range II from range I via various route.

2.3 The Response of Catastrophe to Synergy Field

Catastrophe can be divided into two categories according to the causes: endogenous
catastrophe and exogenous catastrophe; the occurrence of catastrophe is sure to
influence the dynamic evolution of the field. De Menezes and other scholars
(2004) divide the complicated system into two types with impulse law as the
limit: endogenous drive and exogenous drive (De Menezes and Barabasi 2004).
Endogenous catastrophe is produced self-organizing by the interior dynamic evo-
lution; in the process of catastrophe, the time-space symmetry makes prediction
possible. While exogenous catastrophe comes from outside the field, and it is
unpredictable. Meanwhile, the response of field to endogenous catastrophe and

exogenous catastrophe usually presents different regularity. Sornette and other scholars (2003) put forward the mean field theory to discuss the response of the activity of mnemonic complicated system to endogenous impact and exogenous impact (Sornette and Helmstetter 2003), which establishes the foundation for further theoretical research on synergy field catastrophe dynamics.

The research of Zhang J.W. and other scholars (2002) points out the multiple and fractal nature of business volume of the Internet (Zhang et al. 2002). Internet is the technological foundation and core resource for the expansion of E-commerce activities. Therefore, this paper adopts multifractal random walk to describe the activity of E-commerce synergy field. Suppose the earning ratio of synergy field on the time dimension Δt $r_{\Delta t}(t) = \epsilon(t)\sigma_{\Delta t}(t) = \epsilon(t)e^{\omega_{\Delta t}(t)}$, $\epsilon(t)$ is the white Gaussian noise, $\epsilon(t)$ and $\omega_{\Delta t}(t)$ not related, $\omega_{\Delta t}(t)$ is approximate Gaussian distribution,

$f(\omega) = \frac{1}{\sqrt{2\pi}\sigma_\omega} e^{\frac{(\omega - \mu_\omega)^2}{2\sigma_\omega^2}}$, μ_ω and σ_ω^2 are the mean value and variance of ω.

$$E[e^{2\omega}] = \int_{-\infty}^{+\infty} e^{2\omega} f(2\omega)d\omega = e^{2u_\omega + 2\sigma_\omega^2} \tag{14}$$

Field of mathematical expectation of endogenous catastrophe:

$$E_{endo}[\sigma_{\Delta t}^2(t)|\omega_0] = E_{endo}[e^{2\omega_{\Delta t}(t)}|\omega_0] = e^{2E_{endo}[\omega_{\Delta t}(t)|\omega_0] + 2Var[\omega_{\Delta t}(t)|\omega_0]} \tag{15}$$

The model of multifractal random walk is shown by the style of memory kernel function:

$$\omega_{\Delta t}(t) = \mu_{\Delta t} + \int_{-\infty}^{t} \eta(\tau)K_{\Delta t}(t - \tau)d\tau \tag{16}$$

Simultaneous Eqs. (15) and (16):

$$\int_0^{+\infty} K_{\Delta t}(t) K_{\Delta t}(t + \tau)d\tau = \lambda^2 \ln\left(\frac{T}{t + e^{-3/2}\Delta t}\right) \tag{17}$$

With Fourier transform and convolution theorem, after the reverse catastrophe, the above formula becomes

$$K_{\Delta t}(t) \sim K_0 \sqrt{\frac{\lambda^2 T}{t}}, \Delta t \ll t \ll T \tag{18}$$

Suppose at t = 0, synergy field is interference with a endogenous catastrophe or exogenous catastrophe whose intensity is ω_0, the perturbation noise of the field is from $\eta(\tau)$ to $\eta(\tau) + \omega_0\delta(\tau)$, take it to formula (18): $\omega_{\Delta t}(t) = \omega_0 K_{\Delta t}(t) + \omega_{\Delta t}(t)$, so the mathematical expectation of exogenous catastrophe of conditional volatilities is:

$$E_{exo}[\sigma_{\Delta t}^2(t)|\omega_0] = E_{exo}[e^{2\omega_{\Delta t,endo}(t)}|\omega_0] = \overline{\sigma_{\Delta t}^2}e^{2\omega_0 K_{\Delta t}(t)} \tag{19}$$
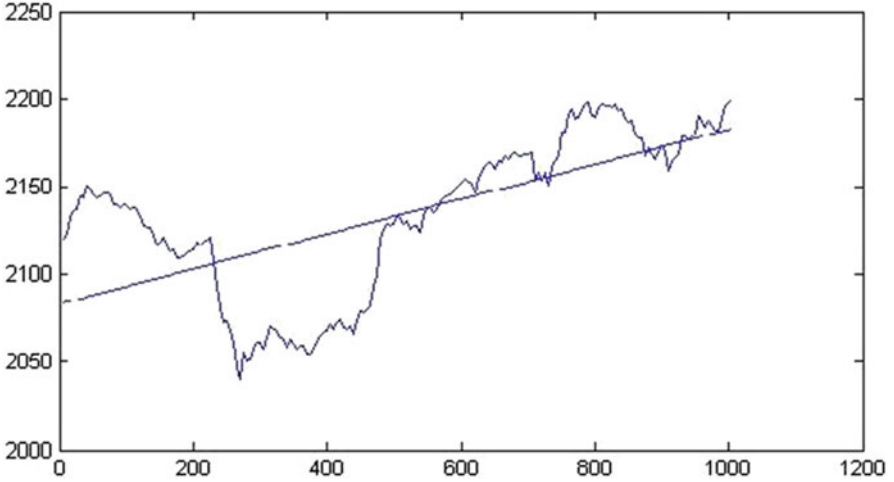
Fig. 3 The influence of exogenous disturbance on electronic information index

$\overline{\sigma_{\Delta t}^2}$ is sample variance, and it has nothing to do with t. When formula (20) is substituted in formula (19), it can be find out the margin difference of conditional volatilities, $\Delta t \ll t \ll T$.

$$E_{exo}\left[\sigma_{\Delta t}^2(t)|\omega_0\right] - \overline{\sigma_{\Delta t}^2} = \overline{\sigma_{\Delta t}^2}e^{2\omega_0 K_{\Delta t}(t)} - 1 = 2\overline{\sigma_{\Delta t}^2}\omega_0 K_0\sqrt{\lambda^2 T/t} \qquad (20)$$

Formula (20) reveals that the influence of exogenous disturbance on the field is linear, and the scale relationship between margin difference and t is not much obvious.

Shanghai stocking index of electronic information and the index of electronic payment reflect the overall developing situation of domestic E-commerce business, and they can reveal the real conditions of field when it is influenced by endogenous disturbance and exogenous disturbance. To analyze the response of the above the index to the raise of interest rate in November 2010 and the United States' announcement of new round reimpression of bank notes. Calculating the daily volatility after the occurrence of exogenous disturbance with 5 min high frequency data, it reaches the sequence of average volatility $\overline{\sigma_{\Delta t}^2}$ and conditional volatility $E_{exo}[\sigma_{\Delta t}^2(t)|\omega_0]$ when Δt is 1 day, and finally reaches the accumulated allowance of conditional volatility. By scale distribution of the 4 accumulated allowances analyzed, it can be found out that they meet formula (21):

$$\int_{s=0}^t \left(E_{exo}\left[\sigma_{\Delta t}^2(t)|\omega_0\right] - \overline{\sigma_{\Delta t}^2}\right) ds \sim t^{1/2} \qquad (21)$$

As a comparison from Figs. 3 and 4, there was similar volatility relationship between the accumulated allowance after the Wenchuan Earthquake in 2008 and t for
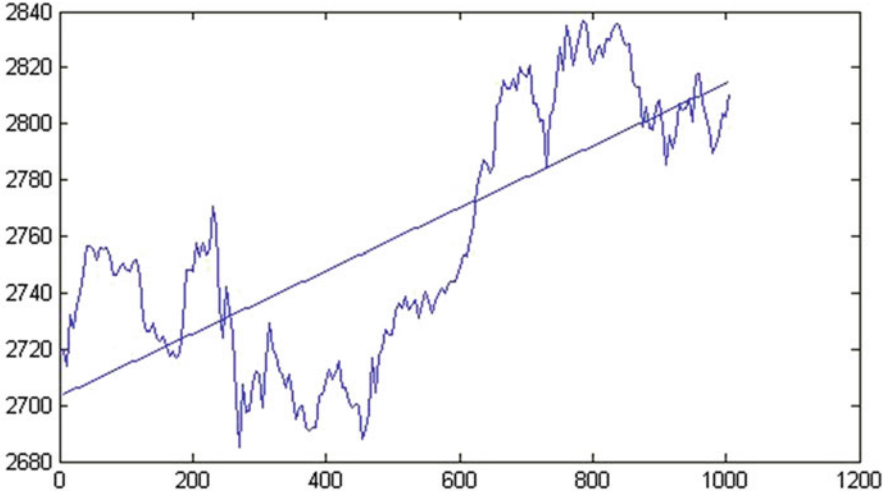
Fig. 4 The influence of endogenous disturbance on electronic information index

the electronic information index; it is thus clear that the disturbance on stock market brought by Wenchuan Earthquake also belongs to exogenous disturbance and the new round reimpression of bank notes belongs to the endogenous disturbance.

Introducing the parameter $q = 1/2 \ln \overline{\sigma_{\Delta t}^2}$ and substituting to formula (15),

$$E_{endo}[\sigma_{\Delta t}^2(t)|\omega_0] = \overline{\sigma_{\Delta t}^2} \exp\left[\frac{2(\omega_0 - \mu_{\Delta t})\mathrm{Cov}[\omega_{\Delta t}, \omega_0]}{Var[\omega_0]} - \frac{2\mathrm{Cov}[\omega_{\Delta t}(t), \omega_0]^2}{Var[\omega_0]}\right]$$

$$= \overline{\sigma_{\Delta t}^2}\left(\frac{T}{t}\right)^{\alpha(q)+\beta(t)} \tag{22}$$

$\alpha(q) = \frac{2q}{\ln(Te^{\frac{3}{2}}/\Delta t)}, \beta(t) = 2\lambda^2 \frac{\ln(te^{\frac{3}{2}}/\Delta t)}{\ln(Te^{\frac{3}{2}}/\Delta t)}$, When $\Delta t < t \ll \Delta t e^{\frac{|q|}{\lambda^2}}, \beta(t) \ll \alpha(q)$, formula (20) turns to:

$$E_{endo}[\sigma_{\Delta t}^2(t)|\omega_0] \sim t^{-\alpha(q)} \tag{23}$$

Research results have been achieved by applying multifractal random walk theory to the research on financial market (Saakian et al. 2011). The field's responsive pattern to endogenous disturbance can be proved by high frequency data as well. To calculate the volatility $\sigma_{\Delta t}^2(t)$ with the time interval Δt, the sampling frequency of the data should be higher than $1/\Delta t$ set the high frequency data of the rate of return to be $r_{\delta t}(i \delta t)$, and when $t = n\Delta t(1, 2 \ldots)$, the volatility is

$$\sigma_{\Delta t}^2(t) = \sum_{i=1}^{\Delta t/\delta t} [r_{\delta t}(t - \Delta t + i\delta t)]^2 \tag{24}$$

while the average volatility $\overline{\sigma_{\Delta t}^2}$ is the sample average of $\sigma_{\Delta t}^2(t)$. Any s corresponds to an disturbance whose size is $e^{2s}\overline{\sigma_{\Delta t}^2}$; $E_{endo}[\sigma_{\Delta t}^2(t)|\omega_0]$ will be reached if all the moments at which the disturbance whose size is $e^{2s}\overline{\sigma_{\Delta t}^2}$ occurs, and shift the time of subsequent conditional volatility to the moment when the disturbance occurs. Writer did research on the condition of conditional volatility attenuation when Δt is 40 min and 1 day with the 5 min high frequency data of Shanghai stocking index of electronic information and the index of electronic payment. Then writer worked out the scale index that different s corresponds to, discovered that there is very good linear relationship between $\rho(s)$ and s, the slope varies with Δt What is interesting is that the volatility of a multi-trader micro-market based on the Ising Model presents the similar pattern (Zhou and Sornette 2007).

By combining the field theory in natural science and the research on E-commerce system, this paper puts forward the theory of "synergy field", which is creative exploration of the modal of E-commerce system. It is discovered by research that there is interaction between various particles in synergy field; synergy field possesses the time-space symmetry of basic particles. With the combined influence of global symmetry and partial symmetry of field, it presents complicated movement prospect of catastrophe. Field equation explains the gauge invariance of synergy field and further reveals the profound connotation that the spontaneous breaking is an important cause of the catastrophe of E-commerce system.

3   Conclusions

This paper further describes the basic process of the catastrophe of synergy field. Bifurcation is the important premise and fundamental mechanism for catastrophe. In the aspect of dynamics, this paper analyses the response of field to the endogenous and exogenous disturbances, describes the synergy field activity of Economic Space by adopting multifractal random walk, obtains quantitative describing model of the response made by the field facing different kinds of disturbances, and do some empirical researches by making use of index related to the E-commerce of Shanghai Stock Exchange.

References

Chen FX, Zhang L, Hu J (2005) The establishment and demonstration about the field of knowledge diffusion. Stud Sci China 23:253–257
De Menezes MA, Barabasi A-L (2004) Fluctuations in network dynamics. Phys Rev Lett 92:02870
Saakian DB, Martirosyan A, Hu CK, Struzik ZR (2011) Exact probability distribution function for multifractal random walk models of stocks. Europhys Lett 95:28007

Scully MO, Svidzinsky AA (2010) The lamb shift – yesterday, today, and tomorrow. Science 328(5983):1239–1241

Sornette D, Helmstetter A (2003) Endogenous versus exogenous shocks in systems with memory. Phys A 318:577–579

Yang CN (1987) Square root of minus one, complex phases, and Erwin Schrdinger. In: Kilmister CW (ed) Schrdinger: centenary celebration of a polymath. Cambridge University Press, Cambridge, pp 53–64

Zhang JW, Chen HW, Chen CJ (2002) Study on source of multifractal nature of internet traffic. J Softw 13:470–474

Zhou WX, Sornette D (2007) Self-organizing Ising model of financial markets. Eur Phys J B55:175–181

Study of Micro-Blog by Self-Efficacy-Based Value Adoption Model

Ge Zhu, Feifei Yang, and Tingting Chen

Abstract As an emerging mobile internet application, micro-blog has been developed from 2009 in China. This paper studies the factors that influence college students on adopting the micro-blog through the self-efficacy-based value adoption model (SVAM). This paper collects 457 valid questionnaires, analyzes the structural equation modeling by LISREL. The hypothesis model is confirmed in which include seven latent variables: Micro-blog's Self-efficacy (MS), Functional Value (FV), Social Value (SV), Emotional Value (EV), Perceived Value (PV), Attitude (AT) and Intention (AI). The results indicate that the model has a higher explanation (64%) for adoption intention of micro-blog than pure value adoption model (VAM) or other models. It cannot be ignored that the influence of self-efficacy on consumer's attitude.

Keywords Micro-blog • User acceptance • SVAM • Self-efficacy

1 Introduction

Micro-blog is an emerging mobile Internet service mainly provided by Sina in China, which is also regarded as a modified Twitter. Registered users posting, sharing, and obtaining instant information within 140 characters of text based on their individual network community, as well as users's image, music, video files and location-based information. In China, micro-blog is called Weibo in Chinese and it is becoming very popular among youth people. According to the 29th Statistical Survey Report on Internet Development in China released by CNNIC (China Internet Network Information Center), by the end of 2011, the Internet users

G. Zhu (✉) • F. Yang • T. Chen

School of Information Management, Beijing Information Science & Technology University, Beijing, People's Republic of China

e-mail: zhuge01@gmail.com; eagle9711@hotmail.com; CTting@gmail.com

in China break through 500 million, and the users of micro-blog reached 250 million which increased 290.0% compared with the previous year and the proportion of the micro-blog up to 48.7%. Besides, Since the mobile applications produced for various platforms to access micro-blog, the micro-blog mobile phone user also made a contribution in the fast growing micro-blog user process. The proportion of micro-blog mobile phone user increased from 15.5 to 34% during the year 2011. In China, there are many Chinese-based micro-blog services including Sina, Tencent, Sohu and NetEase. Sina micro-blog service is one of the most representatives which occupies 57% of the whole micro-blog users, 87% of the whole total micro-blog activities (CNNIC 2011), and has more than 300 million registered users as of February 2012.

Micro-blog has been deeply influencing the mode of transmission of information even China society. However, what factors have influenced the adoption or rejection of micro-blog in youth people? This paper focused on college students to investigate the impact of consumers' self-efficacy and perceived value on their attitude and intention of using micro-blog service based on the social cognitive theory and the value model.

2 Literature Review

TAM, TPB and VAM are regarded as the mainly models of researching the consumers' acceptance of Internet and mobile Internet services. According to the investigation, the previous empirical research on consumers' acceptance adopted TAM frequently which proposed by Davis et al. (1989). In TAM, it considered two factors, Perceived Usefulness (PU) and Perceived Easy of Use (PEOU), to analyze the acceptance and adoption of new technology in an organizational setting (Davis et al. 1989; Van der Heijden 2004). After that, researchers added other influence factors on the basis of TAM and then proposed Extended TAM. Moreover they discussed the issues of consumers' acceptance on emerging IT/IS technology as well, which achieved good results (Moon and Kim 2001; Venkatesh et al. 2002; Chan and Lu 2004). However, in recent years, when applied TAM to consumer market where consumers can decide their purchasing behavior, its explanation of consumers' adoption is not good, which reveals the limitation as user adoption model (Hsu and Chiu 2004).

In comparison, the TPB model, which is also developed from the TRA model like TAM, achieves a better effect when researching consumers' acceptance and adoption on a new product or service in the consumer market (Pavlou 2006; Taylor and Todd 1995). When studying the same object, TPB model performs much better than TAM in explanation of consumers' adoption intention. Besides, Decomposed TPB and Extended TPB also show better performance in explanation of consumers' adoption intention than pure TPB model (Chau and Hu 2001).

According to the value theory, Kim et al. proposed the consumer value acceptance model (VAM), and conducted the empirical research on the mobile

Internet (Kim et al. 2007). The result showed that consumers' perceived value (including perceived benefits and costs) had a significant impact on consumers' adoption intention. Moreover, according to the comparative studies, the explanation of consumers' intention of VAM was far higher than the TAM (Kim et al. 2007). Similar VAM was also taken by Kleijnen et al. to study the customers' acceptance on mobile service (Kleijnen et al. 2007). However, the results showed by the two literatures were not satisfied in interpreting the adoption intention of consumers with $R^2 = 0.36$ and $R^2 = 0.39$. Zhu et al. proposed the Self-efficacy-based Value Adoption Model (SVAM) and studied the consumers' acceptance on new mobile applications such as mobile payment and mobile auction (Zhu et al. 2010, 2011). The results showed that the explanation of consumers' adoption intention was much higher than the pure value which reached to 0.71.

Therefore, this paper will adopt SVAM to study the micro-blog consumers' acceptance considering with the basic characteristics of micro-blog, and confirm the model with structural equation modeling.

3 Model Assumptions

Value driver as a basic variable is often used to study the consumer selection and decision issues. Creating value is often considered as the mission and goal of enterprises or organizations. According to Zeithaml's definition, perceived value means consumer make a comprehensive evaluation of the product utility based on the perceived benefit and cost (Zeithaml 1988). Sheth, Newman and Gross studied the consumers' behavior of smokers and stated that the selection of consumers was on the basis of multiple-value function which included functional value, social value, emotional value, perceived value and conditions value (Sheth et al. 1991). Kim et al. defined the perceived value as benefit and cost (Kim et al. 2007). Sweeney and Soutar put forward the framework of four-dimensional theory: functional value, perceived barrier, emotional value and social value (Sweeney and Soutar 2001). Taking into the account of the cost of mobile internet access continue to decrease, the cost of micro-blog access is no longer a main obstacle to mobile users (Zhu et al. 2010). Therefore, this paper will set functional value, social value and emotional value as the pre-variable of perceived value.

Function value means the effective perception to the functionality, practicality, natural attribute of alternatives (Zeithaml 1988). This is similar to the perceived usefulness of the TAM theory, but the latter is proposed based on the technology acception in organizational setting. This paper summarizes that the functional value of micro-blog mainly comes from its value of personal media platform, social networking platform, instant communication tool and its entertainment features. Besides, as the platform of posting, sharing and obtaining information, micro-blog has a better performance of instantaneity and interactivity than the traditional media and even the other network media. Therefore, the assumption and functional value scale are proposed in FV1-6 as follows.

H1:   The functional value is positively related to the perceived value.
FV1: Micro-blog has the function and value of entertainment.
FV2: In short, micro-blog is a valuable media platform.
FV3: Micro-blog makes it easy to find people who have the common interests and concerns as themselves.
FV4: Micro-blog makes it easy to find and participate in the current hot topics discussions.
FV5: Micro-blog makes it easy to share graphical emoticons or attach own image, music, video files.
FV6: Micro-blog has strong functional value.

According to the definition given by sociology of consumption, consumption is not only the traditional meaning that purchases products but also includes the consumers' self-expression, for example, the status symbol. Sweeney and Shouter believed that the social value meant the ability that product enhanced the self-concept in social context (Sweeney and Soutar 2001). Venkatesh also added the social value factor in TAM to analyze the views of other people for consumers using a new system. The results showed that people preferred to experience the products or services that can win the appreciation of other people. Therefore, as a new network service, the social value of micro-blog will influence its value of perception. The assumption and social value scale are proposed as follows.

H2:   The social value is positively related to the perceived value.
SV1: People using micro-blog can get more respect in society.
SV2: People using micro-blog generally have a good social image.
SV3: People using micro-blog can get more appreciation than unregister users.
SV4: In short, registering and using micro-blog is fashion and trend.

Previous studies show that there is a close contact between emotion and behavior. The emotion value is an attribution of product that can evoke people emotional status (Sheth et al. 1991). Compared with the external motivation of functional value, emotional value emphasizes more on the internal emotional motivation. In the work-related environment, the usefulness is a key factor that affect users' acceptance. However, in the personal consumer market, emotional value was proved to be much explanatory on reflecting the willingness of acceptance (Van der Heijden 2004). So, the assumption and emotional value scale are proposed as follows.

H3:   The emotional value is positively related to the perceived value.
EV1: Micro-blog can help people eliminate solitude and loneliness.
EV2: Micro-blog can release people's stress and anger emotion.
EV3: Micro-blog can help people find the sense of group belonging and identity
EV4: Micro-blog can help people find the sense of achievement when the number of followers continues to increase.
EV5: In short, micro-blog can satisfy people's emotional demand to some degree.

Referring to the Kim et al. (2007) and Zhu et al. (2010), this paper defines the perceived value as consumer comprehensive evaluation of the perceived benefit

and cost. VAM shows that the perceived value significantly impacts on the mobile Internet acceptance intention (Kim et al. 2007). Here, this paper assumes that perceived value of micro-blog will affect the consumers' attitude. The specific assumptions and perceived value scale are shown as follows.

H4:   The perceived value is positively related to attitudes.
PV1: Compared with cost, micro-blog can bring more benefit.
PV2: Compared with efforts, it's worth to register a micro-blog account.
PV3: In short, micro-blog can create a variety of value for consumers.

Self-efficacy is the core concept of social cognitive theory. Bandura (1986) defined it as a self-judgment to the ability of organizing and implementing behaviors in order to achieve the expected effect (Bandura 1986). Schwarzer and Schmitz (2005) stated self-efficacy was self-perception to their disposing capacity (Schwarzer and Schmitz 2005). Self-efficacy has a better performance in explanation when applied to specific object in different field. Zhu et al. (2010, 2011) have confirmed that self-efficacy of mobile applications has a comprehensive influence on perceived value and attitude. Here, this paper puts forward the micro-blog self-efficacy (MS) and define it as consumers' self-perception to their micro-blog using ability. As a new application of mobile Internet, the study on self-efficacy of micro-blog is still in blank, so the scale design can be given by referring to the literatures which studied the similar application. According to the framework of SVAM, the self-efficacy of micro-blog will affect the perceived value as follows.

H5: The self-efficacy of micro-blog is positively related to the functional value.
H6: The self-efficacy of micro-blog is positively related to the social value.
H7: The self-efficacy of micro-blog is positively related to the emotional value.
H8: The self-efficacy of micro-blog is positively related to attitude.

The settings of measured variable for micro-blog self-efficacy (MS) are proposed as follows.

MS1: I am confident that I can post my micro-blog successfully.
MS2: I am confident that I can master the skills of using micro-blog.
MS3: I am able to upload photos and video by micro-blog.
MS4: In short, I am confident that I can flexibly use micro-blog.

It has been confirmed by a large number of literatures that attitude can effectively explain the adoption intention as an intermediate variable, especially based on the TRA and TPA model (Van der Heijden 2004; Moon and Kim 2001; Venkatesh et al. 2002; Chan and Lu 2004). Zhu et al. (2010) pointed out that when people using the information technology voluntarily, they might judge the intention of using by attitude; when using the information technology coercively, they probably decided their intention of using by considering the organizational standards though it is under their individual control (Zhu et al. 2010). Thus, in the information service market where consumers can make choice freely, the attitude of consumers is an essential condition variable to explain the adoption intention. According to the

Fig. 1 Self-efficacy-based value adoption model (SVAM) and hypotheses
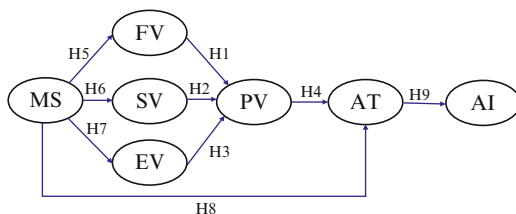
author's previous research, the attitude of consumers towards mobile commerce directly determines the adoption intention. Therefore, this paper assumes that attitude is the premise variable for the adoption intention.

H9: Attitude t is positively related to the adoption intention.

The measured variables for attitude (AT) and adoption intention (AI) are listed as follows.

AT1: I have positive attitude to the development of micro-blog.
AT2: I believe that micro-blog will be widely popularized in the future.
AT3: I don't care about micro-blog, and it's not necessary for me.
AI1: I may register a micro-blog account if I know it earlier.
AI2: If I have registered micro-blog, I will continue using it.
AI3: In short, I will use micro-blog in the future.

Based on the above analysis and assumptions, this paper proposed the Self-efficacy-based Value Adoption Model (SVAM) for micro-blog (Fig. 1).

4 Results and Analyses

The design of questionnaire adopts the seven scale of Likert. To ensure the reliability and validity of the questionnaire, subject scales refer to the measurement standard which adopted by relevant literatures. The questionnaire of this paper specifically refers to the author's previous research (Zhu et al. 2010, 2011). Some of questions are modified and augmented a little in order to match the micro-blog features. The analysis results show that the measures have good reliability that the coefficient of Cronbach's alpha is greater than 0.7. Besides, the average variances extracted also beyond 0.5 which indicates the latent variable has acceptable convergent validity.

We mainly investigated the college students and collected 457 valid question-naires both online and on-site. The sample statistics (see Table 1) shows that male is more than female, age distribution primarily concentrates in the age of 20–25, and most respondents are undergraduate degree. Besides, 65.2% respondents have registered a micro-blog account, but login by mobile phone only occupies 34.1% of the total investigation. Results indicate that micro-blog has a higher popularity among college students but lower percentage of registers login by mobile phone.

Table 1 Sample descriptive statistics

Measure	Item	Frequency	Percentage (%)
Gender	Female	181	39.6
	Male	276	60.4
Age	<20	27	5.9
	20–25	352	77.0
	>25	78	17.1
Education	Undergraduate	371	81.2
	Master	50	11.0
	Ph.D.	36	7.8
Micro-blog	Register	298	65.2
	Unregister	159	34.8
Total		457	100

Fig. 2 Analysis results of structure equation model of SVAM
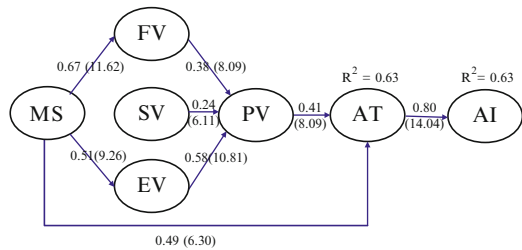
In this paper, the proposed hypotheses are tested by structural equation modeling (SEM) by using the LISREL 8.80. A set of goodness-of-fit indices of the structural model are obtained such as: Chi-S/df,RMSEA,NFI,NNFI,CFI,IFI,GF are corresponsively 3.81, 0.079, 0.86, 0.88, 0.89, 0.89, 0.84. Compared with the recommended threshold (Zhu et al. 2010), although part of the corresponding value are little lower than the experienced recommended threshold, there is no index can evaluate the model alone. Therefore, the model should be accepted as comprehensively evaluating the various goodness-of-fit indices.

It should be noted that the initial model contains self-efficacy (MS) to social value (SV) correlation validation, but its coefficient is not ideal, and even has the negative effects to the whole model. So in the further validations, this paper deletes the relationship between self-efficacy (MS) to social value (SV) correlation. And the final analysis results show that this model is feasible.

As shown in Fig. 2, in addition to H6 with a T-value below 2, the other hypotheses are all set up. The functional value, social value and emotional value of micro-blog are important roles in perceived value formation, the path coefficient are so high that reach to 0.38, 0.24 and 0.58 respectively. Therefore the assumptions of H1, H2 and H3 are valid. The emotional value of micro-blog has a greater impact on perceived value comparing with the usefulness of its function and the appreciation brought by its social value. Therefore, the emotional need fulfilled by micro-blog is more attractive. Besides, the perceived value is positively related to the attitude with the

Fig. 3 Analysis results
of VAM model
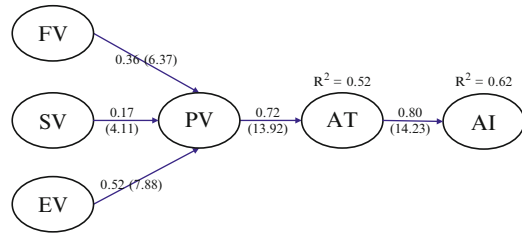
coefficient of 0.41, which indicates assumption H4 is valid. Attitude significantly impacts on intention that coefficient reaches 0.80. Thus, the assumption of H8 is valid.

The results show that the micro-blog's self-efficacy has a significant influence on functional value and emotional value other than social value, the path coefficients reached to 0.67 and 0.51 respectively and the T-value are also beyond the experienced recommended threshold. Thus, the assumptions of H5 and H7 are valid. Moreover, the self-efficacy of using micro-blog directly impact on the attitude to micro-blog as well, path coefficient reaches to 0.49, the T-value equals to 6.30, which means assumption of H8 also is proved.

In order to investigate the influence of self-efficacy of SVAM, this paper provides a comparing VAM model in Fig. 3. Results show that SVAM is better than VAM in explaining of adoption intention with $R^2 = 0.63$ and $R^2 = 0.62$. Especially, self-efficacy have a strong compact on attitude with $R^2 = 0.63$, which is more higher than VAM with $R^2 = 0.52$.

In conclusion, the explanation (63%) of SVAM for adoption intention of micro-blog is higher than the model on the basis of ATM and TPB as well as VAM. The main reason is that the SVAM is based on self-efficacy. The evaluation of consumer's ability in using micro-blog influences on both their judgment to the value of micro-blog and the attitude of micro-blog. Therefore, self-efficacy-based value adoption model is feasible and reliable in explaining the adoption intention of micro-blog as well as other mobile internet applications. Besides, the simple but operable model structure is also an advantage of SVAM.

References

Bandura A (1986) Social foundations of thought and action: a social cognitive theory. Prentice-Hall, Englewood Cliffs

Chan SC, Lu M (2004) Understanding internet banking adoption and use behavior a Hong Kong perspective. J Global Inf Manage 12(3):21–43

Chau PYK, Hu PJ (2001) Information technology acceptance by individual professionals: a model comparison approach. Decis Sci 2(4):699–719

CNNIC (2011) China internet development statistics report. China Internet Network Information Center (CNNIC)

Davis FD, Bagozzi RP, Warshaw PR (1989) User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 35(8):982–1003

Hsu MH, Chiu CM (2004) Internet self-efficacy and electronic service acceptance. Decis Support Syst 38:369–381

Kim HW, Chan HC, Gupta S (2007) Value-based adoption of mobile internet: an empirical investigation. Decis Support Syst 43:111–126

Kleijnen M, de Ruyter K, Wetzels M (2007) An assessment of value creation in mobile service delivery and the moderating role of time consciousness. J Retail 83(1):33–46

Moon JW, Kim YG (2001) Extending the TAM for a world wide web context. Inf Manage 38(4):217–230

Pavlou P (2006) Understanding and predicting electronic commerce adoption: an extension of the theory of planned behavior. MIS Q 30(1):115–143

Schwarzer R, Schmitz GS (2005) Perceived self-efficacy and teacher burnout: a longitudinal study in ten schools, Research paper. Freie Universitat Berlin, Berlin

Sheth JN, Newman BI, Gross BL (1991) Why we buy what we buy: a theory of consumption values. J Bus Res 22:159–170

Sweeney JC, Soutar G (2001) Consumer perceived value: the development of multiple item scale. J Retail 77(2):203–220

Taylor S, Todd PA (1995) Understanding information technology usage: a test of competing models. Inf Syst Res 6(2):144–176

Van der Heijden H (2004) User acceptance of hedonic information systems. MIS Q 28(4):695–704

Venkatesh V, Speier C, Morris MG (2002) User acceptance enablers in individual decision making about technology, toward an integrated model. Decis Sci 33(2):297–315

Zeithaml VA (1988) Consumer perceptions of price, quality and value: a means-end model and synthesis of evidence. J Mark 52(2):2–22

Zhu G, Sangwan S, Lu T-J (2010) A new theoretical framework of technology acceptance and empirical investigation on self-efficacy-based value adoption model. Nankai Bus Rev Int 1(4):345–372

Zhu G, Ao S, Dai J, Li C (2011) Value-based adoption model of mobile payment and empirical research. J Beijing Inf Sci Technol Univ 126(1):42–48

Consumer Satisfaction, Trust and Loyalty in C2C Electronic Commerce

Liefa Liao and Jianjin Zhong

Abstract This paper focuses on loyalty, trust, satisfaction and the way in which they interact in marketplace and electronic seller in consumer to consumer electronic commerce. It was found that buyers increase loyalty to electronic marketplace (EM) when they perceived satisfaction to EM and increase loyalty to seller when they perceived trust in seller, marketplace's satisfaction is influenced by trust in marketplace and trust in seller is influenced by satisfaction to seller, while trust in marketplace do not significantly positive influence buyer's loyalty to marketplace directly, satisfaction to seller do not significantly positive influence loyalty to seller directly. We also find that buyer's trust in marketplace also builds buyer's trust in sellers in marketplace. In addition, buyer's trust in sellers increases buyer's satisfaction to marketplace, buyer's loyalty to seller increases buyer's loyalty to marketplace. Seven of the nine hypotheses are supported in the research.

Keywords Electronic marketplace • Trust • Satisfaction • Loyalty

1 Introduction

In recent years, the transaction in consumer to consumer (C2C) electronic commerce have risen sharply and most recent academic research only shows the significant correlation between the consumers' satisfaction and trust (McKnight et al. 2004; Carlos et al. 2006), consumers' trust and loyalty (Reichheld and Schefter 2000),

L. Liao (✉)
School of Computer Science, JiangXi University of Science and Technology, Ganzhou, China
e-mail: liaolf@126.com

J. Zhong
School of Economics and Management, JiangXi University of Science and Technology, Ganzhou, China
e-mail: zjj699@126.com

consumers' satisfaction and loyalty (Balabanis et al. 2006; Danaher et al. 2003; Anderson and Srinivasan 2003) in B2C e-commerce. The relationship between satisfaction, trust and loyalty was not explored fully in C2C marketplace in China.

In this paper, we explore the relationship of consumers' satisfaction, trust, loyalty to seller and marketplace in C2C transaction, and the paper is organized as follows. We propose the study model and hypotheses based on the related research firstly, and then we discuss the research methodology and process. Finally, we discuss the research direction based on the results of the study.

2 Research Hypothesis

Figure 1 shows the research framework which includes three parts: The relationship (includes H1, H2, H3) between consumer's satisfaction (short as PSS), trust (PST) and loyalty (PSL) to seller; The relationship (includes H4, H5, H6) between consumer's trust (PMT), satisfaction (PMS) and loyalty (PML) to market; The correlation between consumer's perception to market and consumer's perception to seller (includes H7, H8, H9).

2.1 Consumer's Satisfaction, Trust and Loyalty to Seller

Kennedy et al. (2001) shows a relationship between consumer satisfaction and consumer trust with the increases of satisfaction. Therefore, the trust to seller is a consequence of the sellers having the ability to satisfy the requirements of consumers. In the e-commerce, trust of the user increases when the user perceived that the system was usable and have a positive influence on user satisfaction (Carlos et al. 2006).

In C2C e-business, although EM provides security mechanism for transaction security, moral hazard behaviors still exist because there are a variety of sellers in
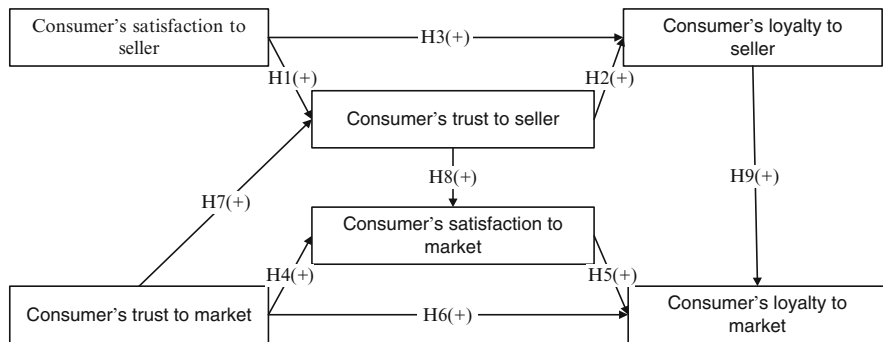
Fig. 1 Research framework

the transaction. Consumers are difficult to confirm whether the sellers are integrity to complete the transaction, existing some uncertainty and risk. For this reason, the consumer's trust to seller mainly comes from the perception of transaction process. Studies show that the consumer's satisfaction is based on the consistency between expectation and actual perception. And the expectation comes from a series of service process emotions. When consumers perceive the sellers' ability, integrity and kindness during the trade, it will increase consumer trust. Thus, satisfaction can be used as tool influencing the trust (Oliver 1980). Based on this, we assume that:

Hypothesis 1: The consumer's satisfaction to seller positively influences its trust.

The foundation of human being's interaction and exchange is trust (Gundlach and Murphy 1993). In the same time, the most important psychological state is trust in an uncertain environment (Rousseau et al. 1998). However, uncertainty is an important characteristic of e-commerce. Consumers are lack of related network technical knowledge, related information about how the electronic market agent works. Research shows that the key reason why consumers are opposed to purchase goods via the Internet is lack of trust. And the trust is considered as the center attribute of the relationship's generation, formation and maintain between two sides (Sirdeshmukh et al. 2002). Due to lack of trust, 49% of consumers oppose to do online shopping (Carlos et al. 2006). Consequently, the consumer's trust in C2C e-commerce turns out to be a key factor influencing online behavior (Jarvenpaa et al. 2000).

In the B2C e-commerce, consumer's loyalty directly depends on consumer's trust, which is the key factor of prompting shopping online (Quelch and Klein 1996). In the C2C marketplace, the seller is anonymous, and seller's credit rating is in different level. Beside, there exists reputation manipulation phenomenon in feedback system (Dellarocas 2006). At the same time, consumers switching to other sellers have to face transaction risks, uncertainties and other perceptual switching costs, comparing to current seller, result in higher costs converting to other sellers. Based on this, we assume that:

Hypothesis 2: The consumer's trust to seller positively influences its loyalty.

The previous positive experience will promote consumer's satisfaction and the consumer's satisfaction will also promote the loyalty (Hong-Youl and Perks 2005). Satisfaction with the product or service makes consumers assume that the subsequent transaction will be effective, which will make consumers more reluctant to change to other sellers. Based on this, we assume that:

Hypothesis 3: The consumer's satisfaction to seller positively influences its loyalty.

2.2 Consumer's Satisfaction, Trust and Loyalty to Market

C2C marketplace provides support service for buyers and sellers during transaction, and its ability and credibility are very important factors for its success. Marketplace can enhance its credibility through multiple channels and provide useful information

for transaction. According to relational exchange theory, the consumer's previous trust directly affects its satisfaction perception (Sirdeshmukh et al. 2002). Therefore, the cumulative perceived trust will influence the perceived satisfaction, reduce anxiety, and reduce transaction risks and uncertainties (Gwinner et al. 1998). In the same time, people will match up their beliefs and actions based on the cognitive consistency theory. When short of trust beliefs, people's satisfaction tendency will become lower. Based on this, we assume that:

Hypothesis 4: The consumer's trust to market positively influences its satisfaction.

In the traditional market, satisfied customers are more inclined to repeat buy goods and provide word of mouth (Yoon and Kim 2000; Anderson and Sullivan 1993). And the study on American consumer satisfaction index also indicates that consumer's loyalty is influenced by the consumer's satisfaction positively (Mithas et al. 2006). Although in the traditional market, a number of consumers will be loyal to the service providers with which they are not satisfied, for instance, consumers can't change to use the unsatisfied services in the bank and communication industry. But in e-commerce, unsatisfied consumers can easily choose to change service provider, service channel and other websites just clicking the mouse.

In C2C e-business, the consumer's transaction process includes six steps. First, visit electronic market website and inquire the commodity. Second, check the information of related seller's credit. Third, complete purchase process through electronic market website system. Forth, complete payment process through the electronic payment system. Fifth, receive the commodity send by the online seller. Sixth, reply feedback information or complaint to the marketplace after receive commodity. In this process, the marketplace provides security mechanism and related services (e.g. feedback system, third-party payment services) for the trade. Consumers need to experience the transaction to perceive the satisfaction to electronic marketplace. On the other side, when switching to other market, consumers not only need to visit the website system, but also need to perceive the services offered by the website. It takes lots of time, thus increasing the consumer's switching costs and reducing the switching possibility. Based on this, we assume that:

Hypothesis 5: The consumer's satisfaction to market positively influences its loyalty.

Hypothesis 6: The consumer's trust to market positively influences its loyalty.

2.3   The Mutual Relationship Between Consumer's Perception to Market and Consumer's Perception to Seller

Electronic marketplace acts as the third party between buyers and sellers, whose main functions are to apply Internet technology into collecting, handling and

spreading information, increase the trust and reduce the information asymmetry between the buyers and sellers (Pavlou et al. 2007). In the transaction process, the marketplace provides a reliable safety trading environment, builds fair and open rules and process, identifies, estimates and eradicates problematic trading parties and encourage friendly transaction norms (Pavlou and Gefen 2004). From the logic of transaction transfer (Stewart 2003), the trust to market affects the trust to seller. Based on this, we assume that:

Hypothesis 7: The consumer's trust to market positively influences consumer's loyalty to seller.

In addition to the seller's own effort, the market rule is also an important element in the establishment of consumer's trust to seller. And the consumer's trust to seller also illustrates the effectiveness of market guarantee mechanism. Consequently, it may cause the consumer's perceived satisfaction to market. Based on this, we assume that:

Hypothesis 8: The consumer's trust to seller positively influences the consumer's satisfaction to market.

In the trading process, consumers need to draw support from the online market services (e.g., QQ offer good communication channels). When choosing another marketplace, consumers may not acquire the function and information provided by the market platform. Therefore, the consumer's loyalty to market may be partly influenced by the market itself. Based on this, we assume that:

Hypothesis 9: The consumer's loyalty to seller positively influences the consumer's loyalty to market.

3 Research Design and Implement

3.1 Variable Measurement

Based on the actual situation for the C2C e-commerce and present research, we use the existing scale. Firstly, we analysis existing scale and translate them into Chinese, then we verify the consistency of the scale content with group discussions. The variable measure is a 5-item Likert scale for the study ranging from "total disagreement" to "total agreement". Customer's satisfaction and trust to seller is based on Gefen et al. (2003) and the loyalty to seller is based on Oliver (1980). The scale of consumer's satisfaction, trust and loyalty to market is similar with the seller's, only making proper adjustment for actual situation in the market. Table 1 shows the item load, variance explained, composite reliability, and discriminate validity of measurement. Therefore, the items corresponding to each variable could be creating an overall measure for each variable (Table 2).

Table 1 Questionnaire items

Latent variables	Items	Load	Variance explained	AVE	Composite reliability	Cronbach alpha
PSS	01. I am satisfied with the commodity	0.855	76.076	0.761	0.927	0.895
	02. My choice to buy the commodity was a wise one	0.891				
	03. My choice to buy the commodity was a correct one	0.884				
	04. I am satisfied with the transaction	0.859				
PST	05. I think the seller can complete the transaction in good faith after I purchase goods	0.804	71.76	0.718	0.91	0.868
	06. I think the seller can provide good service after I purchase goods	0.854				
	07. I feel relieved to trade with the seller after I purchase goods	0.872				
	08. I think the seller is reliable after I purchase goods	0.857				
PSL	09. Compared with other sellers, I always prefer to choose this one	0.878	76.453	0.765	0.907	0.846
	10. Compared with other sellers, I have a weakness for this one	0.88				
	11. Compared with other sellers, I always prefer to choose the goods provided by this one	0.865				
PMS	12. I am satisfied with the electronic marketplace	0.88	77.257	0.773	0.931	0.901
	13. My choice to the electronic marketplace was a wise one	0.898				
	14. My choice to the electronic marketplace was a correct one	0.899				
	15. I am satisfied with the transaction in the electronic marketplace	0.838				

Table 1 (continued)

Latent variables	Items	Load	Variance explained	AVE	Composite reliability	Cronbach alpha
PMT	16. I think the electronic marketplaceplace can complete the transaction in good faith	0.78	73.807	0.738	0.918	0.881
	17. I think the electronic marketplaceplace can provide good service	0.854				
	18. I feel relieved to trade in the electronic marketplacee	0.904				
	19. I think this electronic marketplace is reliable	0.893				
PML	20. Compared with other electronic marketplaces, I always prefer to choose this one	0.902	81.243	0.813	0.929	0.883
	21. Compared with other electronic marketplaces, I have a weakness for this one	0.895				
	22. Compared with other electronic marketplaces, I always prefer to choose the goods provided by this one	0.908				

Table 2 The related decimal of latent variables

	PSS	PST	PSL	PMT	PMS	PML
PSS	0.872					
PST	0.748	0.847				
PSL	0.488	0.549	0.874			
PMT	0.595	0.577	0.346	0.879		
PMS	0.598	0.692	0.394	0.728	0.859	
PML	0.470	0.530	0.587	0.487	0.594	0.902

3.2 Investigation Process

Firstly, we did pre-survey in two universities. Respondents (n = 50) are the students with C2C e-commerce transaction experience. In the process of per-survey, we adjusted the questionnaire expression, increased the clarity of research item and determined the final survey questionnaire. Secondly, we used a snowball method
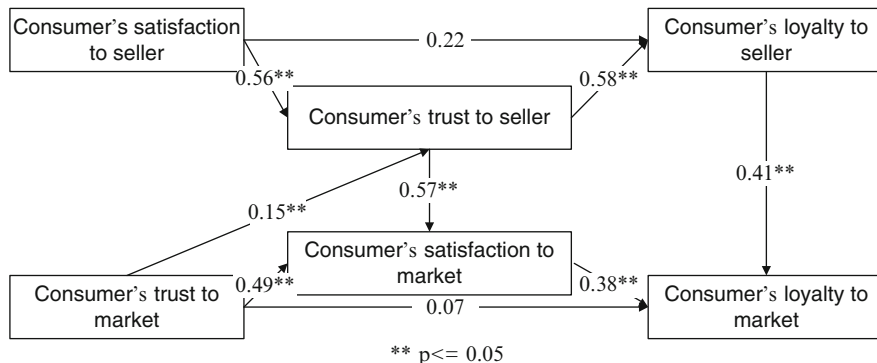
Fig. 2 Sizing chart of the relationship of model factors ($**p <= 0.05$, Remark: Chi-Square $= 393.19$, df $= 199$, P-value $= 0.00000$, RMSEA $= 0.066$, GFI $= 0.852$, NNFI $= 0.930$, CFI $= 0.940$, IFI $= 0.940$)

to implement the research, and the subjects completed the questionnaires based on their shopping experience. Finally, a total of 350 questionnaires was send, and 223 provided usable responses.

3.3   Results Analysis

Figure 2 shows that goodness of fit: $x2 = 393.19$; $P = 0.00000$; CFI $= 0.940$; NNFI $= 0.930$; RMSEA $= 0.066$; These statistics indicate that this model has good degree of fitness. The result shows that 7 of 9 hypotheses are significant and hypothesis 3 and 6 fail to pass significant test. Consequently, the consumer's satisfaction to seller is benefit for the building of the trust to seller, which further produces the loyalty to seller. The consumer's trust to market is benefit for the building of the trust to seller. However, it is different with our expectation that the consumer's satisfaction to seller has no significant relationship with the loyalty to seller. On the other hand, the trust to market influences its satisfaction. And the satisfaction to market influences its loyalty. The loyalty to seller is benefit for the loyalty to market. But, the trust to market has no significant relationship with its loyalty.

In terms of the model explanation capabilities, this model explains the consumer's trust to seller R2 $= 0.582$, explains the loyalty to seller R2 $= 0.315$, explains the consumer's satisfaction to market R2 $= 0.640$ and explains the consumer's loyalty to market R2 $= 0.502$. All these results show that the model can explain the loyalty of C2C e-business consumers in a better way. However, the value to explain the consumer's loyalty to seller is just 0.315, still needed to further explore other important factor affecting the loyalty.

4 Conclusion

The study verifies that the trust and satisfaction both play important roles on loyalty establishment in C2C e-commerce and the different effect on the establishment process of marketplace and seller loyalty. It also shows the mutual correlation between trust, satisfaction and loyalty in C2C e-commerce and explores forming mechanism of the three effectively.

First, consumer's satisfaction to electronic marketplace is the mediating variable between trust and loyalty. But, there is no obvious correlation between trust and loyalty. Second, consumer's trust to electronic marketplace is the mediating variable between satisfaction and loyalty. In electronic marketplace, consumers mainly acquire seller's information through feedback system and communication platform. And sellers are anonymous and can easily modify register name. Satisfactory transaction is difficult to ensure the sellers will complete the transaction with the trading rules in the future. What's more, consumers can seek other sellers easily through net platform and acquire their information of trade history and trade assessment information, which means switching cost for consumer are low. Thereby, it reduces the loyalty. Third, the positive correlation between consumer's trust to electronic marketplace and the trust to seller indicates that in the electronic marketplace with weak licensing laws, the improvement of electronic marketplace's credibility has a significant impact to increase consumer's trust to seller, and also proves the trust transfer in the marketplace. Fourth, the positive correlation between consumer's satisfaction to seller and the trust to market indicates that the improvement of electronic marketplace's credibility is an important factor to estimate market service quality and increase its satisfaction. Fifth, the positive correlation between consumer's loyalty to seller and the loyalty to market indicates that in addition to increase the consumer's satisfaction, electronic marketplace needs to maintain the satisfaction to seller. Otherwise, seller defection can lead to consumer defection and result in a decrease of electronic marketplace popularity. Finally, it can reduce the consumer's loyalty to electronic marketplace.

References

Anderson RE, Srinivasan SS (2003) E-satisfaction and e-loyalty: a contingency framework. Psychol Mark 20(2):123–138

Anderson EW, Sullivan MW (1993) The antecedents and consequences of customer satisfaction for firms. Mark Sci 12(2):125–143

Balabanis G, Reynolds N, Simintiras A (2006) Bases of e-store loyalty: perceived switching barriers and satisfaction. J Bus Res 59(2):214–224

Carlos F, Guinaliu M, Gurrea R (2006) The role played by perceived usability, satisfaction and consumer trust on website loyalty. Inf Manage 43(1):1–14

Danaher PJ, Wilson IW, Davis RA (2003) A comparison of online and offline consumer brand loyalty. Mark Sci 22(4):461–476

Dellarocas C (2006) Strategic manipulation of internet opinion forums: implications for consumers and firms. Manage Sci 52(10):1577–1593

Gefen D, Karahanna E, Straub DW (2003) Trust and tam in online shopping: an integrated model. MIS Q 2007(1):51–90

Gundlach GT, Murphy PE (1993) Ethical and legal foundations of relational marketing exchanges. J Mark 57(4):35

Gwinner KP, Gremier D, Bitner MJ (1998) Relational benefits in service industries: the customer's perspective. J Acad Mark Sci 26(2):101–114

Hong-Youl H, Perks H (2005) Effects of consumer perceptions of brand experience on the web: brand familiarity, satisfaction and brand trust. J Consum Behav 4(6):438–452

Jarvenpaa SL, Tractinsky N, Vitale M (2000) Consumer trust in an internet store. Inf Technol Manage 1(1/2):45–71

Kennedy MS, Ferrell LK et al (2001) Consumers' trust of salesperson and manufacturer: an empirical study. J Bus Res 51(1):73–86

McKnight DH, Kacmar CJ, Choudhury V (2004) Shifting factors and the ineffectiveness of third party assurance seals: a two-stage model of initial trust in a web business. Electron Mark 14(3):252–266

Mithas S, Ramasubbu N, Krishnan MS et al (2006) Designing web sites for customer loyalty across business domains: a multilevel analysis. J Manage Inf Syst 23(3):97–127

Oliver RL (1980) A cognitive model of the antecedents and consequences of satisfaction decisions. J Mark Res 14(3):495–507

Pavlou PA, Gefen D (2004) Building effective online marketplaces with institution-based trust. Inf Syst Res 15(1):37–59

Pavlou PA, Huigang L, Yajiong X (2007) Understanding and mitigating uncertainty in online exchange relationships: a principal–agent perspective. MIS Q 31(1):105–136

Quelch JA, Klein LR (1996) The internet and international marketing. Sloan Manage Rev 37(3):60–75

Reichheld FF, Schefter P (2000) E-loyalty: your secret weapon on the web. Harv Bus Rev 78(1):105–113

Rousseau DM, Sitkin SB, Burt RS et al (1998) Not so different after all: a cross-discipline view of trust. Acad Manage Rev 23(3):393–404

Sirdeshmukh D, Singh J, Sabol B (2002) Consumer trust, value, and loyalty in relational exchanges. J Mark 66(1):15–37

Stewart KJ (2003) Trust transfer on the world wide web. Organ Sci 14(1):5–17

Yoon S-J, Kim J-H (2000) An empirical validation of a loyalty model based on expectation disconfirmation. J Consum Mark 17(2/3):120–134

Design of Digital Campus Based on WebGIS

Pei Yan and Jiao Guo

Abstract WebGIS system on campus is an important subsystem of digital campus, and is the basic platform that provides the location-based information publishing and information services about e office and teaching. The concept of the digital campus is described. On the basis of analyzing the concept, characteristics and basic principles of WebGIS, the digital campus based on the WebGIS technology is proposed, and the overall framework, database and system capabilities are designed. To a certain extent, WebGIS-based digital campus system improves school management and efficiency and achieves the sharing for teaching information and resource.

1 Introduction

Digital Campus is digitalizing all information resources relating to teaching, research, management and life services through computer technology, network technology, communication technology on campus (Xie Hong et al. 2009); integrating all information resources based on scientific and standardized management to form a unified user management, integrated resource management and unified access control; building the school into a campus-oriented and society-oriented virtual university that is beyond time and beyond space (Yu Li and Wang Cheng 2004).

P. Yan (✉) • J. Guo
School of Computer Science and Engineering, Three Gorges University, Chongqing, China
e-mail: peisharp@gmail.com

Digital Campus system will become a new source of information on campus, in which any campus-related information will positioned and linked with the spatial data. Users will be able to query illustrated campus information, and access to the most direct effect.

Currently, most of the digital campuses just stay in the management information system construction stage. Management information systems focus only on the system attribute data input, processing, analysis and management, but almost put aside the relevant geographic information (Liang Xiaotian et al. 2007). It is not enough for comprehensively analyzing system information, especially geographic information. Many data required for the management on campus are built on the basis of spatial data. Establishing WebGIS-based geographic information system on campus will help school leaders and the relevant administrative departments to real-time, intuitively understand the situation around the building in the campus development and planning; parents of students can remotely understand the school through Internet; all students can look up some information on living and learning through the system. Therefore, the digital campus based on WebGIS has great significance for the development of universities.

2 WebGIS and Implementation Technology

2.1 The Concept of WebGIS

Geographic information system (GIS), is the database management system, in which the collection, storage, retrieval, analysis and processing spatial information is computerized (Aronoff 1989). With the aid of computers, GIS will visualize the information with spatial characteristic, provides more intuitive and clear expression for the information users, and has the very strong spatial analysis ability. WebGIS is the product of Internet and GIS technology (Luo et al. 2004). Through WWW, GIS has expanded its functions, and truly become a tool for public people. From any node of WWW, Internet users can browse the spatial data WebGIS site, make thematic map, do various space search and spatial analysis. As a new GIS based on Internet, WebGIS has some features.

1. WebGIS is based on Client/Server model, but most of the traditional GIS are as a stand-alone system.
2. WebGIS exchange information between users and servers with Internet, but even if the traditional geographic information systems can be as a server, its information exchange between the client and the server is usually limited on a single machine or local area network.
3. WebGIS is a distributed system, in which users and servers can be located in different locations and various computer platforms.

3 The Design of System Database

3.1 Collect the Data of System

Collecting data is the core of digital campus, which usually includes both graphics data and attributes data (Wang Jin et al. 2002). Based on digital digitizer function, graphic data collection digitizes existing map. Based on maps and tables of school buildings and site survey, attribute data collection is completed and feeded into related tables. Thus, the graphic information and attribute information input are completed.

3.1.1 The Main Content of Graphic Data

Graphic data required in digital campus that must express the entire campus and its geographical terrain includes the following contents (Li Hua et al. 2004). (1) The main streets and buildings around the campus, such as buildings, roads, playgrounds, green space, water area and so on. (2) Recent general planning on campus, including greenery, renovation or expansion of buildings, planning distribution of roads, teaching area, family area, entertainment area. (3) Distribution plans of school buildings, including office buildings, teaching buildings, laboratory buildings, libraries, university hospital, the Science Museum, canteens, dormitories, family wings and so on. (4) General pipeline maps, including the underground gas pipes, water pipes, sewage pipes, heating pipes and so on. (5) Road maps, including all existing roads, traffic facilities. (6) Distribution maps inside the building, such as the plans of office buildings, teaching buildings, libraries and the university Hospital and so on. The existing topographic maps, plans, and underground pipeline diagrams are digitized and entered into graphics database.

3.1.2 The Main Content of Attribute Data

Attribute database includes personnel files data, statistical data and building attribute information and other data. Staff information database includes the number, name, origin, date of birth, education, professional, job, job title, their sector, wages, home address and telephone number of staff. Student information database includes the number, student number, name, origin, date of birth, their class, professional, registration time, leaving time and dormitory address. Building information database includes building number, building character, number of floors, floor area, building area, date of completion, design units, construction units and prices. Research database includes the number of research projects, project title, project sources, project leader, funding, approval date, concluding date, project identification and Awards. Curriculum Schedule of the annual includes course number, course name,

Table 1 Data forms and formats in layers

Layer code	Layer content	Topological type	Major property project	Mark
BJJZ	Building of teaching and office	Plane	Building name, type, campus code	Yes
JGSS	Staff quarters	Plane	Building name, type, campus code	Yes
XSSS	Dormitory	Plane	Building name, type, campus code	Yes
DL	Road	Line	Road name, type, campus code	Yes
GSGD	Water supply pipeline	Line	Nubmer, campus code	Yes
PSGD	Drain pipeline	Line	Nubmer, campus code	Yes
WXXL	Cable line	Line	Nubmer, campus code	Yes
TXXL	Communication line	Line	Nubmer, length, campus code	Yes
DLXL	Power lines	Line	Nubmer, campus code	Yes
YDC	Playground	Plane	Nubmer, type, campus code	Yes
LVD	Green space	Plane	Nubmer, area, campus code	No
LUD	Street light	Point	Nubmer, campus code	No

teaching time, venue and instructors. Student score database includes the student number, name, course title, scores. Statistical database on campus includes campus area, building area, green area, number of faculty and students. Experimental device management database includes the device number, device name, department, purchase date, price, and usage.

3.2 The Design of Graphical Database

The system divides graphic database into layers based on the topics. The layers include the terrain layer, building layer, building floor plans, road layer, playgrounds layer, green layer, decorative layer, street layer, water supply/pipe layer, drainage/pipe layer, power facilities layer, communication facilities layer and so on.

Division in the graph database layer should also take into account the performance form and format of data in various layers. Functional elements of the campus are clear, each layer display in different colors. Parts of the layers in graphic database are described in Table 1.

3.3 The Design of Attribute Database

The system includes two types of attribute tables in database. One type of the attribute table can be mapized, which corresponds to the map entities, and store the properties of these maps; The other cannot be mapized, which does not contain geographic information and are same as general database files, mainly store students' information and multimedia property.

For mapized attribute tables, the system respectively designs the office house table, teaching house table, faculty dormitory table, student hostel table, road table, pipeline table, sports venues table, and green area table and so on, according to the hierarchical campus information programs.

For non-mapized attribute tables, the system designs the faculty information table, student information table, student dormitories table and multimedia file index table and so on, which are stored in the database.

3.4   Function Design of System

1. System management and data maintenance. System management includes system settings, user management and system information management, in which, user management is responsible for storing and managing the basic information of local or remote user, which provides legal calibration functions for the entire system. Data maintenance will complete add, delete, change, view, search and query, dynamic statistics, map making and charting capabilities for various attributes and geography graphics.
2. Map view and operation. Witch layered management model, the control point, underground pipe network, roads, buildings and other different types of map elements are managed. Map can zoom, drag, and if necessary, roam the map and measurement units switch, and provide print and distance measurement function. Layered management model can focus on some parts, browse and query hierarchically.
3. Online tracking analysis. Online real-time tracking of the system works by monitoring the location of objects based on information from the calculation of its actual geographical location, and publishing the results in the monitoring center network, real-time reflecting the results in the electronic map, and in the browser user can set the frequency of the target display and track paths.
4. Rich spatial query function. Users can complete rapid search, fuzzy query, and bidirectional search for attribute, spatial and multimedia information. It's visual and intuitive to display the results in graphical form.

     The system provides the following query modes: (1) simple and intuitive interaction mode. Query the attribute information of the ground-object selected in the different elements layer or in user-specified query scope. (2) Conditional query and fuzzy query. The system can complete the integrated query based on the conditions provided by users, and then locate on a map, show the query results; (3) Advanced query. Advanced query can be made in accordance with the location and search radius provided by users, and automatically detects the information around queried points, taking advantage of spatial analysis function of GIS server.
5. Line navigation services. Simulate roaming in a real environment, which shows the correct route on the map, according to locating the user's initial position and the needs of users.

4 Conclusions

With the constant expansion of universities, it is more difficult to fully grasp the details of the campus for the management sector. WebGIS technology that will be applied to develop digital campus information system with practical value will help to solve these problems; for other users on campus, it is more convenient to obtain all the necessary information from the campus management information system, furthermore to fully understand the school situation. From this perspective, the system also has certain significance in the digital city construction.

References

Aronoff S (1989) Geographic information systems: a management perspective. WDL Publications, Ottawa

Li Hua, Wu Chaozhong, Yan Xinpin (2004) Design of a campus MIS based on WEBGIS. J Wuhan Automot Polytech Univ 2

Liang Xiaotian, Yu Ling, Peng Yong (2007) Problems and countermeasures in the construction of digital campus' MIS. Libr Inf Serv, 2007-11

Luo YW, Wang XL, Xu ZQ (2004) Design and implementation of a hierarchical component-based WebGIS. Chin J Comput 27(2):177–185

Wang Jin, Wang Linlin, Chen Bo (2002) A campus facility management system based on GIS technology. Comput Eng 6

Xie Hong, Lu Yuefeng, Li Jiansong (2009) Design of digital campus landscape system based on ArcIMS. Geosp Inf, 2009-1

Yu Li, Wang Cheng (2004) Digital campus three dimension simulation system based on virtual reality technology. Comput Simul, 2004-4

Part II
Management Science and Engineering

Scientific Research Input-Output Efficiency Evaluation System for Chinese Universities

Ning Zhou, Xiaowei Zhang, and Xiaoting Han

Abstract With the rapid expansion of university research budget both in amount and scale, the input-output efficiency has become a focus of research interest. Based on the analysis of the latest researches concerned, this paper is intended to investigate the advantages of applying BSC to constructing an input-output efficiency evaluation system for scientific research in universities. In addition, this paper aims to construct a strategy map and an index system by means of BSC. The findings of this paper will provide a model for the scientific research input-output efficiency evaluation for Chinese universities.

Keywords Balanced scorecard (BSC) • Efficiency evaluation • Scientific research • University

1 Introduction

Scientific research in universities is an important part of national innovation system, which has long provided a strong scientific and technological support force for national economic and social development. Scientific research input is a key factor restricting universities' development, whose efficiency level directly influences the development speed of scientific research in universities. With the rapid expansion of university research budget both in amount and scale, the input-output efficiency has become a focus of research interest. Based on the date from "Statistical Yearbook of Chinese Science and Technology", spends on national universities' R&D rapidly increased from 5.826 billion yuan in 1997 to 60.032 billion yuan in 2010, which

N. Zhou (✉) • X. Zhang • X. Han
School of Economics and Management, BeiHang University, Beijing, China
e-mail: zning80@buaa.edu.cn; love_19880822@sina.com; hanxiaoting@buaa.edu.cn

grew at an annual average rate of 19.98% and totally amounted to 314.706 billion yuan for the 14-year period between 1997 and 2010. Consequently, constructing an input-output efficiency evaluation system for scientific research in universities would be significant for optimizing the allocation of science and technology resources, and promoting the construction of innovation-oriented country.

2 Literature Review

As regards scientific research input-output efficiency evaluation in universities, domestic and foreign scholars have made a lot of researches on evaluation methods, index selection and specific application.

Thursby and Kemp (2002) examined the overall and individual productivity of American universities' patent activities by means of DEA method and regression analysis method. Cherchye and Vanden Abeele (2005) investigated the research efficiency of the Economics and Business Management School of Holland University based on non-parameter method, and they found that using micro-date to make efficiency study will get more meaningful conclusion than using macro-data. Benot Godin (2007) followed accounting methods and statistical methods to construct a scientific research input-output framework. Eilat et al. (2008) applied DEA model and BSC to the assessment of R&D projects.

Liu and Zhao (2003) used Cobb-Douglas Model to analyze the key factors which affected the quantity and quality of SCI papers in China. Considering the availability of universities' features and data, Sun et al. (2007) selected scientific research personnel and scientific and technological budget as input indicators, while chose scientific and technological works, papers, authorized number of intelligent property rights and the transfer value of scientific and technological achievements as output indicators. Zhang et al. (2008) chose R&D personnel and expenditure on science and technology in the same year as input indicators, while chose scientific and technological works, papers, authorized number of patent as output indicators. Using the idea of BSC model, Wang and Bi (2009) established a strategy-based performance evaluation system for universities' innovation teams. Fu et al. (2010) analyzed the scientific research input-output efficiency of different universities in several scientific research scale and quality, and the results are different. Liu et al. (2011) constructed a performance evaluation index system of university's Science and Technology Park based on BSC theory, and established a performance evaluation model by use of grey fuzzy theory.

To sum up the above arguments, although scholars have made certain researches on scientific research efficiency evaluation in universities, a comprehensive index system has not been formed yet. Therefore, it's significant to construct an input-output efficiency evaluation index system for scientific research in universities.

BSC constructed by Kaplan and Norton provides a performance evaluation system combining financial indicators with non-financial indicators. Considering

the advantages of strategic oriented, the BSC used to design the evaluation system will focus on the achievements of set objectives, vision and long term goals. As a result, the system could realize the effective combination of strategic target and efficiency evaluation indicators, financial and non-finance indicators, quantitative and qualitative evaluation indicators.

3   BSC-Based Input-Output Efficiency Evaluation System Design

Strategic target of scientific research activities should be first established while applying BSC to scientific research input-output efficiency evaluation in universities, namely to improve input-output efficiency, and make necessary amendments and improvements according to the features and requirements of scientific research activities in universities, then establish scientific research input-output efficiency evaluation system on the basis of target. Evaluation system should be designed following the principles of scientific, comprehensive, strategic dependence, comparability and feasibility, "cause and effect" relationship, and so on.

3.1   Strategy Map of Input-Output Efficiency Evaluation

Strategy map shows the "cause and effect" relationship between the input and output indicators of scientific research activities in universities. As shown in Fig. 1: research funds and achievements of universities directly promote the development of research teams in universities; the ability of learning and growing of research teams directly promote scientific research, belongs to internal process dimension; the participation of scientific researchers, who are responsible, and have administrative abilities, in research platform construction and other management work, will also can optimize and promote internal process; the outstanding performance of research work among the internal process dimension is the primary driving factor in determining the scientific research output and research awards belong to the service value dimension; all the achievements and rewards are the most important determinants among indicators belong to the service value dimension, which can promote universities to make greater contributions to society.

3.2   Input-Output Efficiency Evaluation System

According to the above strategy map and "cause and effect" relationship analysis, the scientific research input-output efficiency evaluation system in universities
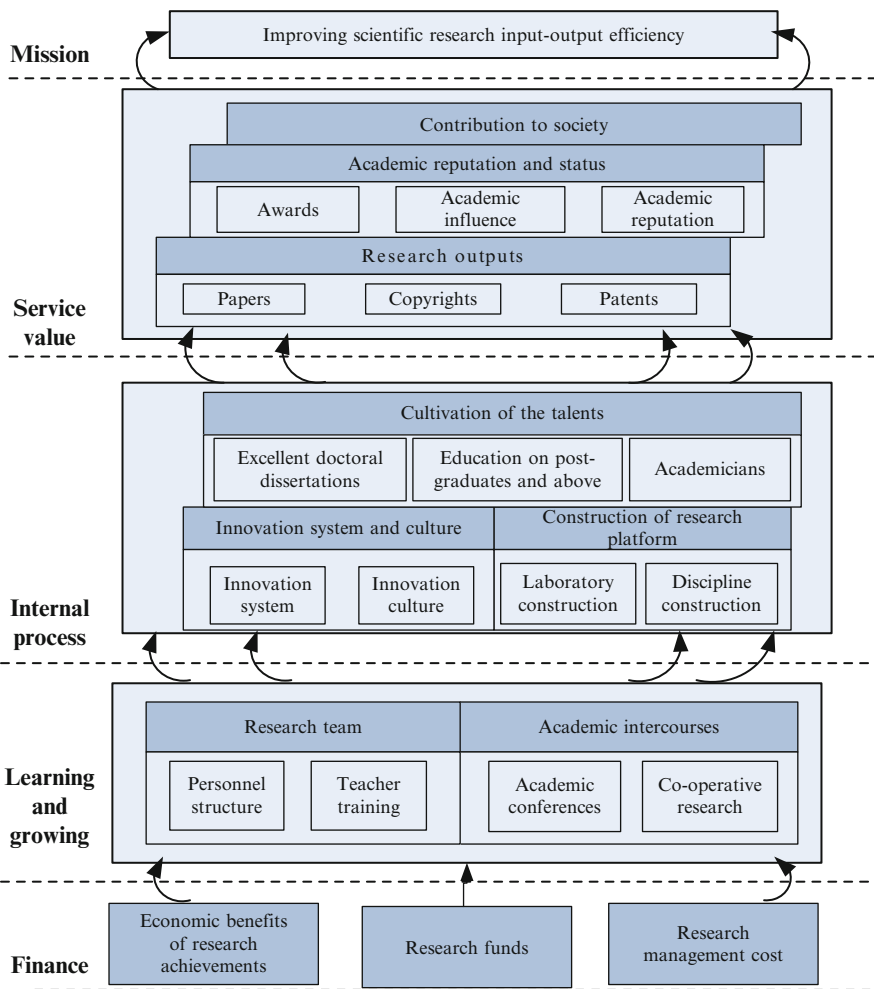
Fig. 1 Strategy map of university's scientific research input-output efficiency evaluation

could be designed from following four dimensions: service value, internal process, learning and growing and finance, and two levels in each dimension. Meanwhile, these indicators will be classified into input or output. As shown in Table 1.

During the practical application, the above index system can be more detailed. Take the "papers", a second-level index, for example, it can cover the papers indexed by SCI/EI/ISTP, the papers published in domestic top level journals as third-level indexes. As shown in Table 2.

Table 1 University's scientific research input-output efficiency evaluation index system

Strategy dimension	First-level index	Second-level index	Classification
Service value	Research output	Papers	Output
		Copyrights	
		Patents	
	Academic reputation and status	Scientific research awards	
		Academic influence	
		Academic reputation	
	Contribution to society	Participation in feasibility studies of significant project	
		Participation in significant policy-making	
Internal process	Science and technology innovation system and culture	Innovation system	
		Innovation culture	
	Construction of research platform	Laboratory construction	
		Discipline construction	
	Cultivation of the talents	Excellent doctoral dissertations	
		Education on postgraduates and above	
		Academicians of the Chinese Academy of Sciences and Chinese Academy of Engineering	
Learning and growing	Research team	Personnel structure	Input
		Teacher training	
	Academic intercourses	Co-operative research	
		Academic conferences	Output
Finance	Research funds input	Research funds	Input
	Research management cost	Administrative management cost	
	Economic benefits of research achievements	Achievements conversion	Output
		Technology transfer	

Table 2 Third-level index for "Papers"

Second-level index	Third-level index
Papers	Papers indexed by SCI
	Papers indexed by EI
	Papers indexed by ISTP
	Papers published in domestic top level journals

4   Conclusion

To sum up, based on the new features and latest trends of Chinese universities' scientific research activities, this paper takes the improvement of scientific research input-output efficiency as a guide, and treats the target of scientific research activities as underpinning. It introduces the BSC into input-output efficiency evaluation index system, and finally establishes the scientific research input-output evaluation index system in universities on the basis of BSC. Meanwhile, it also emphasizes the "balance" between dimensions, which is useful to avoid insufficiency that arising from neglecting some relative respects. In addition, Delphi method and the Analytic Hierarchy Process (AHP) can be used to determine the different weight of each index according to its impact on general objective. Hereby, universities can make clear the focal points of their scientific research, and timely adjust the construction of its own research teams, management process and scientific research platform, in order to improve the efficiency level of scientific research, and provide effective and efficient scientific research services for clients. The findings of this paper will play an active role in promoting the development of universities and national scientific research activities.

References

Cherchye L, Vanden Abeele P (2005) On research efficiency: a micro-analysis of Dutch University research in economics and business management. Res Policy 34:495–516

Eilat H, Golany B, Shtub A (2008) R&D project evaluation: an integrated DEA and balanced scorecard approach. Omega 36:895–912

Fu Y et al (2010) Comparative study of science and technology input-output efficiency of different types of universities. Sci Technol Manage Res 1:204–207

Godin B (2007) Science, accounting and statistics: the input-output framework. Res Policy 39(9):1388–1403

Liu X, Zhao J (2003) Analysis of the factors that effect the quantity and quality of Chinese SCI. Chin Soft Sci 10:87–94

Liu Y, Liu S, Sun G (2011) University of science and technology park performance evaluation based on the Balanced Scorecard and Gray Fuzzy Theory. High Educ Res Pract 1(30):53–59

Sun S, Xiang H, Lan B (2007) Efficiency analysis of research input-output in our country based on DEA. Sci Technol Policy Manage 7:18–21

Thursby JG, Kemp S (2002) Growth and productive efficiency of university intellectual property licensing. Res Policy 31:109–124

Wang Y, Bi K (2009) Construction of university's innovation team performance evaluation system based on strategy. J Harbin Eng Univ 3(30):332–338

Zhang Y, Wu J, Shi Q (2008) Efficiency analysis of university's science and technology input-output and achievement transformation-from the view of value chain. Sci Technol Manage Res 8:133–135

Applications of Data Mining in the Enterprise Customer Relationship Management

Hong-ye Zhao and Shou-xiang Zhao

Abstract In this paper, we have done some basic research on the application of data mining techniques in customer relationship management by learning the theories of customer relationship management and data mining technology.

1 Preface

If a company wants to stand out among peers, he must conduct in-depth study of the customers. And this requires advanced concepts such as customer relationship management. The data which is accumulated in the long-term development of the business is the most valuable wealth. Using data mining techniques to find valuable information to an enterprise, the better the implementation of customer relationship management is the most pressing needs.

2 Customer Relationship Management

Customer Relationship Management is a product of modern science and advanced IT, and its core idea is "customer-centric" (Gu Xiao-wen 2010). Companies can restructure business processes, improve customer service and customer analysis and understand customer requirements more accurately, to improve the relationship between companies and their customers in order to attract and retain more valuable

H. Zhao (✉) • S. Zhao
College of Computer and Information Engineering, Beijing Technology
and Business University, Beijing 100048, China
e-mail: xhlzz@sina.com; zhao_sx4384_cn@sina.com

customers, reduce costs, improve profitability, while enabling customers to maxi-
mize the yield, and to ensure that the status of enterprises in the competition and the
sustainable development.

3 Data Mining

Data mining is an emerging interdisciplinary which is on the basis of machine learn-
ing, computational visualization, pattern recognition, statistics and other disciplines
(He Jin-hua 2011). It is a process which gets the potentially useful information from
a large number of incomplete, noisy, fuzzy random data.

It is application-oriented. And it is not only a simple research for a specific
database, but also an in-depth statistics, analysis or reasoning, using these data
to explore the relationship between the data to complete the transition from
business data to decision-making information. The using in customer relationship
management can help the decision-makers make the right decisions by filtering
the seemingly unrelated data and extracting valuable customer information (Li Zhi
2011).

4 Data Mining Techniques

4.1 Classification

Classify an event or object, is to identify the common characteristics of a group of
objects in the database, and divided them into different classes in accordance with
the classification model. For example, applications of credit card are classified as
low, medium and high risk.

4.2 Regression Analysis

Regression analysis is to reflect the physical database characteristics of the property
values in the time and generate a data item which is mapped to a real-valued
function of the predictor variables. Then find the dependencies between variables
or attributes. Its main research question is the trend characteristics of the data series,
and forecast data series, as well as the correlation between the data.

4.3 Cluster Analysis

Cluster analysis gets together the similar characteristics of users or data items. Its
purpose is to make the similarity within the same category as large as possible, but
the similarity between the categories is as small as possible.

4.4 Association Rules

The association rules are the rules of the relationship that exists between the data items in the database. It is the emergence of some of the items according to a transaction can be exported to others in the same transaction, and that is to hide the association or mutual relationship between the data.

4.5 Sequence Pattern Analysis

The sequence pattern analysis and association rule analysis is similar, but the focus is on analysis of data between the sequences. It can found the knowledge in the database such as "The highest frequency within a certain period of time, the customer to buy goods A, followed by purchase of goods B, and then buy goods C, that is, the sequence A → B → C" (Sierra and Hyman 2011).

4.6 Decision Tree

The decision tree method is a process which finds the attribute fields with the greatest amount of information in the database by using the information theory to establish a decision tree node, then to establish the number of branch according to different values of the attribute field, and Each branch of the subset repeated to establish the tree the lower node and branch.

4.7 Neural Network

The neural network is a new system to simulate the human brain network processing, memory information. It applies to the complex problems which are a large number of non-linear time sequence, multiple targets, incomplete numerical data, processing results-oriented, or needing not explain why.

5 The Application of Data Mining in Customer Relationship Management

5.1 Acquisition of New Customers

The acquisition of new customers is the first step in the life cycle of customer relationship management, is the indicators of business development, and is the key to growing businesses (Sun Ning 2011). Data mining can help companies study the

respondents who visit the corporate website, call toll free, or fill out the application form. And companies can make their marketing activities focus on these people (Xia Wei-li et al. 2007). It can help to complete the division of the potential customer base, in order to achieve the effect of these activities which are the cost of accessing to new customers, effective management and the improving of the return.

5.2 Customer Segmentation

A large customer base is divided into several smaller groups based on certain attributes. Customers in the same group are high similarity in one or a few property values, but there is a big difference between them in the different groups. For example, using technologies in data mining such as decision trees, cluster analysis method, companies can divide into different user groups by the customer's psychology, habits or the frequency of purchase. In order to achieve the targeted customer services and the development of targeted products, and to improve customer satisfaction and lifetime value of the customers.

5.3 Customer Retention and Loss Warning

Customer retention is a process, which maintains established customer relationships to make the customers repeat to purchase products or services. Studies have shown that the cost of a new customer is five times than the cost of keeping an old customer, and improving customer retention by 5%, the profit will be increased by 25% (Zhang Zhe 2011). Therefore, customer retention is the key to business success. Enterprises can analyze the customer transaction history in database, demographic information and other data associated, and do some targeted research on the loss of customers, then create a loss of client model, and develop appropriate plans and programs, to implement different strategies for different customers using neural network technology.

5.4 Cross-Selling

Cross-selling is the marketing of new products or services sold to an existing client. It is not only an effective means to increase profits by expanding sales to existing customers, but also a important strategy to enhance the corporate image, foster customer loyalty, and protect the enterprise sustainable development (Zhong Ruo-wu and Wang Hui-ping 2010). Companies can find the characteristics of the buyers, find the key factors of the buying behavior, and establish a forecasting model to predict and analyze the customer's future buying behavior by using correlation analysis sequence analysis and so on.

5.5 *Customer Profitability Analysis*

"2/8 rule" is the classical theory of customer relationship management, that is 80% of the profits from the 20% of customers. Therefore, companies need to know which customers are profitable, so that more resources are assigned to these customers. Companies can establish the Life Time Value model to predict the profitability of these customers within a predetermined length of time by using the classification analysis techniques and so on.

5.6 *Analysis of Customer Credit*

The customer's credit rating is very important in customer relationship management. Because, when the credit risk and fraud happened, enterprises will face loss of market share, which causes loss of market, customer, competitiveness and reputation. Through data analysis, companies can use the classification analysis, cluster analysis, sequential pattern analysis and other technical to predict the causes and the possibility of customer fraud, the establishment of a fraud rule base can be a timely warning of a variety of fraud, and reduce the loss of business.

6 Conclusion

Using data mining technique, can analyze the historical data of the customers, understand customer preferences, so that enterprises can be more timely and more accurately to provide customers with products and services, and also brought huge profits to the enterprise. Customer relationship management is the important areas of application of data mining, just because of this, it makes the philosophy and objectives of customer relationship management be achieved. The combination of data mining and customer relationship management will promote the development of enterprises, and also will provide customers with better service.

References

Gu Xiao-wen (2010) Data mining technology and customer relationship management applications review. Softw Guide 9(2):145–146
He Jin-hua (2011) The application of the data mining in customer relationship management. Softw Guide 10(10):145–146
Li Zhi (2011) The application of data mining in E-commerce customer relationship management. Sci Technol Inf 11:86–87
Sierra JJ, Hyman MR (2011) Outlet mall shoppers' intentions to apparel: a dual-process perspective. J Retail Consum Serv 7(4):341–347

Sun Ning (2011) The issues of customer relationship management based on data mining. Technol Ind 11(6):59–61

Xia Wei-li, Wang Xin, Wang Qing-song (2007) The data mining in the enterprise customer relationship management. Technol Manage 6:196–198

Zhang Zhe (2011) The study of SME customer relationship management based on data mining. Econ Forum 492(7):202–204

Zhong Ruo-wu, Wang Hui-ping (2010) The customer relationship management system based on data mining. Shaoguan Univ 31(3):29–32

Study on Agile Supply Chain

Zuo-xi Zhao and Shou-xiang Zhao

Abstract Agile Supply Chain is a dynamic supply and demand network of rapid response to environmental changes, which is formed to a number of supply-side and demand-side entities (autonomous, semi-autonomous or subordinate). In combination with the car manufacturer Chrysler to build agile supply chain case, the article explore the agile supply chain strategy.

Keywords Agile supply chain • Chrysler • Strategy

1 Introduction

Along with the rapid development of world economy and the tide of economic globalization, market competition environment are rapidly changing. With the shorter product life cycles, customer requirements for timely system and the transformation of a seller's market to a buyer's market, the response speed of the sudden demand and the elasticity of demand has become an important competitive of enterprise. Thus, in the changing competitive environment, the "agile" would be particularly important.

Agile supply chain management is a new organization and management model (Van Hoek 1997). Responding quickly to market to meeting customer needs is the key factors of the era based on the time competition. Agility is the Requirement to survive in the market environment faced the uncertainty in twenty-first century Therefore, Researching on supply chain agility and exploring the competitive advantages of the agile supply chain and agile strategies, agile performance evaluation system are of great significance.

Z. Zhao (✉) • S. Zhao
College of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China
e-mail: zhaozuoxi126@126.com; zhao_sx4384_cn@sina.com

2   Overview of Agile Supply Chain

In 1995, the council of logistics management firstly defined Agile, Agile is that by providing high-quality supply chain management to obtain and maintain competitiveness and let customer Obtain success (Christopher 2006). "Agile" used to emphasize the supply chain's ability to quickly respond to the market changes and users' needs changes, it means rapid action and immediately to meet customer needs.

Agile Supply Chain is a dynamic supply and demand network of rapid response to environmental changes, which is formed to a number of supply-side and demand-side entities (autonomous, semi-autonomous or subordinate). A supply chain in order to truly agile would have some significant features:

(a) Agile supply chain is required to be able to make a sensitive response to the market, it means Supply chain to understand the real needs and respond to.
(b) Agile supply chain is required to be able to use of IT to share information between all entities in the supply chain, this is in fact created a virtual supply chain.
(c) Partners in the supply chain can make full use of shared information through process alignment.
(d) The fourth element of the agile supply chain is the idea of contacting the various partners linked into a network of Commonwealth (Fig. 1).

3   The Case of Chrysler's Agile Supply Chain

Chrysler manages the supply chain systems Composed by 104 factories in 37 countries around the world, 14,000 suppliers in 200 countries and 13,000 sales outlets, in order to make the system become flexible and creative, Chrysler has taken a series of decision-making:

(a) The Chrysler Group has deployed a large number of systems to automate and streamline their supply chains. For example, transactions among all the respective companies of Chrysler are using a single interface and infrastructure; Chrysler first introduce "powerway" into automobile profession in 2001, which
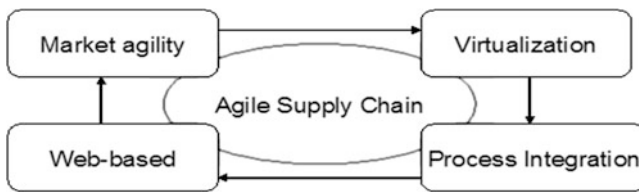
Fig. 1 The schematic of agile supply chain

is one kind of Web actuation quality control system as well as the supply chain cooperation network.

(b) Company collects data from as wide as possible channels and put these data promptly and effectively into use. For example, Chrysler widely and deeply touches all networks about its global supplier's portal business; Company through the overall quantity plan system to collect sale data, and so on.

(c) Because the wireless technology progress, through building the supply chain system, which can monitor the event and respond to those data when the event just happened, Chrysler can capture data generated in the manufacture, production and shipment as it generation. For example, the ABF delivery systems use a series of self-produced wireless applications called NetLink to manage how does the cargo flow in its network.

(d) Through the network, Chrysler contact to sufficient number of partners to clearly figure out when presents out of stock and excess.

Because of this superior supply chain, a number of CIO of Daimler – Chrysler and its respective companies was selected into the list of 100 global chief information officer in a U.S. magazine 《CIO》 this year.

4 The Strategy of Agile Supply Chain

4.1 Information Sharing, Synchronous Operation

Synchronization means that all members of the supply chain must be acting in unison (Dong 2010). In other words, by sharing information and the Unified Process, and Sort all behavior to make the whole supply chain collaboration and "seamless" operation. In the environment of working together, different entities in the network can share information about the real demand inventory and production by network technology in order to achieve synchronous operation.

4.2 Process Reengineering, Reduced Operating Time

It is not difficult to find that large part of the time one end to the other side of the supply chain is the "non- value added time". It cost more time on behavior only increase costs but no benefits to our customers. To shorting the operating time is not necessarily need to accelerate the speed of operation, it can also by streamlining the operation of the process. Business Process Re-engineering are often used to streamline and reshape the organization's activities, its purpose is to achieve the desired results within a shorter time and with less cost (Ma 2006).

4.3 Collaboration with Suppliers, Reduce the Lead Time

The most important way to shortening the lead time is close cooperation with key suppliers (Wang 2010). A powerful method is to use the mode of Vendor Managed Inventory, which model changing the responsibility of management and inventory replenished from the client side to supply side. Through this collaboration to improve the speed of response of the market.

4.4 Reduce Complexity

Complexity in the supply chain is in different forms. From a system perspective, complex systems will reduce its agile response capability. By exploring the root causes of these complexities, we can reduce it. For example, make variety of our products and the actual needs of customer counterparts; by seeking commonality of parts and the generality of the primary assembly in the same series of products to realize simplified, and so on.

4.5 Delay

Delay is the method as long as possible to postpone the process of the final product configuration or delivery of the final location to avoid products provided may not be customers really need (Wang 2011). The delay mechanism includes two kinds, one kind is the delay of the final product molding, and another is on the space-based delay.

4.6 Change the Business Organization, Great Importance to the Process Management

For hundreds of years, business has been based on the breakdown of the labor, various activities in the various functional. But it seems that the functions of the type of business organization can not quickly respond to the changes in market or business environment, those enterprises which able to respond quickly to the changes in customer demand are more likely to focus on the management of the "process". Such as the establishment of an inter-departmental co-operation group, which can exercise changes customer relationship management and supplier relationship management functions.

Process management is essential to improve the agility of the supply chain, The reason is that if the organizational structure is flat rather than vertical, then an alliance of inter-enterprise processes in the supply chain can no doubt enhance the performance of the entire supply chain (Xiangru meng 2009).

5 Conclusion

Because of market volatility and uncertainty of demand, companies must make more rapid responsiveness. The key of rapid responsiveness is that upstream and downstream of the core businesses are agile collaborators. It requires not only the business itself has the agility, but also there are more requirements to improve the entire supply chain agility and collaborating with each partners in supply chain agile response to changes in end- customer demand. Enterprises should be based on customer demand – driven and must make customer satisfaction its main goal, through the supply chain agility strategy to improve the competitiveness of the entire supply chain to gain a pillar of competitive advantage.

References

Christopher M (2006) Supply chain and logistics management. January (in Chinese)
Dong R (2010) Logistics. November (in Chinese)
Ma S (2006) Supply chain management. (In Chinese)
Van Hoek RI (1997) The thesis of agility revisited. Int J Oper Manage
Wang D (2010) Study on agile of the supply chain. (In Chinese)
Wang Y (2011) Business logistics management. March (in Chinese)
Xiangru meng (2009) Logistics management. January (in Chinese)

Safety Management of Oil Depot Based on Fuzzy Analytic Hierarchy Process

Yi Zhou and Xiao-gang Zhao

Abstract In order to further improve the management level of oil depot, Fuzzy Analytic Hierarchy Process (FAHP) is applied in the safety management of oil depot on the basis of the analysis of recent fire accidents in oil depot in China. FAHP is the combination of the traditional Analytic Hierarchy Process (AHP) and the fuzzy mathematics. By analyzing various impact factors of the security management in oil depot, a three-level of safety management hierarchy system is built up to obtain the weight sets and comprehensive weight set order. Consequently, a scientific and rational security guidance for safety management is made up from the results. Through the practical application in Fulin oil depot in Chongqing, China, it is concluded that this method can effectively reflects the priorities within various impact factors on the security management of oil depot, and provides a scientific basis for the security works in oil depot.

Keywords Fuzzy Analytic Hierarchy Process (FAHP) • Oil depot • Safety management • Fire and explosion protection

1 Introduction

Security issues have become the core for daily management in oil depot. Since gasoline or diesel oil has the characteristics of flammable, explosive, volatile, and so on, it is easy to evaporate local accumulation of oil gas in the process of storage and sending or receiving operations which may cause the fire and explosion accidents in the reservoir area and result in the great damages to equipments and facilities (Xu 2009).

Y. Zhou (✉) • X. Zhao
Department of Petroleum Supply Engineering, Logistical Engineering University, Chongqing 401311, China
e-mail: zhouyiasd@foxmail.com

Recently, security incidents were prone to occur in oil depots in China. For example, a flash explosion accident happened on Oct 2006 during the anti-corrosion operations of No. 1 double-disc floating roof crude oil tank of Petrochina, 13 workers were killed and another six were seriously injured; fire and explosion accident occurred in No. 316 storage tank battery of Lanzhou Petrochemical Company on Jan 2010, six people were killed and six was seriously injured. Such fire and explosion accidents arouse a strong response in the community to improve the management level of oil depot. Therefore, it is urgent to find a scientific and rational guidance for security works in the oil depot, and which also becomes the key and focus of safety management.

This paper analyzes the influencing factors during the fire and explosion accidents to establish the safety evaluation index system in oil depot, and proposes Fuzzy Analytic Hierarchy Process (FAHP) to the assessment of safety management in oil depot. FAHP is the organic combination of traditional Analytic Hierarchy Process (AHP) and fuzzy mathematics which can effectively reflect the priorities of safety work in oil depot so as to give an accurate and scientific guidance for safety management.

2 Fuzzy Analytic Hierarchy Process

AHP was firstly introduced by Thomas L. Satty in 1977 (Saaty 1980). It can solve the complex problems through the application of Measure Theory. But the traditional AHP has the following deficiencies during the longtime applications (Song Guang-xing and Yang Li-de 2001): (1) the consistency of the judgment matrix is usually different form the consistency of people thinking; (2) the procedures to check the consistency of the judgment matrix are very complex; (3) when the judgment matrix are not in consistent with each others, it is much troublesome to adjust them into consistency; (4) the criterion to judge the consistency standard of the judgment matrix is CR <0.1 which is still lack of scientific explains. Therefore, some people try to use the combination of the fuzzy mathematics and the traditional AHP to propose a more scientific and simple method, namely Fuzzy Analytic Hierarchy Process (FAHP) (Zhang Jin-jun 2000; Chao Bing-yuan 2005). Such application of FAHP can overcome those deficiencies of the traditional AHP. The concrete steps of FAHP are as follows:

1. Establishment of a hierarchy index system. Through the analysis of the relationship between those impact factors in safety management, the establishment of a hierarchy model is carried out (such as Fig. 1).
2. Establishment of the fuzzy judgment matrix. The relative importance of every two factors in the same level to an upper level is compared to establish the fuzzy judgment matrix $a = [C_{ij}]_{n \times n}$. The specific meanings of C_{ij} using 0.1–0.9 scale are shown in Table 1.
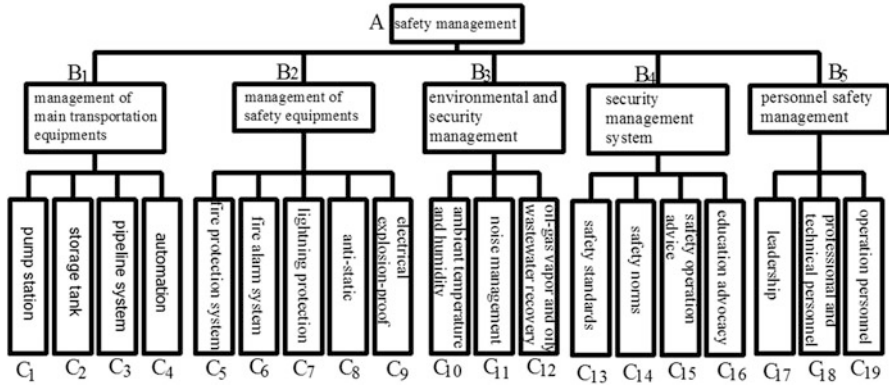
Fig. 1 Hierarchy structure system

Table 1 The specific meanings of Cij in 0.1–0.9 scale

Scale C_{ij}	Meaning
0.5	Two factors with the same importance
0.6	Factor i is slightly important than factor j
0.7	Factor i is obviously important than factor j
0.8	Factor i is strongly important than factor j
0.9	Factor i is extremely important than factor j
0.1,0.2,0.3,0.4	Anti-comparison. If the comparison judgment from factor i to j is C_{ij}, the judgment from factors j to i is C_{ji}, and $C_{ji} = 1 - C_{ij}$

3. Transformation of the fuzzy judgment matrix into the fuzzy consistent matrix. If $r_i = \sum_{j=1}^{n} C_{ij}, (i = 1, 2, \cdots, n)$, based on Eq. (1), the fuzzy judgment matrix is transformed into the fuzzy consistent matrix $b = [r_{ij}]_{n \times n}$, and then it satisfies the consistency and there is no need for consistency checking.

$$r_{ij} = \frac{r_i - r_j}{2n} + 0.5 \tag{1}$$

4. Identification of the weight sets. According to the nature of the fuzzy consistent matrix, the weight sets of each index w_i can be calculated. Because of the high resolution of the fuzzy consistent matrix and the weight set order, the relationship between fuzzy elements and weight sets are select to obtain the weight set vector according to the least squares method (Zhou Yu-feng and Wei Fa-jie 2006), and the specific formula as follows:

$$w_i = \frac{1}{n} - \frac{1}{2\alpha} + \frac{1}{n\alpha} \sum_{j=1}^{n} r_{ij} \tag{2}$$

Where the parameter $\alpha \geq (n-1)/2$. In the practical applications $\alpha = (n-1)/2$, which is of the most emphasis on the importance of the differences between each projects. At this point, Eq. (2) can be rewritten as Eq. (3):

$$w_i = \frac{2\sum\limits_{j=1}^{n} r_{ij} - 1}{n(n-1)} \tag{3}$$

5. Determination of the total order among the index weight vector. In accordance with the hierarchy index system, the total order of index weight vector is obtained and normalized through the calculation from the bottom level to the top one.

3   Application

Combined with the characteristics of safety management in Fulin oil depot in Chongqing, China, a three-level structure system is established. The key to safety management is due to the following four aspects, namely management of the main transportation equipments, management of the safety equipments, environmental and security management, security management system and personnel safety management (B_1 to B_5). Specifically, each aspect contains several impact factors which are all shown in Fig. 1, and C_1 to C_{19} is utilized to represent these above 19 sub-factors.

According to the steps of FAHP, the fuzzy judgment matrix is built on the relative comparisons between the middle level and the target level:

A	B_1	B_2	B_3	B_4	B_5
B_1	0.5	0.6	0.8	0.7	0.9
B_2	0.4	0.5	0.7	0.6	0.8
B_3	0.2	0.3	0.5	0.3	0.7
B_4	0.3	0.4	0.7	0.5	0.8
B_5	0.1	0.2	0.3	0.2	0.5

Then the fuzzy judgment matrix between the middle layer and the target layer is transformed thought Eq. (1) into the fuzzy consistency matrix, and then this matrix is brought into Eq. (2) to obtain the weight set: W = (0.2435, 0.2217, 0.1783, 0.2087, 0.1478).

Similarly, weight sets of layers can be obtained. W_1 = (0.2929, 0.2429, 0.2713, 0.1929); W_2 = (0.2435, 0.1739, 0.2217, 0.2000, 0.1609); W_3 = (0.3905, 0.2619, 0.3476); W_4 = (0.2982, 0.2696, 0.2411, 0.1911); W_5 = (0.2619, 0.3905, 0.3476). Total weight set order of impact factors is then obtained in Table 2.

According to the weight set order in Table 2, the importance of each factor in the safety management of oil depot can be drawn in accordance with the value

Table 2 Total weight set order

	B_1	B_2	B_3	B_4	B_5	
A	0.2435	0.2217	0.1783	0.2087	0.1478	Weight set W
C_1	0.2929	0	0	0	0	0.0713
C_2	0.2429	0	0	0	0	0.0591
C_3	0.2713	0	0	0	0	0.0661
C_4	0.1929	0	0	0	0	0.0470
C_5	0	0.2435	0	0	0	0.0540
C_6	0	0.1739	0	0	0	0.0386
C_7	0	0.2217	0	0	0	0.0492
C_8	0	0.2000	0	0	0	0.0443
C_9	0	0.1609	0	0	0	0.0357
C_{10}	0	0	0.3905	0	0	0.0696
C_{11}	0	0	0.2619	0	0	0.0467
C_{12}	0	0	0.3476	0	0	0.0620
C_{13}	0	0	0	0.2982	0	0.0622
C_{14}	0	0	0	0.2696	0	0.0563
C_{15}	0	0	0	0.2411	0	0.0503
C_{16}	0	0	0	0.1911	0	0.0399
C_{17}	0	0	0	0	0.2619	0.0387
C_{18}	0	0	0	0	0.3905	0.0577
C_{19}	0	0	0	0	0.3476	0.0514

of the weight set. The weight set value of each factor represents its influence on the target level, the greater the bigger. The result of weight set order is as follows: $C_1 > C_{10} > C_3 > C_{13} > C_{12} > C_2 > C_{18} > C_{14} > C_5 > C_{19} > C_{15} > C_7 > C_4 > C_{11} > C_8 > C_{16} > C_{17} > C_6 > C_9$. The results comply with the actual situation of safety management in Fulin oil depot, which shows the feasibility of this method. The priorities in safety management can be drawn from the total weight set order of the assessment result. In addition, the most four important security works can also be obtained, namely equipments management of pump station (0.0713), ambient temperature and humidity control (0.696), safety management standards (0.622) and equipments management pipeline system (0.0661). Therefore, in the daily safety management, these four factors should be focused on to improve the safety management level. In particular, the technical personnel should pay more attention to and regularly check the equipments management of pump station and pipeline system to ensure the reliability, and the detection of safety problems and potential leakages should be carried out in time.

4 Conclusion

According to the characteristics of fire and explosion accidents in oil depot, a three-level of safety evaluation index system is established and FAHP is proposed to obtain the total weight set order. The results show that the most important

safety work in oil depot are equipments management of pump station and pipeline system, ambient temperature and humidity control, safety management standards. Therefore, these four factors should be enhanced to improve the modernization of safety management.

FAHP is based on fuzzy mathematics, thus it can overcome the deficiencies of the traditional AHP. Since it is the combination of AHP and the fuzzy mathematics, this method is not only applicable at the safety management of oil depot decisions, but also provides some reference to other types of decision-making.

References

Chao Bing-yuan (2005) Applied fuzzy mathematics and system. Science Press, Beijing

Saaty TL (1980) The analytic hierarchy process. McGraw-Hill Inc., New York

Song Guang-xing, Yang Li-de (2001) Consistency check and improving method of FAHP. Syst Eng 21:110–116

Xu Hang (2009) Oil depot design and management. Petrochina Press, Beijing

Zhang Jin-jun (2000) Fuzzy analytic hierarchy process. Fuzzy Syst Math 14:80–88

Zhou Yu-feng, Wei Fa-jie (2006) A new method of consistency approach and weight set calculation in uncertain type of fuzzy judgment matrix. Oper Manage 4:28–31

Non-lane-Based Macro Model of Traffic Flow with the Consideration of the Lateral Effects of the Lane Width

Xinhua Cai and Guanghan Peng

Abstract A non-lane-based macro model is proposed by considering the lateral separation effects of the lane width in traffic flow on the basis of car-following model. The extended model can improve the stability of traffic flow. Numerical simulation also shows that the proposed macro model can enhance flux of traffic flow and reproduce the evolution of shock and rarefaction waves by incorporating the lane width effects. In addition, the lateral separation effects have greatly important impacts on equilibrium solution of uniform traffic flow in macro models.

Keywords Traffic flow • Non-lane-based macro model • Lateral separation

1 Introduction

In modern traffic, road conditions have a vital effect on car following behavior. Recently, a non-lane-based car-following model (Jin et al. 2010) and a non-lane-based lattice model (Peng et al. 2011) of traffic flow were presented by considering the lateral effects of the lane width respectively. Above results imply that the lateral effects of the lane width can improve the stability of traffic flow and produce some complex traffic phenomena. But the lateral effects of the lane width did not be investigated in macro models of traffic flow. In this paper, we present a new macro model of traffic flow based on car-following model by taking the lateral effects of the lane width into account. Also numerical simulation will be carried out to verify the influence resulted from lateral separation effects of the lane width on traffic flow.

X. Cai (✉) • G. Peng
College of Physics and Electronics, Hunan University of Arts and Science,
Changde 415000, China
e-mail: xhcai2002@yahoo.com.cn; pengguanghan@yahoo.com.cn

2 The Macro Model

The vehicles were assumed to travel in the middle of the lane in car following theory (Jin et al. 2010; Peng et al. 2011). But in real traffic many vehicles do not run on the center of the lane, specifically in developing countries (Jin et al. 2010; Peng et al. 2011). Therefore, to represent the lateral effects of the lane width, a non-lane-based car-following model with the consideration of the lateral effects of the lane width is proposed with the same form in Jin et al. (2010) as follows:

$$\frac{dv_n(t)}{dt} = a[V(\Delta x_n(t), \Delta \tilde{x}_n(t)) - v_n(t)] + \kappa G(\Delta v_n(t), \Delta \tilde{v}_n(t)) \qquad (1)$$

where $\Delta x_n = x_{n+1} - x_n$ and $v_n(t)$ are the position of car n at time t respectively. a and κ are the sensitivity of a driver and the sensitivity coefficient of response to the stimulus $G(\cdot)$. $\Delta x_n = x_{n+1} - x_n$ and $\Delta v_n(t) = v_{n+1} - v_n$ are the headway and the velocity difference between the preceding vehicle $n+1$ and the following vehicle n, respectively, $\Delta \tilde{x}_n = x_{n+2} - x_n$ and $\Delta \tilde{v}_n(t) = v_{n+2} - v_n$ are the headway and the velocity difference between the car $n+2$ and the car n, respectively. The functions $V(\cdot)$ and $G(\cdot)$ are assumed as follows:

$$V(\Delta x_n, \Delta \tilde{x}_n) = (1 - p_n)V(\Delta x_n) + p_n V(\Delta \tilde{x}_n) \qquad (2)$$

$$G(\Delta v_n(t), \Delta \tilde{v}_n(t)) = (1 - p_n)\Delta v_n(t) + p_n \Delta \tilde{v}_n(t) \qquad (3)$$

where $p_n = LS_n/LS_{max}$ is the effect parameter of lateral separation distance, LS_n is the lateral separation distance between car n and site $n+1$, and LS_{max} is the maximal lateral separation distance which means that the front car is on another lane and has no impact on the following car. LS_{max} can be set as 3.6 m (typical lane width) (Jin et al. 2010; Peng et al. 2011; Gunay 2007). If $LS_n = 0$ ($p_n = 0$), Eq. (1) of the car-following model reduce into full velocity difference model in Jiang et al. (2001), which represents that there was no lateral separation between car n and car $n+1$. If $LS_n = LS_{max}$ ($p_n = 1$), which implies that the car $n+1$ is on another lane and the car $n+1$ has no effect on car n.

In order to obtain the corresponding macro model, we assume that the state of the car n at position x represents the average traffic condition at $\left[x - \frac{1}{2}\Delta, x + \frac{1}{2}\Delta\right]$. And suppose $V(\Delta \tilde{x}_n) \cong V(\Delta x_n) + V(\Delta x_{n+1})$ for simplicity. Thus the above micro variables into the macro ones are adopted with the method in Liu et al. (1998) and Jiang et al. (2002), i.e.

$$v_n(t) \to u(x, t), v_{n+1}(t) \to u(x + \Delta, t), v_{n+2}(t) \to u(x + 2\Delta, t),$$

$$\Delta x \to \rho, V(\Delta x) \to u_e(\rho), a \to \frac{1}{T}, \kappa \to \frac{1}{\tau}, p_n \to p \qquad (4)$$

where ρ and u are respectively the density and speed at the point (x, t); $u_e(\rho)$ is the equilibrium speeds; T and τ are the reactive time and the time of the backward propagated disturbance to travel a distance of Δ. Thus, Eq. (1) can be shifted as

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \frac{(1+p)u_e(\rho) - u}{T}$$
$$+ \frac{(1-p)[u(x+\Delta,t) - u(x,t)] + p[u(x+2\Delta,t) - u(x,t)]}{\tau} \quad (5)$$

By expanding Eq. (5), neglecting the nonlinear terms and combining the conservative equation, we can obtain a new macro model, i.e.

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} = 0$$
$$\frac{\partial u}{\partial t} + [u - (1+p)c_0]\frac{\partial u}{\partial x} = \frac{(1+p)u_e(\rho) - u}{T} \quad (6)$$

where $c_0 = \frac{\Delta}{\tau}$ is propagating speed of the small perturbation. Equation (6) can be rewrote as follows:

$$\frac{\partial U}{\partial t} + [A]\frac{\partial U}{\partial x} = B \quad (7)$$

where

$$U = \begin{bmatrix} \rho \\ u \end{bmatrix}, \ [A] = \begin{bmatrix} u & \rho \\ 0 & u - (1+p)c_0 \end{bmatrix}, \ B = \begin{bmatrix} 0 \\ \frac{(1+p)u_e(\rho) - u}{T} \end{bmatrix} \quad (8)$$

The eigenvalues of the $[A]$ matrix can be set as

$$\det |[A] - \lambda[I]| = 0 \quad (9)$$

The characteristic speeds of Eq. (6) can be obtained as follows:

$$\lambda_1 = u - (1+p)c_0, \text{ and } \lambda_2 = u \quad (10)$$

There are no greater than the speed u because $(1+p)c_0 > 0$. This result ensures that this effect cannot change the anisotropy of traffic flow (Daganzo 1995). Further more, we can obtain the stability condition of Eq. (6) by using the similar method (Jiang et al. 2002), i.e.

$$\lambda_1 < c < \lambda_2 \quad (11)$$

where $c = \frac{d(\rho u)}{d\rho}$ is the kinematic wave speed, When Eq. (11) is violated, traffic will be unstable, which leads to stop-and-go wave occurring. When $p = 0$, the result is same as that of speed gradient model of macro traffic flow in Jiang et al. (2002). Obviously, λ_1 decreases with the parameter p increasing irrelevant to the lateral effects of the lane width, which shows that the stable region of traffic flow increases with the lateral effects of the lane width.

3   Numerical Simulation

For small p, the real overtaking actions do not occur in the simulation. By applying the finite difference method to discretize Eq. 6, we can obtain the difference equations as follows:

$$\rho_i^{j+1} = \rho_i^j + \frac{\Delta t}{\Delta x} \rho_i^j \left(u_i^j - u_{i+1}^j \right) + \frac{\Delta t}{\Delta x} u_i^j \left(\rho_{i-1}^j - \rho_i^j \right) \tag{12}$$

If $u_i^j < (1 + p)c_0$,

$$u_i^{j+1} = u_i^j + \frac{\Delta t}{\Delta x} \left[(1 + p)c_0 - u_i^j \right] \left(u_{i+1}^j - u_i^j \right) + \frac{\Delta t}{T} \left[(1 + p)u_e(\rho) - u_i^j \right] \tag{13a}$$

else $u_i^j \geq (1 + p)c_0$,

$$u_i^{j+1} = u_i^j + \frac{\Delta t}{\Delta x} \left[(1 + p)c_0 - u_i^j \right] \left(u_i^j - u_{i-1}^j \right) + \frac{\Delta t}{T} \left[(1 + p)u_e(\rho) - u_i^j \right] \tag{13b}$$

where ρ_i^j and u_i^j are the density and speed at the point (i,j) respectively, and i, j, Δx and Δt are the space index, the time index, the spatial step and the time step respectively. To investigate congestion and dissipation of the traffic flow by describing shock and rarefaction wave, two Riemann initial conditions (Jiang et al. 2002) are adopted as follows:

$$\rho_u^1 = 0.04, \ \rho_d^1 = 0.18 \tag{14a}$$

$$\rho_u^2 = 0.18, \ \rho_d^2 = 0.04 \tag{14b}$$

where ρ_u and ρ_d are the upstream density and downstream density, respectively. Conditions (14a) and (14b) correspond to shock and rarefaction wave, respectively. The initial conditions of speeds are

$$u_u^{1,2} = u_e(\rho_u^{1,2}), \ u_d^{1,2} = u_e(\rho_d^{1,2}) \tag{15}$$
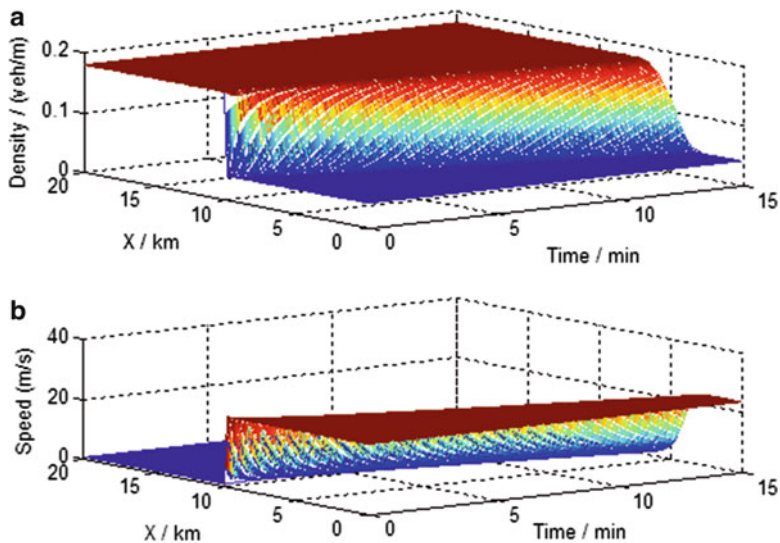
Fig. 1 Shock wave under the initial conditions (14a)

Free boundary conditions are applied here, $\partial k/\partial x = 0$ and $\partial u/\partial x = 0$. The equilibrium speed-density relationships (Del Castillo and Benitez 1995) are adopted as follows:

$$u_e(\rho) = u_f \left[1 - \exp \left(1 - \exp \left(\frac{c_m}{u_f} \left(\frac{\rho_m}{\rho} - 1 \right) \right) \right) \right] \tag{16}$$

where ρ_m, u_f, c_m are the jam densities, free-flow velocity and the propagation velocity of a disturbance under the jam density ρ_m, respectively. The values of the parameters are chosen as follows (Jiang et al. 2002):

$$u_f = 30 \text{ m/s}, \quad \rho_m = 0.2 \text{ veh/m}, \quad T = 7 \text{ s}, \quad c_0 = c_m = 6 \text{ m/s},$$

$$\Delta x = 200 \, m, \quad \Delta t = 1 \, s, \quad p = 0.1$$

The computational results under the two Riemann initial conditions (14a) and (14b) are shown in Figs. 1 and 2, respectively.

From the Figs. 1 and 2, the different Riemann initial conditions lead to different fronts between the congested and free-flow traffic corresponding to shock and rarefaction wave, which means that our model can effectively reproduce the formation and the propagation of shock and rarefaction wave, respectively. The shock wave in Fig. 1 means that traffic becomes more congested, which we often see in rush hours. The rarefaction wave in Fig. 2 shows a queue in the process of dissolution, which is consistent with our daily experiences in real traffic. In addition, the wave front with the consideration of the lateral effects of the lane width is smoother than the wave front (Jiang et al. 2002), which means that the lateral effects of the lane width conduce to smooth the wave front.
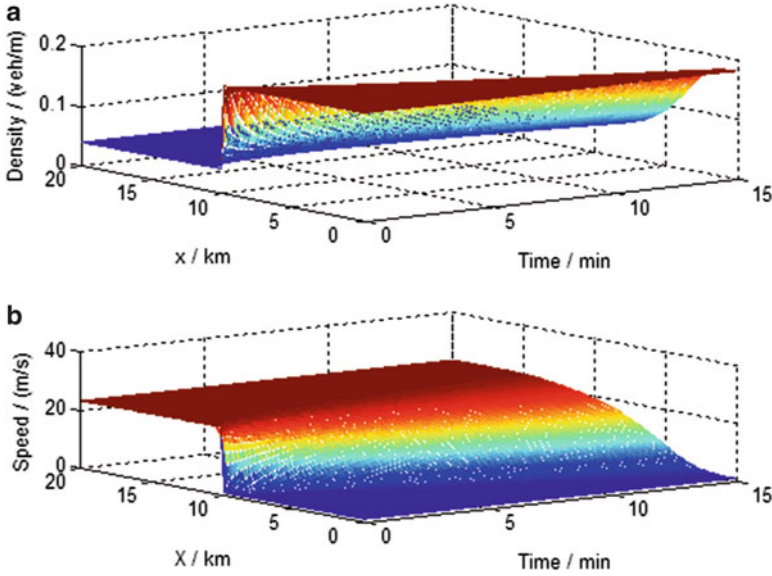
Fig. 2 Rarefaction wave under the initial conditions (14b)

Next, we study the equilibrium solution of the macro models with the consideration of the lateral effects of the lane width. Figure 3 shows the flux-density and velocity-density relationships of uniform traffic flow for $p = 0, 0.1, 0.2, 0.3$. From these curves in Fig. 3a, the peak value of flux at critical density increases with the increase of p, which shows that the lateral effects of the lane width can increase capacity of traffic flow. According to Fig. 3, we can see that the lateral effects of the lane width enhance the equilibrium velocity and flux. In addition, the lateral effects of the lane width have more effect on the equilibrium velocity and flow with the increase of density when the traffic density is below the critical density while it has less impact on the equilibrium velocity and flow with the increase of density when the traffic density is above the critical density. From Fig. 3, an important feature of the fundamental diagram is that the lateral separation effect improves flux density curve at higher density.

4 Conclusion

The lateral effects of the lane width can play a great role on car following behavior. In this paper, a non-lane-based macro model is proposed with the consideration of the lateral effects of the lane width based on car-following model by using the relationship between the micro and macro variables. The results imply that the effects of the lateral separation can effectively stabilize traffic flow. The numerical results show that our model can describe the evolution and propagation of shock and
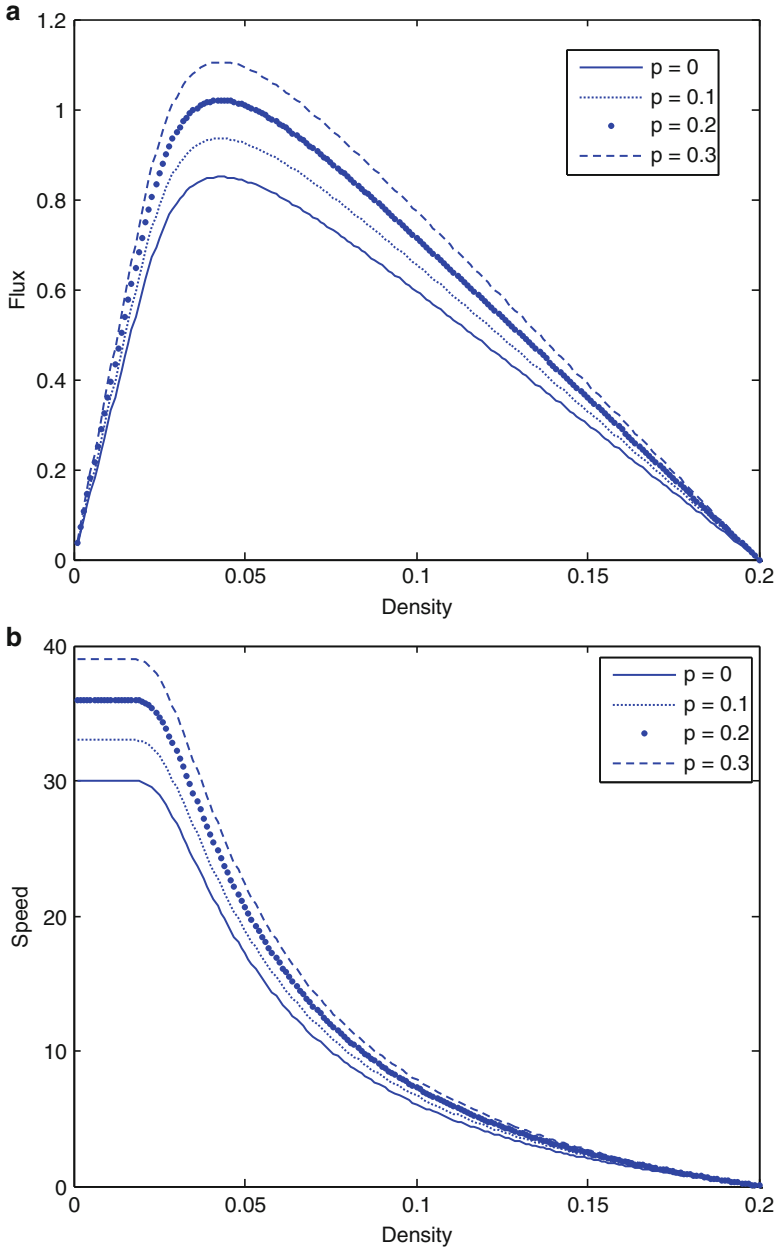
Fig. 3 The flow-density and velocity-density relationships of uniform traffic flow

rarefaction wave and that the lateral effects of the lane width can greatly improve the flux capacity of traffic flow, which shows that the lateral separation caused by wide lane in macro model has a great significance upon the property of traffic flow. Simulation results also verify that our consideration is reasonable and necessary.

References

Daganzo CF (1995) Requiem for second-order fluid approximations of traffic flow. Trans Res B 29:277–286

Del Castillo JM, Benitez FG (1995) On functional form of the speed-density relationship-I: general theory, II: empirical investigation. Trans Res B 29:373–406

Gunay B (2007) Car following theory with lateral discomfort. Trans Res B 41:722–735

Jiang R, Wu QS, Zhu ZJ (2001) Full velocity difference model for a car-following theory. Phys Rev E 64:017101.1–017101.4

Jiang R, Wu QS, Zhu ZJ (2002) A new continuum model for traffic flow and numerical tests. Trans Res B 36:405–419

Jin S, Wang DH, Tao PF, Li PF (2010) Non-lane-based full velocity difference car following model. Phys A 389:4654–4662

Liu GQ, Lyrintzis AS, Michalopoulos PG (1998) Improved high-order model for freeway traffic flow. Trans Res Rec 1644:37–46

Peng GH, Cai XH, Cao BF, Liu CQ (2011) Non-lane-based lattice hydrodynamic model of traffic flow considering the lateral effects of the lane width. Phys Lett A 375:2823

Improvement of Information Transmission Quality of Digital Libraries

Li Jian and Wang Beijun

Abstract This paper analyses influence factors of information transmission quality of digital libraries. The advantages and disadvantages of information transmission mode of digital libraries are discriminated. Strategies of maximize ascension of information transmission quality of digital libraries are put forward. Implement the strategies can effective control service process of digital libraries so that users' expectation on service quality can be realized in the process of using the library. Then, the gap between users' expectations and the service implementation can be maximum narrowed, thus eventually improve users' satisfaction on digital libraries.

Keywords Information transmission • Digital libraries • Digital information • QoS (quality of service)

1 Introduction

A Digital library is a managed collection of distributed digital objects (contents), and services (functionality) associated with the discovery, storage, retrieval, and preservation of those objects in a context of networks, specifically the Internet.

The services of digital libraries are manifest by transmission of digital information resources. The effectiveness of digital libraries mainly embodies in quality of

L. Jian (✉)
College of Computer and Information Science, Southwest University, Chongqing, China
e-mail: lhealthy@126.com

W. Beijun
College of Economics and Management, Southwest University, Chongqing, China
e-mail: wangbeij@swu.edu.cn

the information transmission. Study on the transmission quality of digital library is an important research topic. Quality is the difference value between the user expectations and actual services effect on the user. Improve service quality of digital libraries mainly concentrates on increase of user's satisfaction on transmission quality of digital resources. According to the model of information transfer, there are three major factors should be considered to enhance the information transmission quality of digital library, they are information source, information channel and information receiver. In digital library, the factors actually are the organization certain of information, efficiency of transmission, and user's demand on digital resource.

2 Digital Library and Its Mode of Information Transmission

2.1 Concept of Digital Library

DLF (Digital Library Federation) regards digital libraries as organizations that provide the resources, the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities. (Digital Library Federation's Working Definition 1998) The British library defined digital library as using a digital technology acquisition, storage, access, release information of the library.

2.2 Mode of Information Transmission of Digital Libraries

Digital libraries provide convenient, quick service mechanism to find the exact information for users. Digital libraries composed by the distributed of mass and organized database. The digital content may be stored locally, or accessed remotely via computer networks. A digital library is a type of information retrieval system. Users can access databases and knowledge in the system via unified retrieval interface. By transmitting information with the net, users obtain information or knowledge they needed. Digital libraries widely adopt Internet as network platform. Internet is the most widely used network in the world. It breaks the space-time obstacle of communication in real world, so digital libraries gain the favors of users for its easy-to-use and flexible resource sharing style. However, service type of digital libraries is a kind of 'Best Effect' for they adopt TCP/IP protocol as the basic framework of the Internet, which lack of control. It is disadvantageous to the library users who have specific requirements for information resources.

3   Advantages and Disadvantages of Information Transmission Mode of Digital Libraries

3.1   Advantages

3.1.1   User-Centered Service Mode

A big difference between digital libraries from traditional libraries is easier to take users' demand as the center. The user-centered service mode includes friendly human-machine interface and information space navigation function, rapid transfer of content function, powerful retrieval tool and advanced information processing and analysis tool, readily available guidance, the round-the-clock uninterrupted information retrieval, processing and transmitting.

3.1.2   Digital Information Resources

Digital libraries collect various media of information and store them in combination unity by computer technology. The digital information is managed as areolar tissue by the hyperlink, so that facilitate tectonic interrelated knowledge system.

3.1.3   Networked Information Transmission

Digital library's collections are stored in digital formats and accessible by computers (Greenstein and Elizabeth 2002). The information transmissions by network make digital library transcend time and space, across the collection of information, and accelerate regional boundaries information exchange and feedback speed.

3.2   Disadvantages

3.2.1   Difficult for Users to Put Forward Resources Transmission Requirements

Digital libraries conduct all kinds of information as a united binary symbol processing, and erase differences of all kinds of data as well as treat the data identical. Digital libraries are difficult to transfer information according to different quality level. There are actually different satisfied expectations to different information when users propose information acquisition needs. But the users cannot express their transmission requirements in existing digital library system, and there is no entrance to choose their own satisfactory transmission index in the digital library interface.

3.2.2 Unable to Distinguish Different Service Type

There are various kinds of service types in digital libraries. The distinction is tremendous among different kinds of services and their quality requirements. Transmission of voice and video requires real-time and continuous service. Text download is not very sensitive to time delay, but need high reliability of data. But it is hard to distinguish such service based on the existing Internet environment of the digital libraries.

3.2.3 Lack of Quality Assurance of Resources

Resources in digital libraries are abundant, which include text, images, graphics, audio, and video etc. it is entirely different of network occupy between different forms of media resources. But there is little control on the load of transmission process on existing network, and there is almost no warrant of service quality requirements for different transmission to different users.

3.2.4 Disorder of Resources Transmission

Usage of digital library is still essentially a kind of network application and inevitable competitive for network resources. Contention of network resources includes competition of the terminal system resources (computing resources, storage resources) and network resources (bandwidth resources). Applications of multitask on terminal system cause competition on sharing resource. Applications of multi-user in the network cause competition on network resources. The traditional best-effort network is always in the condition of disorder and uncontrolled competition.

It can be seen so many problems of IP-networked digital library. We must comprehensive thinking about how to ascension information transmission quality.

4 Strategy of Information Transmission Quality Promotion of Digital Library

4.1 Foreground: Users Can Put Forward Information Transmission Quality Requirements

Digital library should try to embody user-centered idea. Ascension of transmission quality is not only reflected in faster speed of data transmission, but also reflected in whether users' needs can be expressed and whether users can be aware of their needs be attached importance in the digital library system.

The users are most care of both service process and service results. So, it should attach importance to the digital library user interface design in the first place. Graphics and 3D technology are adopted to improve intuitive and usability of the user interface. And entrance should be designed for users to put forward information transmission quality requirements through selecting proper index of the information transmission quality. The system operations with the default index when users do not choose. In the second place, it should provide totally function modules on the interface to facilitate users choose appropriate service. The module includes browsing, search and download.

4.2   Background: Resources Are Organized Effectively

In order to improve the information transmission efficiency, it is very necessary to sort and organize information in the background. First, information should be classified scientifically accord with scientific classification system. It is necessary to consider the characteristics of network information on the base of traditional classification, such as UDC or DDC. Second, the classified information should be indexed and bibliography. The appropriate description of information is facilitating the rapid retrieval and extraction information within distributed system. Third, in order to maintain consistency for retrieved information, metadata should be the foundation of bibliographic organization of information.

4.3   Channel: Information Transmitted Accord with QoS (Quality of Service) Requirements

For ascension transmission quality of information, control of transmission process is the most important segment. At present, environment of most digital libraries is the IP network, and information transfers in IP network. So services are easily lack of quality guarantee. It is necessary to accord with QoS.

ITU-T defined QoS as a set of comprehensive indexes used to measure the service satisfaction. (Recommendation E.800 1994) So in the simple sense, qualified service is able to meet the user's application demand, or, can provide a consistent, predicted data delivery service. ISO research network's QoS at first place. QoS is described as a series of service request on data transmission in network in ISO RFC2386. (Crawley et al. 1998) QoS can be concrete as performance indicators as bandwidth, delay, jitter, throughput. QoS reflects the ability in ensuring information transmission and satisfying service requirements of network elements (such as application program, host or router).

4.3.1 Distinguish Different Data Streams to Provide Discriminated Services

When users' requirements are approved, network resource will be allocated based on circumstance. Protocol of DiffServ (Differentiated Services) is invoked immediately to distinguish services. The services are classified, data flows are plastic, buffer is managed and errors are controlled under the protocol. Then different information services are labeled different processing requirements. Data flows of shared characteristics are aggregated, thus a variety of applications are gathered to a limited data stream data flow.

4.3.2 Sequence Control and Quality Guarantee in the Information Transmission Process

- QoS mechanisms of DiffServ and MPLS (MultiProtocol Label Switch) can be adopted to plan throughput and reduce congestion. DiffServ need lots of network units corporate to provide users good QoS, but these units are often highly dispersed. Therefore, although DiffServ is the most feasible QoS scheme on the backbone network currently, it is difficult to provide better QoS guarantee by itself. MPLS can solve this problem through auxiliary DiffServ to realize quality assurance of end-to-end services. Because DiffServ and MPLS are some certain similarly (QoS aspects such as packet classification), it is relatively simple to mapping distinguish service data stream to MPLS conduit. This mechanism is used to improve data flow in a transmission channel of congestion in digital libraries.
- QoS mechanisms of IntServ (Integrated Services) and DiffServ can be adopt to improve the connection reliability. IntServ provides three levels of service guarantee, which includes assurance service, load control service and best-effort service. The three level of service support are lower and lower. The biggest advantage of IntServ is totally customized network service of users according to the actual needs of users. It is considerable flexible. Its disadvantages are complexity of realize and weakness of expansible. So IntServ is difficult used in the backbone. DiffServ's advantage is expansibility in the backbone by contrast. But DiffServ can't handle a specific flow of request because it handles the gathered stream. Therefore, IntServ and DiffServ can be combined mutual and coordinated to provide strict end-to-end QoS guarantee. This model can reduce the error rates including connection failure and transmission failure in digital library so as to improve the reliability of the connection.
- QoS mechanisms of DiffServ and RSVP (ReSource reserVation Protocol) can be adopted to reduce delay and lower error rates. DiffServ identify and mark different package, then classify them and give different priority to the distinguished package. RSVP can provide reserved resources for gathered data flow, thus provide a more flexible end-to-end QoS. In this mode, RSVP is decorated in the edge network, whereas DiffServ is arranged in the core network. RSVP and

DiffServ are unified to reduce time delay and error rate of the communication process in digital libraries.

- QoS mechanisms MPLS and RSVP can be adopted to increase speed and reduce packet loss rate. MPLS give data flow short, fixed-length and unstructured tags and transmit package according to the tags. Meanwhile, in order to realize effective tag distribution and transmitting, RSVP can be combined with MPLS. Label requirements (LR) are added in RSVP message path. Label object (Label) is added in the reserve message. So, hosts and routers with RSVP and MPLS can associate label with specific data flow. Routers can determine appropriate information on reserve resources according to the label of packets. This model can greatly improve the transmission speed of data flow of digital libraries; especially loss rate of digital information can be reduced greatly.

References

Crawley E et al (1998) A framework for QoS-based routing in the internet, RFC2386, IETF

Digital Library Federation's Working Definition (1998) http://www.diglib.org/about/dldefinition. htm

Greenstein DI, Elizabeth TS (2002) The digital library: a biography. Digital Library Federation/Council on Library and Information Resources, Washington, DC

Recommendation E.800 (1994) Terms and definitions related to quality of service and network performance including dependability, ITU-T

Design and Implement of Enterprise Knowledge Management System

Meijing Guan, Jin Shan, Wang Ming, Zhou Hui, and Qingzhang Chen

Abstract Knowledge has become an important business asset with the development of the knowledge-based economy and the need of the flexibly respond to the changing market needs. Take Electric Power Test Research Institute of Zhejiang Province for example. Their knowledge can't be easily managed because of the complexity and the large quantity. It is difficult for the engineers to find the knowledge needed. Therefore, the establishment of a knowledge management system has become a necessity. In this paper, the Knowledge Management System of Zhejiang Electric Power Test Research Institute of Collaborative is introduced, especially the Knowledge Management System Design and Implementation.

Keywords Electric power research • Collaboration • Knowledge management • Enterprise information technology

1 Introduction

With the deepening of the reform of electricity industry and the competition of electricity service market. The existing market share of Zhejiang Electric Power Test Research Institute is being challenged, how to respond to the fast-changing market

In case of Zhejiang Electric Power Test and Research Institute

M. Guan (✉) • J. Shan • W. Ming
College of Information Engineering, Ningbo DaHongYing University, Ningbo, China
e-mail: puddingguan@163.com

Z. Hui
Zhejiang Electric Power Test and Research Institute, Hangzhou, China

Q. Chen
College of Computer, Zhejiang University of Technology, Hangzhou, China
e-mail: qzchen@zjut.edu.cn

appropriately and standardize the operations of enterprises at the same time become a pressing problem. Due to the reality, there are numerous departments, management departments and kinds of professional departments, there are different employees of different levels. How to manage them efficiently becomes an important issue for the leaders. And at the same time, large quantity of knowledge is accumulated during the developing process of the enterprise, their logical relations is very complex. So when the engineers are making designs, the search result they get are often not what they want, the engineers will pay more time to filter these results to get the information they want, it will influence the efficient use of these knowledge. Thus, the establishment of an enterprise knowledge management system is urgent.

In this paper, Theory research and practice is combined, take enterprise knowledge management system of Zhejiang electric power test research institute for an example, noting that only through the construction of enterprise knowledge management in order to effectively promote innovation and development of enterprises. Synergistic combination of gear linkage matrix model and design model, on how knowledge from the enterprise to obtain all the information platform for decomposition analysis, not only deepened the understanding of the theory of the previous, but also put forward including the importance of collaborative knowledge-building awareness, knowledge with the existing combination of management information systems, knowledge and business process integration, knowledge management and competence in five areas such as knowledge-sharing problems, and made a number of ideas to solve problems and cases.

2 System Design

2.1 Ontology Theory and Ontology Modeling

Knowledge of how the information is represented, how it is organized, especially due to the rapid development of the Internet, in the face of the ocean of information, how to organize, manage and maintain information from a mass has become the field of enterprise knowledge management an important and urgent issues to be addressed. Ontology (Ontology) through the concept of the strict definition and concept of the relationship between the concept to determine the precise meaning of the concept that recognized the common, to share knowledge, in this series support the concept of knowledge search, knowledge accumulation, knowledge sharing will greatly enhance the efficiency, so that the true sense of the knowledge reuse and knowledge sharing possible (Haojiang 2006).

Generally speaking, the concept of an ontology model include five kinds of elements such as category, relation, function, axiom and instance, thinking of the application of knowledge management system features, as well as drawing on relevant information, given that the form of a collection of ontology: ontology = (N, F, C, A, R1, R2). Of which, N is the name of ontology; F is N of the paternal ontology; C is composed of a collection of the concept of ontology; A collection of attributes

is the ontology; R1 is the intimate link between the concept, by the C of elements in the relationship between composition; R2 is the correlation between ontology. Knowledge management system in the ontology of knowledge in all there is a certain degree of inter-linked, on the basis of the above system is given in the form of ontology, the paper design of the ontology used to express the association rules between knowledge (Liu et al. 2008).

2.1.1 Ontology-Based Knowledge Base Model for Power Grid Operation

Knowledge base means the business knowledge of the database system; related data and information of power system such as record text and graphic files (such as SVG). Knowledge of business systems by the source can be extracted from the local ontology. Local ontology is the concept of a business system with a description of the terms does not apply to other business systems. Therefore, the ontology needs to be integrated into the overall situation of the local ontology to form a common business system for each of the conceptual model. Ontology is the overall power system operation terms and concepts of the unified description of all business systems knowledge in the field of understanding goes for justice. Goes for the justice of both the goal of information integration is also the basis of information sharing. Ontological link between the overall combined to form a grid to run the knowledge base can store not only the structure of such a static grid of knowledge, to highlight the characteristics of power system operation can be stored in the dynamic knowledge. Ontology knowledge can eliminate the data redundancy and support unified query of the knowledge to implement the share of information (Bian et al. 2008).

The knowledge base of power grid is described by the overall ontology of power sector. Overall ontology is divided into physical ontology and incident ontology. Physical ontology, also known as a static ontology, the power of knowledge to describe the concept of a static model, the concept of property, or bound by the conditions of justice, as well as the relationship between concepts. In the power domain ontology model, the objective refers to the physical ontology, the physical existence of the power system entities, such as components, equipment and a variety of devices to connect topology, also known as equipment, secondary equipment, as well as a variety of metering equipment. IEC61970 series of standards have been well defined within the scope of the vast majority of the data model with some data access interface.

The Definition of Classes

The physical ontology type of the objective world in accordance with the concept of the definition reflects the objective existence of classes of substances or objects, while the relationship between class reveals how their mutual structure. In the CIM model, the relationship between classes may be inheritance, association, aggregation etc.
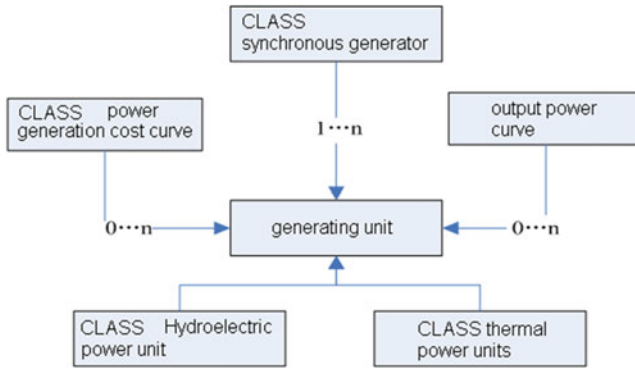
Fig. 1 Generator set model

1. the succession is a more common type (parent type) and a more specific class (subclass) of a relationship between. Subclass level inherits attributes and relationships from its parent.
2. correlation between the classes is a conceptual link. Each of the two roles is related. Role of each association in a direction that the role of target and source class is related classes. The objective of the role of a given class name, you can drive with or without words. Multiplicity of each role there/base, used to indicate how many objects can participate in a given relationship.
3. aggregation is associated with a special case. Classes and types of aggregation indicate that the relationship between the part of an overall relationship. Part of the overall class by class or contains some type of pose, and some is part of the overall class. CIM in the generating units described in class a brief description of the relationship between various types as shown below (Fig. 1).

Types of generating units and the synchronous generator is a simple relationship between the association, a synchronous generator is 0 or a generating unit, a generating unit may contain 1-n of synchronous generator. Hydroelectric generating set classes and categories of thermal power units of the generating units from the upper class to inherit its attributes and relationships. Type of power generation cost curve and the generator set is the relationship between aggregations, a generating set contains a $0 \sim n$ curve of power generation costs.

Collaborative Ontology-Based Workflow Model

We define the expandable workflow machine as a three-tire model: the core layer, the initial expansion layer and application expansion layer, as shown below (Fig. 2).

1. The core layer: workflow process model and workflow involved in the workflow object model constitutes the core layer.
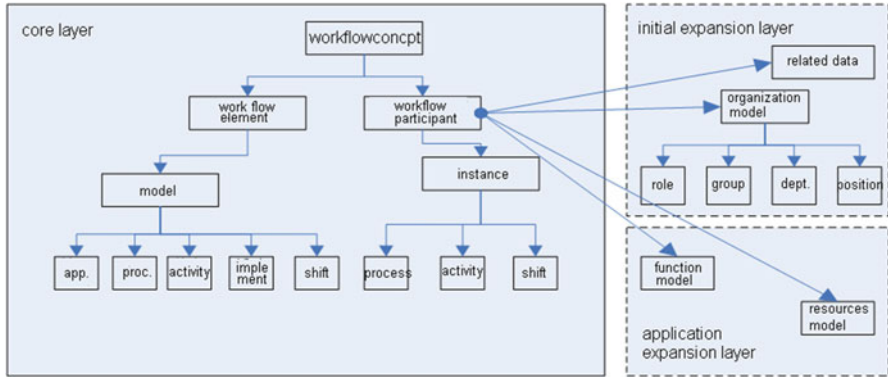
Fig. 2 Workflow meta-model of the ontology level
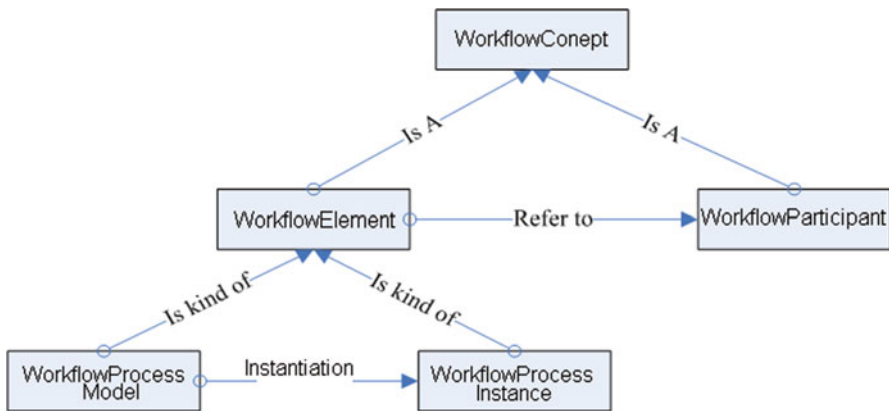
Fig. 3 Workflow ontology model of the core layer

2. The initial expansion layer: organizational models and related data model (or information model) is necessary for the process reference model, but specific applications may be provided by the system, as the initial layer expansion.
3. The application expansion layer: applications such as function model, resource model (product model), and the value chain model should turn to the needs of enterprises. This need can be defined as the application expansion layer, Expansion through the application layer can be personalized to achieve the needs of the enterprise (Fig. 3).

Root for the concept of workflow (WorkflowConcept), this is the most basic class, which includes a sub-class of general-purpose properties. Workflow Process elements (WorkflowElement) and workflow involved in the object (WorkflowParticipant) is a subclass of Workflow Concept. Involved in the process of definition of the concept of Workflow Element (workflowprocess elements) of the subclass,
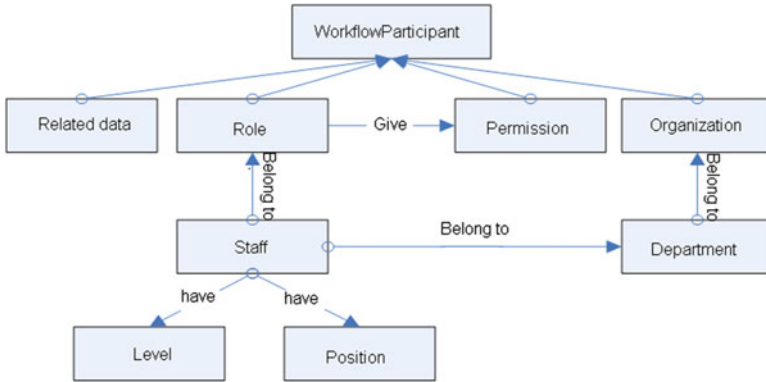
Fig. 4 The initial expansion of layer

Workflow Process (process), Activity (activities), Connector (Connector), Rule (rule), Transition (transfer) and other types of inherited Workflow Element. Routing rules (Rule) contained in the transfer of part of the conditions and terms. Core elements from the workflow process (WorkflowElement) and workflow involved in the object (WorkflowParticipant) component.

The initial expansion layer is composed by organization model and information model (workflow relevant data, WorkflowRelaData). Organization model and information model is to participate in the workflow object (WorkflowParticipant) sub-classes, the work flow through the expansion of participation of target groups. Application layer, including the expansion of the functional model and resource model are, of course, WorkflowParticipant sub-classes, which combines the application of business through succession workflow object class to participate in the work flow to achieve the expansion of ontology, namely the expansion of the work flow model. Workflow Participants (WorkflowParticipant) the definition of the object name, object path, object attributes and object operations, such as object type attributes (Fig. 4).

Workflow participant is used to define all the resources used in the workflow or participate in the workflow. Each instance in it is a class of the enterprise model such as role, application, devices, document etc. A workflow participant has 0 or more attributes, and the attribute include attribute name, data type, default value etc. Operation is defined as the methods of the participants, each method have name and return type, an operation can have 0 or more parameters, and each parameter have its data type and transmission direction.

2.2 Collaborative Matrix and Gear Linkage Model

Any enterprise has seven basic elements: the elements of people, elements of process, elements of knowledge, the elements of client resources, the elements

of the projects, material elements and financial elements; the seven factors was not isolated, there must be strong correlation between them, at the same time these elements in the enterprise organization flow between various departments. As the leaders of an enterprise, their requirements of information is full ranged, for example, the leader want to know about an employee, the leader will concern about: how he finishes all his tasks in the enterprise? How he contributes and accumulates his knowledge in the enterprise? What is his client and supplier like? How is the condition of the project he is responsible for? When the leader concerning about an employee, he will care about not only the archives but all the information mentioned above (Delalondre et al. 2010).

Enterprises have been looking for a method to manage all the elements of all the knowledge in the enterprise in a "three dimensional multi-threaded" platform. "Coordination matrix model" give a suitable solution for the enterprises on this key problem.

Coordination matrix model: an effective solution to enterprise information management needs mesh. Matrix model collaborative enterprises to effectively address the various elements of the various departments are coordinated network structure and operation of "three-dimensional multi-threaded" needs to provide the solution: in the system, if users find an information point, with the associated information for all points information has been found. For example: to find a company's internal sales staff, then with the sales-related: personal financial information (wages, benefits, costs and expenses), he managed client, he wrote the documentation, he subordinates the management, he to participate in the project, he used the company's assets, the organization of the work he has all the information like a network of personal data by employees of the information nodes to be extracted quickly (Francisco et al. 2007).

Coordination matrix model is an effective solution to the "network enterprise information management" demand management platform for collaboration, this is the logic of the ideological dimensions of the system and another system is also a need for power transmission in the need for a model to support "synergy matrix" operations.

Gear linkage model: the effective delivery of the core engine power and coordination to support the operation of matrix. Gear linkage model is the key to the design: The system provides a core engine (similar to "gear school Hopewell Center") to support the various system application modules (each module similar to a gear): knowledge of document management modules, workflow application management modules, human resources application management modules, customer relationship management modules, project management module, financial management and asset management module of the collaborative operation of modules, as long as a one module (gear) rotation can be more agreeable to the Center through the gear drive other modules (gear) rotation. The operation of a module to bring the operation of the other six modules, greatly enhance the function of the module, while improving systems, and meet the requirements of collaborative management of enterprises. (http://carparts.breakeryard.com/cheapcarspares/Gear-linkage.aspx)

Collaborative knowledge management matrix model based on collaboration and interaction model of gear design. Matrix model based on collaborative information network map the characteristics of knowledge management and other strongly correlated with the seven modules. When we find in the system when a document with the documents related to all of the information has been extracted: with this document related to human resources, projects, assets, etc., makes this document all information related to the structure of organized the show (Rawson and McCann 2005).

3   System Implementation

Our system uses Windows 2003 Server Edition and J2EE architecture platform, and it is developed based on the structure of B/C/S three-tier model, the technology is advanced, the implementation is efficient and it is easy to maintain. Modular, process-oriented and systematic design ideas are applied in the development process. System development techniques and the main points include:

JBOSS application server architecture is used; full application of object-oriented design system design; depending on the application type, the use of middleware technologies, improve system reliability and stability; the use of relational database Oracle 10i, a wide range of integrated data management; grading and opening up the use of information, the method of licensing authority, the establishment of system security mechanisms.

3.1   The Implementation of Main Function Model

Establishment of directories for knowledge documents is an important step in the management of knowledge. In this process, we mainly consider three aspects: document directory, positioning systems and search powers. (http:// download.oracle.com/docs/cd/E12529_01/ali65/AdministratorGuide_ALI_6-5/tsk_ documents_editing.html)

3.1.1   Directory Configurations Implementation

Knowledge resources for the content of our institution are complex; systems are wide, through a unified system of classification of knowledge. We designed two kinds of classification of the document directory. One is the classification of the document directory, and the other is multi-dimensional directory; document directory categories are in accordance with the traditional "functional" approach to the management of file management, and multi-dimensional classification of the main directory of the various professional departments in accordance with customary access to classified documents.
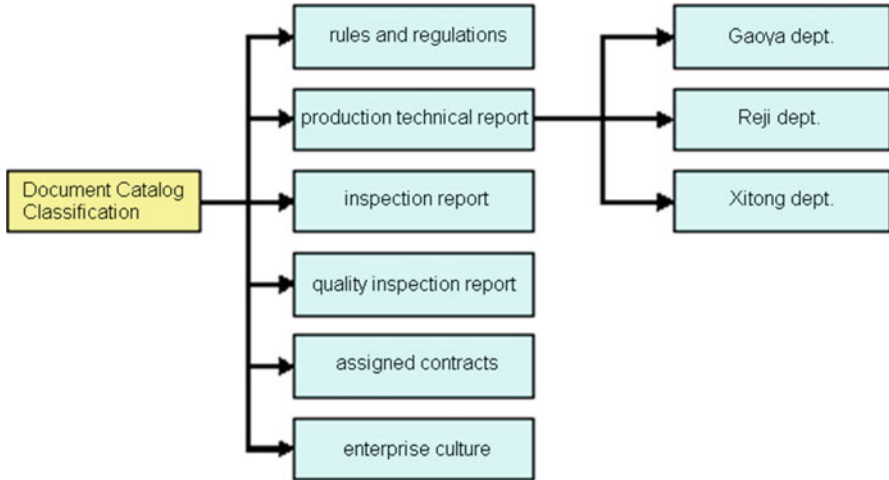
Fig. 5 Document catalog classification chart
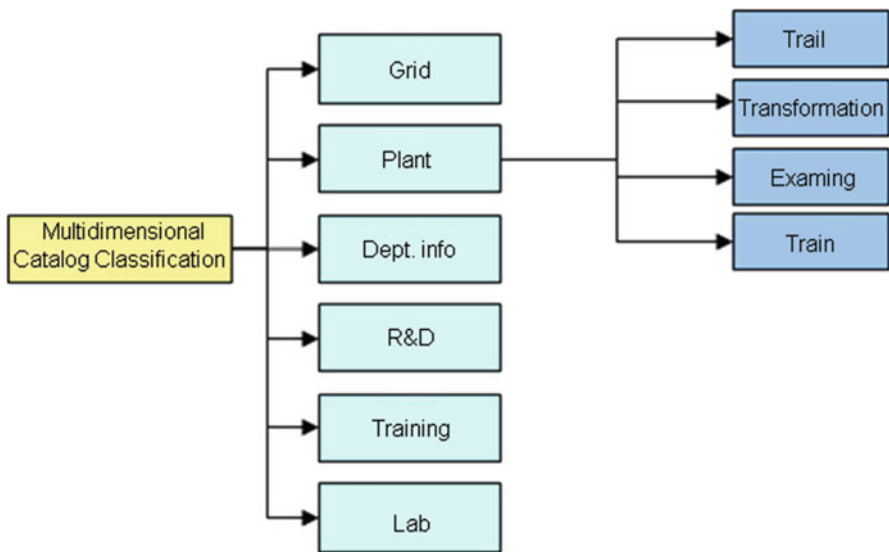
Fig. 6 Multi-catalog classification

Document directory is divided into the first tier classification: the rules of the system, producing technical reports, review (test) inspection report, inspection report, the assigned contract management, community activities, and the corporate culture. The second level is set according to our organizational structure. For example, the institution-level document is in an individual directory, and the documents of different major are placed in different directories too (Figs. 5 and 6).

3.1.2 Custom Configuration of Permissions

When all the knowledge is placed together in a knowledge management system, the applications have to ensure the safety of these knowledge and the documents. We have horizontal and vertical permissions system combination to solve this problem. As a result of a knowledge base for each directory corresponds with the appropriate departments, to find the directory and the corporate sector Knowledge corresponding mapping relations, we can set up through the horizontal authority to control documents, such as financial-related article directory applications mainly in the financial sector, business sector, mainly in the business document applications. In this way, between avoids the possibility of cross-leaked (Bernstein 1999).

Vertical file permissions set to consider two factors and user documentation. Enterprises can set up each document to the appropriate security level. Some common documents, such as corporate culture, rules and regulations, such as commonly used, its security level is relatively low, you can fully share in the enterprise. Like contracts, financial statements and other related confidential documents, only in a small range of content-sharing, the security level higher than it should be set up. Similarly, the system also can set up each user's security level, and the user's security level and the corresponding documents to the security level. In this way, the security system put the definition of vertical clear.

Horizontal and vertical security system is combined together and constitutes a more rigorous, standardized, improved document permissions system. In this way, we can ensure that the information in the document at the same time be fully utilized to minimize the chance of being leaked.

3.1.3 Location Search Implementation

Faced with massive documentation of knowledge, the user can quickly locate, retrieve the contents of the required documents. If there are only hundreds or thousands of documents, the employees can find what they want easily. However, in our institution, from the beginning of MIS use a date, the number of up to tens of thousands of documents. At this time, users need to retrieve the documents they need when they encounter difficulties, can not find what you want, or enter a Mesh search out after too much documentation, it is difficult to filter. To resolve this problem, we proceed from the following three aspects.

First, the increased ease of retrieval. This can be achieved through the establishment of a more complete realization of the naming rules.

Second, by setting the words, a summary of the core of the directory or some other information, allows users to keywords, summary and other relevant documents retrieved.

Third, use multi-dimensional classification method to search positioning document.

Table 1 Collaborative interaction relationship

Collaborative matrix	Knowledge document	Project management	Human resources	Process management
Knowledge document	Document & document collaboration	Project & document collaboration	Staff & document collaboration	Process & document collaboration
Project management	Document & project collaboration	Project & project collaboration	Staff & project collaboration	Process & project collaboration
HR	Document & staff collaboration	Project & staff collaboration	Staff & staff collaboration	Process & staff collaboration
Process management	Document & process collaboration	Project & process collaboration	Staff & process collaboration	Process & process collaboration

3.2 System Testing

1. Into the production plan – click on the "new", after the approval process, the project entered the implementation.
2. In the project, you can see project-related customer information, business statistics, technical reports, the implementation of information and human resources are synergies associated automatically.
3. Found in the system when a process with the processes related to staff, clients, projects, documents and other information can be found in the process.

Test results: The system succeed to meet the collaboration matrix model, knowledge management is the core, and other project management, human resources and process management share their information such as collaborative interaction module, the relationship is shown in the Table 1.

4 Conclusion

The system bases on the ontology theory modeling, using the collaborative matrix and gear linkage design. Matrix model based on collaborative information network map the characteristics of knowledge management and other strongly correlated with the seven modules. When we find in the system of knowledge when a document with the documents related to all the information has been extracted: with this document related to human resources, projects, assets, etc., makes this document all information related to the structure, and organized the show. The design provides a comprehensive knowledge management framework, standardization of the knowledge of the organization entirely.

References

Bernstein B (1999) Vertical and horizontal discourse: an essay. Br J Sociol Educ 20(2):157–173

Bian P, Zhang M, Shi Qi, Zhi L (2008) Research and implementation of East China power grid corporation of enterprise knowledge management. East China Electr Power 36(11)

Delalondre F, Smith C, Shephard MS (2010) Collaborative software infrastructure for adaptive multiple model simulation. Comp Method Appl Mech Eng 199(21–22):1352–1370

Haojiang (2006) Implement of collaborative knowledge management platform. China Power Enterp Manage 6:46

Liu X, Wang H, Yang G (2008) Research and development of ontology based knowledge management system. J Hebei Univ 28(5):550–554

Mesa F, Marin LM, Gámez A (2007) Matrix configuration of collaborative teams for E-Learning in higher education. In: IADIS international conference e-learning, Lisbon, Portugal

Rawson RA, McCann MJ (2005) The matrix model of intensive outpatient treatment. Behav Health Recovery Manag 3:1–37

The Research and Application of City Emergency System for Medium and Small Cities

Dong Zhong

Abstract Medium and small cities are the high proportion in the total number of cities in China, the construction of emergency system for medium and small cities is relatively backward compared with big cities, speeding up the construction of medium and small cities emergency system is imperative. This paper proposes a new design scheme based on the analysis of the traditional medium and small cities emergency system construction, which is more suitable for medium and small cities. The system architecture, the hierarchical structure and business process have been designed in this paper, it provides a good research direction for the construction of emergency system for medium and small cities.

Keywords Medium and small city emergency system • System architecture • Hierarchical structure • Business process

1 Introduction

In recent years, with the rapid development of the social and economy, the issue of city public security becomes the concerning focus, emergencies, such as earthquakes, production security accidents, malignant diseases spreading, natural disasters and other unexpected events occur frequently in domestic cities (Qi et al. 2006), it affected seriously city public security. In order to reduce the incident losses,

D. Zhong (✉)
Department of electronic and information, Hubei University of Science and Technology, Xianning, China
e-mail: Zhongdong0129@163.com

the construction of emergency response system is imperative. At present domestic some big cities like Nanning, Beijing, Tianjin, Guangzhou, Shenzhen, Shanghai and other large and super large cities have built emergency response system, they are responsible for public emergencies (Dong and Chun 2011). Due to smaller scale, financial resources shortage, construction mode and system construction, medium and small cities are relatively backward. Except for the public security and fire department, the information systems of medium and small cities emergency departments are backward, it is necessary to speed up the construction of primary emergency system and to provide better service to the people (Shoupeng 2007). This paper proposes a new design scheme based on the analysis of the traditional medium and small cities emergency system construction, which is more suitable for medium and small cities. The system architecture, the hierarchical structure and business process have been designed in this paper, it provides a good research direction for the construction of emergency system for medium and small cities.

2 The System Architecture Design of Medium and Small City Emergency System

In general, the emergency system of big cities is composed of wireless communication system, wireless communication voice dispatching system, integrated response alarm system, voice recording system, wireless data transmission system, video image system, city geographical information system, moving target positioning system, mobile communication system (Jianhua et al. 2005). The emergency system has the command scheduling function, it can achieve the rapid reaction and unified response alarm to natural disasters and accidents, disasters, public health emergencies, sudden public social security incident alarm, resorts, complaint telephones sources, it also can provide communications and information security for the decision of super-serious events (Yue 2006; WeiJun 2004). But usually the traditional emergency systems of medium and small cities cannot achieve the integration of resources, lack of unified management, lack of unified dispatching, the efficiency of response alarm is low.

This paper proposes the new system architecture suit medium and small city emergency system according to the characteristics of medium and small cities in China. Figure 1 is the system architecture diagram designing for medium and small cities emergency system, the system architecture has been divided roughly into three parts: command application layer; Linkage department data layer; Infrastructure layer.

In the paper, the center of emergency system architecture is composed mainly of response alarm system, emergency resources comprehensive search system; accident warning and evaluation system, alarm and information release system, the emergency treatment knowledge base system, emergency response processing
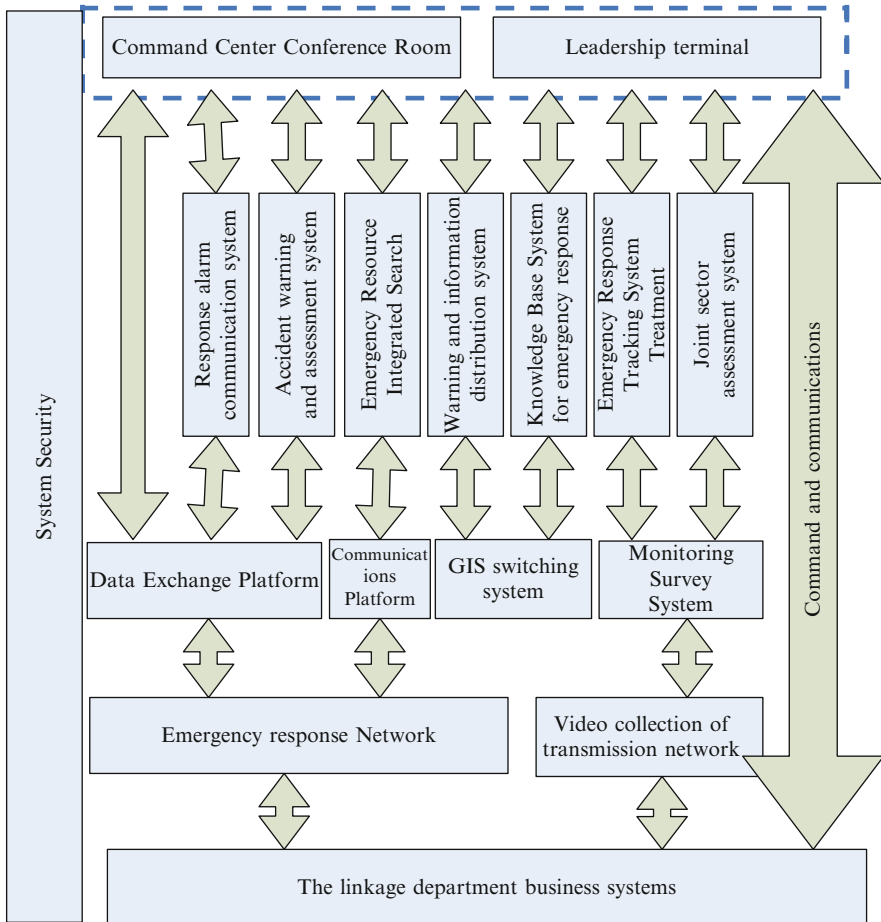
Fig. 1 Small urban emergency system architecture diagram

tracking system, the linkage department evaluation system, it has been showed in Fig. 1, the functions of main systems as follows:

1. The response alarm communication system has the phone scheduling and computer network scheduling function, wireless scheduling function. It is prior to use network scheduling in normal circumstances of public security network, if the network is impassability, the system prompted automatically to use traditional telephone scheduling. It integrated with all communication system to constitute a complex system, achieving convenient and interoperable voice communication function. The response alarm communication system is response alarm call center, it is a wireless scheduling system of all communication terminal, including digital simulation cluster communication equipment, cell phone, public telephone, and so on.

2. Emergency resources comprehensive search system, it can search quickly the data storing in storage device and provide services. It can search specific surveillance video and associated data of government in the Web interface by the ways of lens, time, association events, log and so on. It can provide unified management for the emergency resources of the governments, city leaders can use this system to make decision quickly according to on-site conditions and provide handling for emergency event by mobilizing related resources through the shortest path.

3. Accident warning and evaluation system, this system can provide regulations and system inquires according to the type of emergencies, it can define early warning indicators and assessment criteria of sudden emergencies, and formulated early-warning index and disasters.

4. Alarm and information release system, some information can be posted on the WEB by the system, the related departments can share the information; it can transmit the information to all kinds of medias in the first time, not only releasing information to the public, but also can get the aid of the public.

5. Emergency treatment knowledge base system, the operation data of this system is sensitive and classified information, the professional data must be stored by mutual isolation distribution, important data must be encryption and even physical isolation. It is to complete the collecting, disposal knowledge classification and comprehensive utilization of emergency events.

6. Emergency response processing tracking system, various linkage department personnel and on-site command staff must provide the latest information of dealing with emergencies to government command center report meeting room through the network, or the fixed telephone and mobile phone means. It could be accessible to get the information of all departments and disposal team, even personal working condition, on-site condition, instruction state through the tracking systems.

7. The linkage department evaluation system, this system stressed self-commitment and management, operation is very simple, it is more suitable for departments of not quantitative unit. It can complete the evaluation emergency response work of all departments and promote the linkage department to concern emergency response command system construction, improving the emergency event handling level of departments.

3 The Hierarchical Structure Design of Medium and Small City Emergency System

The paper presents the hierarchical structure design of medium and small city emergency system: the core layer, institutions layer, communication layer, directors, relevant layer. All levels have different data and applications. Its hierarchical structure has been shown in Fig. 2:
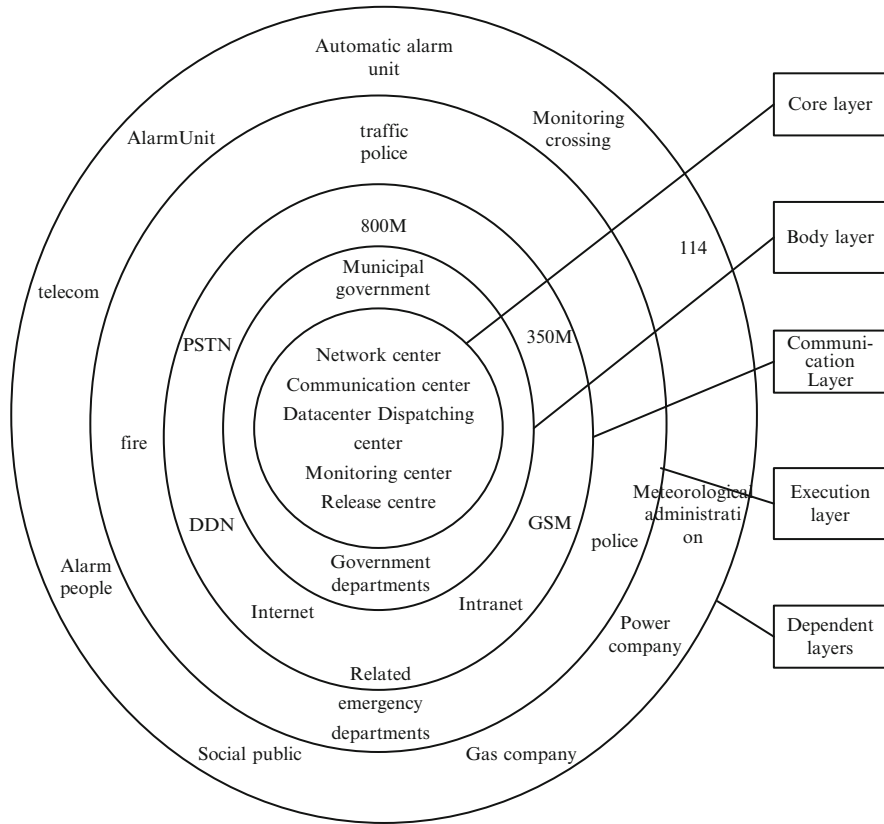
Fig. 2 Small and medium-sized city emergency system structure diagram

1. The core layer is the center of emergency system command center. It is consist of dispatching center, network center, communications center, the monitoring center, data center, release center. Command center can achieve the integrated application of database, comprehensive application of various departments data, real-time monitoring data and image display, the application of electronic map display and so on.
2. Institutions layer, it contains leadership and authority agencies. In the information construction, as a leading unit, it should arrange overall citywide, especially involving the communication system, computer network, GPS system, video surveillance system and information resources.
3. Communications layer, it contains all sorts of cable and wireless communications, Digital Data Network (DDN Data), Public Switched Telephone Network (PSTN), wireless trunking communications.
4. Director layer, the emergency team is an information receiver picker and scheduling command performer in the function. The emergency team is a network

sub-center node to distributed database in the technology. The emergency team is a data update operator in the management.

5. Related layer, it refers the alarm people, alarm unit, meteorology, telecom departments, electric power sector and gas sector and other related with emergency individuals or department.

4 The Business Process Design of Medium and Small City Emergency System

A typical emergency response business process contains six links parts, including call the police, response the police, response alarm, disposal, sources of withdrawal, summarizes. The process has been shown in Fig. 3: when the command center has
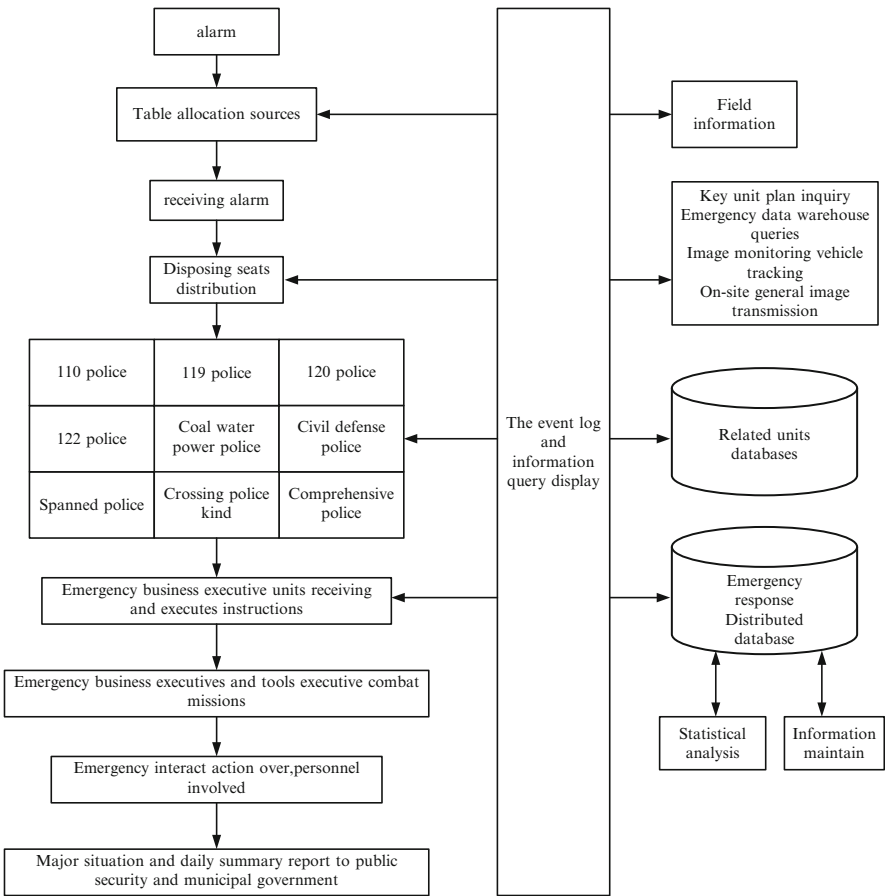
Fig. 3 City emergency system business process

an acknowledge of happening non-emergency event, it will turn to the convenient service center directly and distributing processing; When the convenient service center received emergency for help, it will also turn to the command center directly. In emergency/ non-emergency preparedness, the command center and convenient service center have sharing information. The decision-making command center is usually only one on duty, when encountering major emergency or non-emergency events need coordinate, the attendant report the information to the director of decision-making command center, the director will judge the events and report to the city leaders, then the leader will has a decision. When leaders have arrived to command center for decision-making, the system will provide information collection and leading comprehensive coordination communication (start voice, data and images) platform.

5   Conclusions

This paper proposes a new design scheme based on the analysis of the traditional medium and small cities emergency system construction, which is more suitable for medium and small cities. The system architecture, the hierarchical structure and business process have been designed in this paper, it provides a good research direction for the construction of emergency system for medium and small cities. This system has been applied in Jianli, WuXue medium and small cities, and so on. The system is stability, security.

References

Dong Z, Chun C (2011) The study of scheduling function optimization for trunking communication system in CERS. Sci Technol Eng 18(3):211–213

Jianhua J, Xianhuai Yi, Aiguo L (2005) Public health system present situation and countermeasure thinking. J Chin Diet Hyg Health (4): 1–4

Qi Z, LiQun L, GenMing Z et al (2006) Our country emergent public health event emergency situation of early warning system. Schistosomiasis Prev Mag 18(3): 211–213

Shoupeng D (2007) City emergency system management and development. Digit Commun World (4): 18–20

WeiJun E (2004) Security incident michmash, emergency treatment system analysis. Comput Eng (7): 56–58

Yue W (2006) Construction of the complex information for multiple views – active agents. China Sci 8(5):30–34

An Intelligent Recommendation System for TV Programme

Xiaowei Shi, Weijian Mi, Linping Huang, Yanhua Niu, Daofang Chang, and Yang Zhang

Abstract A personalized recommendation system for multiple users used in TV-Anytime context is presented. The system includes five agents: filtering agent, recommendation agent, profile updating agent, report agent, and interface agent. It also has a user preference database and a database for recommended program information and content. The "like" and "dislike" information of the users is included in the user preference database. The filtering and recommendation agents propose contents based on the ranking of similarity of user preferences and programme metadata. The user interface agent builds and updates the user profile based on explicit feedback, and collects information on user's reaction to the recommended contents and viewing behavior. This system has sensibility and adaptability to the status of itself and outside and can represent the users' potential needs based on implicit feedback and learn potential changes of their preferences, avoiding the limitation of recommendation based on only explicit needs. Experiment results show the recommendation system can recommend contents effectively.

Keywords Multi-agent Recommender • Potential needs • Preference learning • TV-Anytime • Filtering & ranking

1 Introduction

In recent years, the research on personalized recommendation technology on DTV has already reached a variety of achievements, such as Multi-Agent Recommender (Blanco-Fernández et al. 2004), TV-Advisor (Das and ter Horst 1998), and Personal TV (Cotter and Smyth 2000). The recommenders filter/rank Electronic Program

X. Shi (✉) • W. Mi • L. Huang • Y. Niu • D. Chang • Y. Zhang
Department of Logistics Engineering, Shanghai Maritime University, Shanghai, China
e-mail: shixiaowei1990@hotmail.com

Guide (EPG) according to the user' preference, recommend programs which may meet the demand of users. However, basic information of EPG only includes channel name, program title, the start time and the duration time of the program. In addition, several factors which may lead to hard to recommend programs that users will like by the recommendation system are listed as follows: (1) The TV program schedule may be changed at any time; (2) When users choose the program, they always are attracted by the actors, directors, program genre and so on. The TV-Anytime Forum seeks to develop specifications (TV-Anytime) since September in 1999. The specifications enable audio-visual and other services based on mass market high volume digital storage in consumer platforms (TV-Anytime System Description 2002). Metadata as defined by TV-Anytime not only includes detailed information about programs, but also about user preferences and usage history (TV-Anytime Metadata Specifications 2004). TV-Anytime identifies program with Content Reference Identifier (CRID) which is unique for each program, then apply location analysis to obtain program metadata. Through applying location analysis can make the process of obtaining program metadata free from the limitation of TV schedule. The specialization of metadata makes it possible to improve the accuracy and reliability of recommendation quality. A typical TV-Anytime system consists of three parts, that is, a proxy for delivering the information stream of TV-Anytime, program resource, and Personal Digital Recorder (PDR) which belongs to household appliance, and DTV (TV-Anytime System Requirements 2005).

One of important factors that may influence the recommendation quality is the reality and integrity of user preferences. Because of the constant change of user preferences with the time, intelligent recommendation system should learn user preferences dynamically, and update user profile on time. Several recommendation systems have already existed, such as PTV, TV-Advisor, and Multi-Agent Recommender. However, plenty of drawbacks existed in the[1] current recommendation systems. PTV fails to put forward the ideas revise user preference dynamically; TV-Advisor revises users' explicit preferences based on explicit feedback, while it is incomplete to acquire user preference due to impossible for users to express all of their needs by their own; Multi-Agent Recommender revises user preference by combination of explicit feedback and the statistics about the view rates of parameters such as program. However, it is impossible to learn the potential user preferences wholly only according to the basic information such as program.

A self-adaptive recommender based on TV-Anytime is proposed in this paper. The system, as an intermediary, relates users' preference with program resources. This system has sensibility and adaptability to the status of itself and outside. The paper is divided into three big parts in this: Firstly, Architecture of recommendation system is introduced. Then discussion on the detail information about each part of this system: (1) the feature-weight (V, W) user profile structure; (2) the filtering

and recommendation strategy; (3) a method for modeling a user profile dynamically and revising the model according to the explicit and implicit feedback. Eventually, experiment results prove the recommendation system's reliable.

2 Multi-Agent Recommendation System

2.1 System Architecture

The system is a prototype of personalized recommendation system for programs based on TV-Anytime. As shown in Fig. 1, it is comprised of five agents, that is, filtering agent, recommendation agent, profile updating agent, report agent, and interface agent. There exist two parts for content storage, that is, user preference database and a database for recommended program information and content. The function of the agents is presented as follows:

(1) **Filtering agent**. Selecting the program that users may be interested in by filtering the incoming program according to the user profile, then submit the information and contents to the recommendation agent and local storage. At the same time, filtering agents can filter repetition program to eliminate the redundancy information and to save the storage space. (2) **Recommendation agent**. The Recommendation agent can rank the recommended programs according to the like-degree and submit a recommendation list. (3) **Interface agent**. The user interface
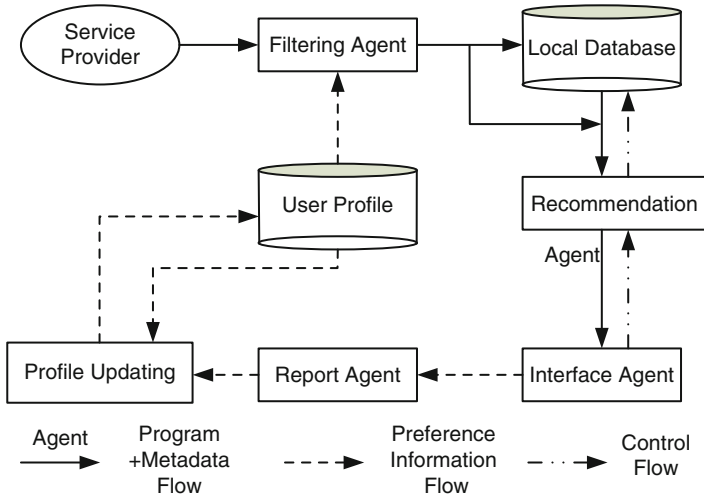
Fig. 1 System architecture

agent builds an interface of user profile according to explicit feedback; Users can provide feedback about the recommendation list and the contents of program through the interface. In addition, it can collect information on user's reaction such as recommended program metadata and program in content database and send the feedback to the report agent. (4) **Report agent**. The report agent provides the feedback to profile updating agent and need to be agreed by the user before sending. (5) **Profile updating agent**. The agent studies the potential change of implicit preference of users further. In addition, combination of implicit feedback with explicit feedback to update the preference profile.

2.2   *Inter-Agent Communication*

The agents in multi-agent system (MAS) rely on inter-agent communication to achieve something which they cannot achieve individually. In the current MAS research, inter-agent communication is based on Knowledge Query and Manipulation Language (KQML) which is a language that is designed to support interactions among intelligent software agents. KQML messages do not merely communicate sentences in some languages, but they rather communicate an attitude about the content (assertion, request). Furthermore, there are a variety of protocol oriented performatives among KQML to support interactions among intelligent software agents. For example, the simplest protocol, namely, client/server protocol, one agent as a client request to another, then the server agent will one-off return the result data (Wu and Sun 2010). In the system, several inter-agent communication pairs exist. For example, when interface agent recommends programs for user, it needs to communicate with the recommendation agent.

3   User Profile

3.1   *User Profile Structure*

User profile is regarded as a vector consists of features and weights. The meaning of feature is generalized, such as certain program genre, actors, etc. Weight is a metric that measures the important degree of each feature. The future is unique, such as the feature as an actor and the feature as a director of Brad Pitt are different at all, and the corresponding weights can be also different. The value of weight ranges from -1 to 1, which above zero represent that user like the feature, otherwise, it represents that user dislike the feature. The larger absolute value means more intense emotion. To simplify the calculation, n features serve as a set of typical features which represent user's emotion. The n features are used to model the user

profile, i.e. weight only includes the first n features in user profile. User profile can be described as the vector sets (V, W), V represents the feature set, which can be shown as follows:

$$V = (v_1, v_2, \ldots, v_n) \tag{1}$$

W represents the corresponding weight set shown as follows:

$$W = (w_1, w_2, \ldots, w_n) \tag{2}$$

3.2 Program Metadata

The description of program relies on the description of user profile, and program P can be represented as follows:

$$P(V) = (u(v_1), u(v_2), \ldots, u(v_n)) \tag{3}$$

Where if feature v is included in the fields of program metadata, then $u(v) = 1$, otherwise $u(v) = 0$.

3.3 The Presentation of User Preference Information and Program Metadata

For the purpose of interoperability, the program metadata and user profile in the system are both represented by The Extensible Markup Language (XML). The W3C's XML 1.0 Recommendation was first issued in 1998. It defines data structure in an opening, self-description pattern. In addition, when it describes the content of data, it can highlight the structure description. XML is widely applied in the field of data exchange due to its greatest strength, that is, the data storage format will not be constraint by its display style ("XML Path Language (XPath) 2.0" 2010).

4 Filtering Agent and Recommendation Agent

4.1 Filtering and Recommendation Procedure

Filtering agent filters incoming programs according to user profile, and recommends and records the program users may like. The function of filtering agent and recommendation agent is shown in Fig. 2. In Fig. 2, a program consists of program
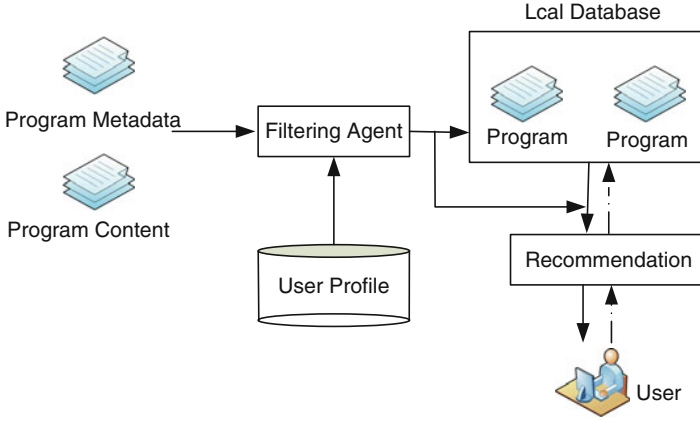
Fig. 2 Filtering and recommendation procedure

metadata and program content. When the program is incoming, the filtering agent calculates the similarity between program and user profile. If the degree of similarity is higher than the similarity degree between storage program and user profile, the filtering agent will execute step 1 in the Fig. 2, the filtering agent will ask recommendation agent to broadcast real-time program directly; if the similarity degree of calculated program is higher than threshold β, the filtering agent will execute step 2 in the Fig. 2, the program is considered that user may like and be stored into the computer memory. The program which meets the user's requirement will be sent to recommendation agent, then the recommendation agent recommend the fist n programs whose similarity degree is ranked from highest to lowest.

4.2 Similarity Measurement

In the classical vector space representation, the similarity between program metadata and user preference refers to their relationship. This paper measures the similarity by comparing the weight of corresponding feature. A commonly used similarity metric is the cosine of the angle between the two vectors. Given a program P and a profile (V, W), the cosine similarity $sim(V, W, P)$ can be calculated as follows:

$$sim(V, W, P) = \frac{(W, P(V))}{\|W\| \bullet \|P(V)\|} = \frac{\sum_{i=1}^{n} w_i \cdot u(v_i)}{\sqrt{\sum_{i=1}^{n} w_i^2 \cdot \sum_{i=1}^{n} u(v_i)^2}} \qquad (4)$$

Where if the similarity between program P and user profile (V, W), that is, $sim(V, W, P) \geq \beta$, then the filtering agent considers P as a program user

likes, otherwise, considers P as a program user don't like. Setting an appropriate value of β which is based on experience can exert influence on the reliability of recommendation.

5 Profile Updating Agent

Two methods are proposed to update user preference in this paper: (1) updating user preference according to usage history of users; (2) updating user preference according to explicit feedback. There is no need to introduce second method. The first method is proposed in this paper. When users select and view the recommended program, they will show the like-degree for the program naturally. Then the system will mine the potential change of implicit preference according to the feedback and update corresponding feature and weight. Suppose certain program P, program duration T, view time of the program t. Then each weight $w(v)$ of feature v in the program can be updated as follows:

$$w(v) := w(v) + \alpha \cdot \left(\frac{t}{T} - \delta \right) \tag{5}$$

Where if v have not shown in the user profile, then $w(v)$ is set as value of 0. α is the learning rate which presents the updating sensitivity of the user profile based on feedback. $\delta \in (0, 1)$ is a threshold and depends on experience, here 0.7 is its value. If user view the program exceed δT, it presents that user is fond of the program and the corresponding weight increase, otherwise, it indicates user don't interested in the program, the corresponding weight in user profile will be reduced. In addition, if user deletes the recommended program directly, the system sets the view time as 0. Eventually, the new user profile is constituted by the highest value of n weights and corresponding feature. Program metadata is modified correspondingly.

6 Report Agent

The user interaction agent provides a chance for revising the feedback. It collects feedback and asks user whether they agree to put the contents into the feedback report. If they oppose the idea, they can revise the report. Sometimes, user regards this as privacy and don't expect this kind of content show on the recommendation list preventing, then user can delete the information; sometimes, some unappealing programs have to be watched for accompanying guest. In this condition it cannot reflect the true preferences of users, and user can delete the information too. User's privacy can be protected while causal behaviors can be deleted in the system. Consequently, updating results of user profile can be more reliable.

7   Experiment

7.1   Experimental Environment

The experimental environment involves 1 PC, 1 TriMedia card, a Multi-agent system and 1 TV set. The PC is working as the host (that is PDR), responsible for depositing program sources. The TriMedia card is used for video decoding (Philips Homepage for Media Processors 2004). The TV set is used for displaying programs. This system runs under PC/Linux platform. The test program contains 150 different program segments, and each program segment lasts 2–10 min. Three users participate in the experiment. The three users is trained with 30 programs. Then they are respectively tested with 20 programs last 6 days. From the perspective of user's view, the effectiveness of filtering is very important. It reflects the degree of personalization. The parameter of Precision measures the filtering ability of relevant program. Precision = RPR/TPR. Where TPR indicates the total number of recorded program; RPR indicates the total number of program which meets the requirement of user.

7.2   Results

There are three users participate in the test. User 1 fails to provide the explicit feedback, his initial profile is empty; user 2 provides the explicit feedback to initial the profile, then his preference is changed suddenly and fail to update the user profile according to the explicit feedback; user 3 also provides the explicit feedback to initial the profile, only the normal potential preference is presented in the test. The experiment result is shown in Fig. 3.
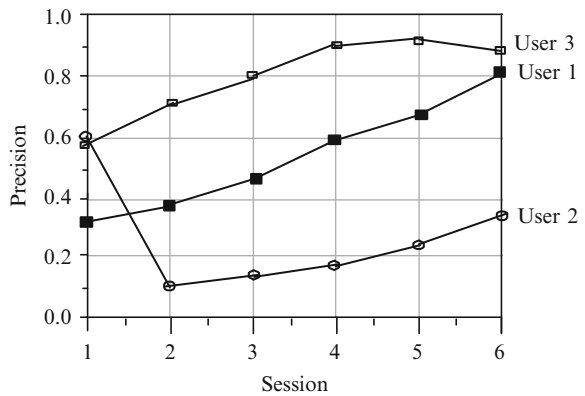
Fig. 3 Experiment result

7.3 Discussion

The experiment result reveals that learning algorithm of user profile can learn the potential demand of users and capture the potential changes of user preference, and recommend personalized programs for user. Compare the three users, user 1 and user 2 learn fast. For user 2, initial preference and the following provided preference are totally different, and the system needs to spend more time to learn new preferences and forgets old preferences. If user changes his preference suddenly and updates preferences on the interface on his own, then the recommendation precision will be better. Comparing user 1 and user 3, the recommendation precision of user 1 is lower than the other user due to fail to provide explicit preferences based on explicit feedback. The result of user 3 reveals the gradually change of user preferences. Therefore the recommendation precision won't always become higher and higher, even closed to 1, that is to say, the recommendation precision cannot reach 100%. The setting of threshold in experiment will influences the learning speed and recommendation precision, but the basic function is invaluable.

8 Conclusion

A personalized recommendation system for multiple users used in TV-Anytime context is presented. The system which includes five agents filters incoming programs according to user preference and program metadata, and recommends the programs which users may like to user. This system has sensibility and adaptability to the status of itself and outside and can represent the users' potential needs based on implicit feedback and learn potential changes of their preferences, avoiding the limitation of recommendation based on only explicit needs. Experiment results show the learning algorithm can learn user preference effectively. It needs further research on the reliability of improving program filtering ability and on how to recommend for groups.

References

Blanco-Fernández Y, Pazos-Arias JJ, Gil-Solla A, Ramos-Cavrer M, Barragáns-Martinez B, López-Nores (2004) A multi-agent open architecture for a TV recommender system: a case study using A Bayesian strategy. In: Proceedings of IEEE six international symposium on multimedia software engineering, MSE 2004, Miami, FL, USA, pp 178–185

Cotter P, Smyth B (2000) PTV: Personalized TV guides. In: Proceedings of the 12th conference on innovative applications of artificial intelligence, IAAI 2000, Austin, Texas, pp 287–295

Das D, ter Horst H (1998) Recommender systems for TV. In: Recommender systems, papers from the 1998 workshop, Technical report WS-98-08. AAAI Press, Madison/Menlo Park, pp 35–36

Philips Homepage for Media Processors (2004) http://www.semiconductors.philips.com/trimedia/

TV-Anytime Metadata Specifications Document (2004) SP003v12 Part A Appendix B, TV-Anytime Forum, August 2004

TV-Anytime System Description Document (2002) (Informative with mandatory appendix B), WD608, TV-Anytime Forum, August 2002

TV-Anytime System Requirements Document (2005) TV036r2, TV-Anytime Forum, pp 10–15, June 2005

Wu X, Sun J (2010) Study on a KQML-based intelligent multi-agent system. In: International conference on Intelligent Computation Technology and Automation, ICICTA 2010, Shengzhen Guangdong, China, pp 466–469

"XML Path Language (XPath) 2.0" (2010) http://www.w3.org/TR/xpath20/, Dec 2010

The Research of Digital Resource Technique in Libraries

Zong-ying Shu

Abstract Digital Library is an extension of traditional library functions. It is information collection, conversion, to describe, and the computer can handle in the form of digital information collection and storage of digital information network to intelligent information retrieval methods and unified search interface to use advanced information processing technology. The Internet, provide a variety of languages compatible with the long-range multi-media digital information services. This paper provides comprehensive coverage of the theories research connected with the provision and management of electronic resources in libraries to show us with comprehensive coverage of the core topics related to electronic resource methods. With the rapid development of computer technology, communication technology and network technology, the construction of information expressway have offered the environment and condition for development of extensive information system, library system. The information resources in libraries are getting more and more rich, the digitized tendency is getting more and more obvious. Some technique such as the information resource integration technology, electronic resource management and cataloging standard technique have enhanced day by day. In order to better serve for the reader and cause readers to refer to the more literature resources in the shortest time, the researchers should pay more attention to library information resource technique.

Keywords Component: library • Electronic resources • ILS • Cataloging standards • Computer

Z. Shu (✉)
Library, Chongqing University of Education, Chongqing, China
e-mail: szy17@163.com

1   Introduction

These developments, coupled with the new expectations of the Internet-savvy user, affected all types of libraries who had to rapidly shift from print-based to electronic resources. Whether the electronic resource comes from a commercial publisher or a local digitization effort, this trend is also rapidly changing library operational and organizational practices. Along with the increase in electronic resource acquisitions, librarians have had to quickly adapt and address an ever complex set of new challenges and changes related to: workflow management and planning; selection and acquisition procedures; copyright and license negotiation; cataloging practices; public access interfaces; and utilization of usage statistics. Libraries must now come to terms with how to better evaluate, acquire, store, and manage this wealth of electronic resources (Anderson 2003; Lee and Boyle 2004). The proliferation of electronic resource management systems (ERMS) presents an additional problem for libraries, that must now develop in-house resource management solutions or acquire one of a myriad of emerging turn-key solutions and implement them in an evolving organizational setting.

Broadly, activities and tools used by a library to manage their investment in electronic products. More finely, electronic resources management refers to several specific management areas (acquisitions, access, workflow, trial, statistics, costs, etc.) that have been defined by the Digital Library Federation.

2   Integrated Library Systems (ILS)

During the past decade, there has been phenomenal growth in the number of electronic resources including electronic journal packages and full text aggregations acquired by libraries. University Libraries projected will become more digital. Though this prediction has yet to come to pass, the Association of Research Libraries (ARL) expenditure trend data (Association of Research Libraries, 2002) showed that academic libraries are "in the midst of a profound shift toward reliance on electronic resources, and this reliance seems to have deepened just within the last year or two as libraries have shed paper journal subscriptions to help pay for online access." Since providing access to electronic resources have become such a major part of the library services, it was crucial for libraries to tackle these new challenges head on.

Libraries provide their users with different ways of accessing its electronic information resources, such as the ILS, a library developed Web site that include listings of available e-resources and a metasearch system (sometimes referred to as federated search). Library users may start an e-resource research session from any of these systems. Many libraries also provide access through other nonlibrary systems, such as the Learning Management System (LMS), an enterprise portal, and so forth. Many of these systems perform authentication. Depending on how the user starts the

research, the user may have different experiences in terms of authentication. These access points should briefly discussed here.

Usually, when a patron tries to access a resource that is part of the library's local collection and restricted, such as an item in an electronic reserves collection, the ILS itself will authenticate the patron. In other words, the ILS will use its own internal authentication mechanism to authenticate the user.

However, if the resource resides outside the ILS, such as an electronic journal, the ILS will simply redirect the user to the resource itself, or to another authenticator, such as a proxy server, which in turn authenticates the user. In this case, the ILS delegates the authentication to another authenticator (White 2006). Some ILS vendors provide a proxy server as an add-on module that is integrated to the ILS (e.g., Web Access Management from Innovative Interfaces, Inc.). These modules authenticate patrons using the ILS's internal patron database.

3 Electronic Resources of Library

Libraries are standard-bound institutions, applying rigorous rules to cataloging, classification, coding, indexing, and authority work. The international descriptive cataloging standard, the Anglo-American Cataloging Rules, 2nd edition, (AACR2) (2002), for example, includes rules governing naming library print, audiovisual, and electronic media in the title statement. Even when no title exists, there are rules for compiling one. But standards are never comprehensive, nor can they be. There are always exceptions that do not fit the rule. AACR2 uses language of sufficient generality and vagueness to leave catalogers with considerable interpretive or subjective latitude.

This flexibility works fairly well for print and microfilm, since these "old" media are relatively stable in their presentation, making exceptions finite. This flexibility becomes a liability in the case of electronic resources, where presentation is far from stable and where name changes occur frequently. The standard makers cannot keep up with the evolving nature of "new" media in libraries. As a result, a small cottage industry of additional guidelines and interpretations flourish between standards editions. The continually updated, Library of Congress Rule Interpretations(LCRI) (Office of Descriptive Cataloging Policy, Library of Congress, 1989), for example, amplifies and explains existing rules and adds new ones for areas not covered in AACR2. CONSER (Cooperative Online Serials) is an international online serials cataloging program run by the Library of Congress. Module 31 of the CONSER Cataloging Manual (2006) is a supplemental standard specifically geared to deal with the complexities of cataloging online serials. Module 31 only applies to individual electronic journals and newsletters, that is, to serials in the narrows sense, and not to databases. While such guides assist catalogers in interpreting standards, they also sanction subjective interpretation and make exceptional applications quasi-standards.

When it comes to the ever-changing field of electronic resources with their multiple sources for description information, the cataloger is often at sea, and must extrapolate from standards applicable to print media or invent totally new approaches for emerging realities. Since its first appearance, AACR2 has gone through a series of revisions and updates. Yet it still lags behind the real world of electronic resources. A so-called third edition, AACR3, is now under development.

Library systems in general become corrupted over time. Standards are applied differently from institution to institution, from one individual to another, and over the life of a system. As new standards and local administrative policies are implemented and displace old standards and policies, catalogs demonstrate a hodgepodge of different standards and policy applications (Kichuk 2000). This is true even in the relatively short time that electronic resources have been widely accessible in libraries A spot check of library catalogs reveals that in many, electronic resource titles receive different treatment, for example, for capitalization, that is only explainable as a change in policy or interpretation. The inconsistency remains embedded in the catalog record unchanged perhaps for the life of the catalog.

Cataloging standards and guidelines can not be systematically applied in the creation of metadata for electronic resource A to Z lists or ERM records, where their weaknesses in relation to electronic resources are even more apparent than in the catalog.

4 Library Technical Services Management

This explosion of electronic resources has created a new workload requirement: management. Management of electronic resources is a time-consuming and, at times, a difficult process. Although the management of electronic resources is often seen as a strictly technical services endeavor, it should been considered a multi-faceted process requiring all areas of the library. The workflow currently in place at JSU is just such a practice. The management of electronic resource workflow is a collaborative effort between public service and technical service librarians. At JSU, the public services librarians collaborate with the technical services librarians in order to provide a substantial collection of electronic resources and to make this collection available to the users as quickly as possible. It is this collaboration that makes the workflow inimitable. The collaborative effort begins with the review process, both trial reviews of databases, e-journals, and so forth and reviews of freely accessible materials such as Websites, extends through the acquisitions, administration, cataloging/bibliographic control, marketing, training, and ends with the assessment of the resources. No matter what workflow a library uses for electronic resources, good communication at each stage in the process is paramount since there is no physical trail of where the item is in the workflow until it gets to the bibliographic control stage.

Fig. 1 Library collaboration
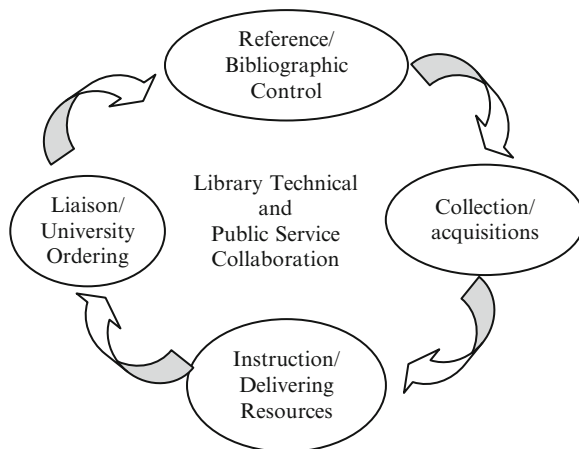
Despite the physical division between "technical" and "public" services in the library, managing electronic resources is one of many shared responsibilities. The mission of the HCL is to provide information services and bibliographic resources to support the scholarly and informational needs of the university community. To meet user expectations better, the public service librarians' responsibilities at JSU have evolved to include four major professional areas (Collantes 1995). All areas are both interchangeable with and interconnected to one another. For example, effective reference is impossible without a sound collection and vice versa. The four areas are: (1) reference, (2) collection management and development of their respective subject areas, (3) instruction, and (4) liaison activities. Liaison activities include a structured liaison partnership between university teaching faculty and library faculty which includes appointment letters, a list serve, and formal communications between librarians and departmental liaisons. An informal relationship between librarians and teaching faculty is also encouraged and expected as part of their liaison activities job responsibilities.

Additionally, these four professional areas for the public service librarians are dependent on and mirror the four technical services professional responsibilities to: (1) Provide the correct bibliographic data and access points; (2) order, pay, and process the materials requested; (3) deliver or notify the public services librarians of new information sources; and (4) process nonlibrary faculty request and inform them of new information sources. (Fig. 1)

5 Conclusion

Electronic resources may take many forms, from e-books or journals to full-text resources from aggregators, or index/abstract databases from publishers. The way in which electronic resources are managed is becoming more distinct from print with

new approaches to planning, tasks, workflow and communication. The planning process encompasses policy-making, budgeting, and staffing. Tasks may include things like setting up trials, license negotiation, authentication, troubleshooting, evaluation, and renewal. Workflow covers the entire process from initial product consideration, making the resource available to patrons, to renewal or cancellation. Communication includes a variety of interactions from local administrators to vendors, IT staff, public service personnel, and users. So it is important to research the topics related to electronic resource management of library, and we should pay more attention in order to promote library digital level.

References

Anderson R (2003) Buying and contracting for resources and services: a how-to-do-it manual for librarians. Neal-Schuman Publishers, New York

Collantes L (1995) Degree of agreement in naming objects and concepts for information retrieval. J Am Soc Inf Sci 46(2):116–132

Kichuk D (2000) Library space study. Unpublished internal report, University of Saskatchewan Library, Saskatoon, pp 148–170

Lee SD, Boyle F (2004) Building an electronic resource collection: a practical guide. Facet Publishing, London

White AC (2006) E-metrics for library and information professionals: how to use data for managing and evaluating electronic resource. Neal-Schuman Publishers, New York

Study on Model of E-Government Public Satisfaction Based on Service-Oriented

Sun Hao

Abstract In order to solve the financial crisis, confidence crisis and management crisis, government has introduced the concepts of market-driven and customer-orientation into management on the guidance of new public management. From the perspective of service-oriented, the contents of public satisfaction and its improvement are the important approach and carrier of reengineering government, providing efficient and convenient services to meet public needs, and building an open, transparent, efficient, information-sharing government. Through reviewing the literature, this article presents an integrated e-government customer satisfaction model, which is a multivariate statistical analysis approach and used to discuss the relations between latent variable and manifest variables. Latent variables includes public expectations, perceived quality, perceived ease of use, perceived usefulness, perceived value, public trust, public satisfaction, government image and public participation, which compose the indicator of e-government customer satisfaction index system, and guide government to promote management efficiency, effectiveness, fairness, participation of e-service.

Keywords Service-oriented • Publicsatisfaction • E-government • Model

1 Introduction

After 70 years of the twentieth century, governments in western countries face financial crisis, confidence crisis and management crisis. Public administration established in the bureaucracy has not effectively solved the real world problems, and can not satisfy the public's request of highly efficient operation and

S. Hao (✉)
School of Management, Hubei University of Technology, Wuhan 430068,
People's Republic of China
e-mail: sunhao15@163.com

responsibility. In this context, the experience, technology and method of the success-ful operation of large number of enterprises have been introduced into the reform of public administration, public administration mode changes from the traditional domination to service. Government management has been more and more pointed to the public interest, adopted democratic and participatory approach, resolved social problems through cooperation with the public and various organizations in the network structure, and taken democratic and scientific decision-making mechanism based on needs of the public and the efficiency.

The development of e-government resulted from the transformation of govern-ment in the world and the strong demand for government reform. The construction of e-government as a pilot project of social information, gets more and more government attention and support. In the guidance of service-oriented, the con-tents of public satisfaction and its improvement are the important approach and carrier of reengineering government, enforcing optimization and reorganization of government organizational structure and workflow; providing efficient and conve-nient services to meet public needs, and building an open, transparent, efficient, information sharing government. Therefore, according to the construction reality of Chinese e-government, it has an important and practical significance to build e-government user satisfaction index system and evaluate quality and level of government services for promoting e-government construction and accelerating transformation to service-oriented government.

This paper proposes an e-government customer satisfaction model based on ACSI Government Model and other domestic scholars opinions, which includes variables of public expectations, perceived quality, perceived ease of use, perceived usefulness, perceived value, public trust, public satisfaction, government image and public participation. These nine latent variables compose the first-level indicator of customer satisfaction index system. Also this paper lists different observation points of each latent variable in details, which compose the manifest variables. These variables constitute the elements of e-government user satisfaction, which can guide government to promote management performance.

2 Reviewing of American Customer Satisfaction Index Model

2.1 American Customer Satisfaction Index (ACSI) Model

The American Customer Satisfaction Index (ACSI) model is a cause-and-effect model with indices for drivers of satisfaction on the left side (customer expectations, perceived quality, and perceived value), satisfaction (ACSI) in the center, and outcomes of satisfaction on the right side (customer complaints and customer loyalty, including customer retention and price tolerance) (Fig. 1).
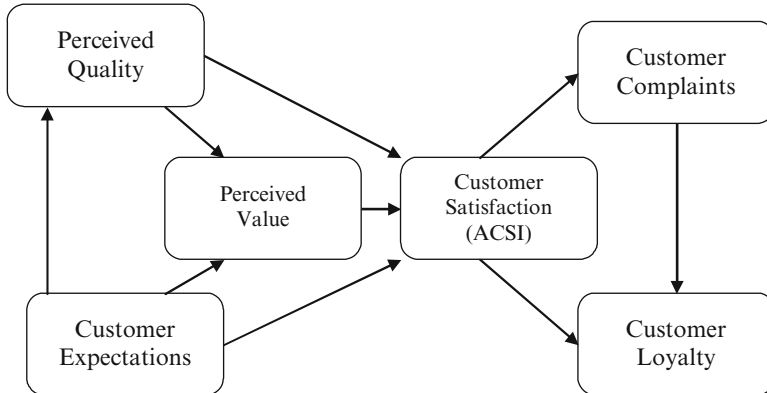
Fig. 1 American customer satisfaction index model

Customer expectation is a measure of the customer's anticipation of the quality of a company's products or services.

Perceived quality is a measure of the customer's evaluation via recent consumption experience of the quality of a company's products or services. Quality is measured in terms of both customization, which is the degree to which a product or service meets the customer's individual needs, and reliability, which is the frequency with which things go wrong with the product or service.

Perceived value is a measure of quality relative to price paid. Although price (value for money) is often very important to the customer's first purchase, it usually has a somewhat smaller impact on satisfaction for repeat purchases.

Customer complaints are measured as a percentage of respondents who indicate they have complained to a company directly about a product or service within a specified time frame. Satisfaction has a negative relationship with customer complaints, as the more satisfied the customers, the less likely they are to complain.

Customer loyalty is a combination of the customer's professed likelihood to repurchase from the same supplier in the future, and the likelihood to purchase a company's products or services at various price points. Customer loyalty is the critical component of the model as it stands as a proxy for profitability.

2.2 ACSI Government Model in Public Sector

In September 1993, President Clinton signed the No. 12862 Executive Order "establishing customer service standard", and ordered federal agencies to develop customer service standards, requiring government departments to provide more choices for customers of public services, resources and means of service provision options. The establishment of customer-first concept in the new public management movement prompts the America customer satisfaction index (ACSI) to be applied in U.S. public sector, and builds ACSI Government model (Fig. 2).
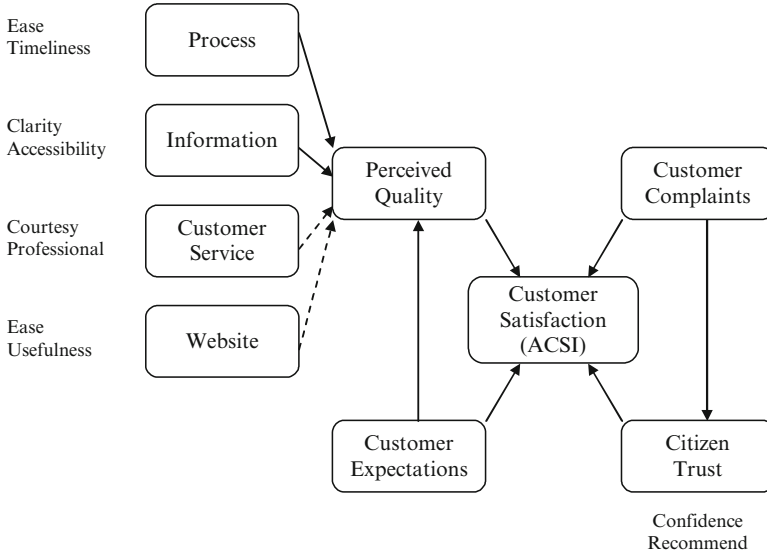
Fig. 2 ACSI government model

The ACSI government model (used for most agencies) is identical to the private-sector model with a key modification in the "outcomes" component. Specifically, the outcome of citizen trust replaces the price-related outcomes found in the private-sector model (repurchase intention and price tolerance).

Because agencies provide services with little or no direct charge to the user, perceived value is not part of the government model. The drivers of expectations and perceived quality, satisfaction, and the outcome of customer complaints appear in both the government and private-sector models.

The results of user satisfaction in public sector are slightly different from private sector. In private sector maintaining customer loyalty and reducing customer complaints are important goals to maintain profits. Satisfaction evaluation and improvement is an important carrier and the key point to improve corporate profits. For most government organizations, the value orientation of public interest determines the main objectives of public sector services is to obtain the trust of customer, to resolve the public confidence crisis, to improve government credibility thus ensuring smooth implementation of all activities and policy.

At the same time, e-government is an electronic online service, which content and form are different from product provided in private sector. First of all, according to the services and product type provided, e-government websites are divided into news and information sites, portals and department main sites, E-commerce and transactional sites, career and recruitment sites. Secondly, private sector, mainly through variables of customization and reliability evaluate perceived quality. While for government organizations, indicators of satisfaction are grouped into four broad categories that are used as input to measure quality (process, information, customer

service, and website). For the outcome of citizen trust, indicators are the degree to which the user customer would recommend the agency's services to others, and the extent to which the user has confidence in relying on the agency in the future. Variable of citizen trust is defined as public confidence and support on online services, which reflects citizens reuse e-government.

The model regards citizen trust as a cause variable of public satisfaction, not as the ultimate outcome variable. This change is because that satisfaction as a psychological measure variable is reflected by external action of the public, which embodies citizens support and reuse e-government, that is, the frequency and effectiveness of external use can reflect the inherent subjective satisfaction.

3   The Model of E-Government Customer Satisfaction—From Perspective of Service-Oriented

3.1   The Proposal of Service-Oriented E-Government

New public management highlights the function change, and thinks that government should be oriented to customer to meet public different needs, and enhance responsibility to the public needs.

Under the guidance of this concept, the construction and implementation of e-government not only improves the quality and efficiency of government services, promotes the change of government management functions, but also fully embodies the basic concept of constructing service-oriented government, which is the important way to build service-oriented government. E-government should pursue from sector interests to the public interests, from attention to organization efficiency to realization of public value and service satisfaction.

Highlighting service function and deepening service concept will placed the public in the core of government management and operation, in which consumer sovereignty of the public and users are strengthened, the public is regarded as the object of access to government services-customer, and public satisfaction as a dominant indicator is used to evaluate e-government services and capabilities.

3.2   The Features of Service-Oriented E-Government

Through the application of network technology, communication technology, E-government achieves electronic, across time and space public service, the public can access to public services anytime and anywhere; network structure integrates originally distributed independent activities to provide "one-step" service for public; at the same time, the powerful information processing capability of background guarantee to offer personalized service to public.

The implementation of e-government is not simply the traditional realization of electronic public services, service-oriented e-government reflects a new cooperation relationship between government and citizens, which realizes the transformation from the traditional domination to modern service. Government no longer through the administrative regulation and approval manages the economy and society; no longer plays the role of producer, supervisors and controller. It provides a good development environment and equal competition conditions for public, provides social safety and public services.

Service-oriented e-government not only pursues lower administrative costs, improvement of service quality, cleanness and efficiency, but also pursues open and transparent goal. There are several reasons explain the importance of Openness and transparency of E-government. Firstly, it makes government information in a timely manner to be known by citizens through a variety of media, which is convenient for citizens to receive services provided by the government; Secondly, it provides a way for citizens participating in administration and policy-making, which is convenient for citizens to participate in various activities related to their interests and express their own aspirations and demands; thirdly, it reduces information asymmetry between government and citizens, reduce corruption and probability of moral hazard and implements effective supervision; Fourthly, it is convenient to assess government service functions and its behavior, to promote administrative efficiency and service capability; Fifthly, it realizes the transformation of power operation from means of single, top-to-down to means of diverse and interactive, which effectively embodies the spirit of public service.

3.3 Customer Satisfaction Model of Service-Oriented E-Government

Influence by western new public management movement, the concepts of customer-oriented, results-oriented, market provision of public services and the public participation in government performance evaluation have gradually applied to the government reform practice. Based on the ACSI model and domestic scholars' opinion, there is an E-government public satisfaction model (Fig. 3).

3.3.1 Latent Variable-Perceived Value

Generally, perceived value refers to the ratio between total benefit for customer and price paid by customers. Since most of e-government service provision is free, it is difficult to measure perceived value. A lot of satisfaction models including ACSI model do not use this latent variable.

This model increases variable of perceived value and expands its content, which reflects not only the ratio relationship between benefit and payment, but also can
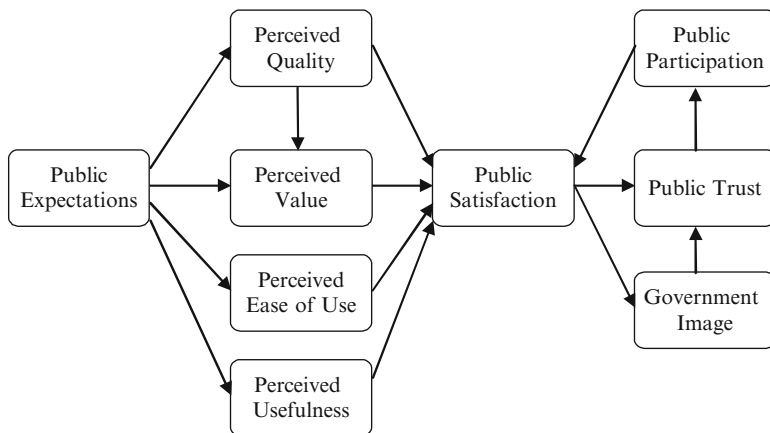
Fig. 3 E-government public satisfaction model

be used to express the perceived effectiveness of e-service. This perception covers not only the overall perception of service quality, but also reflects the degree of coinciding between public expectations and requirements concerned on results of e-services. The extent of achieving policy objectives and user expectations, changing original state and behavior can obtain good consume experience and feelings, also raise e-service satisfaction. As for the free services which can not measure the cost and payment, we can calculate in an indirect way, because the users save time, energy when enjoying high-quality e-services. According to the time value method it can be converted into the corresponding payment of cost and price.

3.3.2 Latent Variable-Perceived Ease of Use and Perceived Usefulness

E-government depends on the Internet to provide the virtual electronic information services for public. The enjoyment and service experience bases on actual usage. Before enjoying the service, there is a process of e-learning and acceptance in which public use network technology, hardware and software. Therefore, based on the theory of technology acceptance model, these two latent variables are introduced in integrated model.

Technology acceptance model (TAM) is the usage of rational behavior theory of Davis to study the degree of accepting new technology and new information systems for users (Davis et al. 1986). It purposes two major determinants:

- Perceived usefulness reflects the extent that a person uses a specific system to increase his job performance.
- Perceived ease of use reflects the extent that a person is easy to use a specific system.

TAM regards that system usage is determined by behavioral intention (BI). Behavioral intention (BI = A + U) is determined by attitude toward using (A) and perceived usefulness (U). Attitude toward using (A = U + EOU) determined by perceived usefulness (U) and perceived ease of use (EOU). Perceived usefulness (U = EOU External Variables) is determined by external variables and perceived ease of use. Perceived ease of use (EOU = External Variables) is determined by external variables. External variables include system design characteristics, user characteristics (including the perception forms and other personality characteristics), task characteristics, the nature of the development or implementation, policy effects and organizational structure, etc., which establishes a link among existing factors in the TAM such as internal beliefs, attitudes, intention, differences among individuals, environmental constraints and controllable interference.

According to TAM, user choosing and actual using of online e-services are determined by behavioral attitude and intentions which are formed by perceived usefulness and perceived ease of use. Once the public has obtained original expectations and gained satisfaction, this use behavior will repeat, the frequency of behavior can be used to measure and improve public satisfaction.

3.3.3 Latent Variable-Government Image

In general, the government image is the comprehensive impression and evaluation of above relevant factors in the public mind through experience of e-services, which is an overall performance of the public satisfaction and recognizing for government. Characteristic of virtualization in e-service determines that the image is more of a subjective experience and evaluation. A good government image can gain public trust, strengthen public confidence, promote public to support and participate in policy activities, and secure implementation of the government service functions efficiently.

Different from ACSI Government Model, this model removes latent variable of public complained, increases latent variable of government image. There are several reasons:

- According to Albert O. Hirschman's exit-voice theory, there is a basic distinction between alternative ways of reacting to deterioration in business firms and, in general, to dissatisfaction with organizations: one-exit-is for the member to quit the organization or for the customer to switch to the competing product, and the other-voice-is for members or customers to agitate and exert influence for change "from within.". Clearly, the complaint as a result of low degree of satisfaction is inferior to two mechanisms of exit and voice concerned on promotion of the recovery and improvement of organizational performance (Hirschman et al. 1970).
- Hirschman believes that it should perfectly combine the market force-exit with the non-market and democratic force-voice during the economic development and organizational performance promotion. User complaints are neither exit nor

voice, whose influence on government only stays in the subjective level, and is lack of substantive action to improve government performance, and to improve public satisfaction.

- According to exit-voice theory, in a monopoly market structure, because organizations do not have competitors, the exit mechanism has nothing threat and voice mechanism is a formality, losing its basis to take effect. Thus, from the subjective level, the public complaint is only regarded as a kind of catharsis in a monopoly e-government environment, which destroys the government image, and confirms that government image as a result of satisfaction index variables is reasonable.

3.3.4 Latent Variable-Public Participation

The model takes public participation as a new latent variable, the reason is that:

- The traditional process of providing public service is monopolized by government, in which the low efficiency and the poor quality of service are difficult to meet the citizens' diverse needs. In the context of new public management movement, people strongly urge to directly involve in the provision of public services, and establish a joint governance model between the state and society to make up with the shortage of government bureaucracy abilities. Thus, the effective provision of public services must introduce public participation and emphasize the cooperation of citizens and government.
- In the service environment, e-government provides services based on Internet. The development trend of network operating environment stresses on interaction and communication. For different user needs, it should customize services and different combinations to realize personalized consumption. E-services is no exception, it also requires government to pay attention to different users' needs, through public participation to realize differentiation service and personalized consumption.
- From the angle of information technology development, the network links government, ordinary citizen, businesses and social organizations a network, in which the mutual communication levels decrease, the cost becomes low, the efficiency becomes higher, the public and various organizations have more opportunities and more effective technical means to participate in the management of public affairs to break traditional bureaucratic monopoly provision situation and enhance the effectiveness and responsibility of services.
- Public participation as a kind of behavior can increase public satisfaction. On the one hand, the process of public participation in public service inputs their own needs and goals, once their proposals and measures adopted, they will get a psychological satisfaction. If such proposals realize need goals, it further strengthens this kind of satisfaction. On the other hand, public participation process is a process of interaction and learning. Through continuous network communication, the public get needed information and information acquisition method, which can increase the accumulation of knowledge and experience

Table 1 E-government public satisfaction index system

Latent variables	Manifest variable
Public expectations	The public's overall expectations of website and its services
	The expectations of the extent of Meeting the public demands of website information and e-service quality
	Expectations of website Information availability and e-services reliability
	Expectations of meeting personalized needs of website
Perceived quality	Reliability of information
	Usefulness of information
	security of Website
	Professional of services
	Interactivity of services
	Attitude of services
Perceived ease of use	Degree of easily understanding website structure
	Degree of easily learning Website operation
	Degree of easily Accessing Website information
Perceived usefulness	Degree to meet personalized needs of website services
	Degree to improve website efficiency
Perceived value	Efficiency of public obtaining e-services
	Effectiveness of public obtaining e-services
Public trust	Support of the public to e-services
	Recommendation of the public to e-services
Public satisfaction	The overall satisfaction of the public obtaining e-services
	Degree of satisfaction Compared with the expected quality
	Degree of satisfaction Compared with the desired quality
Government image	Degree of government liability
	Degree of Government responsibility
	The fairness of the Government
	The open and transparency of the Government
	Management efficiency of Government
Public Participation	Public participate in public administrative activities through e-government
	Public participate in general activities through e-government

to promote the public satisfaction. In addition, in the wave of democratic participation, public change the status of passively accepting public services and in-effectively supervising and restricting the government behaviors, who in a gesture of ownership actively involves in the provision process of public services, and enjoys the satisfaction with supporting and improving democracy.

Through analyzing these latent variable, E-government public satisfaction index system from perspective of service-oriented as follows (Table 1).

4 Conclusions

E-government emphasizes the service-oriented for customers and citizens, so the performance management and evaluation regard user satisfaction as a core indicator from perspective of consumer, which fully reflects benefit, availability, effectiveness, fairness, participation, coverage and quality of the e-services to citizens.

The model introduces variables of perceived ease to use and perceived usefulness, preserves variable of perceived value, and takes variable of government image as a new mental experience and activity for public obtaining services. Taking into account the interactive nature of the network, the model adds new variable of public participation, and regards it as the outcome of public psychological satisfaction. Public expectations, perceived quality and public confidence and other indicators are reserved and its definition is consistent with the opinion of domestic and foreign scholars. Especially, public participation not only is a latent variable of e-government user satisfaction, but also is the behavior result of psychological satisfaction, through which people involve in the public affairs and management process. More and more public determinations and goals embody in the policy-decision and provision of e-service, in which government is not the monopoly service provider, open and transparency is the main pursuing target.

References

Davis FDA (1986) Technology acceptance model for empirically testing new end-user information system: theory and results. Sloan school of management, Massachusetts Institute of Technology

Gong Sha-sha (2009) Study on establishment and evaluation of e-government public satisfaction model, vol 2008. University of Electronic Science and Technology (in Chinese)

Hirschman AO (1970) Exit, voice, and loyalty: responses to decline in firms, organizations, and states. Harvard University Press, Cambridge, MA

Jiao Wei-ling (2007) To establish the evaluating model of e-government citizen satisfaction measurement index of China. Magazine of Information, October 2007.10

Liu yan (2006) Evaluation theory, approach and application of e-government public satisfaction. Ph.D. thesis, University of Defense and Technology (in Chinese)

Liu yan, Chen Ying-wu (2006) The empirical study on e-government customer satisfaction index model. Syst Eng (May) (in Chinese)

Numerical Simulation for EMU Air-Conditioner Vent System and the Flow Inside the Carriage

D.P. Wang, F. Liu, and G.P. Xu

Abstract By using standard k-ε turbulence model and wall-function method, the new design scheme of air-conditioning system of the high-speed EMU was simulated numerically, and the calculation result was evaluated. The result showed that the velocity distribution of the air flow in the carriage was uniform, but the difference between the maximum and minimum temperature around passengers exceeded 3°C. In order to reduce the difference, the position of the guide plate in the main air channel was adjusted and the result of the temperature field was improved that the temperature difference was less than 1°C. The research result would be a helpful reference to the optimal design of air-conditioning system and the comfort of environment in the carriage.

Keywords CFD numerical simulation • Air-conditioning and ventilation system • High-speed EMU

1 Introduction

With the velocity of the high-speed EMU has approached 380 km/h in China, the domestic innovative technologies and the design have been grown up gradually after digesting and absorbing advanced ideas from developed countries (Zheng et al. 2010). In the specific process, it is a systematic project on how to have the capacity of independent innovation through digestion and absorption, and the air-conditioning ventilation system, as one of the subsystems, requires a lot

---

D.P. Wang, professor, engages in research of train aerodynamics.

D.P. Wang (✉) • F. Liu • G.P. Xu
School of Mechanical Engineering, Dalian Jiaotong University, 116028 Dalian, China
e-mail: lf198187@163.com

of meticulous and scientific research work. More requests for the surrounding environment and interior comfort of the CRH series EMU are brought forward, correspondingly, it is very important to study the air-conditioning ventilation system (Zhao et al. 2011).

It is difficult and costly to study the airflow organization inside of the train because of many factors including solar radiation, rejection of heat of passengers and complex structure inside the carriage (Zhang et al. 2006). However, when we use the method of CFD numerical simulation to analyze air-conditioning ventilation system, massive manpower, physical resource and financial resource can be saved, and the efficiency of product development can be increased. It has very high superiority compared with the traditional method of only depending experiment. By using the method of CFD numerical simulation, the investigation gave a numerical simulation on the air-conditioning system, analyzed the effect of air-conditioning, aiming at the new design scheme of air channel and the layout of passenger seats, and ameliorated the air channel according to the analysis result.

2   Model and Computational Method

2.1   Calculation Model

It firstly requires a physical model not only according with the physical property of practical flow field, but also adapting to the air dynamical flow field computation. According to the layout and the correlative dimension in new design system of the high-speed EMU, the 3-D model is set up by the Solidworks software after being predigested rationally. The length of the carriage is 24.7 m, the width is 3.0 m, and the height is 2.5 m. The specific model shows in Fig. 1, including exterior wall of the carriage, windows, passengers, the main air channel, the air channel for heating (located at both sides of the carriage), the air channel for emission, Air return duct, wiring cabinets, (there are 2 on forward and back), and the toilet.
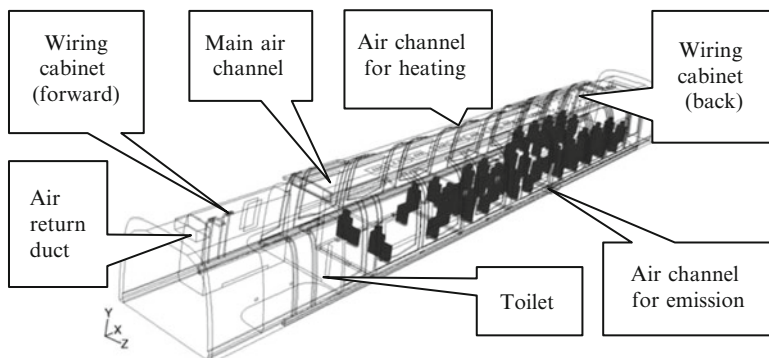
Fig. 1 The calculation model of the air-conditioning system
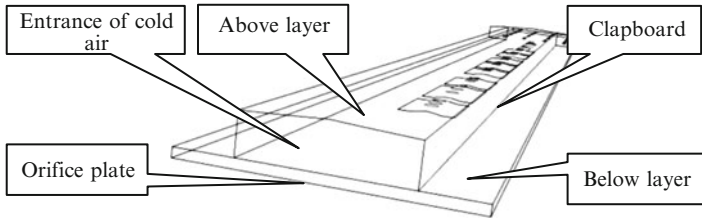
Fig. 2 The calculation model of the main air channel
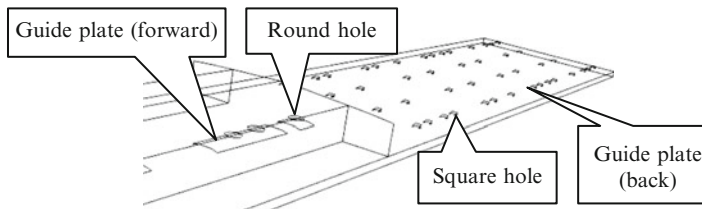
Fig. 3 The calculation model of the main air channel (*back*)

There are two ways of sending air in the air-conditioning system: one is sending air by the main air channel which locates at the middle of the top of train, the other one is sending air by the air channel for heating which locates at each side of the main air channel. The main air channel contains two layers above and below. The two layers are separated by clapboard on which there are some holes. After going into the air channel, the cold air flows from the above layer to below through the holes, and then goes into the carriage through the orifice plate at the bottom of the below layer. The model of the main air channel is shown in Figs. 2 and 3.

There are 28 round holes in the front of the clapboard and 46 square holes in the back of the clapboard. The guide plate lies below the round holes under the bottom of the air channel, and above the air channel each square hole owns a square guide plate.

In the new design scheme, the number of passengers in the carriage is 33, and there are 3 passenger compartments in the carriage, each number of which is 4, 6, and 6. The operation condition, considering the summertime, is that 75% of the cold air enters the carriage through the main air channel, while 25% of it enters the side air channel from the below part of the carriage. Table 1 shows the specific calculation condition.

2.2 Mesh the Model

The precision of grids can affect the veracity and the convergence property of the result. For this reason, the grids are denser in the place of air-delivering inlet and air-returning outlet. Flow and heat exchange are calculated by wall-function method

Table 1 The calculation condition

Item	Value
Temperature in the carriage (°C)	25
Temperature out of the carriage (°C)	35
Number of passengers (person)	33
Temperature of the cold air inlet (°C)	13
New air quantity (m³/h)	1,200
Air-returning quantity (m³/h)	3,600
Supplying air quantity (m³/h)	4,800
Heat transfer coefficient of body (W/(m²K))	1.2
Heat transfer coefficient of window (W/(m²K))	1.8

Fig. 4 The results of quality of equiangular skews and equisize skew on meshing
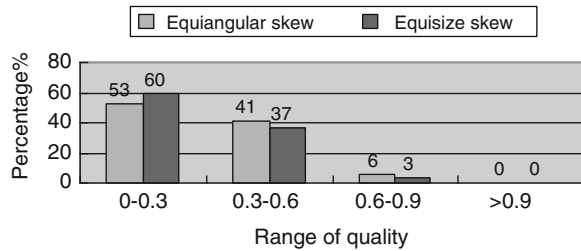
in the place of viscidity sub layer nearby the wall surface, so there are no grids nearby the wall surface and the first node connected with wall surface is disposed in the region of blooming turbulence. The model is meshed with tetrahedron elements, the number of which is about 4,000,000. Figure 4 is the histogram which shows the result of the equiangular and equisize distortion. It can be seen that the quality of grids is between 0 and 0.6, and it can be considered that the quality of the most grids is good (Huang and Bi 2008).

2.3 Calculation Method

The air flow in the carriage is considered to be a kind of three-dimensional, uncompressible, steady and turbulent one. The related governing equations include continuity equation, momentum equation, energy conservation equation, and standard k-ε turbulent model. These equations are shown in the reference (Wang 2005).

The calculation is done by the software of Fluent, using segregated solver and SIMPLE solve method (Yu and Che 2010). The convection term is calculated by second order upwind scheme and the viscosity term is calculated by second central differencing scheme. The calculation meets the hypothesis that the change of the
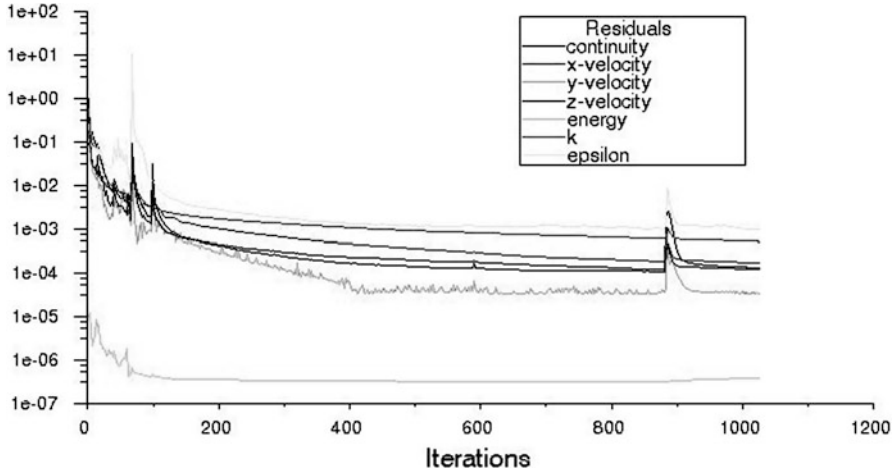
Fig. 5 The computation residual curve

flow density only impacts the buoyancy existing in the momentum equation. The calculation is converged on the step of 1,025, and the relative errors of the flow between inlet and outlet is 0.058%. Figure 5 is the computation residual curve.

3 Method Analysis of Calculation Result

3.1 Distribution of Temperature and Velocity in the Main Air Channel

On the above layer of the main air channel, the cold air enters at a higher velocity, and then the velocity descends gradually (He and Jin 2009). When the air gets into the rear of the air channel, as the area of longitudinal direction section decreases, the velocity becomes some higher. In the below layer of the main air channel, the foreside air goes from the above layer into the below one through the small holes on the clapboard, and then distributes uniformly in the air channel with the effect of the guide plate, at the same time, the cold air from the rear of the air channel firstly goes through the guide plate which locates in the above layer of the air channel, and then goes into the below layer after reflecting by the rear of the channel. Figures 6, 7 and 8 show the distribution of velocity of the main air channel.

In practical situation, there are adiabatic materials adhered to the wall surfaces of the air channel (Stathopoulou and Assimakopoulos 2008). For this reason, the wall surfaces of the air channel are set to be adiabatic. In the air channel, the temperature of the cold air retains about 13°C.
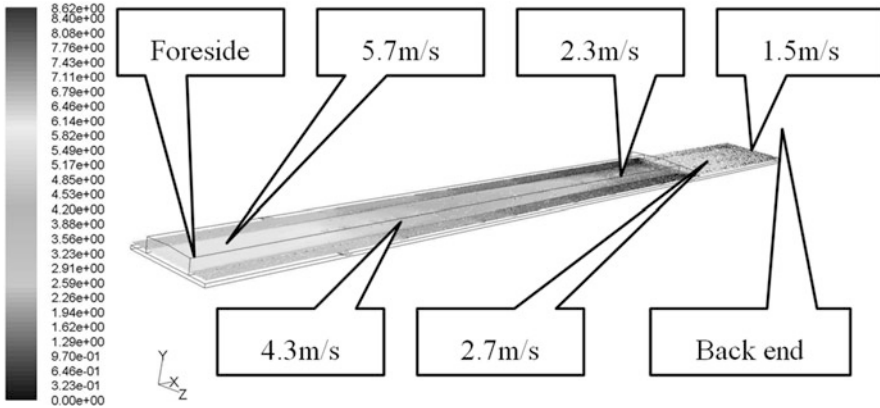
Fig. 6 The distribution of velocity on the above layer of the main air channel
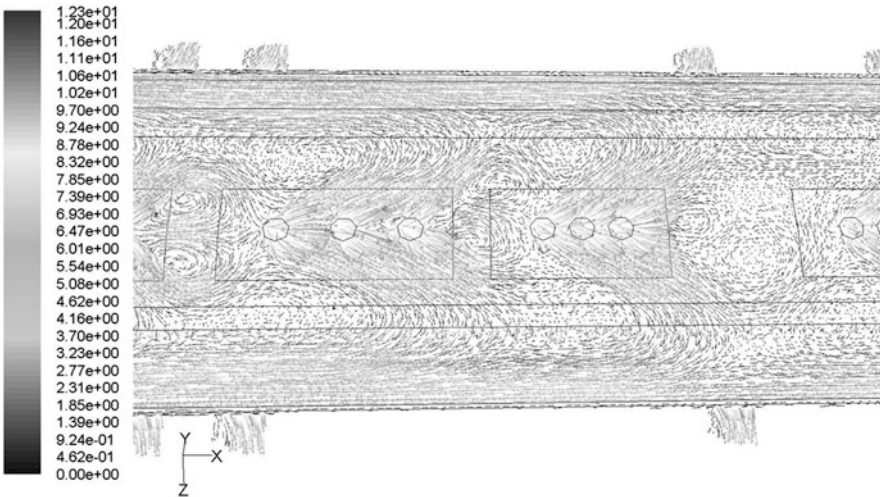
Fig. 7 The distribution of velocity in the below layer of the main air channel (*front*)

3.2 Distribution of Temperature and Velocity in the Carriage

The distribution of temperature around people is not very uniform, and the highest temperature is 25.3°C while the lowest temperature is 22.2°C. The difference of temperature is 3.1°C. The temperature goes a little higher in the compartments along the direction of the body of train, and arrives at the highest out of the compartments. Figure 9 shows the temperature distribution of the section at the height of the body of passengers. The velocity of the wind around passengers is 0.01–0.29m/s, which measures up the comfort of passengers. Figure 10 shows the velocity distribution of the section y = 0 in the carriage.
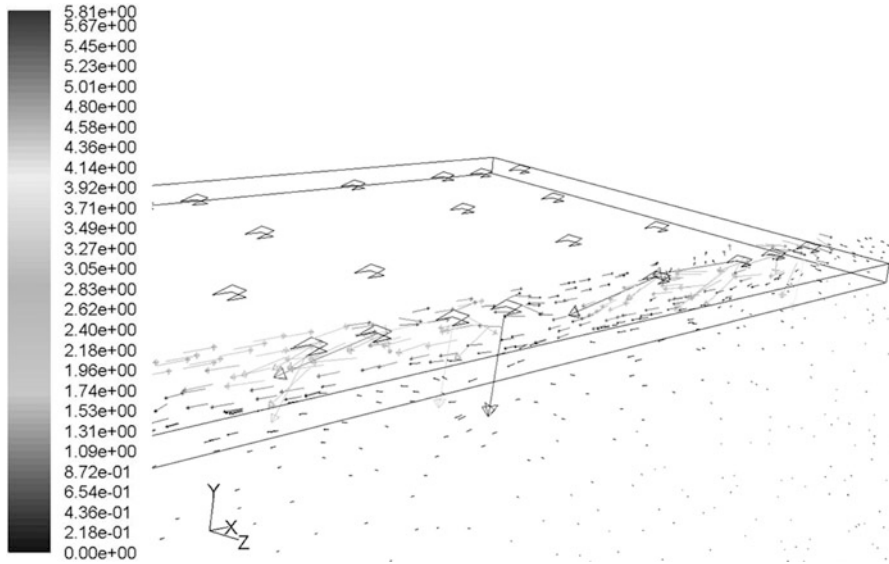
Fig. 8 The distribution of velocity in the below layer of the main air channel (*back*)
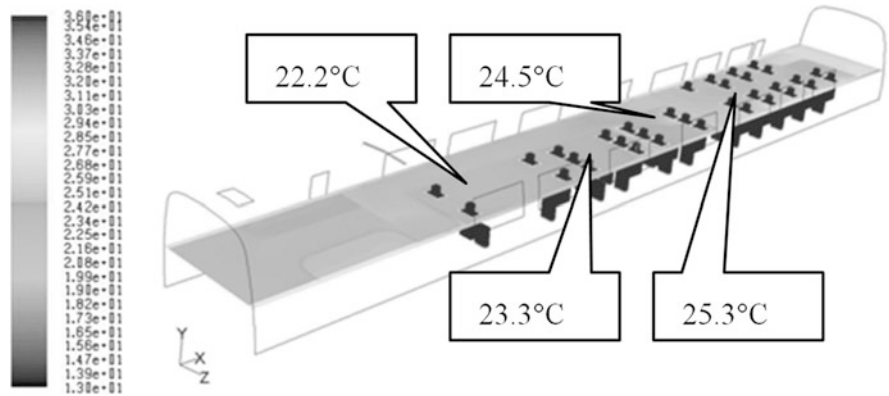
Fig. 9 The temperature distribution of the section at the height of the head of passengers

4 Improvement About the Main Air Channel

It can be seen from the above analysis that, in the new design scheme, the quantity of supplying air descends from front to back, in addition, the density of passengers is rare in the front part and dense in the back part, which result in that the temperature increases gradually, and increases by 3.1°C at the end of the carriage. It can be known from analyzing the structure of the air channel that round holes corresponding to the guide plate are in the same plane (15 cm to the above clapboard), so it
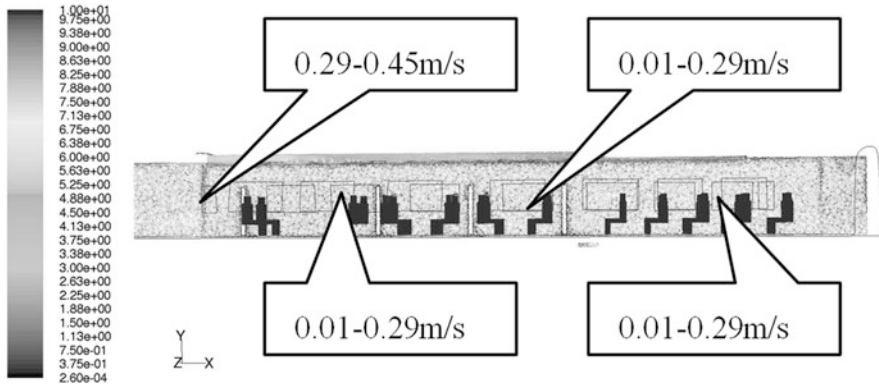
Fig. 10 The velocity distribution of the section y = 0
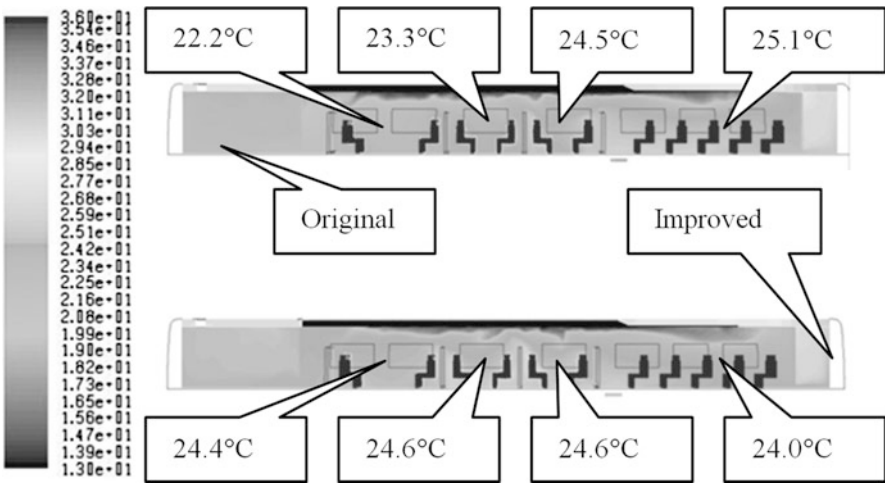
Fig. 11 The temperature distribution contour of the lengthways symmetry plane in the carriage in two situations of the main air channel improved and not

can be considered to control the rationality of the quantity of supplying air from the main air channel by adjusting the vertical height of the guide plate (Mu et al. 2009).

According to the practical experience in factory, the two guide plates in front of the below layer of the air channel are moved up by 5 cm, and the last one at the bottom is moved down by 2 cm. Then the improved model is calculated by Fluent, and it can be found that the temperature distribution in the carriage has been improved to some degree. The highest temperature around the passengers is 24.6°C and the lowest is 24.0°C, and the temperature difference is below 1°C. Figure 11 shows the temperature distribution of the lengthways symmetry plane in the carriage in two situations of the main air channel improved and not.

5 Conclusion

In summertime conditions, we make the numerical simulation of the train air-conditioning ventilation system and the flow field inside the train by the use of the technology of the simulation, and make analysis and evaluation on the air flow inside the carriage directly and vividly (Wang et al. 2004). The following conclusions are:

- In the upper air channel, the cold wind speed gradually decreases from the start of the inlet; and the air flow of the lower is relatively intense.
- In the original design scheme, the velocity of the wind around passengers is between 0.01 and 0.29 m/s which is good. The temperature around the passengers is between 22.2 and 25.3°C, which is not very comfortable to passengers.
- After changing the height of the guide plate under the main air channel, the distribution of air supply can be adjusted to make the temperature inside the carriage more uniform, and the temperature difference is less than 1°C.

References

He YZ, Jin LN (2009) HVAC design of the main sports complex of Shenzhen Universiade. J HV AC 39(8):1–4

Huang XH, Bi XP (2008) The application of computational fluid dynamics in design of the vehicle cooling wind tunnel. J Acad Armor Forces Eng 29(11):1281–1285 (Beijing, China)

Mu GD, Zhan SH, Zhao ML (2009) Displacement ventilation and CFD numerical simulation. J Heat Vent Equip 137(215):36–39

Stathopoulou OI, Assimakopoulos VD (2008) Numerical study of the indoor environmental conditions of a large athletic hall using the CFD Code PHOENICS. J Environ Model Assess 13:449–458

Wang FJ (2005) The analysis of computational fluid dynamics, vol 1, 1st edn. Tsing Hua University Press, Beijing, pp 7–10

Wang G, Hu ST, Lian LM (2004) Optimization of outlet style for a new carriage duct. J HV AC 34(2):9–12

Yu XB, Che YQ (2010) Numerical simulation of influence from cabin thermal environment with door gap. J Mach Des Manuf 5:133–135

Zhang DC, Weng PF, Zou SH (2006) Research of thermal comfort in railway passenger cars. J China Railw Soc 28(8):35–40

Zhao WZ, Cao YN, Wang DP (2011) Evaluation on numerical simulation for high-speed train air conditioning and ventilation system based on UIC standard. J Dalian Jiongtong Univ 32(1):7–10

Zheng BL, Wu XG, Gu CZ (2010) Goals and direction of CAE analysis technology of aerodynamics on high-speed train in China. J Comput Aided Eng 19(2):10–12

The Study of Scientific Development of China's Foreign Trade from the Perspective of Dependence on Foreign Trade

Xue Xuandeng

Abstract Dependence on foreign trade reflects the extent, which a country's economy contact with the whole world economy, and to some extent, reflects the importance of foreign trade in the national economy. The concept of scientific development is the important guideline of our country's economic and social development. This paper analyzes the status, cause and effect of our country's dependence on foreign trade in recent years, and provides some effective measures to promote the scientific development of China's foreign trade from the perspective of dependence on foreign trade.

Keywords Foreign trade • Dependence on foreign trade • Scientific development

1 Introduction

Since the reform and opening, the scale of China's economic and trade development has made great development and the total trade increased year by year, which has been a significant role in promoting of China's economic development. At the same time, China's dependence on foreign trade is continually improving, and this trend becomes more apparent especially after the accession to WTO. The rising dependence on foreign trade has also brought a lot of trade risks.

X. Xuandeng (✉)
School of Economy, Henan University of Science and Technology, Luoyang, China
e-mail: xue69519@163.com

2 The Status of China's Dependence on Foreign Trade

2.1 The Increasing Dependence on Foreign Trade

Since China joined the World Trade Organization in 2001, the pace of the participation economic globalization has accelerated, the degree of integration into the international community has deepened gradually, and foreign trade has been playing a prominent role in promoting economic growth increasingly. From the Table 1 we can know, China's foreign trade increased from 509.65 billion dollars in 2001–2008 to 2.56 trillion dollars. With the continuing expand of China's foreign trade, foreign trade dependence has also been increased except for the decline in exports and dependence on foreign trade in 2009 because of the world financial crisis. In generally, our dependence on foreign trade is growing year by year, continuously increased from 0.44 in 2001 to 0.70 in 2007, slightly lower in 2008, but still reached 0.60.

2.2 The Eastern Reign Is Higher Than Central and Western Reign

In order to facilitate the analysis, this paper selected Shandong Province, Anhui Province, Qinghai Province as a representative of eastern, the central and western regions. Table 2 shows, in 2008, for example, the GDP of Shandong Province was 3.107206 trillion Yuan, import and export amounted to 1.111859 trillion Yuan, and foreign trade dependence was 35.8%, while dependence on foreign trade of Anhui Province and Qinghai Province was 16.0 and 5.0 separately. Thus, there is also a great imbalance for our dependence on foreign trade in the region. That is to say, eastern region is economic development, the import and export trade, which is in the leading level of foreign trade. Is also more developed, and the dependence on

Table 1 China's dependence on foreign trade data during 2001–2009 (billion)

Year	Total import and export	Domestic GDP	Dependence on foreign trade
2001	5, 097	11,620	0.44
2002	6, 208	12,707	0.49
2003	8, 510	14,166	0.60
2004	11, 546	16,494	0.60
2005	14, 219	22,281	0.64
2006	17, 604	26,464	0.67
2007	21, 737	31,050	0.70
2008	25, 633	42,835	0.60
2009	22, 072	49,100	0.45

Source: China Statistical Yearbook 2009 and Ministry of Commerce Website

Table 2 The contrast of dependence on foreign trade among Shandong, Henan, Qinghai in 2008, (unit: million)

Region	GDP	Import and export	Dependence on foreign trade
Shandong	31072.06	11118.59	0.358
Anhui	8874.17	1416.64	0.16
Qinghai	961.53	48.43	0.50

Source: China Statistical Yearbook 2009

foreign trade in general is higher than the Midwest. The Midwest regional industrial is backwardness, industrial configuration is unreasonable, so correspondingly lower dependence on foreign trade in the eastern region.

2.3 Significantly Higher Than the World Average

With the respective of the current average level of the world, the world's dependence on foreign trade reached 0.41 on average, of which the average level of developed countries reaches to 0.384, and developing countries is 0.51. The average dependence on foreign trade in China was 0.57, which is higher than the world level. Compared With dependence on foreign trade of other world powers (take 2007 as an example), China's dependence on foreign trade is not only higher than 0.309 of India and 0.219 of Brazil of developing countries, but also higher than 0.23 of United States, 0.304 of Japan and other developed countries.

Chinese foreign trade dependence on the average annual growth rate is 7.82%. Compared with other countries, our dependence on foreign trade is not only higher than India, Brazil and other large developing countries, but higher than the U.S., Japan and other major developed countries. China's dependence on foreign trade was 11% in 1979, which is lower than India, the United States and Japan's foreign trade dependence, while it exceeded these three great powers in 1990. Take the year of 2007 as a example, the Chinese dependence on foreign trade reached 0.70, is far higher than the United States and Japan, even more than twice higher than India and Brazil of the developing countries. Such a high foreign trade dependence is not common at the international level, So the study of Chinese foreign trade dependence have a good guide and alert significance for the healthy development of Chinese economy, foreign trade, and scientific development.

3 The Cause Analysis of the Status of Chinese Dependence Foreign Trade

3.1 The Promotion of Economic Globalization

The coordinate construction of the International trade, international finance, and international technology provides a system for the protection of economic

globalization, promotes the cross-border movement of the factors of production, and creates the conditions for China's export trade. China, as one of the largest developing countries benefits most during the process of economic globalization. In the 1990s of last century, economic globalization is speeding up, and the cross-border investment in global world optimizes and allocates the sources, all of which become an important driving force of world trade growth. The growth of international trade is almost two times higher than the world's average economic growth. The average world dependence on foreign trade continues to increase. According to the estimation of World Bank, by 2015 this proportion may rise to 60%. In some newly industrialized countries, dependence on foreign trade in general greatly improved because of the implement of the export-oriented strategy in the economic take-off stage. Since China joined the WTO in 2001, the degree of opening further increased, the development with the world economy has become increasingly inseparable, and foreign trade dependence is substantial increased. Therefore, the improvement of our dependence on foreign trade is consistent, and inseparable with the process of economic globalization.

3.2 The Expanding Scale of Foreign Trade

In the twenty-first century, it provides a good historical opportunity for the development of China's foreign trade to join the WTO and deepen economic globalization. The sustained rapid economic growth of China also provided a good macroeconomic environment and the solid material foundation for foreign trade. Since the 1990s of twentieth century, the export growth rate of China is faster than the GDP growth rate. Therefore, the relative growth rate of foreign trade and GDP is much higher than the world average level, which has led to increasing dependence on foreign trade in China and is higher than the world average level. Since accession to WTO, China's GDP increased from 1162.01 billion $ in 2001 to 49106.1 billion in 2009, and the average annual increase is 9.7%. At the same time, the imports and exports of China increased from $ 509605 billion $ in 2001 to $ 22072.2 billion $ in 2009, with the 18.3% of the average annual growth.

3.3 Domestic Industrial Structure Is Irrational, and the Proportion of Tertiary Industry Is Lower

Generally speaking, when the second industry is domination and third industry lags in a country' GDP, the products of import and export is mainly the labor-intensive products or the simple processing technology-intensive products, which

Table 3 The comparison of dependence on foreign trade among U.S., Japan, India and Pakistan, in 2007 (unit: billion U.S. dollars)

States	Import and export	GDP	Dependence on foreign trade
American	31828.8	138438.3	0.23
Japan	13338.6	43837.66	0.304
India	3619.5	11709.7	0.309
Brazil	2872.3	13135.9	0.219
China	21738.3	32508.3	0.692

Source: World Bank, "World Development Indicators 2009"

will be included in the total foreign trade, so the trade dependence on foreign trade calculated by trade is relatively high; When the tertiary industry's share in GDP is higher, The main benefit of foreign trade is through the capital, technology and service trade, etc., rather than import and export of commodity, so the dependence on foreign trade is relatively low calculated by trade. Table 3, China's domestic industrial structure dominated by processing and manufacturing, and the proportion of tertiary industry is low. At present, the proportion of tertiary industry in China's GDP is more than 40%, which is one of the reasons that cause the high dependence on foreign trade in China.

4 Increased Dependence on Foreign Trade of China's Economy

4.1 Positive Impact

4.1.1 Promote the Upgrading of Domestic Industries

Increased dependence on foreign trade promoted the upgrading of domestic industries, and thus promoted China's economic growth. On the one hand, it Promoted the technological progress of related industries in China imports of badly needed by advanced technology and key equipment importing advanced technology and key equipment badly needed; On the other hand, through active participation in foreign trade activities, domestic enterprises can better understand the international market information, and thereby enhanced the ability of imitation and innovation. In addition to, the technical "spillover" effects of foreign enterprise also contributed to the upgrading of domestic industries. Increased dependence on foreign trade, especially in developing countries is also one of the engines to accelerate the development of a country's economy (Zeng et al. 2008). According to development economics, foreign trade is the "engine" for economic growth. Economically backward countries need the help of foreign trade or economic openness to exploit

the advantages of factors of production itself, thus gets economic development funds from the international market relying on foreign trade to realize the industrialization gradually, and adjusts the industrial structure and export structure timely to urge the national economy through foreign trade to the track of industrialization.

4.1.2   Accelerating the Process of Integration into the Global Integration

China's increasing integration into the trends of world economic development and plays an increasingly important role in the contemporary international division of labor. With the rapid development of science and technology, international division of labor will become increasingly greater in breadth and depth, and China's economic impact on the world economy will be growing. Therefore, with the oncoming force of economic globalization, the substantial rise of China's dependence on foreign trade is a necessity. Increased dependence on foreign trade has brought a most direct consequence that is China's economic development is jointing with the international standards.

4.1.3   Raising the Employment Rate and Playing Comparative
Advantage of China

The massive growth of foreign trade provided a large number of labor positions, made the population of hidden unemployment in rural areas can get a job, increased the national income, and improve the overall level national welfare. The improvement of employment rates plays an important role for the social stability. Meanwhile, because of the abundant labor resources of China, providing the world with a lot of labor-intensive products gave full play to China's comparative advantages and accelerated China's pace forward to largest trading and economic power.

A system for the protection of economic globalization, promotes the cross-border movement of the factors of production, and creates the conditions for China's export trade. China, as one of the largest developing countries benefits most during the process of economic globalization. In the 1990s of last century, economic globalization is speeding up, and the cross-border investment in global world optimizes and allocates the sources, all of which become an important driving force of world trade growth. The growth of international trade is almost two times higher than the world's average economic growth. The average world dependence on foreign trade continues to increase. According to the estimation of World Bank, by 2015 this proportion may rise to 60%. In some newly industrialized countries, dependence on foreign trade in general greatly improved because of the implement of the export-oriented strategy in the economic take-off stage. Since China joined the WTO in 2001, the degree of opening further increased, the development with the world economy has become increasingly inseparable, and foreign trade dependence is substantial increased. Therefore, the improvement of our dependence on foreign trade is consistent, and inseparable with the process of economic globalization.

4.2 The Negative Impact

4.2.1 The Impact of Domestic Economic Security

National economic security refers the most fundamental economic interests of a country can not be harmed. As the continuous improvement of our dependence on foreign trade, the fluctuation of external demand is becoming increasingly the risk affecting China's economic operation we have to face. The impact, the dramatic changes of world economy bring for China's economy, which relies on foreign demand increasingly will be difficult to predict. Because of high import dependence on the high-tech products, when the world economy is to significant volatility, China will suffer the economic, scientific, and technological losses. Because of the import markets of certain important commodities are more focused, Once the country-specific export ban on China's exports caused by political or other, the loss will be immeasurable. The rapid economic development, added to China's economic growth is still the dominant extensive mode, and energy consumption is more than the output of China's ability, all of these causes China has to import energy and raw materials. If the world economy takes place more volatile, China's economy will be affected definitely (Peng Zhang et al. 2008). Especially the current import and export of China mainly concentrated in a few developed countries and regions, If the economy of these nation and region is instability or deliberate obstruction, we can not effectively prevent the risks. For example, in recent years, the dependency of China's imported crude oil is rising. More than half of these imported oil is from the Middle East. About 4/5 transport of offshore oil has to go through the Malacca Strait, the majority ships used are rental foreign tankers, once blocked because of this, China's oil security will be seriously threatened.

4.2.2 Domestic Industrial Structure Is Not Conducive to Be Optimized

Since reform and opening, the structure of China's foreign trade has been significantly optimized. The exports of primary product have been less competitive, and the exported competitiveness of manufactured goods is rising with the proportion of total exports has increased to 91.2%. However, as far as the industrial products concerned, the competitive exports of China are still concentrated in products with high labor-intensive. The competitiveness pattern still showed strong labor-intensive products, and weak capital and technology intensive products. In addition to labor-intensive products, China's foreign trade, other elements-intensive products are basically deficit. Labor-intensive products not only make up this deficit, it also appears surplus of China's foreign trade in general. Thus, Although China has a certain improvement of industrial structure in recent years; The structure of goods in foreign trade were optimized; It appears breakthrough in high-tech exports, However, generally speaking, the competitive products of China in the international market is still only labor-intensive. This continuous expansion of low value-added

industry is not conducive to the upgrading of domestic industrial structure in China, and Increased the difficulty to improve the advantage target of foreign trade competition. The larger proportion in China's exports are the textile and garment, footwear, labor-intensive products like toys, and these industries have emerged the oversupply situation in the international market.

4.2.3 Foreign Trade Frictions Tend to Be Frequent

At present, China's foreign trade concentrates in few countries and regions, and the import and export of three main trading partners of United States, Japan, and the European Union accounted for about half of the total. Thus, China has inevitably entered the era of international economic friction, which can be best proved by the Sino-US anti-dumping, trade barriers and intellectual property issues and the pressure the United States has frequently put on RMB in recent years. Thus, the development of China's foreign trade and the rising dependence on foreign trade has made China's foreign trade into the international economic friction times inevitably. How to response correctly and rationally has become an urgent task.

5 Measures to Promote the Scientific Development of China's Foreign Trade

5.1 To Develop Trade in Services

At this stage, China's industrial structure is still irrational. The negligence of the development of modern service industry for a long time causes the development of financial, trade and other modern service trade lags. Tertiary industry in GDP, accounts for only 1/3 or slightly more, while developed countries are more than 70%. The share of services sector of our country in GDP is lower, which is also the reason for high rising dependence on foreign trade in China. Service trade should be developed rapidly and we should improve its proportion in national economy and foreign trade. By doing this, part of the surplus labor and capital can be attracted to the development of service trade; The industrial structure and the structure of foreign trade of China can be optimized; The sources of interest can be increased; The high dependence due to processing trade can be reduced, thereby the dependence on foreign trade in China can be reduced, at the same time, the share of the output value of service sector in GDP can be promoted, and the proportion of numerator can be Increased in order to lower the dependence on foreign trade of our country.

5.2 To Promote the Transformation and Upgrading of Processing Trade

The domestic processing enterprises are only simple processing and assembly, and the usage of the domestic raw materials and components is low during processing trade, which makes the fast growing of the processing trade not promote other industries, and the industrial structure of China has not been upgraded and progress by development of the processing trade. Once the human resources of China are loss, foreign trade development will be seriously affected. Therefore, when the policy to develop the processing trade formulated by government, it should be driven from the role of starting their own industries to reduce the negative nature of the processing trade, Promote the processing trade to gradually shift from the simple processing and assembly to the upgrade of procurement, manufacturing, Retailing service, post-sale service, research and development, and information consultant, and Promote the expansion of domestic demand through the processing trade in order to enable the processing trade to be the link between the home industry chain and the international industrial chain (Qiang Fu 2007).

5.3 To Optimize the Structure of Imported and Exported Products

The exported products of China are dominated by the labor-intensive industrial products like shoes, clothing, toys, etc., and such products are highly dependent on specific markets, and are constrained by it. We should implement the strategy of the diversification of foreign markets and exported commodity's structure to lower the dependence on the specific trading partners and specific imported and exported commodities and to ensure the stability of national economy (Qiyun Wang et al. 2008). As China's economic growth in the past did not change the extensive economic growth mode of "high input, high consumption and low efficiency". Now our natural resources are in the relatively poor state, and many mineral resources are imported from several exported countries of mineral resources, therefore our economy is highly dependent on the external environment. At the same time, a large number of resource products are used for export to increase foreign exchange, and the results of doing like this lead to many precious raw material of our country are out flown. Therefore, at the time of the promotion of energy conservation and improvement of resource utilization, it is necessary to control the export of natural resources to reduce the dependence on foreign trade.

5.4 To Change the Mode of Foreign Trade and Increase Direct Foreign Investment

If the production plant was established in the local target market through foreign direct investment by China's enterprises, which can not only avoid trade barriers, but also can reduce the favorable balance, and balance China's international payments, and then to maintain the stability of the RMB exchange rate. The currency of our country can be maintained at a reasonable level to reduce the pressure of inflation due to the inflow of additional money. With the International development of enterprises, we can make the domestic industry chain extend out and promote the transformation and upgrading of our national economic system (Wei-xian Xue et al. 2008).

5.5 To Speed Up the Investment of Science and Technology to Develop High-Tech Industries

Restricted by the level of domestic technology development, the dominant exported products of China are only labor-intensive products, and high-tech products are few. Meanwhile, the imported products in China are dominated by the high-tech products with high price, which further enhance China's dependence on foreign trade. On the one hand, it can reduce the dependence on foreign imports of high technology to accelerate the scientific and technological investment and develop high-tech industries; On the other hand, we can create our own brands in the world in order to take the initiative advantages in foreign trade. As far as China's dependence on imported raw materials, such as oil, etc. concerned, we should further increase the investment of science and technology, develop the alternatives as soon as possible and reduce imports to avoid the phenomenon of controlled by others.

6 Summary

In short, the dependence on foreign trade is a "double-edged sword" for the improvement of China's economic development. On the one hand, it has made our country more actively involve in international economic development and improved the economic status and influence of our country; On the other hand, it has also brought new risks for China's economic development. China, as a developing country, must pay attention to the issues of dependence on foreign trade to ensure the healthy and scientific development of foreign trade.

References

Zeng-e J et al (2008) Inquiry for the reason of Chinese high ratio of dependence on foreign trade and research on the countermeasure. North Econ Trade J 9:35–37

Peng Zhang (2008) The impact of trade dependent degree on China's economic growth. J Int Trade (April):16–22 (references)

Qiang Fu (2007) New analysis of degree of dependence upon foreign trade in China. World Econ Study 4:44–47

Qiyun Wang (2008) Low consumption policy: root cause of the rapid rise of China's foreign trade dependence. J Hebei Univ Econ Trade 29:42–46

Wei-xian Xue, Dong Wei-wei (2008) Grey relevance analysis of influencing factors for China's dependence on foreign trade. Int Econ Trade Res 24:4–8

The Design of University Library Expert Advice Platform

Luo Ying

Abstract University's virtual reference system (VRS) as the universities library's participation in teaching and research work in the network extensions plays a very important role and position. Based on the study of its connotation, service methods and the present domestic and international development situation and related building systems technology, analyzed the system requirements, features and particularity, designed a library VRS based on university. This system took to enhance service practicability and information acquisition accuracy as the essential starting point, put forward a method based on the analytic hierarchy process (AHP) experts scheduling model. For traditional systems shortages, it put forward the questions wizard module, in which the AJAX technology was applied to realize. AJAX will put clumsy Web interface into strong interactivity application, user in the process of asking will not be refreshed the page, be interrupted operation, and get more quickly response, reduce operating time. Another it used participle technology to take ask and retrieval well together, greatly reducing the probability of repeated questions, and it is quite good to avoid the existing defects when traditional system ask cross-disciplinary questions. Finally summarized the system design, and gave more suggestions to further perfect.

Keywords Virtual reference • AJAX technology • Participle technology • AHP method

L. Ying (✉)
Library Information and Automation Department of Chongqing, Three Gorges University, Wanzhou, Chongqing, China
e-mail: luoying716@gmail.com

1   Practical Significance of the System

Reference library services has always been one of the traditional reference service are the "one to one" model, by time, technology, communication means, the quality of personnel and other conditions, its scope is small, does not influence large. In the network environment, library reference services, content and form have undergone tremendous changes, not only provides us with a wealth of information resources and advanced means of information exchange, and the library, computer systems, databases and even readers of the Advisory between the solution as a whole, the formation of a new information service pattern – Virtual Reference Service.

Virtual reference service, also known as digital reference service (Digital Reference Service, DRS), is based on Internet (or Web) to help service mechanism. Internet users can ask questions to the consultants, and consultants are the answer online feedback to the user. Compared with the traditional reference service, the service object, service mode, service tools, services, great changes have taken place. It effectively transcend time and space limitations, "not to meet face to face" way, so that librarians and clients no matter where, as long as the site log to provide advisory services to interactive consultation, which is all online users are equal. The questions which be answered can provide more information for more readers.

University Library is the dissemination of knowledge and important body of information services is an important distribution center for information resources. As the school's documentation and information center, It has an important position in teaching, scientific research. Today, networks are increasingly developed in the traditional library should be a gradual transition to the Modern Library considering re-location services model, in order to achieve comprehensive, multi-level efficient, convenient, accurate and comprehensive information service.

We therefore based on the university, proposed the construction of expert advice platform, intended to take greater advantage of library resources, more effectively organize the human resources experts, universities, and fully tap the expertise of experts to provide students with a deeper teaching service, to provide an effective platform to facilitate interaction. The establishment and completion of this platform will realize the Internet between students and experts to build bridges of communication, help students to more deeply grasp the scientific and cultural knowledge, and is able to research and teaching experts play a very good supporting role.

2   The System Design

Platform using the groups are primarily college students and subject experts, so the whole system design aspects in which both embody the characteristics of its services and features in college. The enhancement university reference service availability and accuracy of access to information as a fundamental starting point

Table 1 The difference between the general VRS and the university VRS

	Service type	Content by the advisory	Consultants work
General Virtual Reference System	Public library	Classified according to industry	Full-time
University Virtual Reference System	University library	Divided by subject	Part-time

of virtual reference platform development, content and services means the colleges and universities while the specificity of virtual reference service in order to make better use of books, experts and other resources for teachers, students of teaching and research services.

2.1 System Requirements

Table 1 the difference between the General VRS and the University VRS Shown in Table 1, in the public library, the reference content is usually classified according to sectors, and in universities should be divided according to the precise subject; Secondly, the role of consultants as subject specialists from the school, as they usually are busy teaching and research work, answering only spare time for consultation work, the expert's available time throughout the consultation process is a very important factor.

2.2 The System Features

The system should address the following three aspects highlighted: (a) Professional excellence and not exhausted by increasing the difficulty of matching the expert knowledge; (b) The expert's available time in the consultation process can be considered a very important factor; (c) Select the expertise of existing counterparts, those who can answer questions as quickly as possible is the system to solve a substantive problem. Based on the above analysis, should be introduced questions Wizard module in the system, and highlighted the importance of the work of expert scheduling module.

2.3 Questions Wizard

In order to collect enough information to enhance the accuracy of the extracted set of experts, the increase of the questions asked during the Wizard module. Questions from the guide are usually the role of the wizard. In addition, it helps the system to collect information, to solve the two problems, Firstly, cross-disciplinary. With

Table 2 The system compared with other similar systems

	Questions dealing with cross-disciplinary	Avoid duplication of questions
The system	Allows the user the flexibility to select one or more subject categories, cross-disciplinary issues can be properly classified	A question guide as a bridge into the question the problem retrieval process, to achieve the organic combination of questions and retrieval
Other similar systems	Not take into account the existence of cross-disciplinary questions, only to achieve a simple classification problem	In the form of a text to remind the user library in the FAQ before asking queries, questions and search are two separate processes

the development of science and technology, interaction between all disciplines, the phenomenon of mutual penetration has become more evident in the university system in the virtual reference users can not help but encounter the case of interdisciplinary questions, the choice of disciplines and even the experts will search FAQ scheduling accuracy will have a very significant impact. Secondly, the repeated questions. Questioner raised the same concerns in question, often with temporal locality, that is the same question at some time receive special attention. Multiple users repeated the same question concerns experts and the system will increase the workload of maintenance personnel (Table 2).

Here, the system uses AJAX technology to solve and segmentation, Ajax is short for Asynchronous JavaScript and XML, the asynchronous JavaScript and XML. In layman's terms is a request by the client asynchronously call server data without page refresh operation to achieve the technology. It is the realization of the core technology of Web2.0, the clumsy Web interface into strong interactive Ajax applications. Its advantage is very easy to implement asynchronous interaction and dynamic refresh Web pages for users and server interaction frequent occasions. Here the use of AJAX allows users to question the process does not refresh the page, do not interrupt operations, and are responding more quickly to reduce the operating time (Shi Mei et al. 2000).

2.4 The Expert Scheduling

Experts scheduling process shown in Fig. 1, the system based on available information, assigned to a specialist for the user to answer their questions. On the one hand, university virtual reference systems usually require an expert working in a busy teaching and research work, only in his spare time engaged in reference services. On the other hand, given a good reference system is based on the shortest response time for user questions the primary objective.
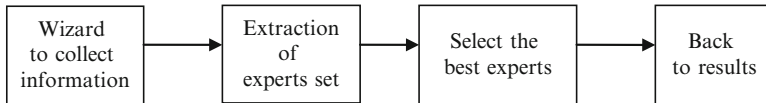
Fig. 1 Experts scheduling process diagram

2.5 *The Consultation Process*

Combining the basic function of reference consultation service and universities features, it include the following two working mode:

(a) asynchronous work patterns. The most basic way of consultation: FAQ, mail advisory, form consulting, etc. In this system, to bring FAQ and form consulting the two mode organic combine together, let them for customer service. In questions into the FAQ guide can reduce unnecessary workload that repeated questions bring experts and database maintenance personnel.
(b) synchronous work patterns.

Is a real "face to face", also called real-time reference. In the system through the text to complete a short time talking to the reference service, because Java is a complete sense of the object oriented language and its support for safety and synchronization mechanism, here TCPSocket realize the use Java online reference service, make the whole system has a good stability and security.

In summary, the university virtual reference system flow is shown in Fig. 2. Firstly, ask the user through the wizard to retrieve the problem, if find the answer, the result will be returned to the user; otherwise, under the guidance of the wizard in question, complete the information questions, and enter the expert scheduling. Secondly, the expert scheduling this session, an expert assigned by the system to answer user questions and the requirements of the user, respectively to form consulting and real-time consultation. In the matter to the experts, the experts may be due to some reason can not answer user questions, the questions can be forwarded to other experts, or application problem levy answer. Experts to answer questions, will collate the information into the Q&A database, and returns the results to the user. Thus, end of the first Q&A round of consultation.

3 The Key Technology

The entire system using JAVA language and SQL database development. And using the following technology to solve problems which be encountered in the development process:

(a) Compared to traditional synchronous interactive web application page, asynchronous interaction with the server will be using AJAX technology, mainly in
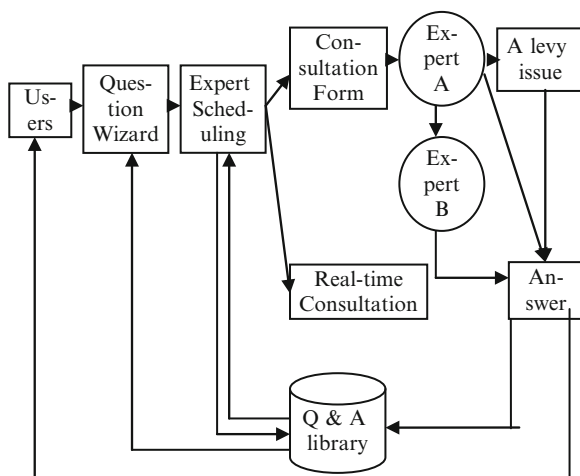
three areas: (a), without refreshing the entire page, communicate with the server
within the page; (b), the use of asynchronous communication with the server,
without interrupting the user's operation, with more rapid responsiveness; (c),
application of only a few pages and a majority of interactions within the
page, without switching the entire page. So, Ajax makes Web applications
more dynamic, and provides rich expressiveness Ajax UI components. The
core technology is the XMLHttpRequest object, which is a unique place Ajax
technology. Ajax update without refresh the page the effect is achieved by the
XMLHttpRequest object.

(b) The Chinese word segmentation be used to search database natural language.
(c) To determine the scheduling expert's importance in universities VRS, an
    important role in the proposed AHP which based on hierarchical analysis
    approach to scheduling expert.
(d) The overall system design process, design patterns introduced the idea of the
    implementation details of the physical separation of the business layer as much
    as possible in improving the expense of reusable code, while enhancing the
    maintainability of the system.

4   The System Outlook

Now, some large network technology company in their website design and devel-
opment of interactive search systems, such as Baidu, known system, Sina love to
ask system are currently the most leading intelligent interactive search technology,
which fully embodies the humane application of the product philosophy, for the
majority of Internet users to provide a new search service. New Age libraries as
information resources on land, only the absorption of the former's advantages and

experience, advanced technology can learn from others in the network environment, the traditional library service model to re-locate and achieve comprehensive, multi-level efficient and convenient, accurate and comprehensive information service.

(a) In an interactive, collaborative, participatory implementation under the premise of the current search page automatically refresh, multi-angle selection problem of subject classification, the expert scheduling practical and reasonable to do a university reference service availability and access to accurate information resistance, followed by further consideration of the students answer each question each expert, each user when in doubt ask both can excel in their own issues with the known answer;

(b) Provides a form of high-quality education consulting, Consultation Form and other online text consultation, but the way the system is not enough rich sync contact, synchronize contact is the future direction of reference the mainstream of development, the power of software development should be strengthened as soon as possible achieve synchronous browsing, video chat and whiteboard features such as synchronous consultation methods.

(c) The formation of reference Union. Conform to the current international trend of development of reference system, the system expands to more colleges and universities in the joint reference system collaboration, expanding the role of reference and influence.

References

Alex Cho, Xue Xin (2005) Intelligent answering system based on web design and implementation [J]. Comput Eng Sci

(U.S.) Deepak Alur, John Crupi, Dan Malks forward, Liu Tin North such as translation (2005) J2EE core patterns [M]. China Machine Press, Beijing

Dengfa Yun (2003) A virtual reference service practice and exploration [J]. Libr Sci 2

Huang Min, Lin Haoming, Yang Zongying (2003) The distributed collaborative virtual reference systems, service oriented Digital Library Conference, Shanghai Jiaotong University, Oct 2003

Larry Netbit, Larry Schankman (2002) Virtual collaboration: the VID experience http://vid.sshe.edu/vid2002-12.htm1

Liao Lili, Sun Yong, Chen Chui (2007) University of expert Virtual Reference System operation mode [J]. Comput Eng

Luo Lili (2002) Library to explore the feasibility of virtual reference services closed. Book Inf 2

Mazhu Jun (2002) Library construction of the virtual reference desk [J]. J Inf 11

Qi Ying Jian, Qi Yinghui et al (1998) Computer supported cooperative work and the system model [A], the first national conference on CSCW. Electronic Industry Press, Beijing

Rui-Min Shen, Liu Yun (1999) Automated answering system based on WWW [J]. Comp Eng 9

Shi Mei, Xiang Yong, Yang Guangxin (2000) Waiting for: "Computer Supported Cooperative Work Theory and Application". Publishing House of Electronics Industry, Beijing

Shiyuan Chun (1998) Distributed computer model of the collaborative work environment [A], set the national papers, Beijing

Wang Yuan, Lai Maosheng (2004) The introduction of online full text library database, a comprehensive evaluation model [J]. Inf Sci

Zhao Feng (2002) Virtual reference (VR) development and application of ways [J]. J Acad Libr 3

Zhu Duo-hui, Wu Guangjun, Zhang Yufeng (2004) AHP-based supplier evaluation model [J]. Comput Appl

Research on Charge Administration System of Machine-Room Based on Network

Hu Ning and Xu Bing

Abstract This paper introduces a kind of general designing plan on charging administration system of machine-room based on network. The system will be realized function of on- line charge administration on Window 2003 platform. This paper puts emphasis on expounding designing idea and designing plan of the system; it sets forth a general designing plan on charging administration system of machine-room based on network, designs structure of this system adopting Visual C++ 6.0 technology so as to accomplish charging administration system based on network. The fact proves design of systematical function is sophisticated with wide applicative prospective.

Keywords Network • Machine-room • Charge • Administration system

1 Introduction

At present, each university opens subjects relevant to computer and encourages students to initially study resorting to computer resources. However, how to administrate computer in machine-room conveniently and which strategy we should use to charge have become a barricade against students' activity (Chen Kangjun and Li Lei 2009). Form the applicative perspective of elevating convenience, simplicity, security of data, for the purpose of solving this problem, this paper lays out a design and realization of charging administration system of public machine-room based on large- scale database. Based on database of all-in-one campus card, it can records usage time users spend on a students computer by card or password, after data processing, judgment and so on to accomplish function of charge (Feng Hogqi et al. 2001) .

H. Ning (✉) • X. Bing
Center of Network and Educational Technology, Chongqing Three Gorges University,
Wanzhou Chongqing, China
e-mail: 15151231@qq.com

2 Design Idea of System

Charge administration system of machine-room based on network satisfies the need of education, which relieves the workload shouldered by administrators in machine-room largely, enhances working efficiency and realizes autonomous charge.

First of all, it must satisfy the need of synchronous transmission of data to charging system of machine-room, when connecting with platform of all-in-one card. That is staff information, group information, role information, card information and etc. can be got form backing database of all-in-one card.

3 Solving Plan of System

Comparing with current charge administration system in machine-room, we can apply this system to working environment of various machine-room managements, adapting to need of different administrative models. Integrating with various working models and choosing different models in accordance with current choose: model of account number of access control, model of solar access control, model of account number, model of reservation.

Model of account number of access control: users who enter machine-room are admitted by access control, meanwhile students who want be on- line must log on client. This way can elevate security of this system in order to avoid any means to escape fee and protect students' profit (Guo Quansheng and Che Yuping 2001).

Model of solar access control: in accordance with on- line environment, we can adopt model of access control. For example, we can adopt this model to simplify on- line identification when students have class.

Model of account number: under the environment of shortage of administrative charge in machine-room, we can adopt software identification in order to decrease investment to access control or can adopt this model when access control works wrong.

Model of reservation: this system is applied to group reserved on- line or computer training, computer exam and so on.

3.1 General Design

Primary functions of system

Server: providing service for client, card terminal procedure, teacher managing procedure in terminal (TCP) 1,024, (TCP) 7,070 AND (UDP) 1,024 (Li Yubo 2003).

Client: check students' and clients' identification and password.

Card terminal: students can enter by card, validity of card is judged.

Teacher managing terminal: it is convenient for teachers or administrators to administrate and supervise machine-room.
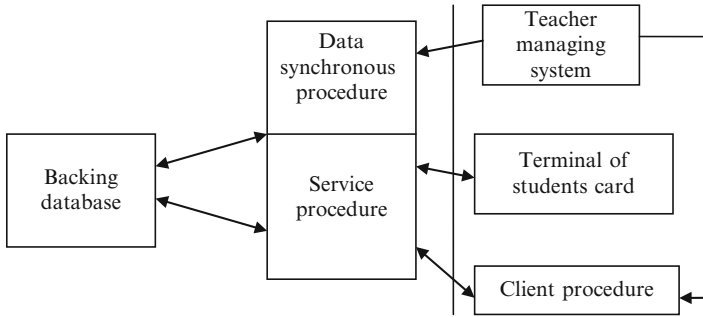
Fig. 1 Integral structure of system
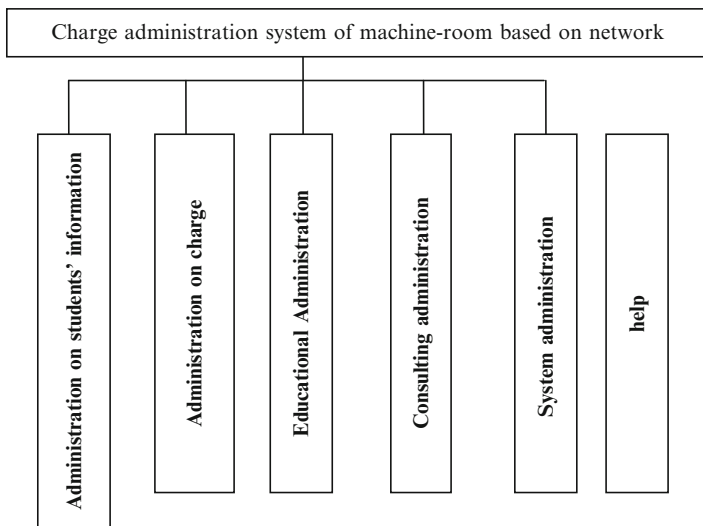
Fig. 2 Main functional module of system

Synchronous tool terminal: it mainly charges data of synchronous transmission of all-in-one card platform and database of machine-room. General Structure of System (Fig. 1)

The Relationship between Functional Demand and Systematical Model

Card terminal is mainly in charge of communicating with server to ensure the validity of someone's with card identification.

Teacher managing terminal firstly connects with server to ensure whether server operates properly, then connects backing database.

Client identifies the rightness of users' account number, password.

Teacher terminal has long- distance control of client machine. Primary Functional Model of System (Fig. 2)

Management on students' information: mainly realizing data import and dealing of students' card.

Charging management: mainly setting students' on- line expense per 1 h.

Educational administration: mainly setting group reservation, anonymous reservation and syllabus reservation.

Module of consulting administration: module of reporting form covering consulting number of card, consulting charge and consult card cost. This module realizes consultation of this system to search basic condition of each card and daily charge and realize criteria query and fuzzy inquiry.

Systematical management of module: systematical management module is in the charge of charging setting, preferential setting, systematical diary, setting of card cost, setting of user and so on. This module is used to set charging standard, administrator and administrators' limited right, providing some functions such as initialization of system and recovery of data backup

Module of help: this model is used to introduce function of system, develop information and so on.

Design of Operation

Combination of module of operation

Service and client: client can be connected with service normally so as to appear normal logging interface.

Service and teacher terminal: teacher terminal can be connected with service so as to appear interface logged by administrator.

Service can be and car terminal: the connected situation of card terminal: successful in connection.

Operation control:

The time of client terminal much match with that of service.

4 Integral Design

4.1 Structure of System

Environment of software: win 2003 + VC 6.0
Environment of background: SQL SERVER 2000
Structure of system (Fig. 3)

4.2 Design of Server

Server mainly provides service for the following modules. Service procedure primarily adopts select I0 model to handle with request of initial service from student server, student card terminal, teacher managing terminal. Server at first sends request I0 accepted to chain table, then handles with I0 request by another routine.
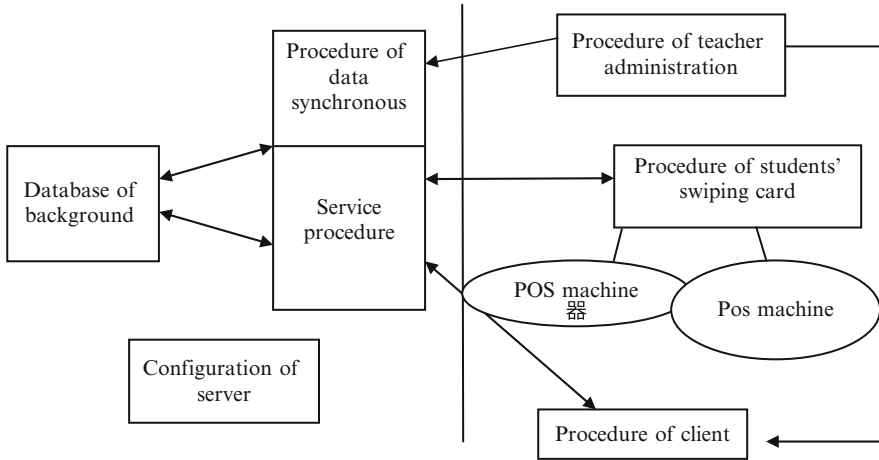
Fig. 3 Figure of system structure

After starting service procedure, it can serve system. After starting system, server will start autonomously and then supervises and waits for I/0 request from client, card terminal and teacher terminal in terminal 1024, 7070.

4.3 Design of Teacher Terminal

Teacher terminal mainly provide a facility for administrators to maintain machine-room and scheme syllabus. The primary functions are following:

Establishing administrator's information:

Adding administrative data information and maintaining administrators' information.

Administration on limits of authority:

Allocating different limits of authority to different levels of administrators so that maintain and manage system better and safer.

Establishing machine-room and information of machine:

In each machine-room, a logical correspondent relationship in database. That means corresponding machine-room in the database to real machine-room. Adding machine to each machine-room, matching with reality. Number of identification of machine is the accordance.

Educational administration: it mainly covers group reservation, anonymous reservation and syllabus reservation. In the terms of anonymous reservation, anonymous and free of charge logging way help instructors to have class and administrator

to maintain machine, what is more, it also guarantees some students and instructors who forget to bring valid identified card (all-in-one card) can be on- line; group reservation aims to facilitate to class and syllabus reservation.

4.4   Design of Card Terminal

The aim of card terminal identifies information of card of students after logging on client. They can enter machine-room to log on computer, once they can be identified as validity (Wei Chengguo et al. 2008). It has following functions: identifying validity of card. Such as invalid card such as lost card, declined card; checking identification by swiping card when students are on- line; when students are off-line, they can close account in card machine.

4.5   Design of Client

It offers interface logged by users. Its function is fix time according to on- line time. Client loggings are divided into four types: common logging, which signifies check identification by swiping card before on-line, after identifying, they input account number and password; after off- line, they swipe card, needing charging. Reservation logging (divided into two kinds), anonymous logging without swiping card and inputting account and password. They just enter system directly after starting, are limited within on-line time; real- name logging need to swipe card so as to check user's identification requiring inputting account password constrained to on- line time. Administrator logs in by his account number and password without swiping card released from time, for purpose of helping administrator to maintain computer, renew system and so on.

4.6   Design of Synchronous Terminal

Synchronous data is located in server to improve efficiency of system. Primary function of synchronous data is transmit data of all-in-one card to database of machine room; data resources of all-in-one card, logging name, password and ORACLE database terminal (Xie Yinzhong and Zhang Baozhou 2010); getting data from database of all-in-one card, after comparing and handling, which can deposited in database of machine room. If it in database of machine room, it will renew condition; if not, it will insert data into database of machine room.

4.7 Design of Main Data Form Is Following

This system mainly designs four tables. The structure of User's for form can be seen in Table 1, the structure of Name of card can be seen in Table 2, the structure of Group form can be seen in Table 3, and the structure of Role form can be seen in Table 4.

Table 1 Users' form: __ user

Name of filed	Type	Width	Illumination
id	bigint	8	Users' name ID
name	varchar	50	Users' name
number	varchar	50	Student number
sex	varchar	50	Gender
birth	varchar	50	
disc	varchar	50	
gid	bigint	8	Group ID
rid	bigint	8	Role ID

Table 2 Name of card: _card

Name of field	Type	Width	Illumination
id	bigint	8	
number	varchar	50	Number of card
flag	int	4	Condition of card 1 valid card not 1 invalid card
status	int	4	
uid	bigint	8	1—userID
balance	float	8	Sum of users' card
login_name	varchar	50	Logging account
pwd	varchar	50	Logging passwaord
free_time	float	8	
deposit	float	8	
cr_time	datetime	8	
deadline	datetime	8	
reserve_1	varchar	50	
reserve_2	varchar	50	

Table 3 Group form: _group

Name of field	Type	Width	Illumination
id	bigint	8	Group ID
lev	int	4	Level of group
id_top	bigint	8	
id_sec	bigint	8	
id_thi	bigint	8	
id_four	bigint	8	
name	varchar	100	Name of current group
tname	varchar	100	Name of the whole group
dis	varchar	100	
reserve_1	varchar	50	
reserve_2	varchar	50	

Table 4 Role form: role

Name of field	Type	Width	Illumination
id	bigint	8	
name	varchar	50	Name of role
dis	varchar	50	

5 Conclusion

Charging system of machine room based on network is a sophisticated system with extension and transferability. Thus interface of system is friendly, administrator is convenient to operate, information output is integral. Taking VC++ 6.0 as explorative tool, this software realizes module of lock screen by calling systematical function. At present, this system has already fulfilled administration and charge of machine room in university, primarily accomplishing open administrative model of machine room, laying foundation for improving utility ratio of public machine room and unmanned administration.

References

Chen Kangjun, Li Lei (2009) Design and realization of charging system of CAN. Railw Norm Univ [J] 30(1):203–205

Feng Hogqi, Pan Hexin, Kong Jian (2001) Charging design and realization based on proxy server. Jiangsu Petrochem Eng Coll [J] 113(1):102–125

Guo Quansheng, Che Yuping (2001) Design of charging administration system of internet information. Commun Technol 26(3):136–137

Li Yubo (2003) Achievable goal and administrative means of open computer lab. Exp Technol Adm 10:56–57

Wei Chengguo, Sun Yaru, Liu Dean (2008) Model of CAN comprehensive charging system [J]. Northwest Univ J Sci Edit 29(1):35–38

Xie Yinzhong, Zhang Baozhou (2010) Applied research of virtual appratus based on VC++platform. Comput Surv Control 18(1):237–239

Research on Informatizational Platform of Tourism in Three Gorges Reservoir Based on Web

Xu Bing, Zeng Jiping, and Hu Ning

Abstract With increasing development of tourist industry, tendency toward informatizational platform of tourism become more and more conspicuous. In view of problems existing in informatization of Three Gorges reservoir tourism and users' demand on informatizational platform of tourist industry, the paper does research on informatizational platform of Three Gorges reservoir tourist industry based on .Net and introduces basic structure and primary function of informatizational platform of Three Gorges reservoir tourist industry and expounds extensively realization of .Net three layer model structure, user's consultation and data administration. It is of significance that enhances practicability and universality of tourist informational system.

Keywords Component • .Net • Three Gorges reservoir • Informatization • Three layer structure

With the development of computer and network, innovational change has taken place on administrative means and logical means of tourist industry (Dai Changying et al. 2009). One important task of informatization of tourist industry is construct informatizational platform of tourist platform. Not only can it enhance labor efficiency, save human resources, but also make administrant work rapid, exact, which is one of way to realize high technology and optimization of administration of tourism and satisfy rapid development of tourism. Three Gorges has long history

X. Bing (✉) • Z. Jiping • H. Ning
School of Computer Science and Engineering, Chongqing Three Gorges University, Wanzhou 404000, China
e-mail: xbbing@cta.cq.cn

and brilliant culture and rich tourist resources, thus informatization of Three Gorges reservoir tourism exerts positive influence on domestic and international tourist industry (Freenam et al. 2007).

1 Current Problems Existing in Informatization of Three Gorges Reservoir Tourism

The Problems of informatization of Three Gorges reservoir tourist industry resulted from history, geography, economy besides some important reasons:

1.1 Corporation Informatization Lacking in Integral Strategic Plan

The management of Tourist Corporation in reservoir maintains weak, system of informatization of tourist management still in initial stage (Huo Hong et al. 2005). Informatization of the majority of tourist corporations falls into dispersion, low-level, and short- sighted implementation of strategy, which is hard to connect with strategy tightly. Account of underestimating significance and strategic status of tourist information resources, the majority of Tourist Corporations, especially the medium and small- sized only purchase short- term economic profit ignoring the great influence of information resources of long- term strategy on development of Tourist Corporation.

1.2 Strategic Inappropriate Construction of Informatization

Due to lack of integral strategic plan, technology corresponding to situation of Tourist Corporation of reservoir needs further research. The development of informatization of tourist industry of reservoir is still restricted with technologic promotion, manufacturer of informational technology's spur and new applied structure sought from technology and system (Liao Guoqiong et al. 2008).

1.3 Delay of Fundamental Work of Exploring Information Resources

Information of tourist resources got by Tourist Corporation in reservoir only stays in surface, which lacks efficient, deep utilization (Meng Xiaofeng et al. 2002). Operating administration of informatization is imperfect, operation of system is

lacking in informational resources. Therefore, it is difficult to transfer informational resources into income and treasury efficiently, which, to some extent, constrains the development of tourist industry in reservoir.

1.4   Informatization of Tourism and Disjunction of Business

Part of tourist corporations regard informatization of Tourist Corporation as showground of information technology and product to decorate network appearance, ignoring huge influence of Tourist Corporation on strategic development. It will result in huge waste and devaluation of tourist information resources and facility and low informatization.

2   Analysis of Demand of Users

In respect of demand of informatization of tourist industry in three gorges reservoir, it divides users into following kinds by comprehensive analysis:

Common users signify the users who are required to get or scan relevant information in system of multi- media. This kind of users only is not required to participate in renewal of system and maintenance of data and has no right to modify limits of authority. Therefore, the primary aim of common users is consult data resources and get more information about media tour.

Registered users (including administrator of tourist spot, travel agency, hotel, transportation department, travel department) who has wider limits of authority and enjoy more services than common users. In order to induce more users to administrate and maintain multi- media data-base, registered users are deployed more limits of authority, who can distribute and modify relevant information.

Senior users (administrators of system) who administrates registered users' information, renews the data of data-base, optimize the whole data of system, and legally verify the files uploaded by users. Senior users can change content of database and information of registered users according to their own requirements, in addition, cancel illegal users and restrict other users' authority.

Administration of users' data covering following functions:

(a) Addition and adjustment newly data in time
(b) Cancellation out- dated multi- media data.
(c) Modification wrong multi- media data and division the information uploaded by users more exactly.
(d) Change more function according to demand.

3 Structure of Informatization of Three Gorges Reservoir Tourism

3.1 Net Three Layer Structure

B/S model (browser/server) is the result of development of c/s model, based on Web technology, dividing traditional servers of C/S model into data-base server and Web servers.

B/S structure is a typical distributive information system on Intranet. The advantage of B/S model is (1) user's operation is simple, user side of B/S model only need to be installed by a common- used browser providing user friendly interface, such as internet explorer instead of extra users' training. (2) B/S structure is flexibility. Permitted by network environment, browser and servers are increased against influence of network resources. (3) Due to easy exploration, installation and maintenance B/S structure only need to explore and integrate the applied platform of servers in order to eliminate exploration, installation and expense of maintenance.

In terms of current technology, it is relatively easy and cost lower price to apply network by means of LAN of B/S structure and apply data-base by means of Internet/intranet model. It is a kind of one-time exploration, which can visit and operate common data-base from different places, different personnel and different inserted mode, such as LAN, WAN, Internet. It can protect data platform and administration of limits of authority and guarantee data-base of server (Muller 2005).

This system mainly adopts B/S system model. The whole system is divided into three levels, which are presentation layer, business logical layer, data-base level (Fig. 1). We use a server to realize all functions or use Web server and Data-base server, which can relief load of server and elevate efficiency of implementation.

3.2 Module of .Net of Three- Layer Is Constituted of Presentation Layer, Business Logical Layer and Data-Base Visit Layer (Fig. 2)

Business logical layer is the core part of system, undertaking procession task of business logical including HTTP requirement of procession interface layer, offering function of business to presentation layer and visiting data-base by data-base layer. The specific procession is analytically transfer HTTP requirement sent by receiving presentation layer of Web server and invokes corresponding logical procession; this logical procession communicates with data-base layer, visit data-base, and returns result to Web server, finally sends the result to browser of user side in the form of XML OR HTML by Web sender (Xu Bing et al. 2007).
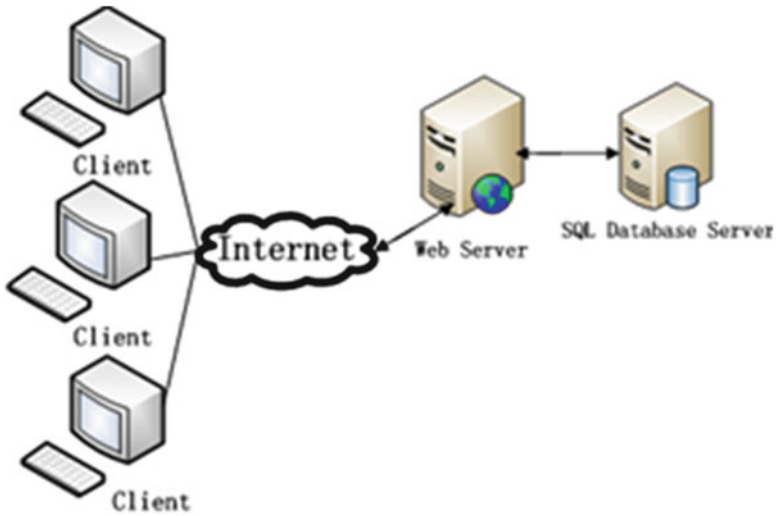
Fig. 1 Hardware structure of platform
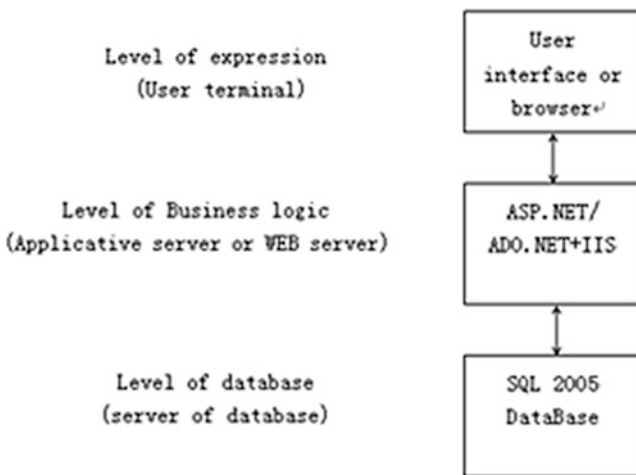
Fig. 2 Three-level structure model of platform

Data-base layer at low layer is mainly used to realize communication with data-base, consult, insert, modify and cancel data of data-base. ADO and .NET are used for interface of data-base and SQL 2005 data-base is used for structure, which handle requirement of data. This system is based on Windows 2003 Server+IIS+SQL 2005 + .NET structure, adopting ASP.Net structural procession. Data-base adopts SQL 2005 as well as ADO .NET as exploring technology of data-base interface.

4 Primary Functions of Informatizational Platform of Tourist Industry in Three Gorges Reservoir

In light of users' demand on informatization of tourism and function on administration of system, we get main functions of module of informatizational platform of tourism in three gorges reservoir (Fig. 3).

(a) The function of users' consultation
    Travelers can consult the information about scenic spot (including word, picture and video), transportation information (train, plane, car and so on), travel agency (introduction to travel agency, routine, address and phone number), information about hotel (introduction to hotel, rank, bed, telephone and so on).
(b) Administration of travel policy
    Administrating current travel policy of state, such as addition, modification, cancellation, distribution and so on
(c) Informational administrative functions of travel spot
    Record of scenic spot resources including number of spot, name of spot, contact person, picture, video, fare and so on; modification of spot resources; cancellation of spot resources; consultation on spot resources.
(d) Informational administration of travel agency
    Record of travel agency resources including number of scenic spot, name of travel agency, contact person, picture and some other information; modification of travel agency information; cancellation of travel agency; consultation on travel agency
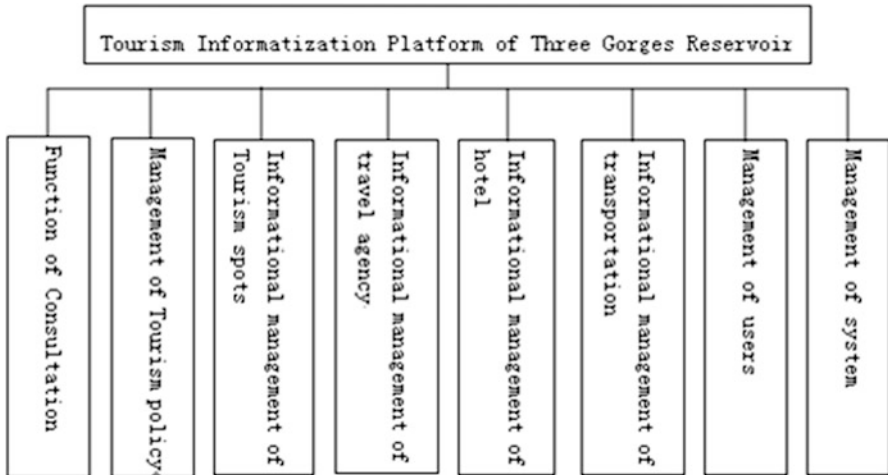
Fig. 3 Main functional module of tourism informatization platform of Three Gorges reservoir

(e) Administration of hotel information
   Record of hotel resources including hotel number, name of hotel, contact person, quotation and some other information; modification of hotel information; cancellation of hotel information; consultation on hotel information
(f) Administrative functions of transportation resources
   Administration of train information; administration of plane information; administration of other transportation resources
(g) Function of administration of travel routine
   Record of travel routine information covering routine number, name of routine, quotation and other information; Modification of travel routine information; cancellation of travel routine information; consultation on travel routine information
(h) Administration of user
   Record of user information including user's name, password, limits of authority and other information; modification of user's information; cancellation of user' information; consultation of user's information and other functions
(i) Administration of system
   Initialization of system, backup of data-base, restoration and other functions

5   Conclusion

By use of .Net technology, exploring informatizational platform of three gorges reservoir tourism based on browser/server (B/S) mode is the inevitable tendency toward exploration and administration of future tourist industry, which has bright prospect and has deeply influence on traditional mode of tourism. Applying system of mode of server/browser (B/S) rationalizes systematic structure; the function of user side transferred to server, which means user sided need only one browser. In this case, installation, allocation and operation of user side are simpler. User side can ensure stably user's requirement can be satisfied. Adoption of Asp .Net technology makes applying procession more convenient, faster, and higher efficient. The establishment and application of informatizational platform of Three Gorges reservoir tourism have become the innovational way to develop tourist industry in reservoir and elevated competition in international and domestic tourism, which worth extending.

References

Dai Changying, Zhang Guangzhi. Model of trusting assessment of web service. Comput Proj 200(09)
Freenam A, Jones A (2007) .NET secure procedure [M]. Tsinghua University Press, Beijing

Huo Hong, Xu Baowen (2005) Latent hazard of web information system and countermeasure of network administrator [J]. Comput Eng Appl, Beijing 1:151–153

Liao Guoqiong, Liu Yunsheng (2008) Real time insertional data-based recovery strategy based on real time diary [J]. Comput Eng Appl, Beijing 2:180–182

Meng Xiaofeng, Lu Hongjun, Wang Haiyan et al (2002) SG-WRAP: aschema-guided wrapper generator [C]. In: Proceedings of 18th international conference on data engineering, Beijing, pp 331–332

Muller C (2005) Mechanisms of knowledge evolution for web information extraction. Comput Sci, Chongqing 3847:202–214

Xu Bing, Yang Xi, LiaoYoucheng (2007) Systematic structure of multi-media data-base of three Gorges reservoir tourist industry. Comput Knowl Technol, Anhui 3:1489–1491

Part III
Theory and Practice of Modern Management

A Study of Model for Optimizing Supply Chain Inventory with Controllable Lead Time

Yao Tang and Bin Yang

Abstract This paper has put forward the importance of shortening lead time in the competition of modern society and researched the decision-making problem in a supply chain system consists of one supplier and one retailer. Basing on the hypothesis that the decision-making power of determining the lead time is controlled by the downstream, which completely undertakes the expenses for shortening the lead time, this paper has established a Stackelberg model of this supply chain system, and given out some corresponding algorithms for reference. Finally according to the analysis of cases, this paper has confirmed the practicality and effectiveness of the model, and proved that shortening the lead time reasonably can reduce the supply chain inventory cost effectively.

Keywords Lead time • Supply chain • Inventory cost

1 Introduction

At the end of the 1980s, academia had put forward a new competition pattern which was called "time-based competition". As a powerful weapon of "time-based competition", lead time was also paid more attention by people. Shortening lead time can effectively reduce the inventory and funds backlog of the enterprise, enhance the service level and competitive ability. Basing on inventory continuous inspection strategy, Liao and other (1991) have proposed controllable lead time inventory model, in which lead time is the unique decision variable. Subsequently, Ben and other (1994) have seen order quantities as another decision variable and expanded model, for considering the joint decision making of both of lead time and

Y. Tang (✉) • B. Yang
Logistic Research Center, Shanghai Maritime University, Shanghai 200135, China
e-mail: tangyaocool@163.com

order quantities. Basing on previous research, this paper has added a new decision variable—order point, in order to analysis the joint decision making of three of lead time, order quantities and order point. Furthermore, being familiar with the reference Liao and Shyu (1991) and Ben Daya and Raouf (1994), this paper has also assumed that lead time decisions is controlled by the downstream, which bears completely the expenses for shortening lead time.

2   Cost Components Analyses

2.1   *Supply Chain System Components*

There is a single-product and two-levels supply chain system, which is constituted by one supplier and retailer. In this system, suppose that lead time consist of three independent operating periods, and each operating period has both a minimum duration time a_i and a maximum duration time b_i. If $L_0 = \sum_{j=1}^{n} b_j$ represents the retailer's the longest lead time, and L_i represents the length of lead time, the component 1 to component i of which is under the adequate time to catch up with the schedule, so there is $L_i = \sum_{j=1}^{i} a_j + \sum_{j=i+1}^{n} b_j = \sum_{j=1}^{n} b_j - \sum_{j=1}^{i} (b_j - a_j) = L_0 - \sum_{j=1}^{i} (b_j - a_j)$. And the lead time can be shorten by increasing a certain fee. The fee which is produced by the i-th part shortening a unit of time is c_i, in non-descending order, that is $c_1 \leq c_2 \leq \cdots \leq c_n$. If $R(L)$ represents the catch-up schedule costs which are generated by shortening lead time for L in one order, there is $R(L) = c_i(L_{i-1} - L) + \sum_{j=1}^{i-1} c_j(b_j - a_j), L \in [L_i, L_{i-1}]$ (Li Hua et al. 2004).

2.2   *Retailer Inventory Costs*

Retailer annual expectation costs consist of three parts: annual order costs, annual inventory costs and annual catch-up schedule costs.

1. Annual order costs. The retailer's cost per order is A. Annual demand is D. The amount of each order is Q. Therefore the annual order costs is $\frac{DA}{Q}$.
2. Annual inventory costs. The annual inventory cost of the unit of the retailer's goods is h_r. The annual average inventory level is $\frac{Q}{2} + k\delta\sqrt{L}$. So the annual inventory costs is $h_r(\frac{Q}{2} + k\delta\sqrt{L})$.

3. Annual catch-up schedule costs. Known lead time $L \in [L_i, L_{i-1}]$, the catch-up schedule costs of each order is $R(L) = c_i(L_{i-1}-L) + \sum_{j=1}^{i-1} c_j(b_j - a_j)$. Because this paper studies the inventory optimization problem that the retailer bears the annual catch-up schedule costs completely, so the retailer's annual catch-up schedule costs is $\frac{DR(L)}{Q}$.

So the retailer's total annual inventory costs can be shown as follows:

$$C_r(Q, L) = \frac{DA}{Q} + h_r\left(\frac{Q}{2} + k\delta\sqrt{L}\right) + \frac{DR(L)}{Q} \tag{1}$$

2.3 Supplier Inventory Costs

Supplier inventory costs consist of two parts: stocking costs and inventory holding costs.

1. Stocking costs. Each stocking cost of the supplier is S, Annual demand is D. Each product capacity is mQ, so the annual stocking costs is $\frac{DS}{mQ}$.
2. Inventory holding costs. Annual holding costs per unit of commodity is h_s, annual average inventory is $\frac{Q}{2}[m(1 - \frac{D}{P}) - 1 + \frac{2D}{P}]$, so the inventory holding costs is $\frac{Q}{2}[m(1 - \frac{D}{P}) - 1 + \frac{2D}{P}]h_s$.

So the supplier's total inventory costs can also be shown as follows (Ouyang et al. 2004):

$$C_s(Q, m) = \frac{DS}{mQ} + \frac{Q}{2}\left[m(1 - \frac{D}{P}) - 1 + \frac{2D}{P}\right]h_s \tag{2}$$

3 Modeling

The retailers need to determine the most economical unit Q, the best lead time L, and the suppliers need to determine the optimal delivery batch m. This problem can be attributed to a class of Stackelberg leader-follower game problem, which is, the retailers seek most economical unit and the best lead time when the inventory costs is minimized, meanwhile the suppliers seek the optimal delivery batch through predicting the possible choices could be made by the retailers. Therefore the leader-follower game problem modeling composed by the suppliers and the retailers is as follows (Jiang Liangkui and Wang Chuanxu 2008):

$$\min C_s(Q, m) = \frac{DS}{mQ} + \frac{Q}{2}\left[m(1 - \frac{D}{P}) - 1 + \frac{2D}{P}\right]h_s$$

$$s.t. \quad C_r(Q, L) = \frac{DA}{Q} + h_r\left(\frac{Q}{2} + k\delta\sqrt{L}\right) + \frac{DR(L)}{Q} \tag{3}$$

4 Model Solving Steps

Seek the second-order partial derivatives of Q and L in the Eq. (3) separately we can obtain that:

$$\frac{\partial C_r^2(Q, L)}{\partial Q^2} = \frac{2D}{Q^3}[A + R(L)] > 0 \tag{4}$$

$$\frac{\partial C_r^2(Q, L)}{\partial L^2} = -\frac{1}{4}k\delta h_r L^{-\frac{3}{2}} < 0 \tag{5}$$

From Eqs. (4) and (5) we can know that:

(a) For any given economical order batch Q, $C_r(Q, L)$ is a concave function of the lead time L.
(b) For any given lead time $L \in [L_i, L_{i-1}]$, $C_r(Q, L)$ is a convex function of the economical order batch Q. So the best lead time L inevitably occurs in endpoints of the range.

Again we seek the partial derivative of Q in the Eq. (3) and obtain that:

$$\frac{\partial C_r(Q, L)}{\partial Q} = -\frac{D}{Q^2}[A + R(L)] + \frac{h_r}{2} \tag{6}$$

Make the Eq. (6) equal to zero then the most economical unit Q can be obtained:

$$Q^* = \sqrt{\frac{2D}{h_r}[A + R(L)]} \tag{7}$$

For the suppliers, they seek the optimal delivery batch through predicting the most economical unit Q^*, the best lead time L^* of the retailers. As m is a positive integer, we can know from the Eq. (2), when $m = m^*$ is the optimal solution, it inevitably satisfies these following inequalities simultaneously:

$$\begin{cases} C_s(Q^*, m^*) \le C_s(Q^*, m^* + 1) \tag{8} \\ C_s(Q^*, m^*) \le C_s(Q^*, m^* - 1) \tag{9} \end{cases}$$

Follow the Comprehensive analysis above; the optimal solution algorithm can be given (Chang et al. 2006):

1. For each lead time $L_i (i = 0, 1, \cdots, n)$, Q_i can be figured out with Eq. (7);
2. For each set of (L_i, Q_i), we can figure out the corresponding $C_r(L_i, Q_i)$ when $i = 0, 1, \cdots, n$;
3. Make $C_r(Q^*, L^*) = \min_{i=0,1,\cdots,n} C_r(L_i, Q_i)$, and (Q^*, L^*) would be the optimal solution;

4. Put (Q^*, L^*) into the Eqs. (8) and (9) to figure out the value range of m, assume that there are k possible solutions in the value range, and the number i possible solution is m_i;
5. Put m_i into the Eq. (2), and make $C_s(Q^*, m^*) = \min\limits_{i=0,1,\cdots,k} C_s(Q_i, m_i)$, and m^* would be the optimal solution, under this optimal solution the total inventory costs of the supplier would be $C_s(Q^*, m^*)$.

5 Analysis of Cases

Assume that the basic parameters of the supply chain system are: $D = 500$, $A = 180$, $h_r = 18$, $k = 1.65$, $\delta = 6$, $S = 1,600$, $P = 2,000$, $h_s = 16$. The lead time consists of four parts and the specific values are shown in the Table 1 (Liu Lei and Tang Xiao-wo 2005):

According to the optimal solution algorithm in the model solving steps the following table can be given:

Known from the Table 2, when the retailer's lead time $L^* = 12$ weeks and the economical unit $Q^* = 102$, the annual total inventory costs can be minimized to $C_r(12, 102) = 2447.76$ Yuan. Put this optimal solution (L^*, Q^*) into the Eqs. (8) and (9), then obtain the value range of the optimal delivery batch $12.6 \le m^* \le 13.6$, which means when $m^* = 13$. When the optimal solution is $m^* = 13$, the annual total inventory costs would be $C_s(102, 13) = 8151.32$ Yuan. Still the annual total inventory cost of the whole supply chain system is $C_{sum}(12, 102, 13) = 10598.38$ Yuan.

Table 1 The operation constitutions in the lead time (unit: week)

Constitutions in the lead time (i)	Normal work time b_i	Adequate catch-up schedule time a_i	Catch-up schedule costs per unit time c_i
1	3	1	3
2	3	1	8
3	6	3	18
4	2	1	36

Table 2 Retailers' inventory costs under different lead time

i	L_i (week)	$R(L_i)$	Q_i	$C_r(L_i, Q_i)$
0	14	0	100	2,466.76
1	**12**	**6**	**102**	**2,447.06**
2	10	22	106	2,470.35
3	7	76	119	2,618.10
4	6	112	127	2,729.11

6   Conclusion

The decision-making problem of the lead time and the order quantity in a supply chain system consists of one supplier and one retailer has been researched in this article. As a delivery strategy under consideration, the retailer was considered to determine the lead time and fully undertake the expenses for shortening the lead time. The best lead time under which the total inventory costs of the supply chain system is minimized, the corresponding most economical unit and optimal delivery batch had been obtained through modeling analysis. According to the specific analysis of cases the practicality and effectiveness of the model has been proved. This model can also be further extended to the decision-making problem in a supply chain system consists of one supplier and two retailers.

References

Ben Daya AM, Raouf A (1994) Inventory models involving lead-time as decision variable. J Oper Res Soc 45(5):579–582

Chang H-C, Ouyang L-Y, Wu K-S (2006) Integrated vendor-buyer cooperative inventory models with controllable lead time and ordering cost reduction. Eur J Oper Res 170(2):48–495

Jiang L, Wang C (2008) Supply chain decision-making considering delivery time crashing cost sharing. J Shanghai Marit Univ 29(2):48–51

Li Hua, Li Yi-qiang, Xu Guo-hua (2004) Research on lead time model in supply chain distribution. J Ind Eng Manage 18(3):112–114

Liao CJ, Shyu CH (1991) An analytical determination of lead time with normal demand. Int J Oper Prod Manage 11(9):72–78

Liu Lei, Tang Xiao-wo (2005) A study on venders' lead time decision model. Syst Eng 23(8):42–45

Ouyang L-Y, Wu K-S, Ho C-H (2004) Integrated vendor-buyer cooperative models with stochastic demand in controllable lead time. Int J Prod Econ 92:255–266

Human Resources Management Based on the Teacher's Empowering

Belferik Manullang, Sri Milfayetty, and M.S. Kons

Abstract The objective of this research were to describe and to see the correlation between normative commitment, spiritual intelligence, pedagogical attitudes, essential scientific, and ability competence. Survey used as a method in this research. Data collected by questionaire.Thirty teachers among 60 teachers got randomly as a sample. Data obtained were descriptive and correlation analysis. The results of the research showed that the condition of the teachers were relatively below average for each variable. Data aquisiton showed that there were significant correlations among the variables. It suggested that the teachers need to recharge, educate and train, due to empower them. The school need to maintenance good atmosphere such as strong leadership, as well as the school policies to support the empowering program.

Keywords Normative commitment • Spiritual intelligence • Pedagogical attitudes • Essential scientific and ability competencies

1 Introduction

The essence of education is developing human qualities, that we say as human character. Education is called humanize also, because humans have had the potential of humanity. Plato said that if you ask what is good education, in general, the answer is easy, education makes good men, and that good men act nobly. The end of education is character. Harrel (2004) said that "In education, character or attitude is everything". Your character or attitude to day, determine your succes tomorrow.

Teachers' normative commitment means the loyalty of teachers to carry out their role as educators in developing student's character. As educator, teachers required

B. Manullang (✉) • S. Milfayetty • M.S. Kons
Graduate School, State University of Medan, Medan, North Sumatera, Indonesia
e-mail: gubes_bm@yahoo.co.id; milfayetty@yahoo.com

to have a good character. Teachers need to be a model how to behave base on pedagogical value due to create good school atmosphere. Learning in the classroom and outside are expected to develop students' character, as a person, as a member of the community and as citizens.

Teacher has strong position in developing students character. They need educational tools to transform pedagogical value to be character of the students. The tools are, such as a sense of affection toward students, sincerity in the line of duty, modeling, reward and educational punisment. Life is preserved as a model teacher, nature of the firm but to educate as well as other pedagogical properties.

There were some indications showed that some teachers did not meet the pedagogical standards. It caused the transformation of pedagogical value uneffectively. As a result, some students were in a bad manner such as cheating, buliying, fihgthing, etc. In these situations, teachers need to develop their normative commitment, due to transform the pedagogical value while interact with the student.

Normative commitment of teachers is the level of duty and loyalty to the institution on the basic of considerations of conscience. According to Colquitt (2009) and Slocum and Hellriegel (2009), normative commitment of teachers, could be improve by supported through ability competencies, scientific mindset essential, pedagogical attitudes and spiritual intelligent. Ability competencies are knowledge and skills that become a part of their life. The teachers felt a duty as a teacher can not be separated from their personality. They have to aware to their role as a teacher. They have to be satisfied, happy and dignity as a teacher. Competencies are effective to improve teachers normative commitment. Essential scientific is the mindset that prioritizes the essence of truth on which base on action. There are scientific technological mindset, theoretical or scientific mindset and essential scientific. To illustrate the difference of these mindset, we can interpret the meaning of teacher as a profession. The technological scientific mindset tend to think how to make a lot of money from the work as a teacher. From the side of theoretical scientific mindset, teachers tend to think how to work theorethically. If the teacher has reached the level of essential scientific, it will effective to support normative commitment and ability competencies. Pedagogical attitudes of teachers are the properties that contain values such as education, affection, sincerity, patience, firmness, and motivation. If these properties have been owned by the teacher, it would be effective to improve their normative commitment and their ability competencies.

2   The Problem Formulation

(a) How did the prospensity of spiritual intelligence, pedagogical attitude, essential scientific, ability competencies, and normative commitment?
(b) Were there significant coorelation between spiritual intelligence and habitual pedagogical, spiritual intelligence and essential scientific, spiritual intelligence and ability competencies, habitual pedagogical and ability competencies, essential scientific and ability competencies, ability competencies and normative commitment?

Fig. 1 The model of
normative commitment.
X1: Spiritual intelligence,
X2: Pedagogical attitudes,
X3: Essential scientific,
X4: Ability competencies,
X5: Normative commitment
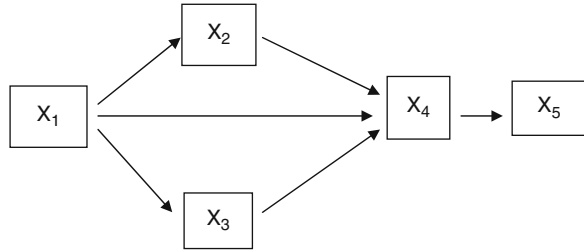
3 The Normative Commitment Model

Normative commitment of teachers is the level of duty and loyalty of the teacher to the institution on the basic of considerations of conscience. Intelligent Sipiritual refers to the properties of noble and humanitarian values. The spiritual intelligence is the intelligence that use to solve problems of meaning and values. This intelligence will position our lives meaning in the context of a broader and richer.

Slocum (2004) said that competencies are interrelated cluster of knowledge, skills and ability that needs by individual to be effective. Competencies of teachers are the ability to perform professional duties as a teacher. Mindset is the essential truth of scientific theories prioritizes the fundamental and through reasoning, with pure thought, the main character suprarasional. Habit is daily pedagogical teacher while interact with students. Behaviors are expressed as an educational tool, patience, dignity, compassion, exemplary, motivation, which educates firmness, sincerity (Fig. 1).

Normative commitment, spiritual intelligence, pedagogical attitudes, essential scientific and ability competencies are aspects of personalities. Commitment consists of continuance commitment, affective commitment and normative commitment. The spiritual intelligent consists of intellectual intelligent, emotional intelligence and spiritual intelligence. Pedagogical attitude consist of thought, action and habits. Scientific mindset consist of practical scientific, theoretical scientific and essential scientific. While the competencies consist of knowledge competencies and skills competencies and ability competencies. The best character in commitment is normative commitment. The best character in inteligence is spiritual inteligence and the best character in attitude is pedagogical attitudes. The best character in the scientific mindset is an essential scientific, and the best character in competencies are ability competencies.

4 The Objectives of Research

The objectives of this research were,

(a) To analysis the level of prospensity of normative commitment, spiritual intelligence, pedagogical attitudes, essential scientific and abilties competencies

(b) To analysis the corelation between spiritual intelligence and habitual pedagogical, spiritual intelligence and essential scientific, spiritual intelligence and ability competencies, habitual pedagogical and ability competencies, essential scientific and ability competencies, ability competencies and normative commitment.

5 The Method of Research

The research was conducted at Junior High School number 17 in Medan. The population were 60 teachers and got 30 teachers randomly as a sample. Data colleced by questionnaires to all five variables: normative commitment, essential scientific, pedagogical attitudes and spiritual intelligence and ability commitment.

Data analysis was performed by using descriptive statistics.

(a) Description of normative commitment, spiritual intelligence, pedagogical attitudes, essential scientific and abilties competencies
(b) Testing of the coorelation between spiritual intelligence and habitual pedagogical, spiritual intelligence and essential scientific, spiritual intelligence and ability competencies, habitual pedagogical and ability competencies, essential scientific and ability competencies, ability competencies and normative commitment.

6 Analysis and Discussion

6.1 The Results of Research

The description of the data for five variables in this research presented as mean, median, mode and frequency distribution histograms of each variable perform in the table below (X1) = Spiritual intelligence (X2) = pedagogical attitudes, (X3) = essential scientific, (X4) = ability competencies and X5 = normative commitment (Tables 1 and 2, Fig. 2)

Hypothesis 1
Ho: Spiritual intelligence has no significant correlation with pedagogical attitudes
H1: Spiritual intelligence has significant correlation with pedagogical attitudes

Testing criteria is Ho be accepted, H1 rejected, if th ≤ tt. Ho rejected, H1 be accepted if th > tt. Apparently th > tt, that is 3,68 > 1,70. Ho is rejected, H1 accepted. It meant that there was significant relationship between Spiritual intelligence and pedagogical attitudes

Table 1 Descriptive analysed

Analysis	Data of variable				
	X_5	X_1	X_2	X_3	X_4
Empirical data					
Number of data	30	30	30	30	30
minimum	12	12	11	11	11
Maximum	25	20	22	20	20
Range	13	8	11	9	9
Number class interval	7	5	6	5	5
Range of class interval	2	2	2	2	2
Mean	16, 93	16, 10	16, 53	15, 77	15, 43
Deviation standard	2, 82	2, 43	2, 78	2, 85	2, 94
Median	16	15, 5	15	15, 5	15, 5
Modus	16	15	16	15	15
Theoretical data					
Minimum	10	10	10	10	10
Maximum	50	50	50	50	50
Mean	30	30	30	30	30
Deviation standard	1, 67	1, 67	1, 67	1, 67	30

X1: Spiritual intelligence
X2: Pedagogical attitudes
X3: Essential scientific
X4: Competencies ability
X5: Normative commitment

Table 2 The level of percentage

No	Variable	Percentage
01	Normative commitment	67%, below of the mean
02	Sipiritual intelligence	74%, below of the mean
03	Pedagogical attitudes	57%, below of the mean
04	Essential scientific	60%, below of the mean
05	Ability competencies	67%, below of the mean

b. Correlation analysis

Var	r	T_h	t_t	Summary
r_{21}	0,570	3.68	1,70	Significant
r_{31}	0,430	2,53	1,70	Significant
r_{41}	0,540	3,40	1,70	Significant
r_{42}	0,338	1,90	1,70	Significant
r_{43}	0,493	3,00	1,70	Significant
r_{54}	0,382	2,12	1,70	Significant

Hypothesis 2
Ho: Spiritual intelligence has no significant correlation with essential scientific
H1: Spiritual intelligence has significant correlation with essential scientific

Testing criteria is Ho be accepted, H1 rejected, if th ≤ tt. Ho rejected, H1 be accepted if th > tt. Apparently th > tt, that is 2,53 > 1,70. Ho is rejected, H1
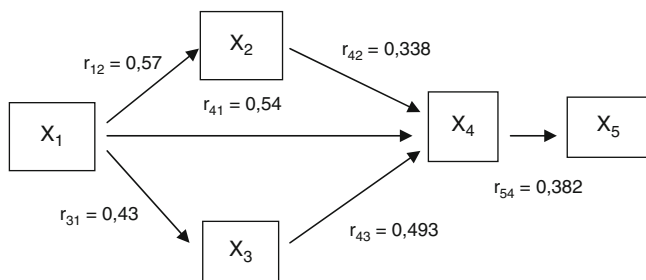
Fig. 2 The Model of Normative Commitment with variable. X1: Spiritual intelligence, X2: Pedagogical attitudes, X3: Essential scientific, X4: Ability competencies, X5: Normative commitment

accepted. These meant that there was significant relationship between Spiritual intelligence and essential scientific

Hypothesis 3
Ho: Spiritual intelligence has no significant correlation with ability competencies
H1: Spiritual intelligence has significant correlation with ability competencies

Testing criteria is Ho be accepted, H1 rejected, if th ≤ tt. Ho rejected, H1 be accepted if th > tt. Apparently th > tt, that is 3,40 > 1,70. Ho was rejected, H1 accepted. These meant there was significant relationship between spiritual intelligence with ability competencies

Hypothesis 4
Ho: Pedagogical attitudes has no significant correlation with ability competencies
H1: Pedagogical attitudes has significant correlation with ability competencies

Testing criteria is Ho be accepted, H1 rejected, if th ≤ tt. Ho rejected, H1 be accepted if th > tt. Apparently th > tt, that is 1,90 > 1,70. Ho is rejected, H1 accepted. These meant there was significant relationship between pedagogical attitudes and ability competencies

Hypothesis 5
Ho: Essential scientific has no significant correlation with ability competencies
H1: Essential scientific has significant correlation with ability competencies

Testing criteria is Ho be accepted, H1 rejected, if th ≤ tt. Ho rejected, H1 be accepted if th > tt. Apparently th > tt, that is 3,00 > 1,70. Ho is rejected, H1 accepted. These meant there is a significant relationship between essential scientific and ability competencies

Hypothesis 6
Ho: Ability competencies has no significant correlation with normative commitment
H1: Ability competencies has significant correlation with normative commitment

Testing criteria is Ho be accepted, H1 rejected, if th ≤ tt. Ho rejected, H1 be accepted if th > tt. Apparently th > tt, that is 2,12 > 1,70. Ho is rejected, H1 accepted. These meant there was a significant relationship between ability competencies and normative commitment.

7 Discussion

7.1 Descriptive Discussion

Normative commitment, spiritual intelligence, pedagogical attitudes, essential scientific and ability competencies are the traits of teachers as educator. This research found that most of the teachers condition in these five variables were still below the mean

1. Normative Commitment
   Normative commitment showed loyalty of the teacher to school culture, attachment to working system, optimism for the future through the school, pride in the institution of employment, a sense of comfort at school, feeling at home at school, positive perceptions about school places to work, feeling togetherness, mutual respect and a deep sense of mutual aid in the life at school.

   Data analysis showed that teacher's normative commitment below average, there were 67% and an above average were 33%. Data showed that more than 50% of teachers have not been convinced the work suited as a teachers. This condition showed that commitment to become a teacher still needs to reinforce. In general the perform of the teacher more focused on personal or family needs, if compared with their fidelity to make school as the foundation of their dedication. Working as a teacher for self more interest in school.

2. Spiritual Intelligence
   Spiritual intelligence conscience means participation in performing their duties as teachers. Professionalism is realized to a greater interest or caring. The teachers with spiritual intelligence means seeing his duty to care the children, responsibility to society or nation, not merely to satisfy his own. The research found that spiritual intelligence of the teachers 74% below average. This meant that the spiritual intelligence of the teachers should be improved in a good condition. This spiritual intelligent should improved and maintained continuously in good condition, so that, could fix the meaning of teacher professionalism. The teacher should be ready to take on the task of preparing future generations, as a form of universal nature and human values that exist within them.

3. Pedagogical attitudes
   Habits behavior is teacher pedagogical properties such as educating, patient, compassion, sincerity, reinforcement, educating and modeling. The research found that 57% teachers had pedagogical attitudes. These condition indicated that education has not been successfully constructed with pedagogical attitudes,

as a conditio sine quanon for good educational process. Instead, all teachers in a deliberate and conscious effort to maintain pedagogical attitudes because only in the conditions such that, education process can be carried out effectively.

4. Essential Scientific
   Mindset is the essential truth of priority to basic and comprehensive theory of reasoning, with pure thought, the main properties suprarasional. The research showed that 60% was below average which meant working with pragmatic ways of thinking. One example of pragmatic thinking was to perceive the work of teachers revenued. Teachers should made their work as their devotion such as scientific essential mindset.

5. Ability Competencies
   The best competencies of teacher is ability competencies. On this understanding of teaching, competencies means an arts, that should be a part of teachers' life. In this condition, teacher would get satisfaction, happiness and dignity. The research found that teachers' competencies ability about 67% below average. This condition indicates that the majority of teachers performing still not good.

7.2 Correlation Discussion

If the spiritual Intelligence had got better, then the pedagogical attitudes should had better. If the spiritual Intelligence had got better, then the essential scientific should got better. If the spiritual Intelligence had got better, then the ability competencies should got better. If the habitual pedagogical had got better, then the ability competencies got better. If the essential scientific had got better better, then the ability competencies should got better. If the ability competencies had got better, then the normative commitment should got better. For that, we need to improve spiritual Intelligence, pedagogical attitudes, essential scientific and ability competencies.

8 Conclusion and Suggestions

8.1 Conclusion

1. The results showed that the majority of teachers normative commitment were below average, spiritual intelligence teachers below average, pedagogical attitudes remained largely below average, essential scientific remained largely below average, the ability competencies remained largely below average.

2. The spiritual intelligence with habitual pedagogical, spiritual intelligence with essential scientific, spiritual intelligence with ability competencies, habitual pedagogical with ability competencies, essential scientific with ability competencies and ability competencies with normative commitment had significant correlation.

8.2 Suggestions

1. The schools leadership were expected to improve spiritually intelligent, pedagogical attitudes, scientific mindset essential, ability commitment due to increse normative commitment of the teachers.
2. The school policy should considered the direction of the teacher development in spiritual intelligence, pedagogical attitudes, scientific mindset essential, ability commitment. It needs supervisory ini everyday school life. The supervisor should observe, analyze and help teacher to improve their performance. This method is very important because the formation of character as educators must be committed intentionally, continuously and sustainable.

References

Colquitt JA (2009) Organizational behavior, improving performance and commitment in the workplace. McGraw-Hill, New York

Harrel K (2004) Attitude is everything. Collins Business, New York

Slocum JW, Hellriegel D (2009) Principles of organizational behavior. Cengage Learning, New Zealand

Slocum M (2004) Innovation in performance excellence: eight paradigms to performance excellence (8PPE). In: 4th, TRIZ future conference, ETRIA, Italy

The Evaluation Based Study on the Innovation Strategy of Electric Vehicles

Maria Ikram and Yan ZhiJun

Abstract Energy saving and environment protection is the major issues in China which speeds up the electric vehicle industrialization sector. The proposed study is a survey based evaluation study. Taking Beijing as a centre, the paper proposes the evaluation of EVs from the following several aspects: performance of the electric vehicles from different aspects; evaluation from different issues relating to its uncertainties; the evaluation with respect to its market trends; evaluation from its acceptance by its users; Furthermore the results from the evaluation study will be helpful to discuss the key points on the innovation strategy of electric vehicles relating to the development areas, policies, drawbacks and further improvements for solving this major issue in China.

Keywords Evaluation • Electric vehicles • Innovation strategy

1 Introduction

Two major issues that China is facing for its social and economical development are energy saving and environment protection. The worst situation is that China's most of the oil consumption is on fossil fuel vehicles which exhaust emissions and is one of the major source of the urban air pollution, that's why electric vehicle transport industry is boosting up in China for solving these major issues. The promotion of electric vehicles cannot only reduce the consumption to save more fossil fuel, but can also reduce or even generate no emissions. Our proposed paper is the evaluation based study on the innovation strategy of electric vehicles.

M. Ikram (✉) • Y. ZhiJun
School of Management Sciences and Engineering, Beijing Institute of Technology, Beijing, China
e-mail: mariabit@126.com; yanzhijun@bit.edu.cn

By considering the sudden raise in the world oil prices and environmental issues (Wu Feifei and Zhou Qian 2011) China's EV transport industrial sector is giving much importance to innovation strategy to make economical and social advances (Fu Xiang et al. 2006). Further more Beijing being a capital has given much promotion to this innovative technology for the country's development (ZhangXiang et al. 2008; Xiang Zhang et al. 2009; Robert Earley et al. 2011). Jonn Axsen (2010) demonstrates that the social assessment is an important factor for promoting new technologies. Yiming Pan et al. (2011) adopted the AGT evaluation method to compute evaluation values for promoting electric vehicles, while Fu xiang et al. (2008) demonstrate the innovation strategy of electric vehicles based on SWOT analysis.

2 A Detailed Comprehensive Evaluation of Electric Vehicles

A comprehensive evaluation of electric vehicles uses the questionnaire technique for evaluation and proposes the innovation strategy of EVs, solving issues related to energy saving and environment protection.

In our proposed study, Beijing is selected as representative because being a capital; it is more representatives and a strong supporter of EV industry.

The following is the detailed evaluated information which we have obtained from our survey.

2.1 Respondent's Information

In the survey about 40% of respondents were students; about 10% of questionnaires were filled in by employees, 10% by teachers and other 10% by electric vehicle dealers which will actually evaluates the markets trends of the electric vehicles.

2.2 Information About Performance Aspects of the Electric Vehicles

Figure 1 shows the information about performance aspects of the EVs. For the evaluation of electric vehicles from their performance aspects we have defined eight parameters for which the highest two scores are found for "environment friendly" and "fuel economy".
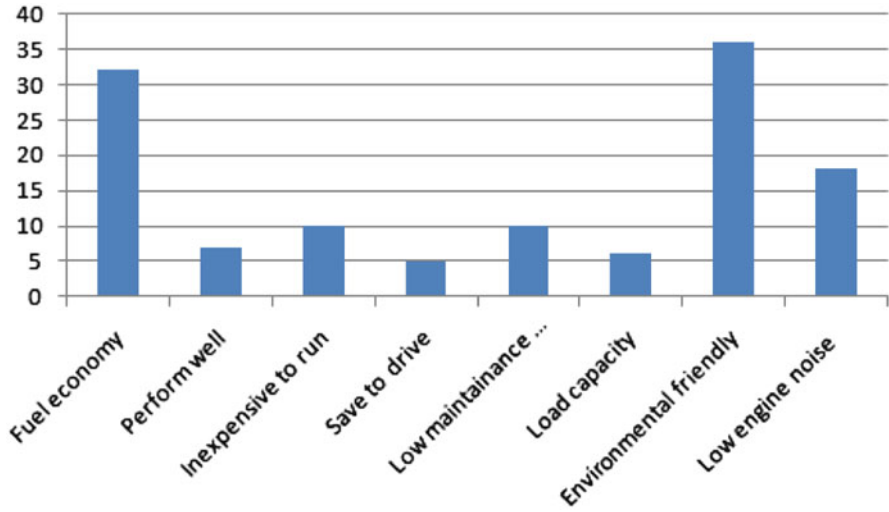
Fig. 1 Information about performance aspects of the EVs
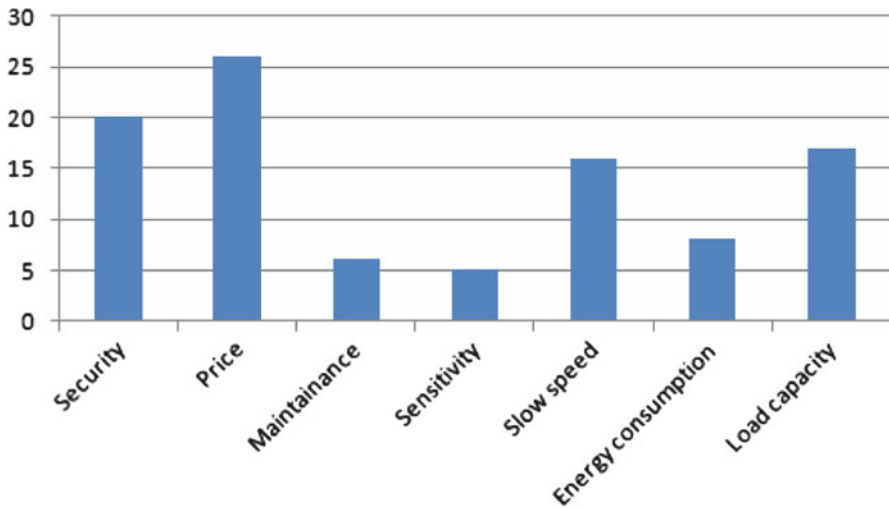
Fig. 2 Information about the drawbacks of EVs

2.3 Information About the Draw Backs of Electric Vehicles

There are some uncertainties from some perspectives as well, which is shown in Fig. 2. Our analysis shows the highest scores for "price". After then the "security" issues have also been rated relatively high.
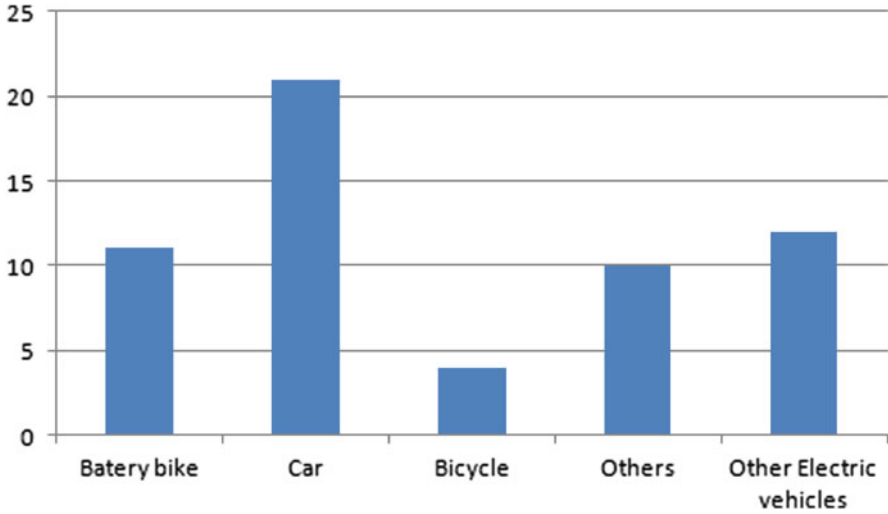
Fig. 3 Information about the value of acceptance for EVs

2.4 Information About the Value of Acceptance for Electric Vehicles

The Fig. 3 shows the information about the value of their acceptance from the user community. For this the second highest information is evaluated for "battery bikes" and "other electric vehicle".

2.5 Information About Sales and Market Aspects of Electric Vehicles

The Fig. 4 shows the information of electric vehicles from their market value; a relatively high ratio was computed against their sales aspects in 1 month.

3 Key Points on the Innovation Strategy of Electric Vehicles Based on Evaluation Study

Based on the evaluation study here we will discuss the key points, which should be concerned for carrying out the strategy.
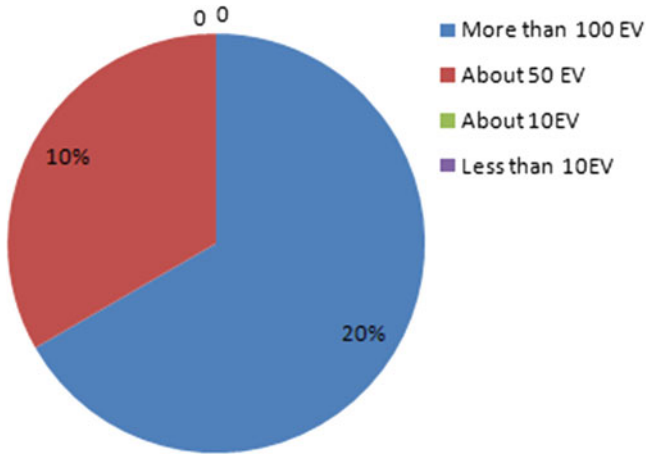
Fig. 4 Information about market and sales aspects of EVs

1. By observing the high evaluation for fuel economy and environment friendly issues, electric vehicles are best to use for solving the major issues. Government of China should pay special attention to this innovation and its management.
2. By observing the high evaluation against the price and security from their drawback aspects, these issues should be taken seriously, which are creating the means of barrier for the EV development; by solving these properly EVs will be more used at domestic level also.
3. Related to the market trends, the government should promote the market of this innovation, so as to overcome the gap between China's and overseas EV market.
4. There is needed a great investment for vehicle's demonstration, because social acceptance plays an important role in the development of any technology. Government should arrange special programs for its demonstration, its reliability issues for greatens its value of acceptance.

4 Conclusions

By having an overview from the analysis which we have derived from our survey, we can say that electric vehicles are the most feasible way to solve this major issue in China for economical and social development. The innovation strategy of Electric Vehicle seems to be an interrelated with government, EV industry, social acceptance, and its market trends in which each one is having equal importance and plays an important role for the development and further improvements for this innovative technology.

References

Fu Xiang, Hu Binxiang, Wang Yuning (2006) A research on the location strategy of China's EV industrialization. In: Proceedings of the 3rd international conference on innovation and management, conference on innovation and management, published in [J] Shanghai auto, pp 210–214

Fu xiang, Hu Binxiang, Wang yuning (2008) The research on the innovation strategy of electric vehicle demonstration. International seminar on future information technology and management engineering, Leicestershine, UK, pp 321–324

Jonn Axsen (2010) Interpersonal influence within car buyers, social networks: observing consumer assessment of Plug –in Hybrid Vehicles (PHEVs) and spread of pro-societal values. Committee in Charge, University of California

Robert Earley, LiPingKang, FengAn, Lucia Green-Weiskel (2011) Electric vehicles in the context of sustainable development in China. CSD19/2011/BP9, Commission on Sustainable Development Nineteenth Session New York, Beijing

Wu Feifei, Zhou Qian (2011) China and foreign countries electric vehicles industry current situation and development countermeasures. Management and Service Science (MASS) international conference, Wuhan, pp 1–4

Xiang Zhang, Jianzhong Yang, BoSun, Jia Wang (2009) Study on the policy of new energy vehicles in China. Vehicle power and propulsion conference, Dear born, MI, pp 35–40

Yiming Pan, Yuanbiao Zhang, Zhang Zhang, Qinglang Cao (2011) Striving for synthetic benefits: is your country suitable for the widespread use of electric vehicles. Management and Service Science (MASS) international conference, Wuhan

ZhangXiang, WangJia, YangJianzhong, CaiZhibiao, HeQinglin, HouYuanzhang (2008) Prospects of new energy vehicles for China market. HEVC08(Hybrid and Ecofriendly Vehicle Conference) proceedings CD, IET, University of Warwick, Coventry, UK

Construction of the Prediction Model of University Library Lending

Bao Sun, Liqin Tian, and Jiangwei Feng

Abstract Library lending is the books which a university library has lent or the readers which a library has served in a school year. Library lending has a strong association with the number of readers. The author has investigated library lending and the number of readers of a university library in China within 18 school years. It is found by the author that the sample data of library lending and the number of readers fit the simple linear regression model. The scatter plot between library lending and the number of readers indicates that there is a linear relationship between the two variables. The result of F-test and correlation coefficient test proves that a strong linear relationship exists between library lending and the number of readers. The result of F-test is acquired by running the function of regression of data analysis of Microsoft Excel. Then Based on the theory of the simple linear regression analysis, the prediction equation is obtained. The author gives the methods of calculating confidence intervals of the mean library lending and prediction intervals of a single library lending in the school year 2011–2012.

B. Sun (✉)
Library, North China Institute of Science and Technology, Langfang, China
e-mail: sunbaoadam@163.com

L. Tian
School of Computer Science and Technology, Qinghai Normal University, Xinin, China

Department of Computer, North China Institute of Science and Technology, Langfang, China
e-mail: tianliqin@tsinghua.org.cn

J. Feng
School of Accountancy, Hebei University of Economics and Trade, Shijiazhuang, China

Department of Finance and Accountancy, Hebei Commerce and Trade School, Shijiazhuang, China
e-mail: fjw836@126.com

1 Introduction

It is necessary for a university library administrator to predict the library lending in terms of changes of librarian and device. So it needs to construct the prediction model of library lending. There are many methods for predicting a variable such as time series analysis (http://www.statsoft.com/textbook/time-series-analysis/) and average prediction method (http://www.crowdcast.com/blog/2010/04/02/not-your-average-prediction-market/). These two kinds of prediction theory merely involve the change of the variable itself, whereas the prediction of library lending should involve the influence of other important factors such as the readers and the books belonging to a library.

Linear regression analysis is an important prediction method that has been used in many fields of nature science and social science such as business management, market policy decision, financial analysis (http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm). Simple linear regression analysis is mainly used to discuss the relationship of two variables and to predict one variable through the other.

It is found by the author that the change of library lending depends on the change of the number of readers. Then the library lending can be predicted through the number of readers based on simple linear regression analysis.

2 Construction of the Prediction Model

The prediction model is mainly used in the university in China and its data come from a university library. The library discussed in this article belongs to North China Institute of Science and Technology.

2.1 Collection of Library Lending and the Readers

The author has collected the library lending from August of one year to July of the next year and the number of readers for 18 school years, i.e. from August 1994 to April 2012. All data are displayed in Table 1.

Table 1 indicates that the increase of library lending relies on the increase of the reader numbers.

2.2 Drawing a Scatter Plot

The number of readers is used as the variable of horizontal axis and similarly library lending is used as the variable of vertical axis. The relationship between library lending and the number of readers is displayed in Fig. 1.

Figure 1 indicates that there exists linear relationship between the two variables.

Table 1 Library lending and the number of readers

School year	Number of readers	Library lending
1994–1995	1,869	53,677
1995–1996	1,484	40,737
1996–1997	1,842	45,587
1997–1998	2,024	45,920
1998–1999	2,605	44,343
1999–2000	3,527	62,520
2000–2001	5,709	149,608
2001–2002	6,850	168,494
2002–2003	8,728	191,353
2003–2004	9,873	231,200
2004–2005	10,525	254,829
2005–2006	11,021	285,733
2006–2007	12,315	299,971
2007–2008	13,706	372,197
2008–2009	15,692	369,425
2009–2010	16,678	342,177
2010–2011	17,493	316,390
2011–2012	21,249	

Fig. 1 A scatter plot between library lending and the number of readers
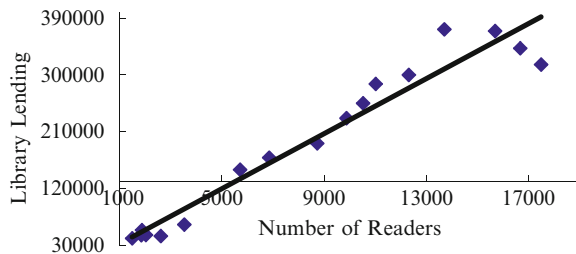
2.3 Acquisition of Linear Regression Parameters

The linear relationship between library lending and the number of readers can be expressed as the following equation (http://www.jerrydallal.com/LHSP/slr.htm).

$$L_i = a + bR_i + \varepsilon_i \ (note : i = 1, 2, \ldots, n) \tag{1}$$

Where L_i is library lending, a and b are linear regression parameters, R_i is the number of readers and ε_i is the residual.

According to the theory of least squares, the least squares estimator of the two linear regression parameters can be acquired (Obremski 2008).

$$\begin{cases} \hat{b} = \dfrac{SS_{RL}}{SS_{RR}} & (2) \\[2mm] \hat{a} = \bar{L} - \hat{b}\bar{R} & (3) \end{cases}$$

Table 2 The analysis of variance of simple linear regression

Source	Sum of squares	df	Mean square	F value
Regression	SSR	1	MSR	$F = \dfrac{MSR}{MSE}$
Residual	SSE	$n-2$	MSE	
Total	SSTO	$n-1$		

Where $SS_{RL} = \sum\limits_{i=1}^{n} (R_i - \overline{R})(L_i - \overline{L})$ and $SS_{RR} = \sum\limits_{i=1}^{n} (R_i - \overline{R})^2$ are the sum of squared values (http://mathbits.com/mathbits/tisection/statistics2/correlation.htm).

2.4 F-test

As for simple linear regression, the total sum of squares, i.e. SSTO can be decomposed as follows.

$$SSTO = SSR + SSE \tag{4}$$

In order to test the linear relationship between library lending and the number of readers, whether b is equal to zero can be tested. The null hypothesis and the alternative hypothesis should be put forward firstly (http://www.stat.yale.edu/Courses/1997-98/101/anovareg.htm).

$$H_0 : b = 0 \quad H_A : b \neq 0$$

If there exists linear relationship between library lending and the number of readers, the following F distribution statistic can be used to test the null hypothesis, i.e. H_0.

$$F = \frac{MSR}{MSE} \sim F(1, n-2) \tag{5}$$

Where $F(1, n-2)$ is F distribution statistic with df of 1 and $n-2$.

The test of regression equation can be realized through the analysis table of variance, which is displayed in Table 2.

Under the significance level assigned previously, the test result can be gained by comparing the computed F value with the critical value of F distribution which we can look up and find out in the statistical tables. If $F \leq F_\alpha(1, n-2)$, then there is not any obvious linear relationship between library lending and the number of readers and H_0 should be accepted; if $F > F_\alpha(1, n-2)$, then there is obvious linear relationship between library lending and the number of readers and H_0 should be rejected (http://stattrek.com/AP-Statistics-4/Test-Slope.aspx).

2.5 Correlation Coefficient Test

Definition 1 The correlation coefficient between library lending and the number of readers indicates how strong the linear relationship between the two variables is.

$$r = \frac{SS_{RL}}{\sqrt{SS_{RR}SS_{LL}}} \tag{6}$$

Where r is the sample correlation coefficient between the two variables, $-1 \leq r \leq 1$.

To test whether the linear correlation coefficient between library lending and the number of readers is positive, the null and alternative hypotheses are written as follows.

$$H_0 : \rho = 0 \quad H_A : \rho \neq 0$$

Where ρ is the population correlation coefficient.

The t distribution can be used to perform this test about the linear correlation coefficient when the sample size is small ($n \leq 30$). The value of the test statistic t is calculated as follows.

$$t = r \sqrt{\frac{n-2}{1-r^2}} \tag{7}$$

From the statistical tables we can consult and find out the critical value of the t distribution under the significance level α which is assigned previously. If the value of the test statistic t calculated through the sample data is much greater than the critical value of the t distribution found out in the statistical tables, i.e. $t > t_{\alpha;n-2}$, then we can conclude that there is obvious linear relationship between the two variables; on the contrary if $t \leq t_{\alpha;n-2}$, then we can conclude that there is not any obvious linear relationship between the two variables (http://mathworld.wolfram.com/CorrelationCoefficient.html).

2.6 Attaining the Prediction Value

The linear relationship between library lending and the number of readers can be depicted as the following equation.

$$L_i = \hat{a} + \hat{b} R_i \tag{8}$$

If we can get the number of readers in the coming school year, the following equation can be used to predict the prediction value of library lending in the same school year.

$$L_0 = a + bR_0 \tag{9}$$

Where L_0 is the prediction value of library lending in the coming school year and R_0 is the number of readers in the same school year (http://onlinestatbook.com/chapter12/intro.html).

2.7 Attaining the Confidence Intervals

If the number of readers in the coming school year has been given, the variance of the mean library lending can be written as follows.

$$\sigma^2\left(\hat{L}_m\right) = MSE\left(1/n + \left(R_0 - \overline{R}\right)^2 / SS_{RR}\right) \tag{10}$$

The confidence intervals for the mean library lending in the coming school year can be depicted as follows.

$$\left(\hat{L}_0 - t_{\alpha/2;n-2}\,\sigma\left(\hat{L}_m\right), \hat{L}_0 + t_{\alpha/2;n-2}\,\sigma\left(\hat{L}_m\right)\right) \tag{11}$$

Where $t_{\alpha/2;n-2}$ is the critical value of t distribution which we can look up and find out in the statistical tables and α is the significance level in t-test (Devore 2004).

2.8 Attaining the Prediction Intervals

If the number of readers in the coming school year has been given, the variance of a single library lending in the coming school year is composed of the variance of average estimating and the variance of regression model.

$$\sigma^2\left(\hat{L}_0\right) = MSE\left(1 + 1/n + \left(R_0 - \bar{R}\right)^2 / SS_{RR}\right) \tag{12}$$

The prediction intervals for a single library lending in the coming school year can be depicted as follows.

$$\left(\hat{L}_0 - t_{\alpha/2;n-2}\sigma\left(\hat{L}_0\right), \hat{L}_0 + t_{\alpha/2;n-2}\sigma\left(\hat{L}_0\right)\right) \tag{13}$$

Where $t\alpha/2;n-2$ is the critical value of t distribution which we can look up and find out in the statistical tables and α is the significance level in t-test (http://www.weibull.com/DOEWeb/confidence_intervals_in_simple_linear_regression.htm).

3 Procedure of Computation Practice

3.1 Equation of Simple Linear Regression Model

Using the data of Table 1, we can get the parameters of simple linear regression model.

$$\hat{b} = 21.811282$$

$$\hat{a} = 192{,}597.705882 - 21.811282 \times 8{,}349.470588 = 10{,}485.044931$$

Then the equation of simple linear regression model between library lending and the number of readers can be written as follows.

$$\hat{L}_i = 10{,}485.044931 + 21.811282 R_i \tag{14}$$

3.2 Practice of F-test

Firstly the null hypothesis and the alternative hypothesis should be put forward as follows.

$$H_0 : b = 0 \quad H_A : b \neq 0$$

The null hypothesis states that the slope is equal to zero, and the alternative hypothesis states that the slope is not equal to zero.

After putting the observed sample data of library lending and the number of readers into several continuous cells within two columns of Microsoft Excel, we can get the result table of the analysis of variance by running the function of the regression of data analysis of Excel. The output table from Excel appears in the following Table 3.

Because the p-value of F-test is equal to 9.947710×10^{-11}, the probability of rejecting the null hypothesis, that is, H_0 is $1 - 9.947710 \times 10^{-11} = 0.99999999990$ and there is strong evidence that the regression coefficient b is not equal to zero.

Table 3 The analysis of variance between library lending and the number of readers

Source	Sum of squares	df	Mean square	F value	Significance F
Regression	2.391025×10^{11}	1	2.391025×10^{11}	247.290783	9.947710×10^{-11}
Residual	14,503,320,969	15	966,888,064.594376		
Total	2.536058×10^{11}	16			

3.3 Practice of Correlation Coefficient Test

The rationale of correlation coefficient test is the theory in Sect. 2.5.

Step1: State the null and alternative hypotheses
The null and alternative hypotheses can be written as follows.

$$H_0 : \rho = 0 \quad H_A : \rho > 0$$

Step2: Select the distribution to use
The sample size n is 17 and is less than 30, that is, $n = 17 \le 30$. So the t distribution should be used to perform the test about the linear correlation coefficient between library lending and the number of readers.

Step3: Determine the rejection and non-rejection regions.
The significance level α is previously assigned 0.1%, that is, $\alpha = 0.1\%$. From the alternative hypothesis we know that the test is right-tailed. Hence,
Area in the right tail of the t distribution $= 0.001$

$$df = n - 2 = 17 - 2 = 15$$

From the t distribution table, the critical value of t based on 15 df is 4.07277, that is, $t_{\alpha;n-2} = t_{0.001;17-2} = 4.07277$. The rejection and non-rejection regions are $(4.07277, +\infty)$ and $(-\infty, 4.07277]$ respectively.

Step4: Calculate the value of the test statistic.
The value of the test statistic t can be calculated as follows.

$$t = r \sqrt{\frac{n-2}{1-r^2}} = 0.970985 \sqrt{\frac{17-2}{1-0.970985^2}} = 15.725482$$

Step5: Make a decision
The value of the test statistic is greater than the critical value, i.e. t $= 15.725482 >$ $t_{\alpha;n-2} = 4.07277$ and it falls in the rejection region. Hence, we reject the null hypothesis and conclude that there is a strong positive linear relationship between library lending and the number of readers such that as values for the number of readers increase, values for library lending also increase (Mann 2004).

3.4 Prediction Value of Library Lending

In accordance with the regression equation between library lending and the number of readers, the library lending in the coming school year which is written as L_0 can be predicted as follows based on the number of readers in the coming school year which has been collected previously in Table 1, i.e. $R_0 = 21,249$.

$$L_0 = a + bR_0 = 473,952.984804$$

3.5   Confidence Intervals of Library Lending

From the given information in Tables 1 and 3, the standard deviation of the mean library lending $\hat{L}_m$, that is, $\sigma\left(\hat{L}_m\right)$ can be computed as follows.

$$\sigma\left(\hat{L}_m\right) = 19{,}416.177626$$

To get the confidence intervals for the mean library lending in the coming school year, the confidence level α must be assigned previously. Here based on the result of F-test and the linear correlation coefficient test, the confidence level is previously assigned 0.001,i.e. $\alpha = 0.001$. From the t distribution table, the 15 df t critical value for a $100(1 - \alpha)\% = 99.9\%$ confidence level is 4.07277, from which we determine the desired confidence interval to be

$$473{,}952.984804 \pm (4.07277)(19{,}416.177626) = 473{,}952.984804$$
$$\pm 79{,}077.625749 = (394{,}875.359054, 553{,}030.610553)$$

3.6   Prediction Intervals of Library Lending

From the given information in Tables 1 and 3, the standard deviation of a single library lending
$\hat{L}_0$, that is, $\sigma\left(\hat{L}_0\right)$ can be computed as follows.

$$\sigma\left(\hat{L}_0\right) = 36{,}658.914580$$

Then based on the result as in Sect. 3.2 and under the same confidence level from the foregoing test, the prediction intervals of a single library lending in the coming school year 2011–2012 can be calculated as follows.

$$473{,}952.984804 \pm (4.07277)(36{,}658.914580\,) = 473{,}952.984804$$
$$\pm 149{,}303.327534 = (324{,}649.657269, 623{,}256.312338)$$

4   Conclusions

The author has observed the change of the number of readers and that of library lending for 18 school years in the library of North China Institute of Science and Technology and has found that there is a positive linear relationship between library lending and the number of readers. Based on the theory of simple linear regression,

the paper discusses how to construct the prediction model of library lending of university library. The prediction value, the confidence intervals and the prediction intervals of library lending in the coming school year 2011–2012 has been obtained. Then library lending in the coming school year 2011–2012 may be 473,953.0 books, between 324,649.7 books and 623,256.3 books and the mean library lending in the same school year may be between 394,875.4 books and 553,030.6 books.

The prediction of library lending will play a key role in the library administration. The administrators of most libraries are paying more and more attention to library lending which is often mentioned in their annual reports. The prediction of library lending may also have great effects on the annual plan of most libraries.

References

Devore JL (2004) Probability and statistics for engineering and the science, 6th edn. Brooks/Cole, Belmont, San Mateo County, pp 496–554
http://mathbits.com/mathbits/tisection/statistics2/correlation.htm
http://mathworld.wolfram.com/CorrelationCoefficient.html
http://onlinestatbook.com/chapter12/intro.html
http://stattrek.com/AP-Statistics-4/Test-Slope.aspx
http://www.crowdcast.com/blog/2010/04/02/not-your-average-prediction-market/
http://www.jerrydallal.com/LHSP/slr.htm
http://www.stat.yale.edu/Courses/1997-98/101/anovareg.htm
http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm
http://www.statsoft.com/textbook/time-series-analysis/
http://www.weibull.com/DOEWeb/confidence_intervals_in_simple_linear_regression.htm
Mann PS (2004) Introductory statistics, 5th edn. Wiley, Danvers, Essex county, pp 580–642
Obremski T (2008) Pricing models using real data. Teach Stat 30(2):44–48

Unmanned Mountain Station Supervising System Based on C Language

Ma Yuchun, Yuan Nannan, and Li Zhuang

Abstract Computer monitor systems are widely applied in many fields, such as industrial control, environmental monitoring, and intelligent transportation etc. Here is an example for computer supervising system which applied in power monitoring, and the software for Supervisor Computer is developed by KingView Tools, and the slave one is developed by the library functions integrated with 7188 system under Turbo C 2.0. A general serial communication solution is proposed and realized.

Keywords Monitor and control • Supervisor/slave computer • RS-232

1 Introduction

The system is used for supervising microwave stations in Hangzhou, Wenzhou and Ninbo, etc. The supervisor stations generally locate in urban computer room and slave stations in mountaintop near the sea with a very harsh environment (Jing Shaohong1 and Li Xiaolu 2010). There are six supervisor stations, and each of them controls a slave station. Data transmission between supervisors is carried out via microwave channels (as shown by bold black line in Fig. 1). However at a specific time, only one supervisor can poll its slave computer, the slave computer should respond accordingly, and all other supervisors must retain in listening mode. After the supervisor finishes polling its slaves, it will yield the polling privilege to the next supervisor and the polling process is similar to that of the first supervisor. The items that are monitored include working status of the diesel generator, switch power supply, mains, storage battery etc., and environment temperature, humidity and guard against theft. The main hardware system architecture is shown in Fig. 1.

M. Yuchun (✉) • Y. Nannan • L. Zhuang
Hainan Key Laboratory of Embedded Systems, Qiongzhou University, Sanya, China
e-mail: walker_ma@163.com; ynn13976969816@163.com; hitlz@163.com
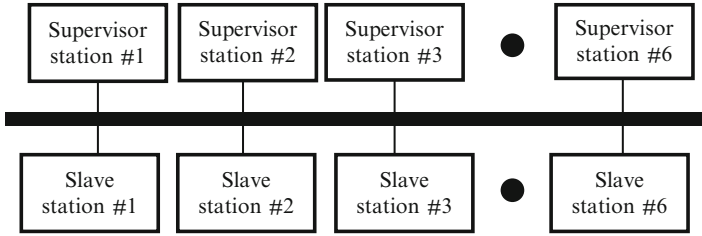
Fig. 1 Main hardware system architecture
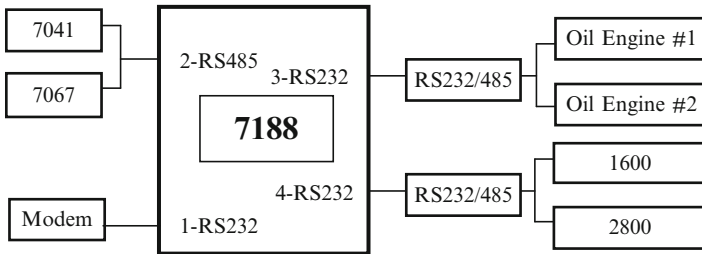
Fig. 2 Hardware architecture of slave computer

2 Hardware Design

Oil engine monitor with double on-line backups is specially designed for working with large diesel oil generator. 1600 and 2800 are also dedicated monitors for original facilities. 1600 is mainly used to monitor the mains and environmental parameters and 2800 is used to monitor two sets of storage battery (each with 12 sections). 7041 is used for alarming digital signal input module, 7067 for controlling digital signals and 7188 for supervisor module. 7188 has four serial ports which include 2 RS-232 ports (only Rx, Tx, and GND signal lines are used), 1 RS-485 and 1 full-signal 9-pin RS-232 port which can be changed to RS-485 by internal jumper. 7188 module has provided many interfaces and can be operated under harsh environment and provide many built-in functions that can help to fast develop reliable supervising system program for field computers. The hardware architecture of the slave stations is shown in Fig. 2. Since all monitoring modules use RS-485 ports, two RS-232/RS-485 conversion ports are required. Port COM1 is full-signal RS-232 port that is connected to a dedicated Modem to transfer data through microwave channels. The supervisor stations are connected to a dedicated Modem via RS-232 port to transfer data through microwave channels.

3 System Software

3.1 Supervisor Software

The supervisor software is developed with KingView 5.1, it is featured with real-time multiple tasks, multi-threading, high sampling speed, high reliability, supporting distributed history database and dial-up network and many ActiveX controls. The software has been extensively used in steel, chemical industry, environment protection, national defense and aeronautical and astronomical fields.

It is easy and convenient to develop real-time computer supervising system using kingView Tools. Firstly it is required to define device and load communication driver. The system, for example, communicates with slave computer through RS-232 port. The facilities are defined in following sequence: Smart module—I-7000 series, I-7188—Modbus RTU Extension, secondly to specify working parameters of RS-232. After that we need to define data dictionary and variable name, data type and relationship between registers. For the variable in data dictionary, the corresponding linear target value can be calculated automatically based on the original values. Finally you can design user interface with primitive controls provided by KingView Tools. "Animate Link" can be built for both string and primitive so as to integrate data variables and data display. Data can be displayed in two ways: text format and animation display such as alarm, green indicating normal and red indicating alarm.

3.2 Slave Computer Software

3.2.1 Slave Computer Main Program Design Scheme

7188 provides extensive library functions which include watchdog, related data sending and receiving functions and time clock interrupt function, etc. In addition it also provides many example programs (Yi Zhou et al. 2010). All of these make it easy and convenient to develop slave computer software and to integrate with modules provided by other companies. Each serial port of 7188 can control two smart devices that can be distinguished by their addresses. The working parameters and module addresses for ICP 7000 series are adjustable and can be set before system installation, while working parameters and addresses for smart modules from other manufactures have been preset and cannot be changed.

Since 7188 has built-in DOS environment, the slave computer program can be written in TC 2.0. The main program should first initialize the system, including serial ports initialization and operating parameters initialization, and install the

Fig. 3 Software architecture
of slave computer
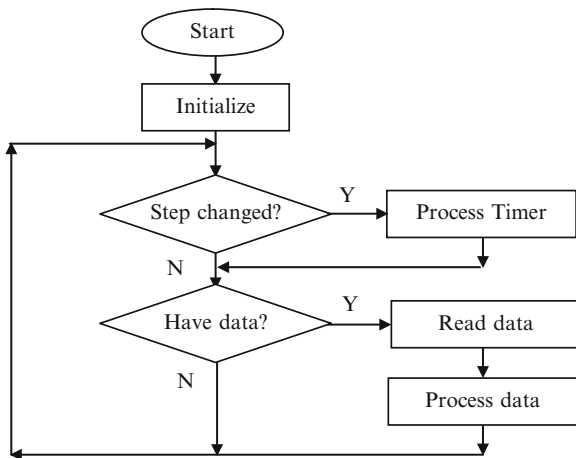
user timer interrupt. Because the interrupt interval is preset as 1 ms, the data receiving and sending routines should not be used within the interrupt service function. The reason for this is because transmitting 1 byte at 9,600 bps, which is typical in a supervising system, will need about 1 ms. If the speed is too high, the data transmission will become unreliable. And typically at least 2 bytes will be transmitted one time; this will result in DOS re-entry error and makes the system unreliable. Therefore time clock interrupt program should be as short as possible and only handle the coordination of different jobs. The main task can be implemented in the main routine. When it's time for polling, it will send inquiring command to the smart modules or devices, and when data is ready; the program will read data and deal with them. The software architecture of slave computer is shown in Fig. 3.

3.2.2 Communication Problems and Solutions

Data communication reliability is the most important for computer supervising system. Industrial computer I/O devices provided by ICP can transmit data steadily and connect reliably with 7188. However data transmission between modules provided by other companies is not as continuous as expected, and the waiting time depends on different modules. Too short will result in data loss and too long may lead to system failure or receiving of the next data package. Data receiving program can be found in function ReadDataDelay of program 1. The first byte of array bData is used to store the number of bytes received and the following is the content of valid data. After many times of debugging, the delay time for the system should be greater than 0×140. If it is less than 0×130, data error rate will be about 1/4 and 1–2 bytes will be lost each time. The delay time between 0×130 and 0×140 is critical value.

Program 1 Definition of Function ReadDataDelay

```
/************************************************/
/* ReadDataDelay:Read data from COM nPort, save */
/* to array bData, nDealy is the waiting steps. */
/************************************************/
  void ReadDataDelay(int nPort, unsigned char
*bData, int nDelay)
  {
     int nData=0; /* count for received bytes */
     int iCount=0; /*waiting steps*/
     while(iCount < nDelay)
     {
  if(IsCom(nPort)==QueueIsNotEmpty)
  {
  nData++;
  bData[nData] = ReadCom(nPort);
  iCount = 0;
    }
   else iCount++; /* waiting for following data */
  }
  bData[0] = nData;
     ClearCom(nPort); /* prepare for next package */
  }
```

4 Comparison with Panasonic PLC System

The large-scale pump station supervising system for Qiantang River, Zhejiang province is a national irrigation works created with a loan from the World Bank. The central computer uses industrial computer, and the front-end computer uses Panasonic PLC, model FP10SH (Gabor Gianina1 et al. 2009). The hardware architecture is shown in Fig. 4. The central computer communicates with front-end computers via two RS-232 serial ports: COM1 is used for sending interrupt signals such as real-time alarm to the central computer; COM2 is programmable port with which the central computer can download PLC program to front-end computer and read/write data. In Fig. 4, from left to right displayed: CPU module, A/D converter, temperature module (RTD), digital input module and digital output module.

All modules are inserted in the motherboard of the system. Main CPU module only has two RS-232 serial ports. In respect of software the central computer doesn't provided configuration software and the front-end computer software is implemented by "ladder diagram". It is difficult for developers without electrical
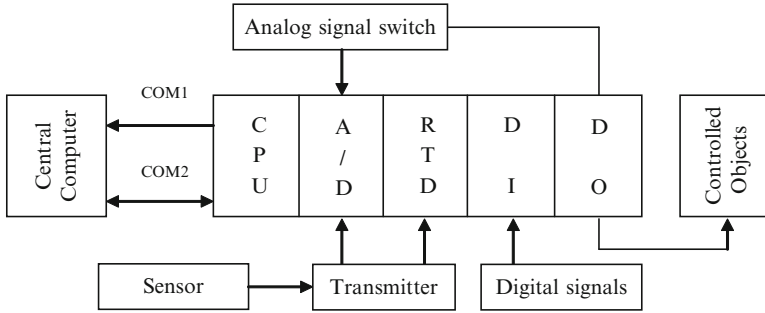
Fig. 4 Panasonic PLC hardware architecture

knowledge to understand ladder diagram. In addition there is no sample program provided with CD. PLC operating system automatically scans "ladder diagram" program repeatedly. But a problem occurred during numeral conversion. After many times of debug, it is found that only one data can be converted each loop of system scan, the multiple-byte conversion cannot be implemented with this system.

On the other hand, it is very easy to integrate using 7000 series. 7188 has four RS-232/485 serial ports. Moreover it can also work reliably under harsh environment, for example when its surface temperature increased to above 50°C. The central computer program can be developed quickly and reliably with configuration software provided by the central computer; the front-end computer software is featured with simple logic and watchdog resetting system by using many library functions and example program.

5 Debug and Conclusions

"Universal and Multifunctional Testing Software on Computer Monitor System"(Ma Yuchun 2011) is designed especially for testing supervising system or related smart devices through RS-232 or RS-485. The software can be used not only as a slave computer to test supervisor program but also as a supervisor computer to test slave computer programs, which have been verified in many supervising systems. 7000 series modules can transmit data steadily during debugging the system. Supervisors provided by other companies have 25 data transmission errors and 7000 series modules have no error when they have been tested continuously for 48 h (testing one time per 5 s). There is no error during the test of supervisor computer (configuration software) with the test software as slave computer. Now the system has passed the debug in laboratory and can operate successfully. Using ICP configuration software and 7000 series modules make it possible to reduce development time and improve system reliability, and the solution of serial communication implemented via Turbo C 2.0 is reliable and flexible. In a word it deserves to be generalized extensively.

References

Gabor Gianina, Zmaranda Doina, Gyorodi Cornelia et al (2009) Redundancy method used in PLC related applications. 2009 3rd international workshop on Soft Computing Applications (SOFA 2009), IEEE Computer Society, July 2009, United States, pp 119–125

Jing Shaohong, Li Xiaolu (2010) Design of host computer monitoring configuration software based on C#. 2010 International Conference on Measuring Technology and Mechatronics Automation (ICMTMA 2010), IEEE Computer Society, March 2010, United States, pp 1000–1003

Ma Yuchun (2011) Universal and multifunctional testing software on computer monitor system: China, SR025314[P].2011-5-3

Yi Zhou, Zhou Zhonghua, Hou Ligang et al (2010) Design and optimization of serial communication system interface module. 2010 International Conference on Computer Engineering and Technology (ICCET 2010), IEEE Computer Society, April 2010, United States, pp 7272–7276

Comprehensive Practice Research of Professional Platform Courses Based on Computer Monitor Technologies

Ma Yuchun, Yuan Nannan, and Li Zhuang

Abstract Computer monitor systems are widely used in many fields nowadays, and the digital input and output module, analog input module which most commonly used in computer monitor systems are simulated by software, including the protocols to operate them. RS-232 to RJ-45 converter is designed to extend the serial communication distance, and RS-232 to RS-232 converter to connect two devices in different serial communication parameters and capture their protocols to database. Finally, a simulated development platform on computer monitor systems is constructed through all those software tools described above. Since computer monitor technologies concerning comprehensive professional courses, applied teaching examples are given accordingly.

Keywords Simulated module • Serial communication • Comprehensive practice • Professional platform course

1 Introduction

Computer monitor systems (CMS) consist of computers which act as controllers, analog input and output (I/O) modules, digital input and output modules, and different sensors etc. In this system, computers are directly involved in every stage such as measurement, monitor and control (Ma Yuchun 2007). Measurement is accomplished mainly through sensors and the corresponding input module to acquire status data of devices being supervised, monitor is done through the analysis to status data and then providing operating reference to operators, control is done through output module with manual or automatic strategy to operate those devices

M. Yuchun (✉) • Y. Nannan • L. Zhuang
Hainan Key Laboratory of Embedded Systems, Qiongzhou University, Sanya, China
e-mail: walker_ma@163.com; ynn13976969816@163.com; hitlz@163.com

being controlled. Input and output modules play a very important role in CMS as described above, to study these modules is very meaningful to the development, test and computer-aided education on CMS.

Most I/O modules are provided with serial communication port. In this paper, 8-chanel digital input module (DIM), 8-chanel digital output module (DOM) and 8-chanel analog input module (AIM) are designed through simulated software, and all these modules can be connected to the controllers, such as embedded controllers or personal computers via RS-232 interface. In addition, the software converters from RS-232 to RJ-45 and RS-232 to RS-232 are developed, the former is used to extend the distance of local serial communication, just about 15 m or so, to Internet, and the latter is used to connect two devices with different baud rate and serial communication protocols can be captured to database. Simulated development platform on CMS is set up based on all those simulated modules and software converters described above, and this platform costs free except computers, so can be widely installed into computer laboratory.

2   Designs of Simulated Modules

In practical projects concerning CMS, data acquisition modules and control modules generally use the RS-485 interface because of needing only two wires, construction convenient at the same time, transmission distance can reach 1,200 m around. Most computers are equipped with RS-232 interface and can be connected to RS-485 modules via RS-232 to RS-485 hardware converter. Consequently, RS-232 interface is default setting in simulated modules.

2.1   Design of DIM

DIM is used to acquire digital status from switches where the change of value is mostly due to the environment factor or analog value out of threshold. Digital status is used to give an alarm or as a control reference to operators. There are eight digital status described with a byte in 8-chanel DIM, 1 for the switch being turned on and 0 for the switch being turned off, as shown in Fig. 1.

Every bit in a status byte can be switched at random under certain cycle set by program. Here take system clock as seeds to generate a 0–7 random number to set corresponding bit 1, and set the others 0. For example, when the number is 3, IN3 is set 1, and the other bits are set 0. Communication parameters of RS-232 can be set in accordance with requirements.

Fig. 1 Digital status in a byte

IN7	IN6	IN5	IN4	IN3	IN2	IN1	IN0

2.2 Design of AIM

AIM acquires analog status of supervised device. At first, physical signals are turned into analog voltage or current through sensors, and then modulated into an appropriate range for analog-to-digital converter to convert into digital numbers. In order to facilitate understanding, 8-chanel AIM is used to acquire temperature value with the range 20–100°C corresponding with byte range 0–255. New temperature T_{new} changes based on old one T_{old}, as in

$$T_{new} = T_{old} + k * step, \tag{1}$$

where k is a coefficient with only two values, when k for 1, the temperature increases, when k is -1, the temperature decreases, but T_{new} is in the range of 0–255; *step* is the breadth changed each time. Both k and *step* can be set via program, including the interval time between changes.

2.3 Design of DOM

DOM is used to simulate industrial control signals to operate supervised device according to digital and analog input. There are eight switches corresponding with eight bits in a byte in an 8-chanel DOM, switch close expressed with 1, open for 0, as shown in Fig. 1. Digital output can be controlled directly via write command transmitted through RS-232 described below.

3 Designs of Communication Protocols

All communication with simulated modules consists of commands generated by the supervisor and responses transmitted by simulated modules. Each module has a unique ID number with the range 1–255 which is used for addressing purpose (I-7000 Remote I/O Introduction 2012). The ID is 01 by default. All commands to the modules contain the ID address, meaning that only the addressed module will respond.

Modules from same manufacturer generally share the same leading character. EOT is a special ASCII character (0×04), frequently used as a control character, here takes EOT as the leading character in read command as following:

EOT	ID	Command	[BCC]	[CR]

Read command is sent from supervisor to acquire the status of simulated modules, where command is "R" (0×52) meaning read, BCC is the data block

Fig. 2 Implementation
of AIM
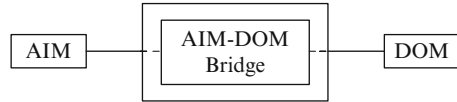
check code, such as XOR, ADD and CRC etc. CR is carriage return character which
used as end of message. There is no BCC or CR code by default in all commands.
BCC is calculated with all bytes except CR.

Response command is sent from addressed module to supervisor, the command
format as following:

STX	ID	Data	[BCC]	[CR]
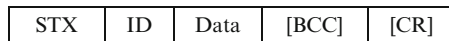
Where STX (0×02) means "start of message", data is byte carried information of
digital or analog status to supervisor.

Write command is sent from supervisor to DOM, used to change its status, the
command format as following:

EOT	ID	Command	Data	[BCC]	[CR]
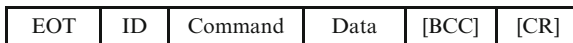
Where command is "W" (0×57) meaning write, data is byte sent to DOM. There are
two functions for DOM, the first is to control output signals directly, and the second
control the temperature trend of AIM, as shown in Fig. 2. Under real environments,
when turn on the heater by DOM, temperature acquired from AIM increases, and
turn off the heater, temperature decreases. Simulated DOM and AIM are isolated,
and there is no relation between them. AIM-DOM Bridge is used to combine AIM
and DOM, when certain bit in DOM is set 1, corresponding temperature in AIM
increases, otherwise decrease, as in (1).

4   Visual Model of Simulated Module

The visual model of simulated module is shown in Fig. 3, where visualization area
is for status display of the simulated module via static and dynamic graphics, for
example, the digital input of a DIM is switched at random according to the period
set by program, input status is shown through hiding a short red line or displaying it.
Digital output of DOM is included in a powered circuit with load, such as lamp and
bell, when certain bit is set 1, corresponding lamp or bell will be lighted or ringed.
There are signal lamps for each module indicating the status of data receiving or
transmitting.

Setting buttons are for man machine interface, used to set port number of RS-232,
baud rate, BCC code, ID of module, timer and if carry CR etc. Timer to DIM is the
period time when switch turned on or off.

Fig. 3 Visual model
of simulated module

Visualization of Module	Protocols Description
	Bytes Received
Setting Buttons	Bytes Sent

Protocols description includes the format of read command, response command and write command, with them programmers need not read user's manual.

Bytes received area displays bytes sent by the supervisor, and bytes sent area displays response bytes sent by the simulated module. The physical module only indicates communication through signal lamps, compared inconvenient with the simulated module which in more ways.

5   Supporting Tools

Supporting tools mainly include RS-232 to RS-232 and RS-232 to RJ-45 converter, the former is used to connect two RS-232 devices with different serial communication parameters, at the same time, take down the protocols between them, and the latter is for remote debug and protocols capture.

5.1   RS-232 to RS-232 Converter

Due to historic reasons, one of the supervisor or controlled device in CMS needs to be replaced, while the other cannot be changed, it may need for converter and analysis of protocols. Since RS-232 is widely used in CMS, here only take it as default settings. As described above, each serial communication protocol includes leading character, module address, block check code and end of message, two devices with different protocol parameters can be connected via RS-232 to RS-232 converter too. For example, the converter received a read command from supervisor device with CR as end message, it can send to controlled device without CR.

As shown in Fig. 4, the supervisor device is connected to COM1 of RS-232 to RS-232 converter, and the controlled device connected to COM2, when the converter received a command from the supervisor device, it can then send to the controlled device with different baud rate and command formats, and vice versa. At the same time, protocols between supervisor and controlled device can be taken down with time stamp; it is very convenient for further analysis on protocols and holding signals.

Fig. 4 RS-232 to RS-232
converter
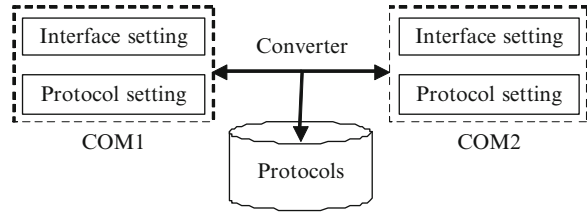
5.2   *RS-232 to RJ-45 Converter*

The RS-232 to RJ-45 converter is used to replace the hardware converter by software one in order to transmit bytes from RS-232 to RJ-45 and vice versa. There is no protocol change but transmit directly. Serial communication parameters of the converter are set according to which device connected to. Due to most supervisors connected to embedded controllers via RS-232, and so did between controllers and input modules, the converter can transmit bytes from local devices to the Internet for remote process. The RS-232 to RJ-45 converter can work in client mode or server mode.

5.3   *Virtual Serial Port*

The RS-232 cable is absolutely necessarily for communication between RS-232 devices except built-in modem. Virtual Serial Port Driver (VSPD) is an advanced utility, which emulates unlimited number of RS-232 serial ports connected via virtual null-modem cable using special driver, which can be easily included in software providing simple and convenient way to create and configure virtual serial ports directly (Virtual Serial Port 2012). Virtual serial ports appear to operating system or any Windows application as "standard" hardware serial ports. Virtual serial ports created by VSPD work exactly the same as real ports supporting all signal lines and baud rate emulation. For instance, using VSPD two virtual serial ports: COM5 and COM6 in a virtual pair can be created so they can be opened by two different applications. Everything that the first application sends to COM5 will be received by the second application at COM6, and all data sent to COM6 by the second application will be received by the first one at COM5. So, with VSPD the RS-232 cable can be omitted.

6   **Application Model**

Using simulated modules and supporting tools, a simulated development platform on CMS can be established via only one computer, and remote RS-232 protocols can be captured too.

Fig. 5 Development model with single computer
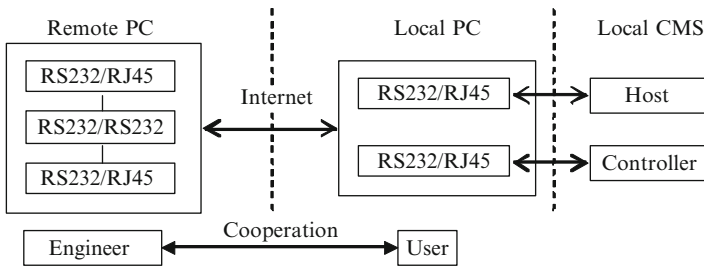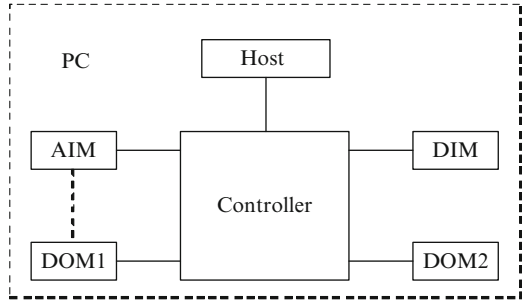
Fig. 6 Remote protocol capture model of serial communication

6.1　Development Model with Single Computer

By building simulated modules around the controller, a simulated development platform is constructed, as shown in Fig. 5, where each module is connected via a virtual RS-232 pair, and the controller is realized by Turbo C program with serials of RS-232 functions (Ma Yuchun and Huang Yinghong 2012). All programs run in a computer.

Before development, it is necessary to design the control logic, for instance, let temperatures of AIM follow changes of DOM1 in order to select the trend of temperatures, and DOM2 follow changes of DIM.

Supervisor programs are written in a high-level language such as Visual Basic or Delphi, and acquire information of modules from the controller via RS-232 interface. At the same time, send control commands to DOM via the controller. Protocols between supervisor and controller may be designed by the programmer, for manufactures have not provided them.

6.2　Remote Protocols Capture

Sometimes an engineer needs to analyze local serial communication protocols between the supervisor and controller from remote PC; such an appropriate solution is shown in Fig. 6. With RS-232 to RJ-45 converter run in local PC, everything sent

Fig. 7 Comprehensive
practice model of
professional platform
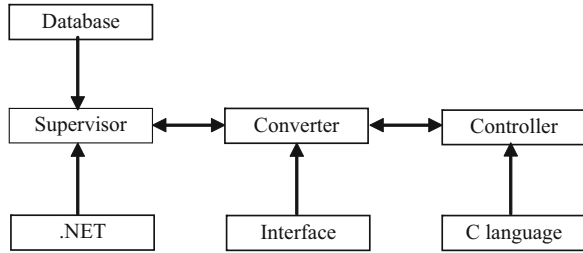
to the controller from supervisor are transmitted to remote PC. These byte streams
reach remote PC via RJ-45 to RS-232 converter and be taken down through RS-232
to RS-232 converter, finally, these data are sent back to controller transparently, and
vice versa. The solution is right for remote debug and maintenance of CMS. The
engineer and user cooperate through Internet, such as negotiate the working mode
of the computer and its IP address.

6.3   Comprehensive Practice Model of Professional Platform

Computer monitor system program includes supervisor and slave program, and
concerning most computer relative courses such as C language (even assembly
language), computer network, computer interface, visual program design (.NET
Series, for example), and database etc. The comprehensive practice model of
professional platform is shown in Fig. 7. C language or assembly language can
be used to write simple program to communicate with single simulated input or
output module, and through this practice, students can master the program tool and
the method to operate computer interface RS-232. Communication with simulated
module will do well to TCP/IP protocols. Based on all these described above,
general slave computer program can be designed and run in controller shown
in Fig. 5, at the same time, communicate with supervisor computer. Supervisor
computer program can write by .NET series, Visual Basic 2008, for example,
and database used to save protocols and system status etc. The same professional
platform courses, since computer monitor technologies is imported, have been more
interesting to students, and will be very helpful for their employment.

7   Conclusions

Computer monitor systems are widely used in many fields. Mostly used modules
such as digital input and output modules, and analogue input modules are simulated
via software, including the protocols to operate them. In addition, RS-232 to RS-232
converter and RS-232 to RJ-45 converter are designed too. With these simulated

modules and supporting tools, and third-party virtual serial port, a free development platform on computer monitor systems can be constructed in the computer lab. All programs and tools are test strictly and can be widely used for simulated development and computer-aided education on computer monitor systems.

References

I-7000 Remote I/O Introduction (2012) http://www.icpdas.com/productsRemote_IO/i-7000/i-7000_introduction.htm

Ma Yuchun (2007) Computer monitor technologies and system development. Tsinghua University Press, Beijing

Ma Yuchun, Huang Yinghong (2012) General application research on GSM Module [J]. Appl Mech Mater 151:96–100

Virtual Serial Port (2012) http://www.eltima.com/

Situation and Problems of Chinese College E-learning Resource Development

Yufang Chao

Abstract E-learning resource is very important to carry out E-learning in china. This paper takes Chinese college for the study, analyses the situation and problems of Chinese college E-learning resource development. Based on the above analysis and feasibility studies, came to this conclusion that there are lacking integration, lacking organized and lacking platform and tolls to integrate resources in china college E-learning resource development. To carry out in-depth E-learning, Chinese government and colleges should renewal education idea, construct development platform, and integrate all kind of learning resources, introduce incentive policy.

Keywords E-learning resource • Chinese College • Situation and problems • Development

1 Introduction

Since the twentieth Century 1980s, with the rapid development of information technology, learning styles also produced change. E-learning is among of new-style ways of learning. E-learning means digital learning, electronic learning, means learning mainly through the internet, is also a kind learning style applied widely in information society (Aldrich 2006).

In 1999, the word "E-learning" is put forward firstly on conference on line-learning in the United States (Brown et al. 1989). At first, the "E-learning" definition is teaching content or learning experience supported or leaded by network electronic technical, what learning and teaching activities can be called E-learning through the network, radio, TV, video and other electronic media (Chan et al. 2006).

Y. Chao (✉)
School of Business, Shandong Polytechnic University, Jinan, Shandong, China
e-mail: chaochuwen@sian.com

Chinese government, the departments in charge of education and colleges has fully aware of the importance of E-learning, and adopt various measures to vigorously develop the E-learning. Although E-learning has achieved great progress in china, but there were many problems. Because colleges are the main position of knowledge education in china, it provides knowledge acquisition, access, sharing channels for students. So this paper takes college for the study, in view of E-learning resources development, analyzes problems of Chinese college E-learning development. Through this research, hoping to promote Chinese E-learning development.

2 E-learning Advantages and Content

2.1 *E-learning Advantages*

E-learning makes full use of IT technology, have the all new communication mechanism and learning environment of rich resource (Mandinach 2003). So, E-learning has many advantages compared with the traditional learning style. Firstly, by using comprehensive images, sounds, animations etc., E-learning provides the diversity of stimulation, and attracts the attention of learners, stimulates learners' interest in learning. So, it can largely increase the interest of learner, enlarge learner range of knowledge and make more friends. Secondly, because E-learning riches multi-channel teaching materials, and improves efficiency to cognitive through learners call more information processing, it can improve learners' cognitive effect. Thirdly, because it is strong interactivity, so E-learning let learners to participate the learning process control. This can improve learning initiative, and make the cognitive process more in line with the learners' learning progress and needs.

In addition, E-learning broke the limited by time and space; learners can plan their learning activities according to their own needs (Russet 2001 and Nick Van Dam 2003). Meanwhile, variety of channel information provided by E-learning is processed by different channels, and no mutual interference. This strengthens the acceptance and quantity to information, strengthens depth to specific information, and helps learner understand well to learning materials.

2.2 *E-learning System and E-learning Resource Important*

E-learning is a system included by E-learning environment, E-learning resources and E-learning style, this system is shown as Fig. 1 (Siemens 2005). E-learning resources are a kind of digital multimedia material and resource. It makes E-learning become possible through learner search and process information in vivid way. And, it is key part to stimulate learner cooperation, creativity. E-learning environ-
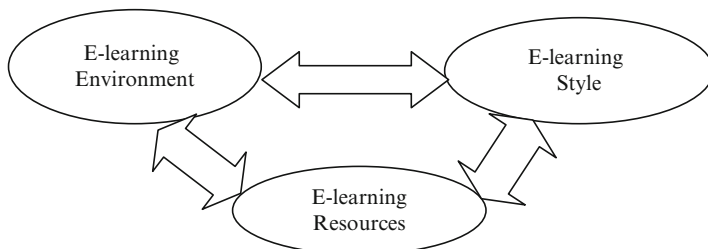
Fig. 1 Three elements of E-learning system

ment is informatization information processing environment based on multimedia computer and network environment. It has characteristics, multimedia-oriented display, Information processing intellectualization, teaching environment virtualization. E-learning style is a kind of style through autonomous learning, cooperative learning, learning by discovery, experiential learning, and that it is mainly based on digital content and in the digital environment (Weiser 1991).

Figure 1 is shows that these three elements restrict each other, and promote each other. In these three key elements, learning resource is the key; learning environment is ideal atmosphere; learning style is the realization form. So, E-learning resources are very important to carry E-learning for China colleges.

3   Situation and Problems Chinese Colleges E-learning Resources Development

3.1   Situation of Chinese E-learning Resources Development

After years of construction, Chinese college E-learning resources have been greatly enriched. "National Knowledge Infrastructure" (CNKI) project led by the Chinese government has the international advanced level of technology of digital library. The project built the world's largest full-text information "CNKI digital library", and officially launched the construction of "Chinese knowledge" and CNKI grid resources sharing platform. The project provides the most abundant knowledge and information resources and the most effective knowledge dissemination and digital learning platform for the whole society to knowledge resources sharing. Now, Chinese colleges have purchased this item engineering services.

In addition, the Chinese Ministry of education constructed the national quality course project. The project use modern information technology, network technology, and integrated national, provincial, schools at all levels of quality courses, formed the support environment and sharing service platform to meet the

Table 1 Human resources managemnent profession E-learning websites

SN	Net name	Website
1	China human resources network	http://www.hr.com.cn/
2	Free sharing networks	http://www.vsharing.com/
3	Valley habitat	http://bbs.21manager.com/forum.php
4	Borui management online	http://www.boraid.com/
5	World executive	http://www.icxo.com/
6	HR Bar	http://www.hrbar.com/
7	Study hard community	http://bbs.qs100.com/

Resource: According to the popularity of the website, the website visit quantity

high-quality curriculum for storage, retrieval, service operation. Development so far, national quality curriculum project has constructed national quality curriculum resource library and resource center. As of 2011 December, Resource Center have teaching video data high up to 34,363, teaching courseware high up to 328,315, teaching case high up to 49,541, electronic teaching plan high up to 310,826. Curriculum Library Center have undergraduate course of 14,345 doors, 5,921 doors of Higher Vocational curriculum. Now Chinese colleges part in the project, many teachers and students make use of this platform to teaching and learning.

In addition, Chinese colleges also recognized the importance of E-learning, and have invested substantial resources for E-learning resources construction. Under the guidance the ministry of education, provinces established excellent curriculum resource sharing system; the University also established its own class excellent course resource sharing system. Chinese college on excellent course construction seriously, greatly increased the E-learning learning resources for learners, and provides a new way of learning.

In addition, on the development of Digital Libraries, Chinese colleges generally constructed the electronic reading room, campus LAN. These expand the E-learning resources. At last, China launched the college teachers' modern educational information technology training project. This project can greatly improve the information literacy of university teachers, and promote teachers make use of multimedia network. That information literacy improving for the E-learning resources development has laid a good foundation.

In addition, some enterprise and people having same interest and hobby co-sponsored some professional E-learning websites. Some human resources management professional E-learning websites and platform is shown in Table 1.

These professional E-learning platforms in Table 1 have numerous E-learning resources. But, their contents are disorder, and difficult to the quality of resources. At same time, these E-learning resource lack professional knowledge management, learners often fail to determine whether the resources they need, causing frequent download information, occupied the learners a lot of time.

3.2 Problems of Chinese Colleges E-learning Resources Development

3.2.1 Lacking Initiative to Develop E-learning Resources

A significant feature of Chinese E-learning resources development is the government leading. Although Chinese colleges also do some work in E-learning resources development, but still some college don't realize importance of E-learning. An investigation by taken in 2007 shows that understanding of most colleges teachers on E-learning is superficial and one-sided. The preliminary investigation by author found that understanding of students on E-learning only is learning through network; understanding of colleges' personnel (College audio-visual center) on E-learning is mostly confined in the school, only think E-learning as a multimedia teaching, distance education, network education. There no think E-learning as the knowledge creation, management and share knowledge.

The E-learning one-sided understanding influence university in resource development input. Most Chinese colleges passive participated to the E-learning, rather than active participated. These Chinese colleges are associated with a state of E-learning resources to establish contact, lack of social integration of E-learning resource initiative.

3.2.2 Lacking Integration to E-learning Resources

Now, most of Chinese colleges have some E-learning resources, especially in excellent courses construction. But, these resources lack integration, can't reach the requirements of E-learning. Because the E-learning system is open; learner can develop through learning E-learning. So, this also affected the efficiency of resources utilization, and hindered further development of E-learning resources. Chinese colleges lack integration to E-learning resources embodied in the following aspects.

Firstly, the colleges' resources are scattered and independent. Taking china colleges' translation courses as example, colleges' excellent translation courses have been development, but they are independent, shown in Table 2.

Xi'an FANYI University, XIAN International Studies University, Sichuan International Studies University in Table 2 has many E-learning resources on translation courses, but there no integration. People don't belong to this college difficult use these E-learning resources. While, Harvard, Yale etc. world-famous universities put the most famous courses on website to the free. And there were new entries into the open course activity, as London Business School etc. These famous universities raised a lot of quality E-learning resources, and these resources are free, and easily integrated, shown in Table 3. China colleges should study universities in Table 3, and open their E-learning resources, integrate these free quality resources.

Table 2 Colleges excellent course for translation

College	URL
XIAN International Studies University	http://www.xisu.edu.cn/jingpin/
Sichuan International Studies University	http://jpkc.sisu.edu.cn/jpkc/
Xi'an FANYI University	http://61.185.224.141/jpkc/kczt.htm
Tianjin Foreign Studies University	ttp://xibu.tjfsu.edu.cn/jpkc/2005/yhfy/
Tianjin Normal University	http://59.67.71.237:8080/wy/
Tianjin University of Commerce	http://www.tjcu.cn/web/wangshangketang/biyi/index.htm
Hunan Normal University	http://202.197.120.40/ec/C41/Course/Index.htm

Resource: corresponding college website

Table 3 Foreign universities open courses

Foreign universities	URL	Cources quantity
Harvard Business School	http://www.harvard.edu/	Hundreds of course
University of California Berkeley	http://webcast.berkeley.edu/courses.php	Hundreds of course
Massachusetts Institute of Technology	http://ocw.mit.edu/OcwWeb/web/courses/courses/index.htm	1,800
University of Washington	http://www.cs.washington.edu/education/course-webs.html	Hundreds of course

Resource: corresponding college website

Table 4 Professional organizations free resources

Company	Courses
Intel	Intel® Teach Skills for Success Course (Intel® Teach Getting Started Course)
IBM	Java train course, AU16 course etc.
Microsoft	Microsoft_SQL Server,
HP	HP China on-line course

Resource: website

Secondly, Chinese colleges lack integration to social resources. Now, with the development of the network, social professional organizations develop many E-learning resources, as Microsoft, IBM, CISCO, Inter and HP etc. These E-learning resources are public benefit course, so are free and open, as shown Table 4.

In Table 4, these E-learning resources are open, free and professional. Integrating these E-learning is of great value to improve Chinese college teaching. But no Chinese colleges integrate these E-learning resources.

Thirdly, there is lack integration to campus resources. Colleges' teachers and students are intellective. So, they can provide a lot of quality E-learning resources. A little college integrates these resources.

Fig. 2 Renmin University of China economic BBS Screenshots

3.2.3 Lacking Organization to E-learning Resources

Although some china colleges attend the construction of learning resources, and build up material database; develop multimedia teaching material and the network courseware; provide exercise, testing and teaching on internet. However, widespread problem in resources construction is that these network course emphasizes display teaching content, all the resources are piling up in the network, resources did not well organized, no structure. By Renmin University of China economic BBS for example, although BBS has many resources, and also makes a preliminary classification, but resources are mostly mixed and disorderly distributed, as shown Fig. 2.

In Fig. 2, all kind of resources are mixed and disorderly. This causes that learners depend on their ability and patience to get resources they went. Learners often have to collect resources on different subsystem or child menu to learn some knowledge. While learners often does not know exactly how much the curriculum resources, how much particular knowledge resources, more do not know how much resources what needed in all resources.

While, due to lack organization, many of resources are repeated, this course that resources database is hard to manage. That resource lacks organized hampers E-learning develop in China colleges, affected resource usage.

3.2.4 Lacking Platform and Tolls to Integrate E-learning Resources

Hardware device and corresponding tolls are necessary to E-learning resources development. Hardware device and tolls are premise to resources development.

Now, through many of companies have developed E-learning resource platform and tolls, for example, Apple Inc developed iTunesU to integrate E-learning resources. Now, Princeton University, Harvard University, MIT higher education

institutions, etc., provide E-learning resources for it. But, many of china colleges have not developed a standardized platform for E-learning resources development.

While, resource software simply provides resources or collects resources, has no corresponding tool to integrate resources. Many repository, in addition to provide a resource search tools, no provide other information integration tools, learners still need other production tools combination information, it is not convenient for E-learning resources utilization.

4 Suggestion on E-learning Resources Development to Chinese College

4.1 Renewal Education Idea on E-learning

To promote E-learning develop, in the United States, from President to congress, from central to the states, from the government to the enterprise, they all attend E-learning and gave great support to E-learning. American government leaders have personally called on congress, state and enterprise to increase investment to E-learning. These attention and support create favorable external environment for development of E–learning, improve people to recognition of the importance of E-learning.

In china, many leaders don't understand deeply to E-learning, don't know trend of the development of education. So, leaders and people don't give full attention to E-learning. It must improve their information literacy, guide they attend E-learning.

Secondly, college should pay attention to E-learning resources development, actively adopt mordent education technology.

Thirdly, college teachers must realize that time has into information age, multimedia technology, network technology and digital technology is necessary, and E-learning will be ideal study form on education in the future. So, now, they should pay full attention to E-learning, and do well work on E-learning resources development.

4.2 Constructing Development Platform

For the platform construction, resource integration view is must be set up. Following the basic principles of integration should be taken in construction E-learning resource development platform. These principles are optimizes structure, innovates management ideal and integrate technology. Platform construction should meet the following requirement:

It should put all E-learning resources into a unified planning, unified development platform and unified technical standard, as far as possible. Only in this way, we can guarantee the platform meet generality, replacement, compatibility and sharing.

```
┌─────────────────────┐   ┌─────────────────────┐   ┌─────────────────────┐
│ Resources collection│──▶│Resources classification│──▶│  Resources storage  │
└─────────────────────┘   └─────────────────────┘   └─────────────────────┘
           ▲                                                     │
           │                                                     ▼
┌─────────────────────┐   ┌─────────────────────┐   ┌─────────────────────┐
│  Resources spread   │◀──│  Resources update   │◀──│   Resources share   │
└─────────────────────┘   └─────────────────────┘   └─────────────────────┘
```
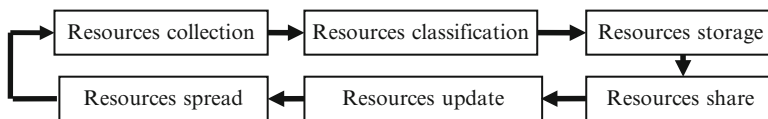
Fig. 3 Platform function figure

It should break resources situation, handle the segmentation investment and return, management and limits of authority, ownership and use right relationship, establish flexible, fair and reasonable sharing mechanism.

It should clean, classify, eliminate resources, and timely to be supplemented, upgrade, the transformation and updating, and make the teaching resources not only complete facilities, also keep pace with the times, maintain novelty and advanced.

This platform must have following function, shown in Fig. 3. The E-learning resource platform should have resources collection, resources classification, resources storage and resources share etc. functions in Fig. 3.

When Chinese colleges construct platform of E-learning resources development, they should reference commercial company development platform, especially well-known companies, for example Apple Inc' iTunesU, NTES's open courses plan etc. these platforms that commercial company has developed are matured, applied widely.

Of course, this platform construction needs lot of investment; single college is hard to undertake this investment. So, it should establish college alliance to develop this platform.

4.3   Integrating All Kind of E-learning Resources

Firstly, college should integrate itself resource; fully improve teachers, students' initiative and enthusiasm, guide them become E-learning resource development subject. Through this, teachers and students can take themselves E-learning resource into the platform, update E-learning resource. Initial stage, colleges should reward the teachers and students contributing their resources to promote E-learning resources development. While, it should shape university culture to support resources share.

Campus cultural of E-learning resource sharing atmosphere is a crucial soft environment. This cultural atmosphere needs schools vigorous advocacy and mobilization. Building Fairness and trust of campus culture can create atmosphere of resources sharing, and can support E-learning resources develop.

Secondly, it should integrate and develop resources between universities. China colleges should establish open spirit, learn famous universities practice, for example, Harvard, Yale etc., go out, and share oneself resources. Resources integration between colleges can abide by steps following, shown in Fig. 4. In Fig. 4, regional

Integration between territory colleges	→	Integration between cross-regional colleges
Integration between abroad l colleges	←	Integration between domestic colleges
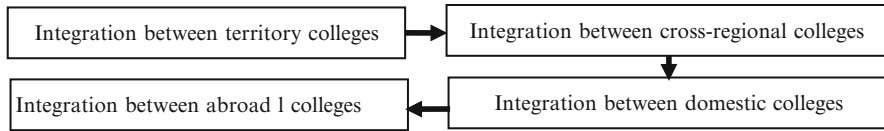
Fig. 4 Resources integration steps figure

universities must first be integrated, and then cross regional universities were be integrated. At the right time, Chinese colleges can integrate foreign college E-learning resources.

The Government should play a leading role in resources integration, encourage colleges integrate E-learning resources. The government can formulate relevant policy to guide college integrate, can also build a platform to integrate all colleges E-learning resources, for example, national level excellent course construction project.

Thirdly, it should integrate and develop E-learning resources of social resources. Lots of companies open public benefit course. We can integrate those resources through network. These companies also willing contribute their resources to college.

5   Conclusion

This paper analyzes the situation and problems on E-learning resource development of Chinese University. E-learning resources development has achieved great progress, and has accumulated much of E-learning resources in Chinese college. But, there were many problems that lacking initiative, lacking integration, lacking organization and lacking platform. Chinese college should renewal education idea, construct development platform, etc.

References

Aldrich C (1996–2006) E-learning in the workshop. TD, Cambridge, England

Brown JS, Collins A, Duguid P (1989) Situated cognition and the culture of learning. Educ Res 18(1):32–42; Article in a conference proceedings, Champaign

Chan T-W, Roschelle J, Hsi S, Kinshuk K et al (2006) One-to-one technology enhanced learning: an opportunity for global research collaboration. Res Pract Technol Enhanc Learn 1(1):3–29

Mandinach EB (2003) The development of effective evaluation methods for E-learning. A concept paper and action plan; effective evaluation methods

Nick Van Dam (2003) Implementation lessons and case studies from Companies that making e-learning work. The e-learning field Book, Macgraw Hill

Russet A (2001) The ASTD E-learning handbook. McGraw-Hill Trade, New York

Siemens G (2005) Connectivism: a learning theory for the digital age. Instr Technol Distance Learn 2(1):3–10

Weiser M (1991) The computer for the twenty- first century [J]. Sci Am 3:94–100

Accurate Train Stopping by Model Following Sliding Mode Control

Renshi Luo, Zhenyu Yu, and Tao Tang

Abstract Accurate stopping control is crucial to ensure safe operation of urban mass rapid rail transit. In this paper, automatic trajectory tracking control of urban rail vehicles during the stopping phase is investigated and a model following sliding mode control algorithm is applied. It is shown that the proposed control algorithm can effectively tackle the external environment disturbances and achieve high tracking performance. Theoretical analysis and numerical simulation also confirm the effectiveness of the proposed methods.

Keywords Accurate stopping • Sliding mode control • Model following • Urban rail vehicle • Automatic train operation

1 Introduction

Accurate stopping control is one of the most important technologies that enable high efficient and cost-effective operation of urban mass rail transit system. One of the challenging issues related to the design of automatic train operation (ATO) system is that the dynamics of the train are coupled with disturbances, which calls for high performance controller. There have been several literatures studying the problem of train automatic operation and various control approaches, such as PID control (Tang Tao and Huang Liang-ji 2003), fuzzy control (Yasunobu et al. 2002), neural-fuzzy control, predictive fuzzy control (Oshima et al. 1988) have been proposed. Although some of these literatures have achieved inspiring breakthroughs, all of them used the model-free control approaches which have several drawbacks, such as being unable to tackle the uncertainties and external disturbance directly, causing the controller

R. Luo (✉) • Z. Yu • T. Tang
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing, China
e-mail: 09120347@bjtu.edu.cn; zhyyu@bjtu.edu.cn; ttang@bjtu.edu.cn

to have complex structure as well as requiring the controller designers have rich engineering experiences. By contrast, sliding mode control has the advantages of order reduction, decoupling design procedure, insensitivity to parameter variations, disturbance rejection and simple implementation. Although having several advantages, sliding mode control hasn't been used in the research of ATO.

This paper investigates distance tracking control of urban rail vehicles during the stopping phase. A braking model of urban rail vehicle proposed by Yu Zhen-yu and Chen De-wang (2011) is considered in which the major external disturbances are explicitly included. Considering the external disturbance due to varying environment, a sliding mode control algorithm is derived. It is shown that the designed controller is able to achieve good tracking performance and stopping accuracy.

The remainder of the paper is organized as follows. Section 2 presents the dynamic braking model. Section 3 presents the approximation model used in the controller design. Section 4 develops the sliding mode controller. Section 5 presents numerical simulation results and finally Section 6 offers the conclusion of this study.

2 Braking Model of Urban Rail Vehicle

Braking system is part of the train control system; it is used to adjust the speed and stop the train and is critical to ensure the safe operation of the train. The main function of the train braking system is to achieve the same braking performance characteristics, which is controlled and managed by the braking controller (Yu Zhen-yu and Chen De-wang 2011). ATO is designed to substitute drivers; it generates the traction or braking command according to the target speed, driving license and track conditions. ATO cannot directly manipulate the train power actuators, such as motors and internal combustion engines, while it control the vehicle through vehicle's own traction control system and braking control system. Therefore, the vehicle dynamics braking model is the motion model including the braking control system (Yu Zhen-yu and Chen De-wang 2011), which is shown below (Fig. 1) where u denotes the command input, a_t denotes the target acceleration, $\hat{a}$ denotes the control acceleration, v denotes velocity and s denotes distance. $F(\cdot)$ denotes the mapping between braking command and target acceleration, it is a linear or non-linear invertible function. T is the system transport delay and τ is system response time. d denotes the environment disturbance which is mainly caused by ramp slope and curve. With the existence of transition slope and transition curve, d is bounded.

In this study, we take a_t as the virtual input and focus on the design of it, while the actual system input u is obtained via $u = F^{-1}(a_t)$ (Luo Ren-shi et al. 2012).

Fig. 1 Braking model of urban rail vehicle
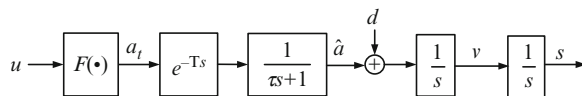
3 Model Approximation

To simplify the analysis, we use the *pade* function in *Matlab* to get the approxima-tion of the transport delay in the Model, which is

$$e^{-Ts} \approx \frac{-s + \lambda}{s + \lambda} \tag{1}$$

where λ is a constant decided by T. Accordingly the approximation model is given as

In order to get the state-space equation expression, we reorganize Fig. 2 and get the equivalent model, which is shown in Fig. 3

Selecting distance s, velocity v, control acceleration $\hat{a}$ and intermediate variable a_z to be state variables, we have the following state space equation

$$
\begin{cases}
\dot{x} = Ax + ba_t + f_d \\[2mm]
x = \begin{bmatrix} s & v & \hat{a} & a_z \end{bmatrix}^T \quad b = \begin{bmatrix} 0 & 0 & -\dfrac{1}{\tau} & 2\lambda \end{bmatrix}^T \\[3mm]
f_d = \begin{bmatrix} 0 & d & 0 & 0 \end{bmatrix}^T \\[3mm]
A = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & -\dfrac{1}{\tau} & \dfrac{1}{\tau} \\
0 & 0 & 0 & -\lambda
\end{bmatrix}
\end{cases} \tag{2}
$$

In the following section, we focus on the distance tracking problem and use model following approach to design a sliding mode controller based on the approximation model.
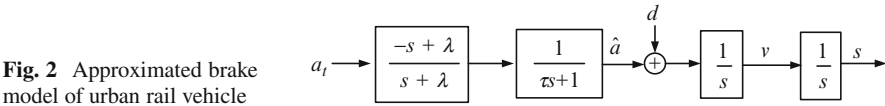
Fig. 2 Approximated brake model of urban rail vehicle
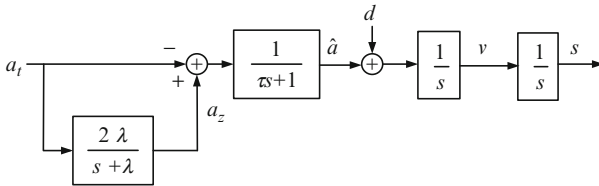
Fig. 3 Equivalent model of Fig. 2

4 Controller Design

There are two main approaches to design a sliding mode tracking controller, which are model following approach and integral-action approach. However, the model following approach avoids the difficulty of performance specification because the reference model will specify the design objectives; the controller is required to minimize the tracking between the model and plant (Edwards and Spurgeon 1998). In this section, we will design a sliding model controller using model following approach.

4.1 Reference Model and Error Dynamics

As the first step, we specify a reference model which generates the desired states,

$$\dot{x}_m = A_m x_m + b_m r$$
$$x_m = \begin{bmatrix} s_d & v_d & \hat{a}_d & a_{zd} \end{bmatrix} \tag{3}$$

where r is the reference input. Denote the state tracking error as follow

$$e = x - x_m \tag{4}$$

then by subtracting Eq. (2) from Eq. (3), the error dynamics is obtained as

$$\dot{e} = A_m e + (A - A_m)x + b a_t - b_m r + f_d \tag{5}$$

It should be noted that in order to use model reference approach the following match condition has to be satisfied

$$A = A_m - bF \tag{6}$$

$$b_m = bG \tag{7}$$

Substituting A and b_m using Eqs. (6) and (7), Eq. (5) can be reorganized as

$$\dot{e} = A_m e + b(a_t - Fx - Gr) + f_d \tag{8}$$

With this error dynamics model, next we will focus on the stabilization problem of Eq. (8).

4.2 Design of Sliding Mode Control Law

For error system Eq. (8), we define the switching function as

$$\sigma(t) = Se(t) \tag{9}$$

Substituting Eq. (10) into Eq. (8), the error dynamics can be written as

$$\dot{\sigma} = SA_m e + Sb(a_t - Fx - Gr) + Sf_d \tag{10}$$

where $\|f_d\| = |d|$ satisfies the bounded condition. Ignore the effect of f_d and make the right-hand side of (10) be 0, then the equivalent control is obtained as

$$u_l = -(Sb)^{-1}SA_m e + Fx + Gr \tag{11}$$

where $S = b^T P$ and P is the solution of the following Riccati equation

$$PA_m + A_m^T P - Pbb^T P + Q = 0 \tag{12}$$

$F = b^+(A_m - A)$ and $G = b^+ b_m$, b^+ denotes the Moore-Penrose pseudo-inverse matrix of b. The nonlinear control is defined to be

$$u_n = -\rho\,\mathrm{sgn}(\sigma) \tag{13}$$

where ρ is the control parameter and if it satisfies $\rho > k\|f_d\| + \gamma$, where $k > 0$ is decided by S and b, then it can be proved that as time goes infinite, $\sigma(t)$ goes toward zero and $\gamma > 0$ ensure that the sliding surface is reached in finite time (Edwards and Spurgeon 1998). However, the *sign* function will cause the chattering phenomenon which is highly undesirable in practical use. There are several methods to avoid chattering; here we choose the following function (Khalil 2007) to replace the *sign* function

$$\phi(\sigma) = \frac{\sigma}{|\sigma| + \omega} \tag{14}$$

When replacing *sign* with ϕ, the ideal sliding motion will not exist, but pseudo-sliding motion, during which $\sigma(t)$ is forced to remain within an arbitrary small boundary (decided by ω) about 0, is still reached according to the theorem that the boundedness of ω ensures the uniform ultimate boundedness of $\sigma(t)$ (Yoshizawa 1958). Further, as $\sigma(t)$ remains within a small boundary about 0, according to Eq. (9), we know that $e(t)$ also remain with a small boundary about 0, and then the accurate tracking is achieved.

In addition, in order to shorten the sliding motion reaching phase, we add one more feedback control, which is given as

$$u_m = -\mu(Sb)^{-1}\sigma \tag{15}$$

where $\mu > 0$. Finally, the virtual control a_t comprises three components, specifically

$$a_t = u_l + u_n + u_m$$

$$= -(Sb)^{-1}SA_m e + Fx + Gr - \rho \frac{\sigma}{|\sigma| + \omega} - \mu(Sb)^{-1}\sigma \qquad (16)$$

and the actual system input is obtained by

$$u = F^{-1}(a_t) \qquad (17)$$

5   Simulation

To verify the effectiveness of the designed sliding mode controller, we conduct a simulation of distance tracking based vehicle stopping control with track curve disturbance in Matlab/Simulink.

The parameters of the braking system are $T = 1.2\ s$ and $\tau = 0.4\ s$. By using *pade* function, the parameter λ corresponding to $T = 1.2\ s$ is 1.6667. The reference model, which satisfies the model match condition, is chosen as (Fig. 4)

When written into state space expression, the reference model is given by Eq. (3), where

$$A_m = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$b_m = \begin{bmatrix} 0 & 0 & -0.75 & 1 \end{bmatrix}^T$$

The initial conditions for the simulation are $s(0) = 0\ m$ and $v(0) = 20\ m/s$. The target stopping position, which is the command input by system supervisor, is chosen to be 300 m. After the target position is given, the corresponding reference input is calculated and here the one corresponding to 300 m is given by

$$r = \begin{cases} -4, & \text{if } v_d \geq 5 \\ 0, & \text{otherwise} \end{cases}$$

The conditions of the track are chosen as follows; the section between 50 and 200 m is a curve section with radius of 200 m and the other sections are straight. Besides we also consider the constraint of controller saturation as well as the distance and velocity measurement noise, of which the control signal a_t has the
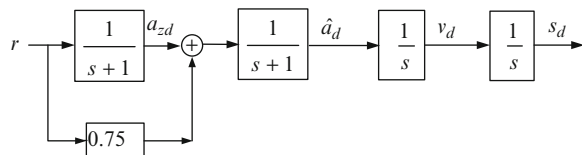
Fig. 4 Structure of reference model
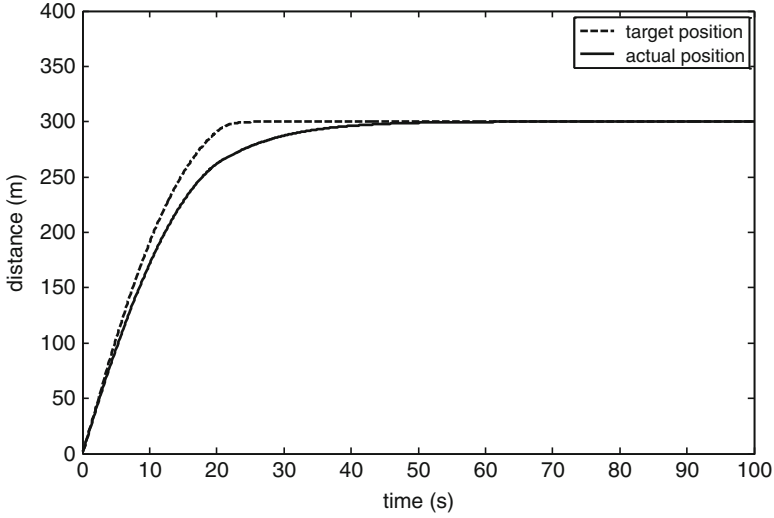
Fig. 5 Target position curve and actual position curve
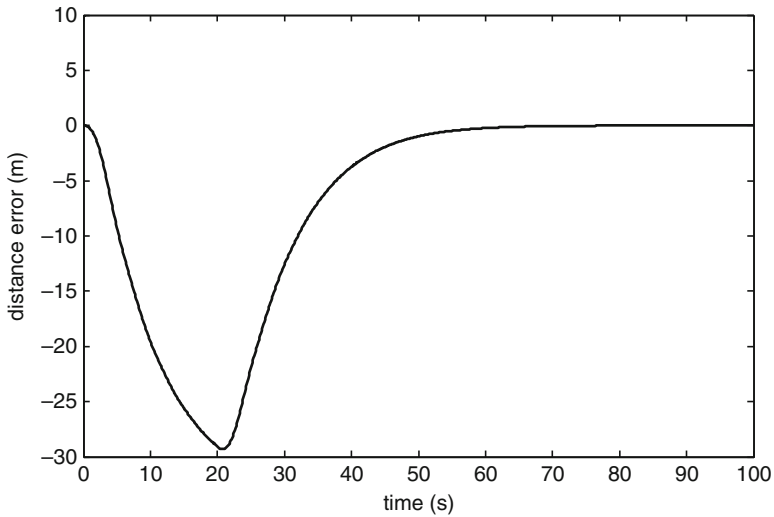
Fig. 6 Tracking error

upper bound 1 and lower bound −1 and the distance and velocity measurement noise are both the uniformly distributed random numbers with the mean 0 and standard variance 0.05. The controller parameters are calculated or chosen as $F = \begin{bmatrix} 0 & 0 & -2.16 & 0.344 \end{bmatrix}$, $G = 0.3$, $S = \begin{bmatrix} 1 & 8.52 & 7.44 & 6.79 \end{bmatrix}$, $\rho = 0.1$, $\omega = 1$, $\mu = 0.2$.

Simulation results are given as follows. Figure 5 shows both the target position curve and actual position curve, which are dashed and solid respectively. Figure 6 is the plot of the tracking error. Figure 7 is the plot of virtual control signal a_t. Figure 8
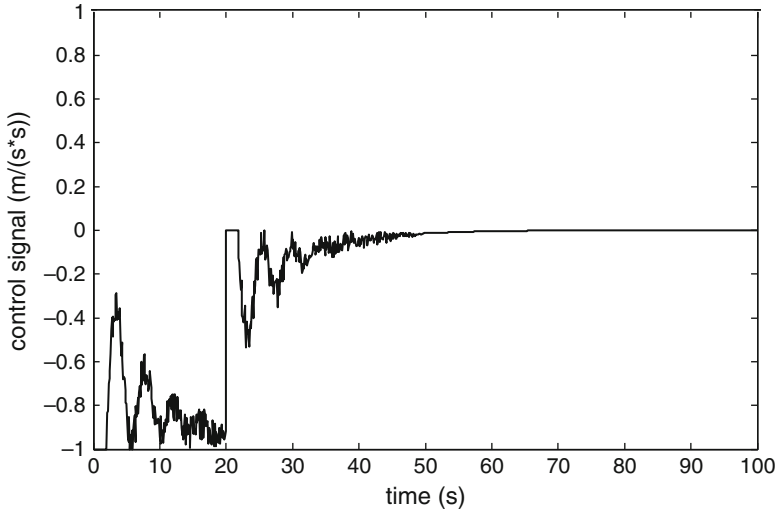
Fig. 7 Control signal a_t
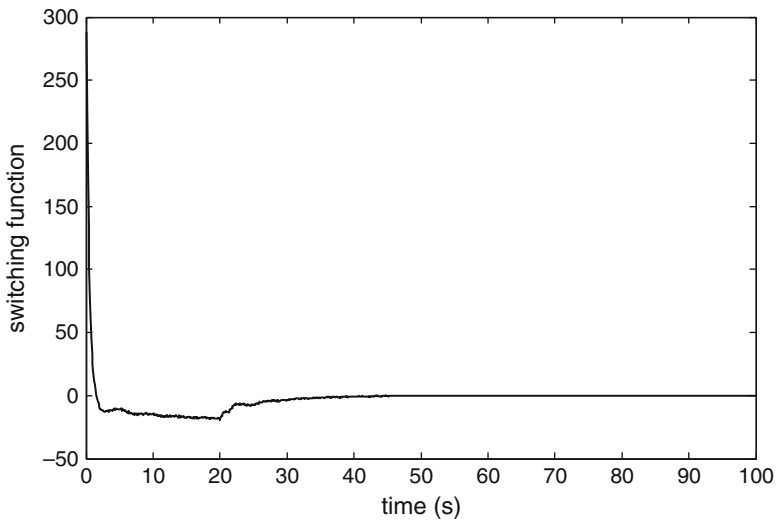
Fig. 8 Switching function

plots the switching function. Finally, Fig. 9 shows the distance-velocity curve of vehicle's actual movement. Note that although the controller saturation constraint and distance and velocity measurement noise are considered and the environment disturbance is unknown, good tracking precision still maintains with high stopping accuracy.
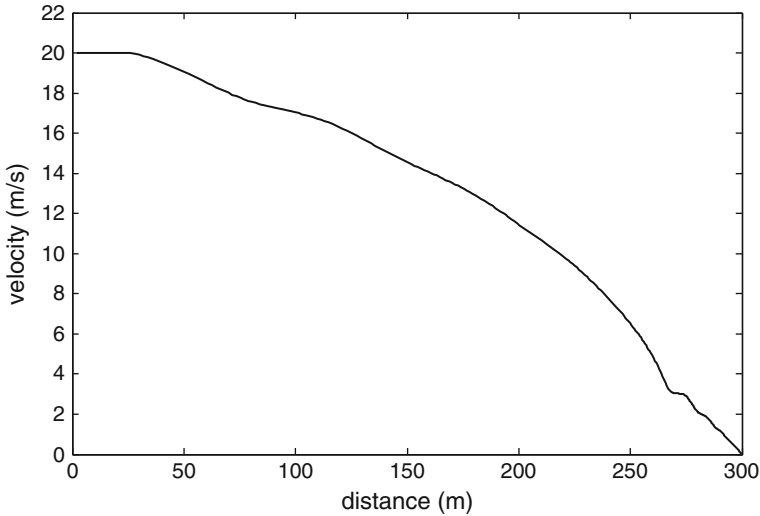
Fig. 9 Distance-velocity curve of vehicle's movement

6 Conclusion

This paper investigated the problem of accurate stopping control of urban rail vehicles. A dynamic braking system coupled with environment disturbance was used for control design. A sliding mode controller that can reject the environment disturbance to achieve high stopping accuracy is proposed. Theoretical analysis and numerical simulation verify the effectiveness of the proposed control algorithm.

References

Edwards C, Spurgeon SK (1998) Sliding mode control: theory and application. Taylor and Francis Ltd, London
Khalil HK (2007) Nonlinear system, 3rd edn. Publishing House of Electronics Industry, Beijing, pp 230–267
Luo Ren-shi, Wang Yi-hui, Yu Zhen-yu, Tang Tao (2012) Adaptive stopping control of urban rail vehicle. J China Railw Soc 34(4):64–68
Oshima, Yasunobu S, Sekino S (1988) Automatic train operation system based on predictive fuzzy control. In: Proceedings of international workshop on artificial intelligence for industrial applications, IEEE Industrial Electronics Society, Hitachi City, pp 485–489

Tang Tao, Huang Liang-ji (2003) A survey of control algorithms for automatic train operation. J China Railw Soc 25(2):98–102

Yasunobu S, Miyamoto S, Ihara H (2002) A fuzzy control for train automatic stop control. Trans Soc Instrum Control Eng 2(1):1–9

Yoshizawa T (1958) Liapunov's function and boundedness of solutions. Funkcialaj Ekvacicj 2:71–103

Yu Zhen-yu, Chen De-wang (2011) Modeling and system identification of the brake system of urban rail vehicle. J China Railw Soc 33(10):37–40

Building Occupant Evacuation Models Inside Buildings – Approach, Progress and Solutions

Rong Xie and Xin Luo

Abstract Occupant evacuation is currently becoming a challenging research direction in the field of emergency management. Approach to evacuation modeling is firstly proposed in the paper, which is mainly consisted of four core components, like spatial model, crowd movement model, individual behavioral model and psychological model. Secondly, the related solutions and development about these models in the recent studies are reviewed. The establishment of general process of evacuation modeling is then presented in the paper. Some existing issues and future directions are finally discussed.

Keywords Occupant evacuation • Occupant evacuation model • Evacuation behavior • Simulation • Emergency management

1 Introduction

With the increasing number of crowded places, people become more and more gathering, causing high pressure on emergency evacuations. Occupant evacuation is increasingly becoming a challenging research direction in the field of emergency management. Occupant evacuation is how to determine an optimal evacuation route for occupants when emergency events occur. For instance, people wish to conduct quantitative analysis, evaluation and emergency plan on evacuation safety performance of large and complex buildings (theaters, stadium etc.), which can

R. Xie
International School of Software, Wuhan University, 39 Luoyu Road, Wuhan 430079, China
e-mail: xierong@whu.edu.cn

X. Luo (✉)
State Key Lab of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan 430074, China
e-mail: mexinluo@hust.edu.cn

provide with optimal escape routes and rescue decision-making when various types of emergency events occur. Currently, researchers and engineers have been doing a lot of work and made a series of useful research results on the aspects of evacuation modeling, behavioral analysis, evacuation simulation, as well as applications. There have been about 30 kinds of evacuation models and corresponding software available (Kuligowski and Peacock 2005), applied to wide applications, such as building design safety analysis, crowd management, safety engineering etc. Evacuation models become important sources for the understanding of evacuation processes. However, most existing evacuation models are either computationally inefficient, or missing some crucial human behaviors in crowds (Song et al. 2006), making evacuation results difficult to comprehensively reflect a more realistic evacuation. Therefore, how to effectively describe evacuation behavior and how to represent decision-making process are the key issues to evaluate whether an evacuation model is successful or not. Under surveys on some effective solutions in the aspects of occupant evacuation modeling that researchers and engineers put forward in the recent years, the paper proposes evacuation modeling method, as well as four core components of modeling, including environmental spatial model, crowd movement model, individual behavioral model and psychological model.

The rest of the paper is organized as follows. Section 2 proposes an approach to occupant evacuation modeling and core components of the model. The establishment of general process of evacuation modeling is then presented in Sect. 3. Some existing issues and future research directions are discussed in Sect. 4.

2   Approach to Occupant Evacuation Modeling

Large-scale occupants' evacuation is a special phenomenon of crowd flow. Bryan (1977) thinks that individual behavior under emergency is a complex decision-making processes, including cognition, recognition, definition and evaluation etc. The individual interaction during evacuation, i.e. occupant-occupant (i.e. interaction with other occupants), occupant-structure (i.e. interaction with enclosing building) and occupant-environment (i.e. interaction with hazards, fire earthquake etc.), directly affects complex group behavior. It is required to represent the three types of intersection in modeling. These interactions can produce congestion, bottleneck, jamming towards exits, affecting an occupant's behavior, and therefore trigger the decision-making process, so that behavior analysis is further complicated by the way in which occupants interact with their surroundings on the three levels of psychology, physiology, and sociology. Hence, evacuation modeling process is required to describe different characteristics of individuals in the different evacuation choices during the applications of decision-making, simulation and analysis. The main characteristics can be summarized as, (1) Spatial characteristics of enclosing building; (2) Motion characteristics of evacuation crowd; (3) Behavioral characteristics of an occupant; (4) Psychological characteristics of an occupant.
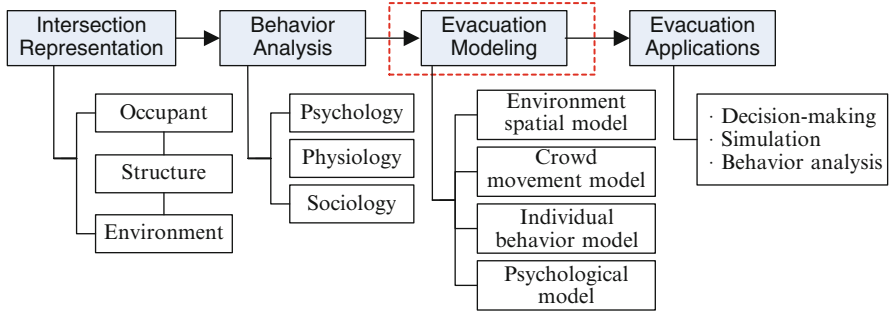
Fig. 1 Basic tasks of occupant evacuation process

Figure 1 shows basic tasks of the whole occupant evacuation process, such as intersection representation, behavior analysis, evacuation modeling and applications. Among them, evacuation modeling provides with systems or methods which can simulate and evaluate these main factors, including environmental spatial model, flow movement model, individual behavior model, and psychological model.

2.1 Environmental Spatial Model

Fahy (1993, 1996, 1999, 2001) proposes an EXIT89 gridded model. The input to the model includes a network description of building, representing room or sections of rooms or corridors with nodes, connecting nodes with arc, while each occupant moves from node to node. This model does not take into account geometry of building partitions, and occupant cannot move in the regional unit of the same building. So, exact locations of occupants are actually ignored in the model. Some models divide building partition into smaller grid cells. Thompson and Marchant (1994, 1995a, b) propose Simulex model, dividing building plane into 0.2 m × 0.2 m square grid. Ketchell et al. (1994, 1995) propose EGRESS model using a hexagonal grid. Relative to the EXIT89 model, these models can more accurately describe geometry shape of architecture space and positions of obstacles in building, and can simulate occupants to move inside grids or network nodes. However, it needs to have sufficient fine grids to describe status of occupant evacuation. For large evacuation area, discretization will yield massive grid data which making real-time processing and simulation application difficult.

Considering these issues of gridded method, some researchers propose topological modeling method (Bruckner et al. 2005; Krieg-Bruckner and Shi 2006; Lorenz et al. 2006; Thrun 2003). Different from gridded method, it does not represent exact location and size of building objects, but topological relationship among objects in building. For example, Thrun (2003) proposes topological method and its procedures, which can divide free spaces into several regions and establish region-based

topology by critical line of grid map. It is a compact spatial representation method. The model is simple and easy to implement real-time updates. But such model is inaccurate, and the found route-choice is not optimal.

Geometry-based spatial modeling is a common method used in the application of path planning of mobile robot. Some geometric modeling methods are MAKLNK Graph (Habib and Asama 1991), generalized Voronoi (Wallgrun 2005), Visibility Graph (Kitzinger 2003) etc. Therefore it is often applied to environment spatial modeling of occupant evacuation. But using geometric method, complexity of time and space will be relatively high. With the increase in the number of obstacles, structural complexity increases exponentially. Therefore, this method is suitable to some open evacuation regions, like squares.

2.2 Crowd Movement Model

Fruin (1971) gives relationship curve between average moving speed and population density by statistics. After this study, there are a lot of researchers make statistics of relationship between walking speed and density in different regions and different spaces. Henderson (1971, 1974) gives probability distribution formula of moving speed and derives relationship between moving speed and people and environmental factors (such as gender, age, forward way, location etc.). This early work opened up a road towards the later research. On the basis of this study, some scholars propose to establish mechanical equation of population movement using fluid dynamics and thermodynamics. Bradley (1993) proposes Navier–Stokes equations using fluid mechanics to describe population movement in high density areas. Helbing et al. (2002) presents Maxwell Boltzmann distribution law using thermodynamics to establish distribution equation of population movement.

Hoogendoom and Bovy (2004) proposes a more complex gas dynamic pedestrian model and establishes relationship between density function and position, velocity, desired speed. The model describes change of crowd density under continuous or non-continuous interaction. This model is microscopic in the sense. But such model ignores effects of obstacles and infrastructure on crowd movement. Also large-scale crowd, calculation time will be more relatively.

Lovas (1994) proposes a stochastic model to simulate pedestrian traffic flow. In the model, pedestrian facility is modeled as a network of walkway sections, and pedestrian flow is modeled as a queuing network process, where each pedestrian is treated as a separate flow object, interacting with the other objects. The process of change is described using equation to calculate the expected values of number of evacuees and evacuated evacuees. Some similar models include EvacSim model (Poon et al. 1994) etc. These queuing network models are simple and require less calculation. But population in queuing network, as a whole, represents the same movement characteristics. On the other hand, interaction and psychology are not reflected in the model yet.

2.3 Individual Behavior Model

Takahashi et al. (1989) present a fluid model to predict and evaluate evacuation time of people from a low level hazard. The assumption of this model is that people move like a fluid with movement equation. Okasaki and Matsushita (2004) propose that movement of each pedestrian is simulated by the motion of a magnetized object in a magnetic field. Magnetic force which acts on a pedestrian from a magnetic pole is basically calculated by equation according to Coulomb's Law. Completely determined by function(s), occupant's movement and behavior will react in a deterministic manner to its influences, and therefore these models will be limited in the same way for each occupant.

Rule-based models assign individual actions and decision based on pre-defined sets of rules which can be triggered in specific circumstances or conditions. Stahl (1982) simulates the perceptual and behavioral responses of building occupants involved in fire emergencies with some behavioral rules in BFIRES model. Levin (1988) proposes that occupants make decisions based on a large set of decision rules which are a function of smoke conditions in the building. Ozel (1991) thinks each object in an intelligent simulation system needs to be associated with logic rules that govern its operation.

The above styles of decision-making process are in simplistic methods. All objects in the system are logically interconnected with each other and with the processes that could influence their behavior. The same decisions are taken under the same circumstances in a deterministic fashion. This is obviously not advantage of denying the possibility of natural variations. CRISP model (Boyce et al. 1998; Fraser-Mitchell 2001) improves it and views occupants as individuals by giving the occupants certain behavioral roles. These certain behavioral activities taking place during evacuation are in a probabilistic fashion. buildingEXODUS (Gwynne et al. 1998; Parke et al. 2003) incorporates deterministic and stochastic approaches depending on the circumstances.

In the recent years, many behavioral models tend to use agent technology to simulate human and social behaviors during emergency evacuations. Thompson et al. (1996) define individual as an agent with certain characteristics like age, gender, social experience and movement capacity. Physical characteristics, such as walking speed, are determined by population density. Individuals may have some basic behaviors under emergency situations (bypass, turn etc.). They can adjust their decisions according to the changes in the surrounding circumstances. Williams (2004) proposes Legion model which views occupants as an intelligent agent with social, physical, and behavioral characteristics. These characteristics make up a profile for each person. These models can accurately represent decision-making process, but weakened the capacity that user control over the model, so accuracy of simulation is unpredictable.

Another agent-based application is called multi-agent, which have been found sufficient to represent complex human behavior and decision making process. Pan et al. (2006) has prototyped a multi-agent system to model emergent human social

behaviors, such as competitive, queuing, herding, and bidirectional crowd flow through simulating the behavior of human agents at microscopic level. In order to study the role of leadership during evacuation process, Murakami et al. (2002) think leaders are presented in many real-world situations and introduce a leadership in simulator. The system describes the behavior of each leader and evacuee as been modified to represent the interaction among leader and other occupants. Pelechano and Badler (2006) also simulate a leader agent who is trained and has complete knowledge about the internal building connectivity and would help others during the evacuation process. The leadership and other human responses during an evacuation process have also been represented by Sugimoto (2006) using multi-agent simulation in virtual space.

2.4 Psychological Model

Thompson and Merchant (1994, 1995a, b) presents Simulex model which can simulate psychological behavior of Evacuees, like response time to fire alarm etc. An ORSET (Occupant Response Shelter Escape Time) model (Sime 2001) is advocated as a framework unifying aspects of psychology, architecture, engineering and facilities management. It might also mention that Gwynne et al. (1998) and Fahy and Proulx (1995) have been researching on evacuation modeling under perspective of occupants's behavior restricted to the psychological and sociological aspects. Kuligowski and Peacock (2005) launches study during emergency evacuation, including personnel decision-making and response to environment, assessment methods of fire impact on occupants, self reaction during fire. Helbing (1992) presents social force model based on self-driven particle system framework, assuming that individuals have ability of thinking and responding to the surrounding environment. Psychological response is quantified to the force, and is added to evacuation model to simulate typical effects of evacuation behavior.

In an emergency situation, human psychological reactions generally make people represent some behaviors, such as tending to choose their own familiar path and environment, passive agreement, and blindly herding behaviors. Current research does not take into account important aspects associated with the psychological attributes of the occupants yet.

3 Discussion

Environmental spatial model represents geographical position of evacuation environment, such as barriers, signage, passage, exits, stairs and elevators. Accurate models can be helpful to effective evacuation route planning and decision making which is the basis of occupant evacuation under emergency. In general, we can establish a spatial network model using a combination of topological method and

geometrical method. The model builds subspace node network using topological method, records geometric information of buildings in the unit of nodes, which can be a good solution. Another effective solution is a data modeling through the use of integration of vector and raster, similar to research in GIS. The method uses vector data accurately represent location and topological relationship between entities. Raster data is used for spatial analysis and operation of entities. Data conversion is achieved through combination of vector data and raster data.

Evacuation is a dynamic process of flow movement which combines a large number of individual behaviors. Crowd movement model describes the effects that impact population density, people walking speed and flow impact on crowd movement.

Individual behavior model aims to describe behavior characteristics during the whole evacuation, including behavior interaction on the three levels of psychology, sociology, physiology. How to properly and effectively describe the interaction between them in the model is the key to evaluate whether evacuation model is successful or not. An agent-based system is particularly suitable for simulating individual decision-making process and behavior to explore macro phenomena. Particularly, multi-agent simulation has been widely accepted as a promising approach to model complex emergent phenomena. We can define human intelligence throughout the evacuation, where occupants are regarded as active agents taking into consideration their response to various hazards and individual behavior. Behavior rules can be designed for each agent in the model, like: (1) An objective/goal; (2) Attempt to maintain a minimum distance with others; (3) Attempt to maintain unimpeded velocity. Also each agent may have its quantitative psychological factors which can be also defined in the model.

Under the establishment of spatial model, crowd movement model, individual behavioral model and psychological Model, route-choice can be planned optimal or near optimal evacuation routes for all evacuees. There are some factors of evacuation route-choice, such as distance from exit, obstacle, personnel distribution density in the direction of escape, determination of multiple exits, which it is necessary to consider in the model.

4   Conclusions and Future Work

This paper presents method of occupant evacuation modeling inside buildings, including spatial environment model, crowd movement model, individual behavior model, and psychological model. Throughout the work done by researchers and their research results in recent years, precise description and quantification of evacuation problem and phenomenon are the key issues in the study. There is no model which can fully quantify various behaviors of people in the evacuation yet. A well-designed behavior model should allow users: (1) To define their own individual characteristics of evacuees. (2) To describe restrain state environment acts on evacuation, and interactions among occupant, structure and environment.

(3) To establish knowledge-base for evacuation decision-making applying to approaches such as heuristic rules. Based on these ideas, the in-depth study of behavior modeling will largely depend on the use of artificial intelligence technology to further improve ability to behavioral rules, knowledge base and intelligent reasoning.

To meet the social needs and new user demands, the study of intersection and integration of psychology and behavior is still in its infancy. How to introduce artificial intelligence technology to further deal with people as individuals with complex behavioral characteristics and certain decision-making capacity and establish intelligent evacuation model shall be a promising research direction. On the other hand, In order to carry out effective evacuation research, it is important to simulate individual behavioral characteristics and flow movement characteristics. Thus, computer simulation is becoming an effective way and prospectus research direction.

References

Boyce K, Fraser-Mitchell J, Shields J (1998) Survey analysis and modelling of office evacuation using the CRISP model. In: Shields TJ (ed) Proceedings of the 1st international symposium of human behaviour in fire. London, pp 691–702

Bradley GE (1993) A proposed mathematical model for computer prediction of crowd movements and their associated risks. In: Proceedings of the international conference on engineering for crowd safety. Elsevier, Amsterdam, pp 303–311

Bruckner BK, Frese U, Luttich K et al (2005) Specification of an ontology for route graphs. Spatial cognition IV, lecture notes in artificial intelligence 3343. Springer-Verlag Berlin Heidelberg, pp 390–412

Bryan JL (1977) Smoke as a determinant human behaviour in fire situation (project people), final report, NBS-GCR-77-94, University of Maryland, College Park, 304 pp

Fahy RF (1993) EXIT89: an evacuation model for high-rise buildings. In: Proceedings of international interflam conference, London, pp 306–311

Fahy RF (1996) EXIT89-high-rise evacuation model-recent enhancements and example applications. In: Proceedings of international interflam conference, Cambridge, England, pp 1001–1005

Fahy RF (1999) User's manual, EXIT89 v1.01, an evacuation model for high-rise buildings quincy, National Fire Protection Association, Quincy

Fahy RF (2001) Verifying the predictive capability of EXIT89. In: 2nd international symposium on human behavior in fire, Massachusetts Institute of Technology, Cambridge, Boston, USA, 26–28 March 2001, pp 53–64

Fahy R, Proulx G (1995) Collective common sense: a study of human behavior during the world trade center evacuation, NFPA J 87(2):59–67

Fraser-Mitchell J (2001) Simulated evacuations of an airport terminal building using the CRISP model. In: 2nd international symposium in human behaviour in fire, Boston, pp 89–100

Fruin JJ (1971) Designing for pedestrians: a level-of-service concept. Highway research record, number 355: pedestrians, Highway Research Board, Washington, DC, pp 1–15

Gwynne S, Galea ER, Lawrence PJ, Owen M et al (1998) Validation of the buildingEXODUS model. CMS Press, London, No.98/IM/29

Habib MK, Asama H (1991) Efficient method to generate collision free paths for Autonomous mobile robot based on new free space structuring approach. In: IEEE/RSJ international workshop on intelligent robots and systems lROS'91, Osaka, Japan, 3–5 November 1991, pp 563–567.

Helbing D (1992) A fluid dynamic model for the movement of pedestrians. Complex Syst 6:391–415

Helbing D, Farkas IJ, Molnar P, Vicsek T (2002) Simulation of pedestrian crowds in normal and evacuation situations. In: Schreckenberg M, Sharma SD (eds) Pedestrian and evacuation dynamics. Springer, Berlin, pp 21–58

Henderson LF (1971) The statistics of crowd fluids. Nature 229(5284):381–383

Henderson LF (1974) On the fluid mechanics of human crowd motion. Transp Res 8(6):509–515

Hoogendoom SP, Bovy PHL (2004) Pedestrian route-choice and activity scheduling theory and models. Transp Res B 38:169–190

Ketchell N, Cole SS, Webber DM (1994) The EGRESS code for human movement and behaviour in emergency evacuation. In: Smith RA, Dickie JF (eds) Engineering for crowd safety. Elsevier, London, pp 361–370

Ketchell N, Bamford GJ, Kandola B (1995) Evacuation modelling: a new approach. In: Proceedings of the 1st international conference on fire science and engineering, Kowloon, pp 499–505

Kitzinger J (2003) The visibility graph among polygonal obstacles: a comparison of algorithms, 66 pp. Available at: http://www.cs.unm.edu/~treport/tr/03-05/Kitzingerthesis.pdf

Krieg-Bruckner B, Shi H (2006) Orientation calculi and route graphs: towards semantic representations for route descriptions. In: GIScience, lecture notes in computer science 4197. Springer-Verlag Berlin Heidelberg, pp 234–250

Kuligowski ED, Peacock RD (2005) A review of building evacuation models. Technical note 1471, National Institute of Standards and Technology, Gaithersburg, 156 pp

Levin BM (1988) EXITT-A simulation model of occupant decisions and actions in residential fires. In: Proceedings of the second international symposium of fire safety science. Hemisphere Publishing Corporation, New York, London, pp 561–570

Lorenz B, Ohlbach HJ, Stoffel EP (2006) A hybrid spatial model for representing indoor environments. In: W2GIS, lecture notes in computer science 4295. Springer-Verlag Berlin Heidelberg, pp 102–112

Lovas GG (1994) Modeling and simulation of pedestrian traffic flow. Transp Res B 29B(6): 429–443

Murakami Y, Minami K, Kawasoe T, Ishida T (2002) Multi-agent simulation for crisis management. In: Proceedings of the IEEE workshop on knowledge media networking. IEEE, Los Alamitos

Okasaki S, Matsushita S (2004) A study of simulation model for pedestrian movement with evacuation and queuing. Available at: http://www.anc-d.fukui-u.ac.jp/~sat/ECS93.pdf

Ozel F (1991) Simulation of processes in buildings as a factor in the object representation of built environments. In: Proceedings of building simulation, Nice, France, pp 250–256

Pan X, Han CS, Dauber K, Law KH (2006) Human and social behavior in computational modeling and analysis of egress. J Autom Constr 15:448–461

Parke J, Gwynne S, Galea ER, Lawrence P (2003) Validating the buildingEXODUS evacuation model using data from an unannounced trial evacuation. In: Proceedings of 2nd International conference on pedestrian and evacuation dynamics conference. CMS Press, Greenwich, UK, pp 295–306

Pelechano N, Badler NI (2006) Modelling crowd and trained leader behavior during building evacuation. IEEE Computer Graphics and Applications 26(6):80–85

Poon LS (1994) EvacSim: a simulation model of occupants with behavioural attributes in emergency evacuation of high-rise buildings. In: Proceedings of the 4th international symposium of fire safety science, Ottawa, Ontario, Canada, pp 681–692

Sime JD (2001) An occupant response shelter escape time (ORSET) model. Saf Sci 38(2):109–125

Song WG, Yu YF, Wang BH, Fan WC (2006) Evacuation behaviours at exit in CA model with force essentials: a comparison with social force model. Phys A Stat Theor Phys 371(2):658–666

Stahl FI (1982) BFIRES-II: a behavior based computer simulation of emergency egress during fires. Fire Technol 18:49–65

Sugimoto Y (2006) Modeling action rules through participatory simulation in virtual space. Master thesis, Kyoto University, Japan

Takahashi K, Tanaka T, Kose S (1989) An evacuation model for use in fire safety design of buildings. In: Proceedings of the second international symposium on fire safety science. Hemisphere, London, pp 551–560

Thompson PA, Marchant EW (1994) Simulex; developing new computer modelling techniques for evaluation. In: Proceedings of the 4th international symposium of fire safety science, Ottawa, pp 613–624

Thompson PA, Marchant EW (1995a) A computer model for the evacuation of large building populations. Fire Saf J 24(2):131–148

Thompson PA, Marchant EW (1995b) Testing and application of the computer model 'SIMULEX'. Fire Saf J 24(2):149–166

Thompson PA, Wu J, Marchant EW (1996) Modelling evacuation in multi-storey buildings with Simulex. Fire Eng 56:7–11

Thrun S (2003) Towards a framework for human-robot interaction. Hum Comput Interact 19(1):9–24

Wallgrun JO (2005) Autonomous construction of hierarchical Voronoi-based route graph representations. In: Freksa C et al (eds) Spatial cognition IV, lecture notes in artificial intelligence 3343. Springer-Verlag Berlin Heidelberg, pp 413–433

Williams A (2004) Go with the flow. Archit J 219(6):42–46

The Research of Data Quality Problems in Power Enterprise Data Integration

Kehe Wu, Cheng Duan, and Yayun Zhu

Abstract After years accumulation of informationization construction, power enterprises have established huge and complex application systems, which include multi-class data related to business like production, operation and management. Along with the demands of developing new business application, the data integration is urgently needed to grasp overall data view; however the complex data integration environment causes the data quality problem particularly outstanding. Under the background of the power industry, this paper first studies the data quality definition and dimension recognition; then proposes a data integration platform architecture and studies the data quality control methods in data integration; elaborates data quality assessment methods and procedures; in the view of data product, designs a total data quality management (TDQM) process finally. In a typical case of application, the results show the practicability of the proposed method.

Keywords Data integration • Data quality • Quality control • Quality assessment • TDQM

K. Wu (✉) • Y. Zhu
School of Control and Computer Engineering, North China Electric Power University, Beijing, People's Republic of China
e-mail: epuwkh@126.com; zyyaaa@126.com

C. Duan
School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, People's Republic of China
e-mail: duancheng1985@126.com

1   Introduction

After many years of information construction, power enterprises have left huge and complex application systems, which include multi-class business data about daily production, operation and management. However, due to the lack of unified data planning, data and application systems are highly interrelated, data storage types and locations, data models, data structures, data forms are different, which together constitute the enterprise multi-source heterogeneous data, the dispersion, heterogeneity and non-uniformity of data make it difficult to find the hidden information. As the information constructions of power enterprises have entered the stage of deepen application, it is urgent to integrate above data to build data centers and realize the accurate and consistent description of data. As the amount of data and complexity of the environment where data in are both rapid increasing, which make data quality issues in the process of data integration become increasingly prominent.

This article studies the definition of data quality, then presents a data integration platform architecture aimed at the characteristics of power enterprise data integration environments, studies the process of data statute and data quality control methods, and then we study the data quality measurement methods and procedures from the perspective of qualitative and quantitative assessments, finally, we view data as a product, design a total data quality management process of power enterprise. The methods proposed in this paper have been put into practical applications, and obtained good results.

2   The Definition of Data Quality

ISO9000 defines quality as the ability with a set of inherent characteristics to meet the requirements. Data is the result of the data generating process, but also products, data have quality issues too. The study on data quality originally appeared in the field of statistics, and it has become a hot spot in computer information field since the 1990s.

At present, the definition of data quality has not yet formed a unified understanding in domestic and foreign academic. Chanana and Koronios (2007) thinks that the data quality is the measure of the gap between data view showed in information systems and the same data in the objective world. Refer to the definition of quality in TQM (Total Quality Management), Wang et al. (1995a), Wang and Strong (1996) consider that data quality should be defined as "the applicability of data usage", which is the basic definition of data quality generally accepted. Data quality and the data application objects are interrelated, therefore, many researchers and agencies directly identify data quality through a set of dimensions or attributes according to data application environments. The commonly used dimensions are accuracy, completeness, consistency and timeliness (Songmin and Tanzheng 2007).

Table 1 The dimision of data quailty

Dimension	Definition of dimension
Accuracy	Degree of data correctness, meet realistic semantic
Integrity	Degree of data completeness and meet constraint
Consistency	Degree of data compatibility in logical relationship
Effectiveness	Degree of conformity of data attributes
Identification	Degree of data have unique identification
Connectivity	Degree of associated data to facilitate the merger
Reliability	Degree of the credibility of data
Metadata	Degree of data described information completeness
Security	Ability of data against damage and illegal access
Privacy	Degree of data privacy, confidentiality
Accessibility	Degree of data ease to access and use
Integration	Degree of data integration easy to use
Correction	Ability of user rectification to misdata
Translatability	Degree of data ease to understand
Correlation	Degree of compactness between data and user needs
Timeliness	Degree of data to meet requirements on timeliness

Data is the foundation of the application, high-quality data is a prerequisite to protect reliable analysis of the results. As the appearance of deepen application systems, which based on data centers and utilize data across multiple business domains, data quality issues are particularly important. Start from intrinsic property, business property and user property, we provide a multi-dimensional definition of data quality for power enterprise by comprehensive analysis and extraction, which is shown in Table 1. Since data quality is multi-dimensional, its definition is dynamic in different application environments, which is also constantly updated and enriched with the development of business and progressing of technology.

3   Data Quality Control

3.1   *Characteristics of Power Enterprise Data Integration*

With the rapid development of power information construction, the architecture of power enterprise information system is continuous evolving. The application mode has developed from traditional mode that divided by business systems, to the mode which based on enterprise data bus to build data sharing, finally to the mode that based on enterprise data centers. The development of power enterprise application system's architecture reflects the information developing process of power enterprise.

While grid is getting larger and more complex, its secure, stable and economic operation, in particular, the optimal operation is not only the responsibility of
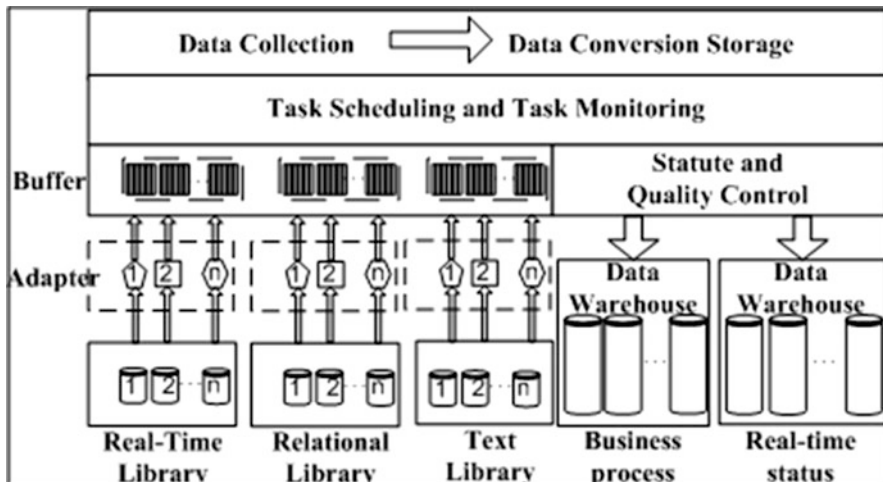
Fig. 1 The framework of power enterprise data integration platform

frontline managers, but also the problems what should be interested by business executors and policy makers. China's smart grid built should pay attention to the acquisition, integration and sharing of production-class and operation-class data (Chen Shuyong et al. 2009). The dispersive form of data in the early application mode is difficult to meet the needs of enterprise's new businesses, it urgent needs to integrate various data sources in legacy systems and format a global data view through the standardization process and dump, so a solution is presented: build data centers by data integration, then provide a unified, transparent data sharing service. However, power enterprises face a very complex data integration environment, including a large number of legacy systems such as SCADA system, production management system, electric energy collection system, ERP System, etc. These systems used a large number of heterogeneous database products and a variety of storage forms, including: (a) Real-Time Libraries: PI, eDNA, etc. (b) Relational Libraries: Oracle, Sybase, etc. (c) Text Libraries: XML documents, TXT documents, EXCEL documents, E language documents, etc.

Through analysis and research, we know that power enterprise data integration platform should contain functions like data standardized transmission, data statute process and data identity storage, it is able to integrate the data in different data types, different response time requirements, different data sources. The platform should also provide high-performance data access scheduling functions to meet a variety of response time requirements. Figure 1 shows a structure of power enterprise data integration platform, which uses various types of data adapters to implement translation and conversion between source data model and target data model, it extracts data from different types of data sources for statute (data filtering,

conversion of data attributes and domains, data merge, etc.), and controls data quality, finally loads data into the data center in accordance with the pre-defined target model.

The main features of platform are as follows: (a) Platform can read large number of data directly from the database application systems, it can access data in terms of regular time incremental, real-time update and real-time service. (b) Platform can be timed for data extraction, transformation and loading, such as daily or weekly. (c) Platform has an automatic scheduling function which can automatically call the data extraction, transformation, loading and updating multi-dimensional data cube task within a specified time window. Besides, it can monitor the change of data sources and create data or event trigger by setting parameters. (d) Platform can record each task's operation detail in order to provide a reliable basis for system maintenance and management. If the task is interrupted by exception, platform can provide alarm message to maintenance personnel to take remedial measures in time.

3.2   Process Method of Data Statute

In data integration, the process of data statute means that according to a unified structure and type to calibrate, converse and process data in different systems, different types and different data structures. Topics include: (a) Uniform coding. Code is the unique identification for enterprise application systems to identify the data sources. Uniform coding is to establish the only criterion of classification and coding architecture in accordance with units or departments for the normalization of information classification, which makes it easy to distinguish the sources of data. (b) Uniform property. Uniform property means the same calibration for property description of the same object. Refer to national standards, power industry standards or industry agreed standards, we could establish uniform standards for the properties of objects, then according to the standards to process data and implement the definition of properties in the same calibration. (c) Uniform type. Uniform type refers to the same property described by the same data type. Objects are described by properties, and data types reflect the belonging of properties in the data classification. Use the same data type to describe the same property in data process could ensure the normalization of data. (d) Uniform data structure. Data structure reflects the internal structure of objects. Here the structure is formed by one or more elements that meet particular relationships, these elements are properties of the object. Uniform data structure is to provide the same definition to the properties of the same object. (e) Uniform data model. Since the lack of a unified planning in early information construction, a lot of similar business systems were developed independently, and these systems used different data models. Uniform data model is to achieve the same definition for the data description of systems (Fig. 2).

Fig. 2 Processing method of statute
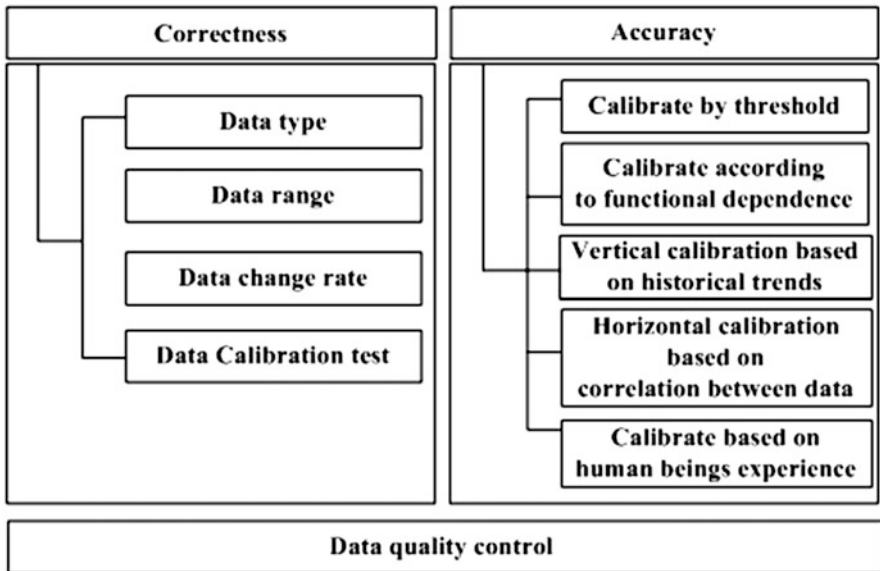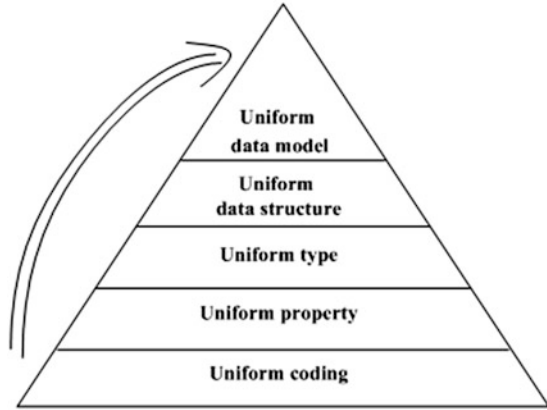
Fig. 3 Improving method of data quality

3.3   Data Quality Control Methods

According to "Garbage in, garbage out", data quality control in the process of data integration plays a crucial role for improving the level of business decisions. In this paper, data quality control methods include two aspects: the correctness and accuracy of data, as shown in Fig. 3.

1. Control of Data Correctness

   Control of data correctness is to ensure the consistency of data value and data symbol content, which includes: (a) Data type: The conversion process to data in accordance with the specific type, length, unit, precision, etc. Such as the type of substation load requires for the numerical, and its unit is KVA (kilovolt-ampere) and so on. (b) Data range: The process to data in accordance with the value range or the value set. Such as the scope of sector performance assessment score is 0 100, etc. (c) Data change rate: Analysis and process the rate of data sequence variation based on the data change regulation. Such as the collection time of grid real-time data (SCADA or EMS) is interval of 5 min in continuous. (d) Data calibration test: Based on prescriptive data standards, data calibration test is to detect and process the correctness of the data, such as the format of substation commissioning date is "YY.MM.DD", etc.

2. Control of Data Accuracy

   Control of data accuracy is to ensure the consistency of data value and the actual value of the data. The main methods are as follow:

   - Calibrate by threshold. Threshold calibration is to measure the accuracy of the data value by judging whether the data within a reasonable fluctuant range. There are two ways to determine the threshold: (a) Take advantage of the expert domain knowledge, institute the data threshold by expert or technical personnel. (b) Summary the data distribution as threshold by statistical methods (Zhang Liang 2009). For example, the maximum and minimum of load threshold can be determined by its fluctuant range in a certain period.
   - Calibrate according to functional dependence. Business data in power enterprises are not isolated, that is, there exists some function relationship between data, some target data can be calculated through this logic or arithmetic function. We can get some constraints by analyzing target data on the functional dependence, and such constraints can be used to calibrate the target data.
   - Vertical calibration based on historical trends. Many grid operation data are continuous and cyclical. Data are changing all the time, but the trend is with the principle of inertia. We can get data change rules via analysis the trend of history data to predict the result of data in a time or time period without the impact of external factors. In this way longitudinal calibration can be achieved.
   - Horizontal calibration based on correlation between data. In addition to the direct functional dependence there may be a fuzzy dependence between the grid business data. That is, the interactions and correlation relationships between data values cannot be directly deduced by the function. We can implement the data horizontal calibration through the consistent test of changes between target data and its relevant data. When the target data value changes, check the changes of its relevant data values at the same time. If the change trend meets the rules, then the target data value is accurate. The rules

and knowledge of consistency test can be determined by business domain knowledge or using machine learning methods to implement historical data mining.

- Calibration based on human beings experience. It is not difficult for expert and technical personnel to find the problems in business data. However, calibration based on human experience requires a lot of professional technical personnel to participate, the workload is pretty heavy, and the efficiency is relative low, so this method may work with other methods to calibrate the key data.

4 Data Quality Assessment and Management

4.1 Data Quality Assessment Methods

We all know that "You can't manage what you don't measure." Generally researchers measure and assess data quality based on dimensions. The International Data Corporation (IDC) suggested that China's large enterprises should assess the data quality from aspects of data accuracy, compliance, completeness, consistency, timeliness and repeatability (The white paper of data integration and data quality market in Chinese enterprises). Wang et al. (1995b), Redman (2005) think that the pursuit of the practical application of zero-defect data is neither necessary nor difficult to achieve, we should select representative measurement of quality indicators according to specific application environments. Data quality measurement methods include simple ratio method, the minimum or maximum method, the weighted average of multiple indicators method (Karr et al. 2006; Yang Qingyun et al. 2004; Michnik and Lob 2009), and decision-making method (Dillard 1992) and so on.

Data quality assessment can be conducted through qualitative or quantitative methods. The qualitative assessment emphasizes the subjective experiences and requirements of relevant personnel, while quantitative assessment uses mathematical methods to measure data objectively. Assessment objects of data quality can be inherent characteristics of the data itself, or the business processes based on the context or users' requirements and so on. The first six dimensions in the Table 1 mentioned are facing to the level of field, record, table or database, they are suit for the objective quantitative measurement. The last ten dimensions are highly subjective, we can use the questionnaire to measure qualitatively. Pipino et al. (2002) proposes a combination of qualitative and quantitative measurement methods. Due to limited space, details will not be discussed here.

We consider that data quality assessment will follow these steps: first, we should make sure the objects (applications, databases, data tables, records or fields, etc.) to be assessed, and then select the KPI evaluation indicators, and make clear the corresponding measurement methods, and finally assess the quality of data, which can be inspected one by one, or be comprehensive evaluated.
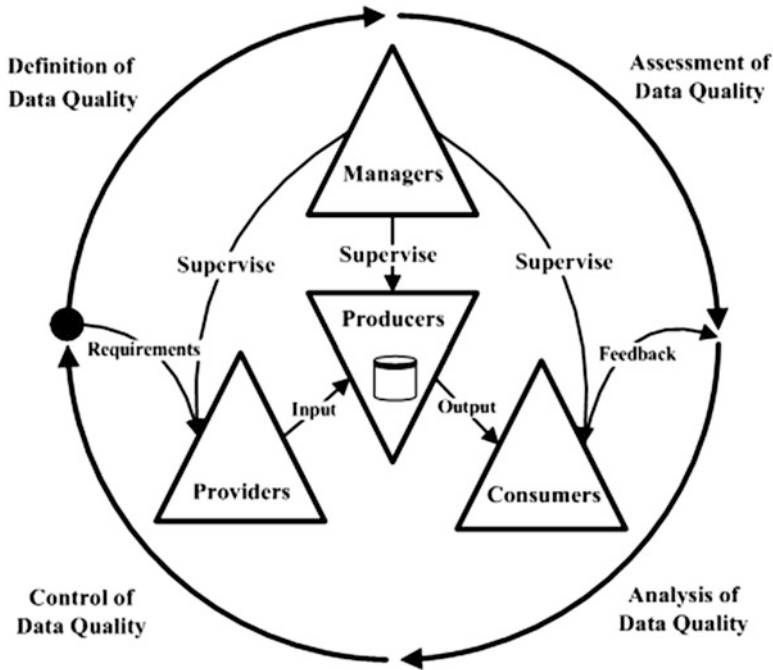
Fig. 4 Process of TDQM in data integration

4.2 Data Quality Control Process

Richard Y. Wang, authoritative data quality researcher of MIT, introduced the TQM technology about traditional manufacturing product into data quality management, and presented a total data quality management TDQM methodology (Wang 1998). After a sufficient summaries of research and practical experience, combines with TDQM methodology, this paper presents a total data quality management processes of power enterprise data integration, as shown in Fig. 4.

In the process, we define the data providers, producers, consumers and managers as four roles. Provider is the company's existing data source system and its associated business responsible officer; Producer is an enterprise-class data integration platform and its application process designer; Consumer is the target applications on the top of data center and its personnel; Manager refers to the analysis tools and responsible officer who monitor the entire life cycle of enterprise data product. Data monitoring can be divided into "upstream" as the data sources, "middle" as the data integration process, "downstream" as the data object application these three stages to implement.

According to the corresponding environment, data quality management of data products needs to formulate appropriate management process, including the

definition, evaluation, analysis and control of data quality, it is a closed loop which in continuous improving process. (a) Definition of data quality: First analysis characteristics of data product, from the view of data providers, producers, consumers and managers to define the requirements of data quality and establish the corresponding data standards. (b) Assessment of data quality: According to the characteristics of data products and data quality requirements, make clear assessment indicators and methods of data quality. (c) Analysis of data quality: According to the assessment conclusion of data quality and feedback information from consumers, survey the causes of data quality problems, generate data quality analysis report. (d) Control of data quality: In the light of data quality analysis report to control data quality from aspects of "upstream", "middle" and "downstream". Control in "upstream" is to improve the data sources, "middle" is to improve the data integration platform application process, and to improve the correctness and accuracy of the data in data center, and "downstream" is to improve the data accuracy of the target application system through the calibration procedure.

5 Conclusion

On the basis of theoretical research about data quality from home and abroad, this paper analyses its definition in-depth, and summarizes the dimensions identification of data quality. For the characteristics of power enterprise data integration, proposes a kind of data integration platform architecture, and describes the data quality control methods in detail. Finally, studies assessment methods and techniques about the data quality, proposes a power enterprise total data quality management processes based on TDQM, which can bring the data quality management process into virtuous circle. In the application of the power grid investment benefit evaluation support system for an enterprise, we integrate a large number of SCADA real-time data, energy metering data, reliability data and investment statistics data into data center utilizing data integration platform, and take proposed methods to ensure data quality. The data center provides reliable data support for the system, thus make the system realize automatic data collection and credible evaluation results. Data are important strategic resources to enterprise; data quality is a critical issue for the sustainable development of enterprises. Hope that the article can provide some useful ideas for solving the data quality issues of power enterprise data integration.

References

Chanana V, Koronios A (2007) Data quality through business rules. In: International conference on information and communication technology (ICICT 2007). IEEE Press, Dhaka, pp 262–265

Chen Shuyong, Song Shufang, LI Lanxin, Shen Jie (2009) Survey on smart grid technology. Power Syst Technol 33(1–7)

Dillard RA (1992) Using data quality measures in decision-making algorithms. IEEE Expert 7:63–72

Karr AF, Sanil AP, Banks DL (2006) Data quality: a statistical perspective. Stat Methodol 3:137–173

Liufei (2008) The white paper of data integration and data quality market in Chinese enterprises. IDC Company

Michnik J, Lob MC (2009) The assessment of the information quality with the aid of multiple criteria analysis. Eur J Oper Res 195:850–856

Pipino LL, Lee YW, Wang RY (2002) Data quality assessment. Commun ACM 45:211–218

Redman TC (2005) Measure data accuracy: a framework and review. Infor Qual Adv Manag Info Syst 1:21–36

Songmin, Tanzheng (2007) Reviews of foreign studies on data quality management. J Inf 2:7–9

Wang RY (1998) A product perspective total data quality management. Commun ACM 41:58–65

Wang RY, Strong DM (1996) Beyond accuracy: what data quality mean to data consumers. J Manag Inf Syst 12:5–33

Wang RY, Storey VC, Firth CP (1995a) A framework for analysis of data quality research. IEEE Trans Knowl Data Eng 7:623–640

Wang RY, Reddy MP, Kon HB (1995b) Toward quality data: an attribute-based approach. Decis Support Syst 13:349–372

Yang Qingyun, Zhao Peiying, Yang Dongqing, Tang Shiwei, Tong Yunhai (2004) Research on data quality assessment methodology. Comput Eng Appl 9:3–5

Zhang Liang (2009) Research on data quality for electric power dispatch data centers. East China Electr Power 37:403–405

Study on Classified Teaching Model of Computer Common Course in Finance and Economics Colleges

Lingbo Liu, Hengnong Xuan, Weimin Wang, Jie Lv, Lang Zhou, Huaili Zheng, and Hui Tang

Abstract   In order to further improve the teaching quality of computer basic course to the non-computer majored, this paper analyzed the current teaching status of computer common course. It gave the basis of the study of classified teaching and a teaching model about classified teaching was presented. It also discussed the teaching goals, students grouping, teaching methods, assessment methods, teacher requirements of classified teaching etc. Classified teaching for the individual differences and different cognitive levels of students can practice the teaching according to his ability. As well, it could fully promote quality education and improve significantly the teaching quality.

Keywords Classified teaching • Computer basic course • Teaching model • Individual differences

1   Introduction

Classified teaching is a teaching model, which based on the specific circumstances of students learning ability, students are divided into different levels and teachers will apply pertinent instruction and guidance for students of different levels in the teaching process. The conception of "Classified teaching" abroad was put forward in the half of the nineteenth century. Its prototype is the Activities of the Branch System, also known as Flexible System, founded by the American educator Harris (Harris, WT) in St. Louis in 1868. Classified teaching abroad in recent years is

L. Liu (✉) • H. Xuan • W. Wang • J. Lv • L. Zhou • H. Zheng • H. Tang
School of Information Engineering, Nanjing University of Finance & Economics, Nanjing, China
e-mail: djbllb@21cn.com

various in forms, theory and practice has become more scientific, personalized and comprehensive (YeLin and LiuWenXia 2008). For the teaching status of computer common course, the teaching method of classified teaching should be studied. It can improve teaching effectiveness, so that each student can be enhanced in learning and we can cultivate the community's overall quality of talent.

2 Computer Common Course Teaching Status

The computer common course is one of the important basic courses for non-computer majors of the higher education institutions, which aims to enhance the information literacy of the non-computer professional students, strengthens students versus computer application capable, and lays the foundation with the use of computer as a tool for their future work.

Computer common course for financial college such as our school is absolutely necessary. In order to improve our students' the skills of computer operations and applications, it has opened two computer public courses: Computer Application and Database Management System Application for non-computer professionals in our school, and student can graduate with a bachelor's degree if he through the National Computer Test Band 2 or the Jiangsu Provincial Computer Proficiency Test Band 2. As the more and more high requirements from community of the college graduates in computer applications and skills, the reform of computer common course is imperative.

Now, there is the phenomenon in class that it is difficult to teach for teacher and the students haven't the desire to learn more. The main reasons of these include:

2.1 Different Levels of Students

Students in different localities have different levels of the basis of computer. Some students in high school have contact with computers and they can operate perfectly the application of basic. But some students are ignorant of the computer and some of them even do not know the basic operation of switch.

2.2 Different Accept Ability of Students

In any course of learning, learning abilities are always different of different individual students. For the same question, some students will be able to understand after listening only one time, and some students have no notion of what the teacher means after a few times.

2.3 A Single Teacher Teaching Methods

To most of the teachers there is an idea: strictly in accordance with the outline, imparting the knowledge as major. All students in the classroom are as a whole and teaching contents are transferred out with the broadcast form. The individual differences of the receiving capacity of the audiences are not taken into account.

For these reasons, although lots of teachers work hard to explain in the classroom, a phenomenon is existed that the top students can't be supplied their learning desire and the poor students are unwilling to listen to. Certainly the teaching can't achieve good results.

3 Study Basis of Classified Teaching

3.1 Theoretical Basis of Education of Classified Teaching

Individualized instruction put forward by great ancient educationists Confucius is the prototype of classified teaching. "Learning control" of American educational psychologist Benjamin Bloom is a theory of teaching about take different teaching approaches to the different students to adapt the diversity among students. Based on the research of the students' characteristics, the former Soviet educators Babanski's "optimization theory of teaching process" takes full account of teaching law, principles, forms and methods of modern teaching and teachers select purposefully the best scheme of teaching process, so that teachers and students can receive the best results by spend less time and less energy (LinYongHeng 2010).

3.2 Accord with the Demands of the Community

Development of information technology make our society more closely combined with computer technology and it is also put forward higher requirements for non-majors in computer operation and application of computer technology. Unit of choose and employ persons hope to appoint the employee whose work would Combined with the professional and he is good at analyzing and solving problems with computer as a tool. By using the teach model "classified teaching", it makes each student's ability improved to cope with increasingly fierce competition in the community.

3.3 Needs of Computer Teaching Reform of University

In the "National education reform and development of long-term planning programs (2010–2020)", it was brought forward the guidance of teachers and student-centered

and giving full scope to the initiative, to promote the healthy growth of all work as a school starting and ending points. Caring for every student and improving their development internally and lively, the rule of education and the discipline of students' body and mental development must be respected and suitable education should be provided for each student (National Education Reform and Development of Long-term Planning Programs 2010).

4 Particular Act of "Classified Teaching"

4.1 Teaching Aim

Integrating with the current goal of education of computer common course in college, there are three alternative schemas for the teaching aim of "classified teaching" (as shown in Fig. 1).
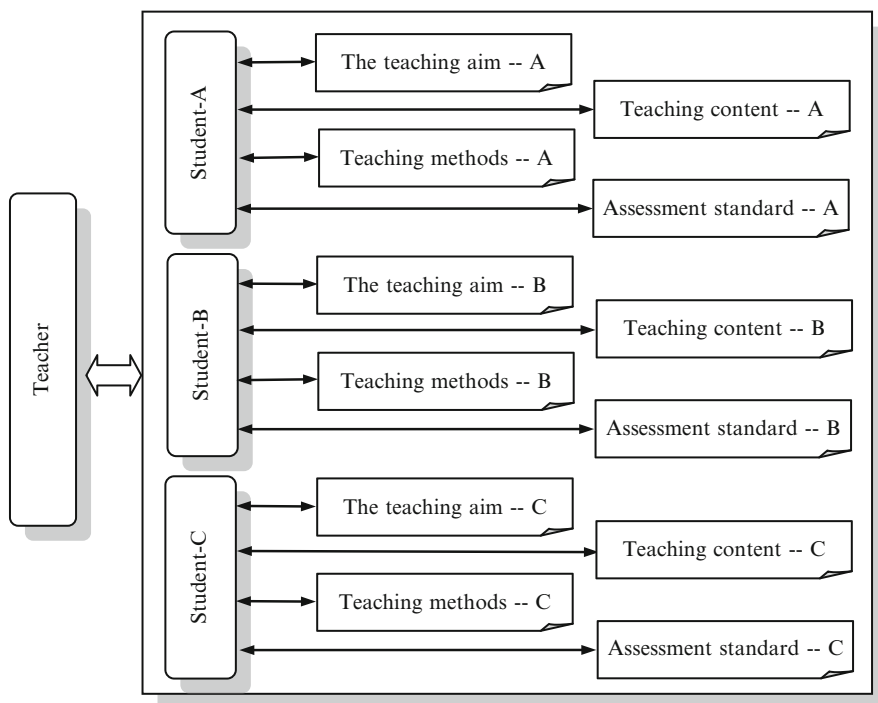
Fig. 1 Diagram of classified teaching

4.1.1 A-Level Teaching Aim

Upon the basis of completion the requirements of the instructional program of computer common course, the much higher request is being put into a higher level for the learning content and learning ability. It needs to fully mobilize the initiative and activity of students, and to train the self-learning ability and sense of innovation of students. It would enable them to discover problems and solve them with computer knowledge and able to combine with specialized subject to be a comprehensive human resources.

4.1.2 B-Level Teaching Aim

It is demanded to reach the required standard of instructional program of computer common course and master the basic computer knowledge, and can solve some problem with their knowledge to further develop good study habits and improve learning ability.

4.1.3 C-Level Teaching Aim

The teaching aim is to perform the requirements of the instructional program of computer common course and cultivate a taste of learning computer courses to improve study habits (LiuLong et al. 2008).

4.2 The Grouping of Students

We should analysis earnestly the individual characteristics of each student and give an overall consideration of each student's intellectual and non-intellectual factors. In the full study, according to the three levels of teaching aims above, students are divided into three groups:

1. A group: In this group, students have good study habits and strong learning capability and they showed a keen interest in computer learning, so we can set an even higher demand on them to achieve A-level teaching goals.
2. B group: This group of students has a good foundation and they also listen to class very attentively. But they are lack of strongly initiativeness of learning and need the power outside to promote. The students in this group are required to achieve B level.
3. C group: Influenced by sorts of factors, the students of this group didn't form a good habit of learning whose computer foundation is comparatively poor, have little interest in learning the curriculum and received very low marks test scores are low, this group of students will lead a walk, required to complete C-level teaching objectives.

4.3 Teaching Content

The teaching content of different groups is different too. After complete the required teaching content of the instructional program, the students of A group should be demanded further in all its breadth and depth to satisfy the curiosity of these students. We will train them to accept the challenge of a difficult problem and to have the spirit of assiduous study. For the students of B group, they are asked to fully understand and absorb teaching content outline required and complete the corresponding operation. As the different cognitive level, for the students of C group, we don't require them to understand immediately the teaching content in the classroom and allow them digest and absorb slowly after class.

4.4 Teaching Methods

When we adopt the teaching model of "classified teaching", teachers can't use a single teaching method and they are required to have various teaching methods. As the students of A group have strong learning capability, teachers do not need too much explanation and only need give them reference-point construction. They are allowed to discuss in groups, even part of the contents can also be self-learning to enhance their spirit of cooperation and self-learning ability. As well as, teachers can give a reading list to them to expand the horizons and broaden their knowledge. For the students of B and C two groups, teaching content is needed to explain in class, especially to pay attention to the accept ability of students of C group. If necessary sometimes, the important and difficult points of teaching content should be repeated.

4.5 Assessment Standard

The students of different groups have different teaching targets and their teaching assessment standard and assessment methods are also different. The assessment method of A group can be more flexible, such as solving problem personal or in group or examining the difficult exercise problems; For the students of B and C group, the mainly assessment method is generally paper-based test, but their papers to be in difficulty difference to adapt to different levels.

4.6 Dynamic Adjustment Grouping

In the teaching process, the grouping is dynamic and changes. According to the development of each student, the student of the C group, who made remarkable

progress in learning ability and learning habits during the section of teaching time recently, can be adjusted to the B group, and vice versa. It is interoperable between different groups. It is a pressure for the students of high groups and they are forced to maintain good learning state. For the students of low groups, it is a driving force and they would have improvement and progress in the process to strive.

4.7 Demands with Teachers

Under the traditional teaching mode, the majority of teachers, while in the process of preparing lessons, focus on the process of study the features of textbook and familiar with the teaching content. In the design process of teaching all the time, the students are treated as a whole. Lesson conducting is the process of the implementation of lesson plans. The teachers always think that they only need spread the knowledge to students. In this series of teaching process, teachers seldom regard the differences and individualistic between students and this is also our traditional teaching habits and limitations.

"Classified teaching" is a complex process. It will put forward high demands and high standards on teachers and the class load of a teacher will be increased a lot. It requires teachers to have a profound understanding the teaching philosophy of "classified Teaching" first and to form a correct teaching ideas in the mind. "Classified teaching" also put forward higher requirements for teachers' teaching ability. Teachers need to keep learning, thinking, analysis, and further improve the teaching ability themselves. In the teaching preparation before class of "classified teaching", firstly, the basic conditions and acceptability of the students must be familiar to the teacher. Secondly teacher should design the teaching process and prepare lesson plans according to the specific circumstances. Finally the teacher is in control of the teaching content and teaching methods of each level student in the classroom, so that the teaching works pretty well.

5 Conclusion

Our university belongs to financial college and most of the faculties are liberal arts classes. Therefore, students of the university have the various levels of the abstract computer theory and the acceptance of computer concept. Implementing classified teaching, we should teach students in accordance with their aptitude facing individual differences of students and aiming at different cognitive levels. It will be taken appropriate educational methods for each student to better adapt to students' interests and differences, so that every student can be improved and made progress for different levels of students. Though fully dig for the potential of every student and promote the individuality development of students, it can bring about an all-round improvement of quality education.

References

LinYongHeng (2010) On class teaching in college english without divide into classes. Edu Circ 2:117–118

LiuLong, LiMing, YuBin, LiuYing (2008) Thoughts about the hierarchical teaching of advanced mathematics. Edu Explor 11:55–56

National Education Reform and Development of Long-term Planning Programs (2010–2020). http://www.gov.cn/jrzg/2010-07/29/content_1667143.htm

YeLin, LiuWenXia (2008) Historical development of layered teaching abroad. Teach Manage Theory Ed 1:159–160

Strategic Human Resource Management, Enterprise Competence and Competitive Advantage – A System Dynamics Model

Weipeng Wang

Abstract There exist strong relations among strategic human resource management based on competency model, enterprise competences and competitive advantages. By building a system dynamics model, we demonstrate the interaction mechanism among the three above. The competence of human resource management plays an important strategic role in the formation and development of enterprise competence system. Likewise, strategic human resource management can promote the complementation and gain of the enterprise competence system, which is of important theoretical significance for the competitive advantage. This article also points out that there exists strong correlation between the competency-based human resource management and strategic human resource management.

Keywords Strategic HRM • Enterprise competence • Competitive advantage • Competency model

1 Instruction

The theory of strategic management maintains that competitive advantage comes from the heterogeneous enterprise resources and competences. Since McClelland (1973) first proposed the concept of competency, many scholars began to deal with the "competency" in the field of HRM, and then gradually formed a theoretical system of competency-based HRM. However, this theory concerns more individual competency and performance, with little consideration of the relations among

W. Wang (✉)
School of Economics and Management, Weifang University, Weifang, China
e-mail: wangweipeng1974@126.com

individual competency, enterprise performance, enterprise competence and the needs of enterprise development. And with the rise of strategic HRM, the traditional research on local and functional optimization began to turn to the important role of HRM in promoting the overall enterprise objective. But most of the studies of strategic HRM were based on Barney's theory (1991), focusing on the resource characteristics of human capital and HRM system, which is closely associated with the sustainable competitive advantage. So it is still a black box for us today that how HRM activities affect enterprise performance and create competitive advantage. On the basis of the review of the related studies on the competency-based HRM and strategic HRM, this paper will analyze the relations between enterprise resources and competence, discussing the relations among employee competency, enterprise competences, competitive advantage performance and HRM from the perspective of enterprise competences.

Furthermore, this paper will deal with the relations between competency-based HRM and strategic HRM, which are the two important branches in the area of HRM nearly 30 years, and point out that we can build and apply an employee competency model based on enterprise strategy for strategic HRM.

2 The Competency-Based HRM and Strategic HRM

2.1 Review of the Former Research

Modern competency management began from McClelland's (1992) pioneering work. Then many scholars made further research on the competency model from different angles. And the competency model was widely used in many areas. For example, the combination of competency model and HRM can improve employee productivity. The HRM activities, procedures, methods, which are based on competency model, are called competency-based HRM by some scholars. Overall, competency-based HRM research accepted the traditional ideas of HRM, discussing the improvement and functional optimization of HRM from the relations between employee competency and their performance. Unlike traditional function-based HRM, strategic HRM discusses the relations between HRM activities and organizational goal from a macroscopic angle of an organization, emphasizing on human capital and HRM activities' important contributions to enterprise performance and competitive advantages. Based on Barney's (1991) theory, Wright and McMahan (1992); Brockbank (1999); Teece et al. (1997) pointed out that human capital had the special characteristics which could bring sustained competitive advantages for enterprise, and thinking that HRM activities could affect organizational performance and competitive advantages through affecting human capital, employee attitudes and behaviors. However, some scholars believed that human capital was not the only source of competitive advantage, HRM system could also bring sustained competitive advantages.

So Becker and Huselid ([2006]) thought that it was still a black box until today for the study of the relations among HRM activities, enterprise performance, and competitive advantages.

2.2 The Differences and Similarities of Strategic HRM and Competency-Based HRM

There are many similarities among human capital, employee behavior and employee competency in the research on the competency-based HRM, only that the former stresses the important role of human capital and employee behavior in enterprise performance and competitive advantages, the latter focuses on the relations between employee competency and individual performance from microscopic view. Until now we still know very little about how enterprise human capital and employee behavior affect enterprise performance and competitive advantages, which leads to the uncertainty of causes and effects, pseudo relations in the empirical analysis of HRM activities and enterprise performance. Also, since the human capital and HRM system can bring a sustained competitive advantage, then what relations exist between the two, and how they interact? These are still an unresolved problem in the research of strategic HRM.

3 The Enterprise Competence and Enterprise Competency System

3.1 A Resource-Based Enterprise Competence

An important cause of the dilemma which the study on the relations between HRM and organizational performance faces is that most of the researches on strategic HRM are based on Barney's ([1991]) theoretical model of resource-based view and analysis framework, focusing too much on the characteristics of the resources in human capital and HRM system, which is closely associated with the sustained competitive advantage. However, the resource-based view is still deficient. For example, there is no agreement about the concept of resource, and the explanations for how to create competitive advantages are deficient. In fact, there are no differences in nature between the strategic resources of the resource-based view and the core competence of the competence view. The enterprise resources include all the assets, competences, organizational processes, enterprise features, information and knowledge, which an enterprise can control and improve the efficiency and effectiveness during strategic formation and implementation. The meaning of human resource living is basically the same as that of employee competency.

Therefore, this paper argues that the enterprise can be regarded as a collection of resources, also a collection of competences. The resource-based view focuses on the identification of strategic resource, yet the competence view on the efficiency and effects of the allocation and use of resources for specific tasks. Because of the multi-dimension of the division of resources and competences, resources or combinations of resources often correspond to enterprise competences. In other words, enterprise competence comes from the combination and configuration of some resources. Considering that the resource-based view studies the characteristics of the resources which are closely related to competitive advantages, but without deeper research on the interaction among different resources and on the links between the different resources and tasks, this article will deal with the relationship between HRM activities, enterprise performance and competitive advantages based on competence view. But this does not affect our learning from the research of strategic HRM of the resource-based view.

3.2   What Is Enterprise Competence: A View of This Paper

An enterprise is a collection of competencies. Some scholars defined enterprise competence as a series of knowledge and skills which could create value for customers but must be rooted in various enterprise processes. This paper argues that the profit indicator of organizational performance depends on the difference between the price a consumer is willing to pay for the products and services and the cost to provide these products and services, while the price a customer is willing to pay depends on the value an enterprise provides for him. Thus, compared with the competitors, organizational performance ultimately depends on whether an enterprise can provide more value for customers at the same cost, or provide the same value for customers at a lower cost. But that depends on the enterprise competence of completing the value chain activities. The formation and development of these competences are not a spontaneous process; they are purposefully cultivated and integrated under the guidance of the strategic target. At the same time, these competences affect the formation and implementation of the future strategy in turn. But the concept of competitive advantage has not been well defined: one view is that competitive advantage is a gained position advantage relative to competitors and potential entrants in the market; another is that competitive advantage is the competence advantage of the more efficient use of resources than the resource owners and other users.; the third is that competitive advantage is an advantage shown in the competition with the opponents. Relying on this advantage, an enterprise can gain more return than the normal revenue rate of the industry. In other words, competitive advantages are a market position which comes from the greater value an enterprise created for the customers relative to the competitors, because it has heterogeneous competences and resources. However, competences, resources and market position must eventually be reflected and proven by enterprise performance. The relations among enterprise strategy, enterprise
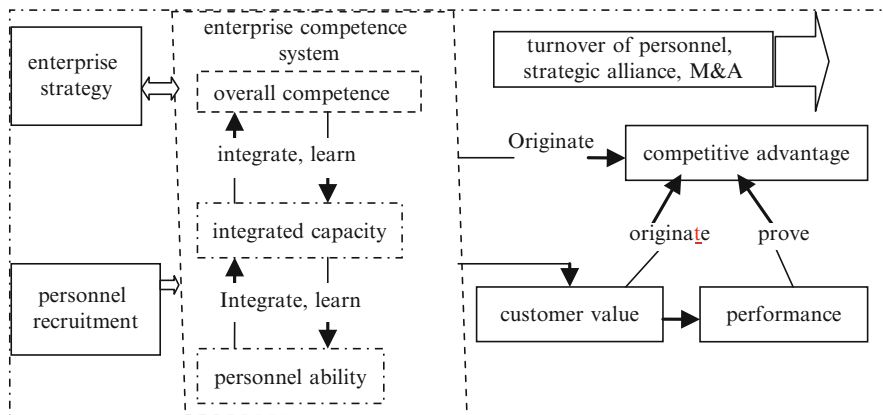
Fig. 1 The enterprise competence system

competences system, competitive advantage, customer value, and enterprise performance are shown in solid line in Fig. 1 below:

3.3 The Enterprise Competence System

Different scholars divided the enterprise competence system for different purposes. This paper will learn from the study of some scholars, and divide the enterprise competence system into the competency of individual level, integration level, and enterprise level. This article argues that it is the combination of employee competency, experience, knowledge, skills, relationship, attitude, judgment and equipment, plant, geographical location, raw materials that form the individual-level competency of accomplishing a particular job. The interaction (complement, alternation, gain, suppression) of the individual-level competency forms the integration-level competency of accomplishing a particular job in the value chain or business process. With the interaction of the integration-level competency, thewhole enterprise value chain enable the competence to create customer value. It should be pointed out that high-level competency is based on low-level competency in the enterprise competence system, and low-level competency is the "raw materials" of high-level competency. But once the high-level competency had been formed, it would affect the formation and integration of low-level competency in turn, so the enterprise competence system would show path-dependent. Meanwhile, the enterprise competence system is an open system, it exchanges competences with other enterprises through employee recruitment and turnover, employee learning, M&A, strategic alliances, which enables the enterprise competence system show the dynamic and developing. In a word, the competitive advantage comes from the organic combination of the competency of different levels and their integrated work.

However, compared with low-level competency, the higher the level of competency is, the more valuable, scarce, inimitable, and non-substitutable it is, and it can bring more sustained competitive advantage for the enterprise. The structure of the enterprise competence system is shown in dotted line in Fig. 1 below.

4 The Enterprise Competence System and HRM

4.1 HRM: The Impact on Enterprise Competence System

Under the direction of strategic target, the ideal enterprise competence system will make the inside operation of the enterprise competence at different levels run in harmony with the outside operation of the competences at different levels, and get the effects of complementation and gain. It can also adjust and develop constantly in accordance with the changes of external environment, especially customer demands. The changes of HRM from the operation-level response pattern to the operation-level active pattern, and then to the strategy-level response pattern and the strategy-level active pattern, meet the needs of the enterprise system from local optimization to global optimization, and from static management to dynamic development. It also highlights the increasing importance of knowledge and HRM in the enterprise competence system and the increasing contribution to the competitive advantage in turn. The operation-level response HRM can achieve the people-thing match by hiring junior employees, managing salary and welfare, providing basic skills training, and then improve the individual-level competency and employee productivity. The operation-level active HRM can improve the quality of HRM and employee satisfaction, promoting the integration of individual-level competency and the formation of the integration-level competency through working process reengineering and total quality management. The strategy-level response HRM concerns the effective implementation of enterprise strategy, promoting the integration of integration-level competency and the formation of the enterprise level competency through shaping and building a strategy-oriented organizational structure, incentive systems, communication mechanism, enterprise culture. After a thorough understanding of the external social environment and internal function areas, the active strategic HRM can shape a customer-centric culture, founding an innovation-centric incentive system, communication mechanisms and organizational structure, building learning-oriented enterprise and promoting the formation of a growing enterprise competence system based on external environmental requirements. There were two great changes in the employee management: the first is from the personnel management to HRM, the second from HRM to the strategic HRM. Compared with the traditional HRM, strategic HRM, from the perspective of enterprises as a whole, achieves the vertical match of human resource activities and external environment or enterprise strategy, the horizontal match of different HRM activities, promoting the complementation and gain of competences at different levels, the changes of
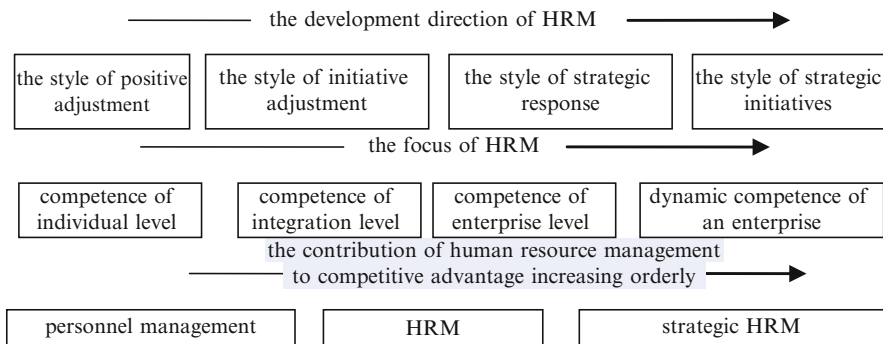
Fig. 2 The development of HRM, level of enterprise competence and the contribution to competitive advantage

the competences from a low level to a higher level, the development of enterprise competence system based on external changes. (see Fig. 2 below)

4.2 The Enterprise Competence System: The Impact on HRM

The level of strategic HRM of an enterprise depends on it's competences of HRM. Liqun Wei (2006) once pointed out that the factors, which affected the vertical match of enterprise HRM activities and enterprise strategy, and the horizontal match of HRM activities, included the ability of human resource managers, the ability and support of senior managers, the knowledge and skills of employees, the nature of strategy, enterprise values and culture, human resource policy, the option of human resource activities, budget constraints and other factors. That explained to some extent that the competence of HRM was affected by some other competence subsystems of strategic management competence, financial management competence. In fact, the HRM competence subsystem may play a more significant role in the formation and development of other competence subsystem: HRM activities are one of the major factor to decide employee competency, yet the employee competency is the main "raw material" to build the enterprise competences; the HRM activities in strategic alliance directly affect the enterprise competence and motivation to acquire knowledge and skills; the integration of human resources after M&A determines the availability to obtain the knowledge and competence of M&A object and the success of M&A activity. The policies of human resource will affect employee turnover, and then affect the speed and extent of the outside transfer of enterprise competence; The Incentive mechanism, learning mechanism, and communication mechanism of HRM will affect the relationship and the integration of the competences at different levels, and then affect the formation of competence at a higher level and the competitive advantage of the whole competence system. The

interaction of the competence subsystem of HRM and other competence subsystem will get a stronger match of enterprise competence subsystem of HRM and other competence subsystems. That is, the stronger the embeddedness is, the greater contribution to the enterprise the competence system makes. At the same time, Li-Qun Wei (2006) pointed out that enterprise competences at different levels such as the competency and support of human resource managers, the competency and support of senior managers, knowledge and skills of employees, enterprise values and culture, human resource policies, the option of human resource activities, are both the foundation of human resource activities and the results of the human resource, As a bridge to connect the current competence and future competence of HRM, human resource activities make the competence of HRM show a dynamic non-equilibrium and path-dependence. In short, the enterprise competence system determines the value created for customers, and then the enterprise performance. As an important component of enterprise competence system, there exists interaction and mutual influence among the competence of HRM and other competences, which has a major impact on the formation and development of enterprise competence system. Compared with the individual competency of human capital, the human resource system including processes, functions, activities, is at a higher level in the competence system, and contributes more to competitive advantages. But this does not deny the importance of employee competency. The employee competency being the "raw material" of competency at other levels, the basis of the competences in enterprise value chain, and an important source to gain access to external competences, so the competency model based on enterprise strategy and the changes of external environment, and the human resource activities based on the competency model, are also of the same strategic significance for enterprise, and can be regarded as an effective way to carry out enterprise HRM.

5   Conclusions

In this chapter, we argue that an enterprise is a competence system, and point out that as part of enterprise competence system, the competence of HRM plays an important strategic role in the formation and development of enterprise competence system. Through the matching of HRM activities and enterprise strategy or the external environment, strategic HRM can promote the complementation and gain of the competency at different levels, encouraging the conversation of a low-level competency to a higher level, the development and innovation of the enterprise competence system based on external changes. HRM activities first affect the enterprise competence system and customer value, and then enterprise performance, which is of important theoretical significance for the empirical analysis of HRM activities and enterprise performance. At the same time, this article also points out that, as the two important study branches in the HRM area nearly 30 years, there exits strong correlation between the competency-based HRM and strategic HRM. If the employee competency model in the competency-based HRM is based on the

analysis of enterprise strategy and external environment, then the competency-based HRM is an effective method to carry out strategic HRM, which is of an important practical significance for the enterprises to carry out strategic HRM.

References

Barney J (1991) Resources and sustained competitive advantage [J]. J Manage 17(1):99–120

Becker BE, Huselid MA (2006) Strategic human resource management: where do we go from here? [J]. J Manage 32(6):898–925

Brockbank W (1999) If HR were really strategically proactive: present and future directions in HR's contribution [J]. Hum Resour Manage 38(4):337–352

Liqun W (2006) Strategic human resource management: determinants of fit [J]. Res Pract Hum Resour Manage 14(2):49–60

McClelland DC (1973) Testing for competence rather than for intelligence, Am Psychol 28(1):1–24

Teece D, Pisano G, Shuen A (1997) Dynamic capabilities & strategic management [J]. Strateg Manage J 18:509–533

Wright PM, McMahan G (1992) Theoretical perspectives for strategic human resources management [J]. J Manage 18(2):295–320

A Neural Network Model on the Forecasting of Inventory Risk Management of Spare Parts

Weipeng Wang

Abstract This paper proposes a neural network-based classification approach to inventory risk level of spare parts. Firstly a fuzzy evaluation of spare parts is made in terms of their availability of suppliers, importance, predictability of failure, specificity and lead time. Then a multilayer feed forward neural network model is established. The Back Propagation (BP) algorithm for training a neural network is used to decide the weights to connections in the model. Choosing a sample of historical data of 100 spare parts and undertaking a BP training stimulation, the model is used to predict the inventory risk levels of 60 spare parts for a well-logging service firm. The forecasting reliability reaches 84%.

1 Introduction

The inventory management of spare parts plays increasingly important role in the modern production and operation management. The management of spare parts shoes many different characteristics and requirements from inventory management of other materials, such as there are so many different types of spare parts, and the prices vary widely; the shortages of spare parts will inevitably lead to shutdown, the loss of business contracts and customer satisfaction, etc. These characteristics lead to a large inventory of spare parts and low inventory turnover rate for many process-oriented enterprises. For many companies of small quality, multiple variety and complete production, there is a wide range of spare parts, some of them are very

W. Wang (✉)
School of Economics and Management, Weifang University, Weifang, China
e-mail: wangweipeng1974@126.com

expensive. Moreover, most parts are not standard parts, must be purchased from specific suppliers, ordering a longer period. In order to avoid unnecessary losses, some companies have to reserve a large number of spare parts, which takes up a lot of liquidity.

The existing research on inventory management focused on the inventory management model and inventory management control strategy in the supply chain, working on some effective demand forecasting method, but with little consideration of the inventory management of spare parts. Chang et al. (2005) built a model of (r, r, Q) to demonstrate the inventory management of dangerous parts, that is, to determine the number of spare parts to order according to the levels of inventory spare parts. But this model considered the inventory risk only from the angle of costs, without considering the uncertainty of demand and supply due to the inherent characteristics of spare parts. Fisher et al. (2001); Hua et al. (2006); Huiskonen (2004); Willemain et al. (2004) described an optimal replenishment strategy for fashion products. After considering the factors of the minimum shortage costs, delaying in delivery and inventory, they built a two-stage stochastic dynamic programming model, finding the optimal order time and replenishment strategy bi the heuristic method. Liu Liwen and Wang Jun (2006) established an inventory management model containing order cycle, marketing strategy and logistics characteristics, but without considering the order of the products of small and uncertain order. Xu Xiaoyan (2006) proposed an inventory management method of spare parts based on the classification of demand characteristics, which could solve the problems of too many kinds of spare parts and lack of the historical data of demand of some of the spare parts, but with little consideration of the uncertainty of demand and supply. For some companies of small quality, wide variety and complete production, it is very difficult to decide the order time and quality, because they have a wide variety of spare parts, and the spare parts have high price, complex source, little qualities, uncertain order cycle, limited historical data of the demand. Due to the uncertainty of the demand, it is very difficult to forecast the demand of the spare parts, which lies in the followings: (1) randomness. Most of the demand of the spare parts is random, once the equipments were out of action, we need to replace the ineffective components with stock spare parts in time, but the demand time is uncertain and intermittent. (2) demand uncertainty. While we can make the regular and overall maintenance plan, because of the difference of the use and maintenance of the equipments, it is difficult for us to forecast when and how many we will generate the demand of spare parts.

There are already some solutions to these problems. For example, we can use ABC classification and AHP to classify the inventory spare parts, and to determine the inventory update cycle, Groston (1972); Aronis et al. (2004) use the method of Bayesian to forecast the demand quality of inventory spare parts and inventory levels. In addition, there are two methods of Croston and IFM, the former concentrating on the forecasting of intermittent demand, and the latter considering the maintenance of the equipment and the autocorrelation of demand. But there are so many factors that affect the stock of spare parts that we have no general forecasting method considering a variety of affective factors. For the enterprises

of small batch, wide variety and complete sets of production, the spare parts have characteristics of wide variety, high price, small-lot, non-standard. In this paper, after considering the source of supply, importance, degree of vulnerability, degree of standardization, supply cycle and other factors, through fuzzy evaluation of spare parts, we can get the initial information of the spare parts, and then build a model of multilayer feed-forward neural network, that is, a neural network model containing the importance, predictability, degree of vulnerability, degree of standardization, supply cycle, and other factors that affect inventory risk. Then we can use this model to predict the level of inventory risk of different spare parts, and then take appropriate strategy of risk management according to different risk levels.

2 Description of the Model

2.1 Preparation for the Establishment of the Model

The conventional inventory management of spare parts in enterprise classifies spare parts according to the properties of spare parts such as physical characteristics or performance parameters. Unlike the conventional classification of spare parts, we re-classify the inventory spare parts based on the purpose of the decision-making of inventory risk, one hand, to prevent a large number of stocks which led to the funds used, the other is to avoid the losses and reduction in customer satisfaction due to out of stock. In order to solve this problem, we will analyse the factors that affect the inventory cost and the approach of risk-averse., and then find the factors that affect inventory cost. There are five: (1) The predictability of spare parts failure: It is very difficult to predict exactly when the spare parts will fail, but we can get the levels of the failure of spare parts from the probability rating of the failure of spare parts. The easier to failure, the greater inventory risk, the spare parts not easy to failure have little effects on inventory. (2) Importance of spare parts: For example, for the logging apparatus is concerned, some belonging to high-end devices, the exports of which are controlled by American government, the cost of imports is high, the cycle of imports is long, which lead to be very difficult to order. The importance of devices affect the inventory costs and inventory risks, the more difficult to import, the greater impacts on inventory. (3) Number of alternatives for spare parts: If we can easily find the alternatives for spare parts, and easily get from markets, then the risk of low stock is low, and then the risk level is also low. Instead, the risk level of spare parts is relatively high. (4) Standardization of spare parts: If the spare parts meet a standard, easy to order, and the order time can be controlled, then the inventory risk of spare parts is low. On the contrary, the non-standard spare parts are difficult to order, and the order time is uncertain, then the inventory risk is high. (5) Length of order cycle: The longer the order cycle, the higher the risk of out of stock, and the higher the inventory risk. Considering these five factors generally, and through the classification characteristics of the approach of neutral network, we can get the

Fig. 1 Structure of neutral network
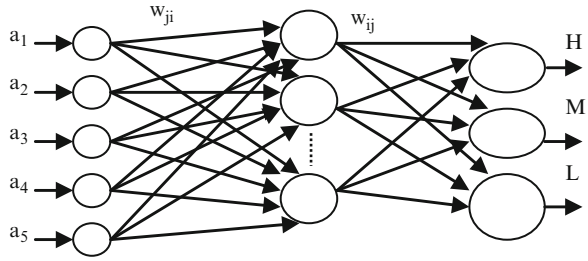
level of inventory risk of the spare parts. For the apparatus seldom used before and then lack of relevant historical data, we will use the method of fuzzy evaluation to classify these factors. There are two purposes: (1) To overcome the difficulties of predicting the demands in the lead time because of lack of the historical data of the demands of some spare parts. (2) To overcome the difficulties of predicting and decision-making of inventory duo to the uncertainty of the demands of spare parts. So in this paper, we classify the spare parts according to some indicators, and then get the risk levels of spare parts (high risk, medium risk, low risk). These indicators include: (1) a_1: predictability of the failure of spare parts (easy, possible, difficult). (2) a_2: importance of spare parts: important, medium important, general. (3) a_3: number of the alternatives for spare parts (easy, difficult) (4) a_4: standardization of spare parts (standard, non-standard) . (5) a_5: length of order cycle (4–20 weeks).

2.2 Structure of the Model

We use the results of fuzzy evaluation on the five indicators as the initial input to construct multilayer feedforward neural network model. As shown in Fig. 1, the neural network is divided into three layers: input layer, hidden layer and output layer; the output of input layer as input of hidden layer, the output of hidden layer as the input of output layer, then the upper and lower layers of model reaches full connectivity, but there is no connection among the neurons of each layer. The input variables of input layer are: the predictability of the failure of spare parts (a_1), the importance of spare parts (a_2), the number of alternative parts (a_3), the standardization of spare parts (a_4), the range of order cycle (a_5). The output variable of output layer can be divided into three styles: high risk (H), medium risk (M), low-risk (L). Here, the I said the neurons of input layer, j said the neurons of hidden layer, k said the neutrons of output layer, w_{ij} said the connection weights between input layer neurons and hidden layer neurons, w_{jk} said the connection weights between hidden layer neurons and output layer neurons.

According to the number of original input variables, there are five neutrons on the input level; according to experiences, there are 16 neutrons on the hidden level;

according to the number of the output variables of output level, there are three neutrons on the output level.

The transfer function of this neutral network is s-shaped function. A s-shaped function is differentiable everywhere, the divided area is composed of non-linear super plane, and the surface is relatively smooth. This classification is more accurate than the method of linear classification, so is the fault tolerance.

3   Determine the Connection Weights of the Neutral Network Model

To take advantage of the multilayer feedforward neutral network model to predict the risk level, we need to know the connection weights w_{ij} between the input level neutrons and the hidden level in the model, and the connection weights w_{jk} between the hidden level neutrons and output level neutrons. To determine the connection weights of the neutral network model, we use BP training algorithm. By the following steps:

- Initialization settings: In the interval $(-1,1)$, we select randomly the connection weights wij between the input layer and the output layer, the connection weights wjk between the hidden layer and the output layer, the threshold value θj, the threshold value γk, and set a higher value sum of squares of the output error (SSE), as 1.
- Select a group of sample values randomly, and assign $P = (a1, a2, \ldots, an)$ to the input variable of the input layer, and assign a target value to the corresponding output variable T of the output layer.
- Use the input sample $P = (a1, a2, \ldots, an)$, connection weight -, threshold value θj to calculate the input value u(j) of each node of the hidden level, and then use u(j) to calculate the output value U(j) of each node of the hidden level by transfer function.

$$U(j) = f\,(u(j)) \tag{1}$$

- Use the output value U(j) of the medium level, connection weights wjk , threshold value γk to calculate the input value γ(k) of each output level, and then calculate the output value Y(k) of each part of the output level by transfer function.

$$y(k) = \sum_{j=1}^{N} w_{jk} U(j) - \gamma_k \tag{2}$$

$$Y(k) = f\,(y(k)) \tag{3}$$

$$K = 1, 2, \ldots, M$$

- Use the target value T(k) of the network, the real output value Y(k) of the network to calculate the error dkt of each node of the output level.

$$e_j^t = \left[\sum_{k=1}^{M} d_k w_{jk} \right] u_h(i) \left(1 - u_h(i)\right) \tag{4}$$

$$d_k^t = (T(k) - Y(k)) \times Y(k) \left(1 - Y(k)\right) \tag{5}$$

$$K = 1, 2, \ldots, M$$

- Use connection weights wjk, the error dkt of the output level, and the output value U(j) of the hidden level to calculate the error ejt of each node of the hidden level.
- Use the error of each node of the output level, the output value uh(k) to correct the connection weights wjk and the threshold value γk

$$w_{jk}(1 + t) = w_{jk}(t) + \alpha d_k^t U(j) \tag{6}$$

$$\gamma_k(t + 1) = \gamma_k(t) + \alpha d_k^t \tag{7}$$

Use the errors of each node of the hidden level, the input value P = (a1, a2, ... ,an) of each node of the input level to correct the connection weights wij and the threshold value θj.

$$\theta_j(t + 1) = \theta_j(t) + \beta e_j^t \tag{8}$$

$$\theta_j(t + 1) = \theta_j(t) + \beta e_j^t \tag{9}$$

$$i = 1, 2, \ldots, N \quad j = 1, 2, \ldots, M \quad 0 < \beta < 1$$

Here, β is the coefficient of learning efficiency.
- Select next sample randomly, and return to step (3), until the total P samples are all be practiced.
- Select one sample from the P samples again, and return to step (3), until the global errors are less than the pre-set minimum value. At last, the process of learning is over.

4 Analysis of Simulation

In this paper, we select the historical data of 160 types of spare parts of a logging service as samples, therein, the data of 100 types of spare parts as training sample, the data of other 60 types of spare parts as test samples.

Table 1 Forecasting accuracy of the spare parts classified by the neutral network mode l(%)

Total training samples	84.00
Classification of the training samples	82.50
Level H	76.80
Level M	85.60
Level L	86.30
Classification of the test samples	85.70
Level H	100.00
Level M	76.70
Level L	84.40

The method of BP training simulation is as the following: selecting randomly one type of spare parts, and then the results of the fuzzy evaluation of the five indicators of this type of spare parts as input value of the input variables of neutral network model, and then the evaluation results of the risk level of this type of spare parts as the target output value of the neutral network model. And then following the steps described by the above to calculate the SSE, and then to calculate the connection weights between the input level neutrons and the hidden level neutrons, and the connection weights between the hidden level neutrons and the output level neutrons in the new neutral network model, until the SSE is less than 0.001, and then the training simulation is over. At last, we can obtain the two connection weights. The forecasting accuracy of the 160 types of spare parts classified by the neutral network model is shown in Table 1.

5 Conclusions

In this paper, we proposed a decision-making method to determine the level of risk of the spare parts based on neutral network, which could provide a practical decision-making tool for enterprises to make full use of the information management platform, and then achieve more accurate and effective management of the inventory of spare parts. The method proposed in this paper is applicable to the inventory management of spare parts of wide variety, complex sources, small batch, non-standard custom or import controls.

The results of this study can be used to predict the risk level of inventory of spare parts, helping enterprises take appropriate risk management strategies to improve the effectiveness of inventory management of spare parts according to different risk levels of spare parts. But with thousands of problems of spare parts management, we need to consider the costs of the risk factors of inventory management. Under the circumstance of the minimal loss of out of stock, our further research is the lowest inventory costs.

References

Aronis KP, Magou I, Dekker R, Tagaras G (2004) Inventory control of spare parts using: Bayesian approach; a case study [J]. Eur J Oper Res 154:730–739

Chang PL, Chow YC, Huang MG (2005) A (r, r, Q) inventory model for spare parts involving equipment criticality [J]. Int J Prod Econ 97:66–74

Fisher M, Rajaram K, Raman A (2001) Optimizing inventory replenishment of retail fashion products [J]. Manuf Serv Oper Manag 3(3):230–241

Groston JD (1972) Forecasting and stock control for intermittent demands [J]. Oper Res Q 23(3):289–303

Hua ZS, Zhang B, Yang J, Tan DS (2006) A new approach of forecasting intermittent demand for spare parts inventories in the process industries [J]. J Oper Res Soc 57:889–898

Huiskonen J (2004) Maintenance spare parts logistics: special characteristics and strategic choices [J]. Int J Prod Econ 71:125–133

Liu Liwen, Wang Jun (2006) Inventory and ordering management with consideration of the supply chain logistics characteristics [J]. Ind Eng Manage 6:19–23

Willemain TR, Smart CN, Schwarz HF (2004) A new approach to fore-casting intermittent demand for service parts inventories [J]. Int J Forecast 20:375–387

Xu Xiaoyan (2006) A demand classification based approach of inventory management for spare-parts and it's application [J]. Theory Appl Syst Eng 2:62–67

Improved Spam Filtering Algorithm Based on AIS

Jue Huang and Changwu Liao

Abstract Arming at the inefficiency in generation of self libraries and gene libraries, it added MI and the frequency of words and improved the auto-update strategy of self libraries and gene libraries. The result of experiments showed that the new algorithm can reduce the false positives rate and at least 10% of the time.

Keywords Artificial immune • Spam • Self libraries • Gene libraries • Algorithm

1 Introduction

Spam as the carrier of advertisement, virus, reactionary content threats normal email systems operation seriously, anti-spam has became the important subject in the world.

In recent years, people begin to resolve project issues by means of Artificial Immune System model (Khorsi 2007; Qing et al. 2009), Andrew Secker proposed AISEC Algorithm model and it has been used to detect spam successfully (Secker et al. 2003). The model used independent memory and mature antibody libraries, gene libraries and self libraries. It's good at variety, adaptation and self-learning. The algorithm has obvious improvement in recall than Bayesian. It can auto turn antibody libraries according to environment change and customer response. The algorithm is good at processing unknown email.

J. Huang (✉) • C. Liao

Department of Computer and Software, Nanjing Institute of Industry Technology, Nanjing, China
e-mail: huangj@niit.edu.cn; liaocw@niit.edu.cn

With the study unceasingly thorough to AISEC, people incessantly improve the model in their project (Prattipati and Hart 2008; Wang and Wang 2006; Zhu and Zhuang 2008; Zhang et al. 2007). A modification (Prattipati and Hart 2008) is proposed to the learning process of the algorithm and to the mutation operator. Wang and Wang (2006) improved the antibody update process of AISEC by Virtual Gene Libraries technology, It use antibody in the correct classification to improve the quality of antibody and don't delete antibody in the wrong classification simply, but to give them hyper mutation, keep the update of antibody conform and coherent. The experimental results show that this improvement can achieve higher classification accuracy rate. Zhu and Zhuang (2008) designed an algorithm to produce detector which is based on weight of gene. The algorithm can achieve higher efficiency of classification and reduce false positive. Zhang et al. (2007) improved the accuracy rate and the speed of classification by incremental learning algorithm. All of these improvements don't involve the generation of self and gene libraries. According to the negation algorithm we can know that the generation efficiency of self and gene libraries will affect the performance of whole spam detect. The paper will give an improvement algorithm to produce self and gene libraries, and improve the auto-update strategy of self and gene libraries.

2 AISEC Model

First, AISEC splits gene sections from spam samples and add them to gene libraries. Original immature antibody is composed of gene libraries randomly. Mature antibody evolves from immature antibody by tolerating with self libraries. After receiving emails, AISEC classifies them and updates antibody gene libraries according to the classified result. Finally, it updates the antibody's lifetime according to user's response and Prolongs the lifetime of antibody participated in the correct classification. It Deletes antibody participated in the classification from antibody libraries directly if the classified result is wrong.

3 The Defect and Improvement of AISEC

The defect of the model is generating gene and self section randomly and can't think about weight. So each section has the same weight and it's not optimum. It will affect system performance.

First, each word has different distribution for classification of mail. Some words have big distribution, but some words have not. They even no any distributions for classification of mail. If we don't think about the fact, just select and put them to self or gene libraries randomly, it will make self or gene libraries too big. It will downgrade the efficiency of classification.

Second, each word has different frequency in email. So, high frequency antibody should be used firstly.

Third, it Prolongs the lifetime of antibody participated in the correct classification and deletes antibody participated in the wrong classification directly. The update mechanism is too single and simple and lacks quantitative analysis.

Base on the preceding analysis, we will add two attributes for self and gene section-MI (Mutual Information) and word frequency. The following is the detail.

- During the generation of new self libraries and gene libraries, we control these size by calculate MI (Mutual Information).
- During the email detecting, we add word frequency for each self and gene and mach high frequency self or detector firstly to improve system detect performance.
- After receiving user response, if false positive appear, we should punish self and gene by tuning the frequency of self and gene.

4 The Generation Algorithm of Self Libraries and Gene Libraries After Improved

4.1 The Generative Process of Self Libraries and Gene Libraries

The algorithm generates self libraries, gene libraries and detector during training. The Fig. 1 shows the generation of self libraries and detector.

4.2 Calculate the Mutual Information

MI (Mutual Information) is a standard of establishing implicative statistic module. The more MI we get, the more probability words co-occurrence. We can take MI differentials as evaluate function. He provides food and clothes for his family.

$$difference\,(W) = \left\| \left| \log \frac{P(W|C_1)}{P(W)} \right| - \left| \log \frac{P(W|C_2)}{P(W)} \right| \right\| \tag{1}$$

P(Cj):the appeared probability of the document j in all document; P(W):the word W appeared probability in all documents; P(W|Cj):the word W appeared probability in document j.
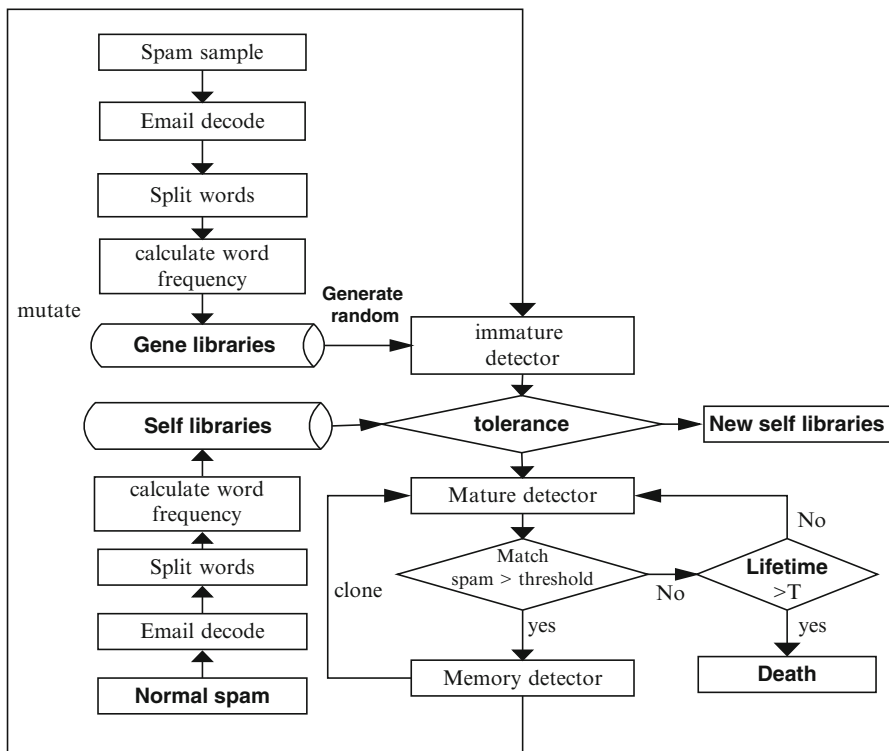
Fig. 1 The generation of self libraries and detector

We calculate MI by the improved formula (1), we need set a threshold. If one word appeared in mail sample, its MI bigger than the threshold, it can be considered as a characteristic item, here we set the threshold is 0.3 (the value of the threshold is connection with mail sample).

5   The Updating Strategy of Self Libraries and Gene Libraries

False positives are generally considered to be more harmful than false negative. We deal with the spam by user synergy stimulation signal. User cooperation mechanism imitated principle of biology immunity T cell. The mechanism improved system's detection performance. If false positive occurs, for avoid happening again, the frequency increment of word appeared in self libraries is equal to the word exposure frequency in the mail * weight; the frequency decrement of word appeared in gene libraries is equal to the word appeared frequency in the mail * weight;

6 Examination Result Analysis

6.1 Examination Samples

The sample used in the system is the same as Gu and Chen (2008). It includes 4,520 spam mails and 4,480 normal mails. We separated these mails into nine groups randomly. It includes seven groups for training and two groups for testing. In order to obtain the objective examination result, we tested many times and calculated average of the test result as the final test result.

6.2 Examination and Result Analysis

Examination one. Size of self libraries influence on system performance.

We will evaluate system performance by precision, recall and accuracy. Precision stands for correct rate of spam and high value shows less normal mails were judged to spam. Recall stands for call back rate and high value shows less spam were judged to normal mails. Accuracy stands for correct rate of all samples.

We can see from the Fig. 2 that if the system has less self detectors, the precision is low, the Recall is high. Because too less self can stand for normal mail character, many normal mails were regarded as spam. Along with the self libraries increasing,
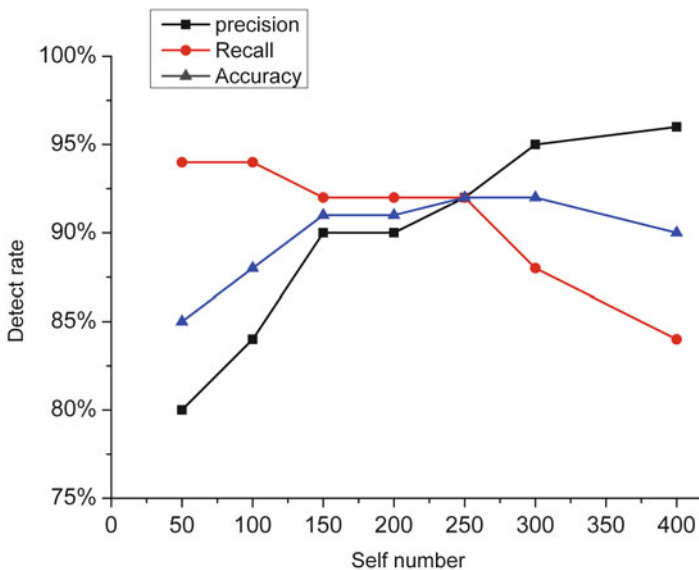
Fig. 2 Size of self libraries influence on system performance

Table 1 Detection speed contrast

Mail num	Self num	Mature detector num	Memory detector num	AISEC spend time(s)	Improved AISEC spend time(s)	Improve percentage (%)
2,000	50	100	50	1,024	904	11.7
2,000	50	100	70	760	668	12.1
2,000	50	100	90	606	528	12.9
2,000	100	150	50	1,166	1,008	13.6
2,000	150	200	50	1,362	1,148	15.7
2,000	200	250	50	1,610	1,318	18.1
2,000	250	300	50	1,928	1,536	20.3

precision will increase and Recall will reduce. After the number of self libraries reaches to 250, precision and recall will reach to best value. If the number of self libraries continues to increase, recall will continue to reduce and precision will continue to increase. If we select too many self, some unrepresentative words have been added to self libraries. When mail is detected, many words match with self, the mail will be regarded as normal email. So, recall will reduce.

Examination two. Improved algorithm compare with AISEC algorithm.

- Detection speed
  After adding frequency of word and MI, the improved algorithm has faster speed than AISEC. In Table 1, system spends more time to detect mail when increasing the number of self or mature detector. If it freezes the number of self and mature detector and just moderately increase memory detector number, it can reduce detect time because of the more representative memory detector. After adding frequency of word to self and detector, when mail detects, it matches high frequency of word firstly, so the improved algorithm will have higher detect speed than AISEC. The experimental results show that the improved algorithm can reduce 10% of the time.
- False positives rate contrast
  In Fig. 3, AISECG stand for improved algorithm, along with detector number increasing, false positives rate will increase, because we add MI, the character will be more representative. Otherwise, the self and gene update policy have been improved, by processing customer response, it punishes the characters take part in error detect. The improved algorithm's false positives rate has obvious reduce.

7 Conclusion

This paper analyzed the shortage of AISEC algorithm module in generation efficiency of self and gene libraries, proposed an improved algorithm: adding word frequency and MI to self and gene libraries, updating strategy of self libraries and gene libraries. The improved algorithm considered sufficiently the importance

Fig. 3 False positives rate contrast
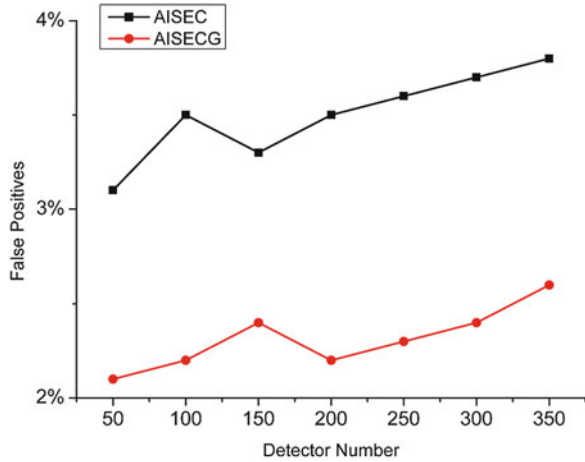
of word frequency and MI in mail detection system. We experimented with this algorithm and the result showed that the improved algorithm can improve at lease 10% detection performance and can improve false positives rate obviously.

References

Gu LH, Chen B (2008) The research and implement on the Chinese anti-spam filtering system based on advanced Winnow algorithm. Me D Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing

Khorsi A (2007) An overview of content-based spam filtering techniques. Informatica 31(3):269–277

Prattipati N, Hart E (2008) Evaluation and extension of the AISEC e-mail classification system. In: ICARIS 2008, LNCS 5132, Phuket, Thailand, pp 154–165

Qing JJ, Mao RL, Bie RF, Gao XZ (2009) An AIS-based e-mail classification method. In: ICIC2009, Ulsan, Korea, pp 492–499

Secker A, Freitas AA, Timmis J (2003) AISEC: an artificial immune system for e-mail classification. In: The 2003 congress on evolutionary computation, California, USA, pp131–138

Wang XW, Wang LM (2006) Dynamic artificial immune classification algorithm for e-mail. J Comput Appl 26:2463–2465

Zhang ZM, Luo WJ, Wang XF (2007) Individual spam filtering algorithm based on immune principles. Pattern Recognit Artif Intell 20:406–414

Zhu H, Zhuang Y (2008) The research of spam filtering technology based on artificial immunity system. Me D Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing

Network Monitoring Model Extension or Reduction of The Mathematical Algorithms

Xu Guanglian

Abstract In practical network, the nodes have capacity, a circulation flow in the node has the phenomenon of hysteresis, making the practical application of network flow theory and the distance between. Nodes in the network based on the inflow and outflow balance theory and the node point and a separate entry points for the method of loops by the node, the establishment of an improved mathematical model of network monitoring N_{IO}. Introduction of network flow state period T, given the network nodes N_{IO} increase and decrease, the network of local compression and expansion algorithm, the network model shows the adaptability and flexibility.

Keywords Network flow • Network model • Node loop flow • Circular flow

1 Introduction

Network flow problems in engineering,   computer science and technology, communications network technology, transportation, regional planning and other fields have a wide range of applications, from the twentieth century, 50 years since many of the scholars through the accumulation of a large number of studies, the network flow method has become more mature theoretical system (Gao 2009). Network Model *N* vertices and directed edges from the composition. With the development of modern science and technology, in particular the system of things set up with a trend accelerated markedly, these network systems urgently need a suitable network

XU GuangLian,1954.5-, An associate professor of mathematics, research: Applied Mathematics, network flow programming.

X. Guanglian (✉)
Electronic Engineering Department, Hebi College of Vocation and Technology, Hebi, China
e-mail: xuglian@126.com

flow model with a monitor. As the model in the vertex N ignore the logistics flow between the loading, unloading, transfer, temporary storage, consumption or value-added and other errors and asked for the new network flow model to compensate for this defect. Network flows model N_{IO} is makes the improvement in the model N foundation (Xu and Ma 2010), both has the model N nature, and can express the node loop which in model N neglects, enhanced the mathematical model precision. Because is restricted length, here only discusses the network model N_{IO} node to increase the algorithm which the reduction and the compression launch.

2 Network Model N and N_{OI}

2.1 Simple Network Model N

Model N uses the edge of the simple weighted directed graph. Network flow theory is based on the model in the N.

Definition 1 Let $N = (V, A, s, t, C)$ is a directed graph, where V is the set of vertices, A is a directed edge (arc) set, $s \in V$ is a source, in-degree is 0; $t \in V$ is a sink, Out-degree is 0. C is a non-negative capacity function on a directed edge (arc). If v_i, v_j $\in V$, called arc (v_i, v_j), v_j for the arc in the head, v_i for the arc tail; arc capacity is denoted by $C(v_i, v_j)$ or $C(i,j)$ or C_{ij}; the flow on the arc denoted $f(v_i, v_j)$ or $f(i,j)$ or f_{ij}. Calls this kind network that with source and sink the capacity network. In $V - \{s,t\}$ vertices is called the transit vertices (Wang 2009); Assigns network $N = (V, A, s, t, C)$, the number of vertices in V as $|V| = n$, the number of arc in arc set A as $|A| = m$.

Definition 2 If the flow f defined on N in the network to satisfy the following condition:

1. Capacity constraints: for each arc $(v_i, v_j) \in A$,

$$0 \leq f_{ij} \leq C_{ij}; \tag{1}$$

2. Equilibrium condition: Regarding besides the source and the sink transit vertices v_i, has:

$$\sum_{(v_i,v_j)\in A} f_{ij} - \sum_{(v_j,v_i)\in A} f_{ji} = 0; \tag{2}$$

It called the feasible flow.

The expression $\sum_{(v_s,v_j)\in A} f_{sj} - \sum_{(v_j,v_t)\in A} f_{jt}$ is called the feasible flow f the current capacity, denoted by $v(f)$.

Note: The vertex of graph theory corresponds to a network is called node.

Fig. 1 Graph with loop flow
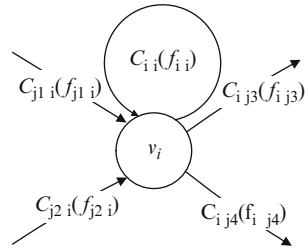
Fig. 2 Node separates
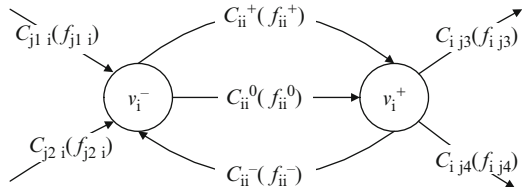
2.2 Network Mathematical Model N_{IO}

Model N is converted to model N_{IO} rules are as follows:

1. multiple arcs (parallel arcs) merge rule:
   If v_i, $v_j \in V$, moreover v_i arrives at v_j to have l the directed edges, then may merge is, namely

$$a_{ij} = (v_i, v_j) = \bigcup_{k=1}^{l} (i_k, j_k), \tag{3}$$

   and $$C_{ij} = C(i,j) = \sum_{k=1}^{l} C(i_k, j_k), \quad f_{ij} = f(i,j) = \sum_{k=1}^{l} f(i_k, j_k). \tag{4}$$

2. In the node loop flow separates the rule:
   Assuming that each node there arc head and tail are the same node v_i arc loop, denoted by $a_{ii} = (v_i, v_i)$. Ring with capacity set to C_{iiloop}; loop with traffic, denoted by C_{iiloop}. Node has the capacity, denoted by C_{ii}, while node in a flow, denoted by f_{ii}. All the nodes are split into a pair of nodes that is v_i^- and v_i^+, which is only one arc from v_i^- to v_i^+. If the node does not have the loop, then the capacity (flow) is 0.Understand the node loop of the change process node is divided into three steps, respectively, in Figs. 1, 2, and 3. In Fig. 2, C_{ii}^0 (f_{ii}^0) says node positive transfer capacity (transfer flow) (Wang 2009).

Theorem 1 *According to the multiple arcs (parallel arcs) merge rule and the node loop flow separates the rule that the two transformation rules, the network model is constructed in the N_{IO} diagram is a simple graph.*

Fig. 3 Separates node in N_{IO}
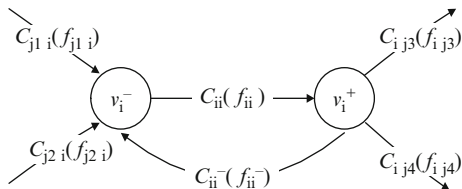
Fig. 4 Adjacency matrix M_{IO}

$$M_{IO} = \begin{bmatrix}
(v_s^-,v_s^+) & (v_s^-,v_s^+) & (v_s^-,v_1^+) & 0 & (v_s^-,v_2^+) & 0 & \cdots & (v_s^-,v_n^+) & 0 & 0 & 0 \\
(v_s^+,v_s^-) & (v_s^+,v_s^+) & (v_s^+,v_1^+) & 0 & (v_s^+,v_2^+) & 0 & \cdots & (v_s^+,v_n^+) & 0 & 0 & 0 \\
0 & 0 & (v_1^-,v_1^-) & (v_1^-,v_1^+) & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & (v_1^+,v_1^-) & (v_1^+,v_1^+) & (v_1^+,v_2^+) & 0 & \cdots & (v_1^+,v_n^+) & 0 & (v_1^+,v_t^-) & (v_1^+,v_t^+) \\
0 & 0 & 0 & 0 & (v_2^-,v_2^-) & (v_2^-,v_2^+) & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & (v_2^+,v_1^-) & 0 & (v_2^+,v_2^-) & (v_2^+,v_2^+) & \cdots & (v_2^+,v_n^+) & 0 & (v_2^+,v_t^-) & (v_2^+,v_t^+) \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & (v_n^-,v_n^-) & (v_n^-,v_n^+) & 0 & 0 \\
0 & 0 & (v_n^+,v_1^-) & 0 & (v_n^+,v_2^-) & 0 & \cdots & (v_n^+,v_n^-) & (v_n^+,v_n^+) & (v_n^+,v_t^-) & (v_n^+,v_t^+) \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & (v_t^-,v_t^-) & (v_t^-,v_t^+) \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & (v_t^+,v_t^-) & (v_t^+,v_t^+)
\end{bmatrix}$$

Proof. According to the simple graphs neither multiple arcs again without the loop of definition (Gao 2009), according to rule 1, the graph of the multiple arcs merge into a arc, namely remove multiple arcs, According to the rule 2, node separately for the in-node and the out-node, then a node loop separately for positive arc, negative two arcs, these two nodes all without loop. Because positive arc and positive transfer arc is multiple arcs (parallel arcs), merged, so the graph without the multiple arcs, so the network graph is no multiple arcs, node without loop to be a simple graph.

Definition 3 network model $N_{IO} = (V^-, V^+, A, s^-, s^+, t^-, t^+, C)$ is a simple graph, where V^- and V^+, the node after the separation of the entry nodes set and the out nodes set, A is a directed edge (arc) set, s^- and s^+ is the source entry and the source out, t^- and t^+ is the sink entry and sink out, the C is a directed edge (arc) on the capacity of the function.

The network model N_{IO} correspondence's directed edge (arc) the adjacent matrix expression is M_{IO}, namely Fig. 4.

Definition 4 In a simple network $N = (V, A, s, t, C)$, all flow f is feasible, and is at the invariable steady state, this condition maintains the time is called flows f feasible one period, expressed with **T**.

Actually the network flow change is with timeliness, this paper mainly research network flow stable state, the relation between does not involve changes of timeliness.

Feasible flow period T_i ($i = 1, 2, \ldots$). There are three qualities:

Property 1 The period T_i to T_{i+1} is the step of conversion, with no intermediate state.

Property 2 The period compared to T_i and T_{i+1}, each possession is not such a long period of time. (Similar to the replacement train schedule time is long range).

Property 3 In two consecutive period in T_i and T_{i+1}, the T_i and T_{i+1} state can be the same or different, but if the state of T_i and T_{i+1} different, must be the two period.

3 Nodes of Network Model N_{IO} Increases and Decreaes Algorithm

3.1 Nodes Increases Algorithm

In the network N_{IO}, if increased a node, the flow state changing, and enter a new state. In the model N_{IO}, set to increase the node v_{n+1}, that is, to add a pair N_{IO} nodes $v_{n+1}{}^-$ and $v_{n+1}{}^+$, will have to increase the entry arc $\{(s^+, v_{n+1}{}^-), (v_i{}^+, v_{n+1}{}^-) \mid i = 1, 2, \ldots, n\}$, to increase the out arc $\{(v_{n+1}{}^-, t^-), (v_{n+1}{}^+, v_j{}^-) \mid j = 1, 2, \ldots, n\}$, the node to increase loop for the $\{(v_{n+1}{}^-, v_{n+1}{}^+), (v_{n+1}{}^+, v_{n+1}{}^-)\}$, expanded the entry node set is denoted by $V_{n+1}{}^-$, the out node set is denoted by $V_{n+1}{}^+$, arc denoted as A_{n+1}, the capacity of the arc set is denoted by C_{n+1}. Algorithm steps:

1. To expand the node set and arc set.
   node set

$$V_{n+1}{}^- = V_n{}^- \cup v_{n+1}{}^-; \quad V_{n+1}{}^+ = V_n{}^+ \cup v_{n+1}{}^+; \tag{5}$$

   arc set

$$A_{n+1} = A \cup \left\{ (s^+, v_{n+1}{}^-), (v_{n+1}{}^+, t^-), (v_{n+1}{}^+, v_{n+1}{}^-), (v_{n+1}{}^-, v_{n+1}{}^+), \right.$$
$$\left. (v_i{}^-, v_{n+1}{}^+), (v_{n+1}{}^+, v_j{}^-) \mid i = 1, 2, \ldots, n, j = 1, 2, \ldots, n \right\}. \tag{6}$$

2. Flow balancing.
   If increases newly a pair of nodes can form an independent connected component, then only need increase source s^+ to $v_{n+1}{}^-$ and $v_{n+1}{}^+$ to the sink t^- flow then, namely has:
   Capacity set

$$C_{n+1} = C_n \cup \left\{ C(s^+, v_{n+1}{}^-), C(v_{n+1}{}^+, t^-), C(v_{n+1}{}^-, v_{n+1}{}^+), \right.$$
$$\left. C(v_{n+1}{}^+, v_{n+1}{}^-) \right\}. \tag{7}$$

   The corresponding arc the flow is:

$$f(s^+, v_{n+1}{}^-), \ f(v_{n+1}{}^+, t^-), \ f(v_{n+1}{}^-, v_{n+1}{}^+), f(v_{n+1}{}^+, v_{n+1}{}^-) \tag{8}$$

Total flow:

$$v(f)_{n+1} = v(f)_n + f(s^+, v_{n+1}{}^-) = v(f)_n + f\left(v_{n+1}{}^+, t^-\right) \qquad (9)$$

If increases newly a pair of nodes do not constitute an independent connected component, then Considers the original node connection, so must maintain the node flows in with balanced which flows out, namely through adjustment, then has:

Capacity set

$$\begin{aligned} C_{n+1} = C_n \cup \{ &C\left(s^+, v_{n+1}{}^-\right), C\left(v_{n+1}{}^+, t^-\right), C\left(v_{n+1}{}^-, v_{n+1}{}^+\right), \\ &C\left(v_{n+1}{}^+, v_{n+1}{}^-\right), C\left(v_i{}^-, v_{n+1}{}^+\right), C\left(v_{n+1}{}^+, v_j{}^-\right), \\ &| i = 1, 2, \ldots, n, j = 1, 2, \ldots, n \}. \end{aligned} \qquad (10)$$

The corresponding arc the flow is:

$$f\left(s^+, v_{n+1}{}^-\right) \cup \sum_{i=1}^{n} f(v_i{}^+, v_{n+1}{}^-) = \sum_{j=1}^{n} f(v_{n+1}{}^+, v_j{}^-) \cup f\left(v_{n+1}{}^+, t^-\right) \qquad (11)$$

All node flow balance:

$$f\left(s^+, v_{n+1}{}^-\right) \cup \sum_{i=1}^{n}\sum_{j=1}^{n} f(v_i{}^+, v_j{}^-) = \sum_{j=1}^{n}\sum_{i=1}^{n} f(v_j{}^+, v_i{}^-) \cup f\left(v_{n+1}{}^+, t^-\right) \qquad (12)$$

Total flow $v(f)_{n+1}$ must the recompilations.

Note: In theory, the program has a lot of trim, for example, re-calculation of maximum flow algorithm, or simply for diversion and other methods, application to the actual situation and optimization.

3. In increases two lines and two rows with N_{IO} on corresponding matrix M_{IO} the adjacency element which constitutes by node $v_{n+1}{}^-$ and $v_{n+1}{}^+$.
4. State change, increasing the node status before T_1, increased to T_2.

3.2 Delete Node Algorithm

In the network N_{IO}, if you delete a node, the flow state periods is changing, and enter another state period. Adjacency matrix M_{IO} used here the nature of the theorem.

In matrix M_{IO}, to calculate conveniently, the geminate node cannot inverted. The M_{IO} matrix formally, besides source s^-, s^+ occupies the first line generally (first row) and sink t^-t^+ occupies follows (tail row), other geminate node subscript serial number relies on the node numeral order.

Theorem 2 *The matrix which and original matrixM_{IO}after a series of lines in the matrixM_{IO}doubling's node (row) exchanges obtains is equal.*

Proof. M_{IO} is one of 2n th-order square matrix, through a series of row (column) exchange, the matrix is obtained M_{IO}'. If the matrix is obtained by elementary transformation, this matrix is equivalent to the original matrix (Wang and Shi 2003).Because in matrix M_{IO} of row (column) exchange is elementary transformation, so income matrix M_{IO}' with original matrix M_{IO} equivalent.

By the Theorem 2, the matrix M_{IO} can be on line (column) exchange method to adjust the position of the nodes, for example may the node which is going to delete adjust v_n^- and in the v_n^+ position carries on the computation.

In the N_{IO} model, supposes must delete a pair of N_{IO} nodes is v_n^- and v_n^+, then must delete the enter arc set for $\{(s^+, v_n^-), (v_i^+, v_n^-)| \ i = 1,2,\ldots, n-1\}$, must delete the out arc set for $\{(v_n^+, t^-), (v_n^+, v_j^-)| \ j = 1,2,\ldots, n-1\}$, must delete node loop for $\{(v_n^+, v_n^-),(\ v_n^-, v_n^+)\}$,after deleting to the enter set of nodes is V_{n-1}^-, leaves the set of nodes is V_{n-1}^+, the arc set records is A_{n-1}, on the arc capacity set records is C_{n-1}.

Algorithm steps:

1. Deletes node v_n^- and node v_n^+ from the node set, deletes their related arc from the arc set.node set

$$V_{n-1}^- = V_n^- \setminus \{v_n^-\}; V_{n-1}^+ = V_n^+ \setminus \{v_n^+\}; \tag{13}$$

   arc set

$$A_{n-1} = A_n \setminus \{(s^+, v_n^-), (v_n^+, t^-), (v_n^+, v_n^-), (v_n^-, v_n^+), (v_i^-, v_n^+),$$
$$(v_n^+, v_j^-) \ |i = 1, 2, \ldots, n, j = 1, 2, \ldots, n\}. \tag{14}$$

2. Flow balancing.

   If the node must try to be delete is an independent connected component, then only need delete source s^+ to v_n^- and v_n^+ to sink t^-, namely has:
   Capacity set

$$C_{n-1} = C_n \setminus \{C(s^+, v_n^-), C(v_n^+, t^-), \ C(v_n^+, v_n^-) \ C(v_n^-, v_n^+)\}. \tag{15}$$

   The corresponding arc the flow is:

$$f(s^+, v_n^-), \ f(v_n^+, t^-), \ f(v_n^+, v_n^-), f(v_n^-, ^+v_n).$$

   Total flow:

$$v(f)_{n-1} = v(f)_n - f(s^+, v_n^-) = v(f)_n - f(v_n^+, t^-). \tag{16}$$

If you want to delete the node is not a separate connected component, connected with the original node, the node will have to maintain a balance of inflow and outflow through the adjustment, there are:

Capacity set

$$
\begin{aligned}
C_{n-1} = C_n - \{ & C\left(s^+, v_n^-\right), C\left(v_n^+, t^-\right), C\left(v_i^+, v_n^-\right), C\left(v_n^+, v_j^-\right), \\
& C\left(v_n^+, v_n^-\right), C(v_n^+, v_n^-) | i = 1, 2, \ldots, n, j = 1, 2, \ldots, n \}.
\end{aligned} \quad (17)
$$

All node flow balance:

$$
\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} f(v_i^+, v_j^-) = \sum_{j=1}^{n-1}\sum_{i=1}^{n-1} f(v_j^+, v_i^-) \tag{18}
$$

Total flow $v\left(f\right)_{n-1}$ to recalculate , remove with a pair of nodes v_n^- and v_n^+ associated with the flow.

3. Deletes two lines and two rows in N_{IO} on corresponding matrix M_{IO} side the adjacency which constitutes by node v_n^- and node v_n^+ the element.
4. Change state period, not delete node of state for T_1, delete the node after network state for T_2.

4   Regional Compression And Expansion Algorithm

4.1   Region k Nodes into One Node Algorithm

If the locally model N_{IO} k-connected nodes, can be combined into a single node, called the merger algorithm compression algorithm. Local model N_{IO} compression, N_{IO} diagram and the corresponding adjacency matrix M_{IO} form of change, but maintain the original state intact.

N_{IO} based network has n nodes, denoted by N_{IO} (n), $| V_n | = 2n$; to merge the k-connected local area network formed by nodes denoted by N_{IO} (k), $V_k \subset V_n$, $|V_k| = 2k$; merged into a node, denoted by v_{n-k+1}, the N_{IO} said the v_{n-k+1}^- and v_{n-k+1}^+, the merged network is denoted by N_{IO} (n − k + 1), $| V_{n-k+1} | = 2(n-k+1)$.

The N_{IO} (k) merge divides into two steps, namely flows out the flow from the node to merge or the inflow node flow must merge, circulation merge.

1. Node out arc, into the arc merger:
   Enters new node v_{n-k+1}^- to enter the arc flow = into regional node flow = the out flow of arcs from regional nodes = out flow of arc from new node.v_{n-k+1}^+

Namely

$$f(s^+, v_{\text{n-k+1}}{}^-) + \sum_{i=1}^{n-k} f\ (v_i{}^+, v_{\text{n-k+1}}{}^-) = \sum_{j=1}^{k} f(s^+, v_j{}^-) + \sum_{i=1}^{n-k} f(v_i{}^+, v_j{}^-)$$

$$= \sum_{j=1}^{k} f(v_j{}^-, t^+) + \sum_{j=1}^{n-k} f(v_j{}^+, v_i{}^-) = f(v_{\text{n-k+1}}{}^+, t^-) + \sum_{i=1}^{n-k} f\ (v_{\text{n-k+1}}{}^+, v_i{}^-)$$

$$(19)$$

Among them, $v_i \in V_n - V_k$, $v_j \in V_k$, $i = 1, 2, \ldots, n-k$; $j = 1, 2, \ldots, k$.

2. Circulation merge

Compression, the new node v_{n-k+1} loop = loop + node region circular flow within the region. Node loop denoted as: $f(v_{n-k+1}{}^+, v_{n-k+1}{}^-)$.Circulation within the region denoted by the node: $\sum_{i=1}^{k} f(v_j{}^+, v_j{}^-)$.Requests the region recycling flow needs to introduce a feasible circulation flow the definition.

Definition 5 A simple network N, if there is an arc set A, flow to meet:

1. $\forall v_i, v_j \in V - \{s, t\}$, satisfied:

$$\sum_{v_i, vj \in V-\{s,t\}} f(v_i, v_j) - \sum_{v_i, vj \in V-\{s,t\}} f(v_i, v_j) = 0; \qquad (20)$$

2. $\forall (v_i, v_j) \in A$, satisfied:

$$0 \le f(v_i, v_j) \le C(v_i, v_j). \qquad (21)$$

Then said that f is a network N internal recycling flow, is also called feasible circulation flow (Gao 2009). Internal recycling flow search method (Sedgewick 2003):

(1) In the simple network N in a flow never to 0 the tail of the node v_i arcs starting in the head along the arc of a flow within the meaning of any one arc forward, until it encounters an already visited node v_i far, that detects a ring, is circulation flow .

(2) Along the ring back to find a minimum flow of the arc, each arc in the ring minus the amount of flow, then the ring at least one arc on the flow 0, (for example, to find out to which the minimum flow arc), then took out a ring.

(3) Repeats (1), (2) two steps until no ring so far, it has at most $|V| = n$ rings.

Now on N_{IO} (k), then has $|V_k| = 2k$, because N_{IO} (k) is also one kind of simple network, its algorithm and the N network is the same. Because the node

loop flow is also one kind of internal recycling flow, altogether has k, if already removed, the region internal recycling flow is:

$$f = \sum_{i=1}^{k-1} f_i \qquad i = 1, 2, \ldots k - 1. \tag{22}$$

$k - 1$ expresses internal recycling class most integers.

Therefore, after the compression, in node loop = local area network circulation, namely:

$$f(v_{n-k+1}{}^+, v_{n-k+1}{}^-) = \sum_{i=1}^{n} f(v_j{}^+, v_j{}^-) + \sum_{i=1}^{k-1} f_i \tag{23}$$

Note: Regional nodes on the arc, to enter the arc flow merger, the new node in flow terms, does not affect other nodes in the state, does not affect the total amount of sources and sinks, to maintain the original operating condition.

4.2   In the Region a Node Expansion is k Node Algorithm

N_{IO} local network if a node can be expanded into k-connected nodes, which need to use a node expansion algorithm. Network N_{IO} expanded, network N_{IO} and the corresponding adjacency matrix M_{IO} form may have changed, but able to maintain the original status quo. Supposes in the original network to have n node, records makes N_{IO} (n),$|V_n| = 2n$. The random node may expand, might as well supposes must the nth node expansion be k-connection node, then the nth node records is v_n, in N_{IO} the expression is $v_n{}^-$ and $v_n{}^+$. Expands k spot composition's regional network, records makes N_{IO} (k),$|V_k| = 2k$, after the expansion network records is N_{IO} $(n + k - 1)$, $|V_{n+k-1}| = 2(n + k - 1)$.

N_{IO} in node v_n, namely $v_n{}^-$ and $v_n{}^+$, must carry on the expansion, has 3 prerequisites:

1. Have been identified within the layout local network;
2. The flow into $v_n{}^-$ = flow into the region = flow out of the region
   = the flow out of $v_n{}^+$
3. Node v_n loop = in region circulation sum
   = region internal node loop sum+region internal recycling flow sum.

If the above three conditions had been reached, the N_{IO} (n) expanded into N_{IO} $(n + k - 1)$ is divided into two steps.

1. The inflows node into arc flow matching for regional flow; inflows Outflow node for an arc flow matching regional flow out, namely:
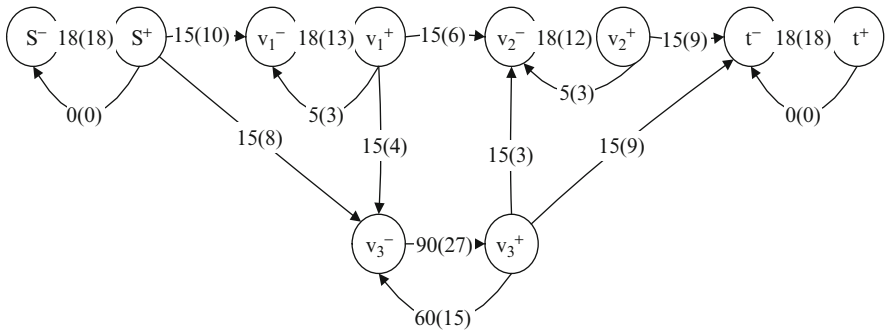
Fig. 5 Compression of graphs
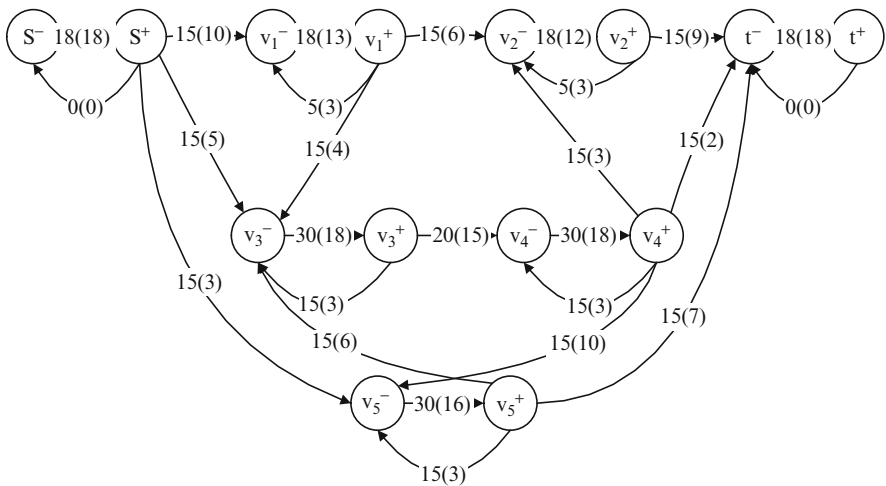
Fig. 6 Extended of graphs

$$f(s^+, v_n^-) + \sum_{i=1}^{n-1} f(v_i^+, v_n^-) = \sum_{i=1}^{k} f(s^+, v_i^-) + \sum_{i=1}^{n-1} \sum_{j=1}^{k} f(v_i^+, v_j^-)$$

$$= \sum_{j=1}^{k} f(v_j^+, t^-) + \sum_{j=1}^{k} \sum_{i=1}^{n-1} f(v_j^+, v_i^-) = f(v_n^+, t^-) + \sum_{i=1}^{n-1} f(v_n^+, v_i^-)$$

$$(24)$$

Among them, $v_i \in V_{n-1}$, $v_j \in V_k$, $i = 1, 2, \ldots n-1; j = 1, 2, \ldots, k$.

For example, to see Fig. 5 compression of graphs, and Fig. 6 extended of graphs.

The prerequisite (2) means that enters the region total quantity and the node receive total quantity is equal, but makes concrete to the region outside some node loop, if enters in the region some node, not necessarily is the match. Has the match plan, and the plan is not only. Below gives an algorithm, one side the solution according to enters the region the order, with one side according to enters the node in addition the order simple match question in order.

Like Figs. 5 and 6, those who enter the region:

$f_0 = 4 = f(v_1^+, v_3^-)$; $f_1 = 8 = f(s^+, v_3^-)$, enters region various nodes the class is:

$g_0 = 4$ (enters v_3^-), $g_1 = 5$ (enters v_3^-), $g_3 = 3$(enters v_5^-), the match result is: $f_0 g_0 = 4$; $f_1 g_1 = 5$; $f_1 g_2 = 3$, Matching algorithm:

```
public class match{
int f[]={4,8};int g[]={4,5,3};int fg[][]=new int[10][10];
public static void main(String args[])
   {match obj=new match();obj.comp();}
public void comp(){
int state=0,i=0,j=0;boolean flag=true;
  for (i=0;i<=f.length-1;i++){
  do{ if (f[i]>g[j]) state=1;
     else if(f[i]<g[j]) state=-1;else state=0;
     switch(state){
case 1: fg[i][j]=g[j];f[i]=f[i]-g[j];
        j=j+1;flag=true;break;
case 0: fg[i][j]=g[j];j=j+1;flag=false; break;
case -1: fg[i][j]=f[i];g[j]=g[j]-f[i];
        flag=false; break; }//case end
    }while (flag); }// for loop end
for(i=0;i<10; i++)   // fg[][] output
  for(j=0;j<10; j++)  if (fg[i][j]>0)
  System.out.println(``fg[''+i+''][''+j+'']=''+fg[i][j]);
}}
```

2. Decomposes the v_n node loop into the region internal node loop and the region internal recycling flow.

In node loop = region internal node loop + region circulation, Namely

$$f(v_n^+, v_n^-) = \sum_{i=1}^{k} f(v_j^+, v_j^-) + \sum_{i=1}^{k-1} f_i. \qquad (25)$$

Node loop capacity = regional nodes loop in the flow capacity sum + regional network circulation capacity.

For example, to see Figs. 5 and 6:

Here: the nodes loop capacity each is 15,15,15; regional network circulation capacity, namely Min{15,30,15,30,20,30} = 15; regional network circulation. min $\{f(v_3,v_4) = 15, f(v_4,v_5) = 10, f(v_5,v_3) = 6 \} = 6$, nodes loops are three, $3 + 3 + 3 + 6 = 15$.

5   Conclusion

The N_{IO} model is establishes according to the network flows theory (Xu and Ma 2010), it can adapt the flow change and the network change. On this model's data may take the research network flow the basis, in the national economic planning and the research, has the vital significance. As a result of the length limit, others, like the network optimizes the circulation emergency to affect questions and so on (Xu and Shang 2010), circulation shift, circulation control, the separate article discussion.

References

Gao SX (2009) Graph theory and network flow theory. Higher Education Press, Beijing, pp 260–314 (in Chinese)

Sedgewick R (2003) Algorithms in C++, Part 5: graph algorithms, 3nd edn. Pearson Education Asia limited/Tsinghua University Press, Edition Science Press, Beijing

Wang SH (2009) Graph theory, 2nd edn. Edition Science Press, Beijing, p 150 (in Chinese)

Wang EF, Shi SM (2003) Higher algebra, 3rd edn. Higher Education Press, Beijing, p 189 (in Chinese)

Xu GL, Ma XL (2010) An improved mathematical model for real-time control of local flows. In: ICEEE2010, the international conference on E-product E-service and E-entertainment, Proceedings volume 8, IEEE Press, Guanglian, Xiaolu, pp 5057–5060

Xu GL, Shang YL (2010) Network system exception deal with the mathematical model. Sci Technol Eng 5(14):123

A Novel Parallel Hardware Methodology for Solving Linear System of Equations

Bowei Zhang, Guochang Gu, Lin Sun, and Xingzhou Zhang

Abstract In this paper, we proposed a parallel hardware methodology employing the modified Gaussian elimination algorithm to efficiently solve linear system of equations (LSEs). Two parallel operators are issued in the hardware-optimized algorithm. Moreover, to be the proof-of-concept, the proposed parallel methodology is implemented to hardware structures in cases to address solving LSEs over GF(2) (primarily are bits operation) and LSEs with floating-point (IEEE-754 standard, 32-bit single precision) coefficient matrix. The corresponding hardware is mainly composed of uniformly distributed basic cells which store and register data, yielding a standalone worst case time complexity $O(n^2)$ opposed to $O(n^3)$ of the software replication. Finally, the given experimental result inosculated with the theory analysis.

Keywords Parallel methodology • Linear system of equations • Hardware-optimized algorithm

1 Introduction

Solving linear systems of equations (LSEs) is a very common computational problem appearing in numerous research disciplines. From a complexity theoretical point of view, the solution of the LSEs is efficiently computable; using the

B. Zhang (✉) • G. Gu • L. Sun
College of Computer Science and Technology, Harbin Engineering University,
150001 Harbin, China
e-mail: zhangbowei1983@gmail.com; guguochang@hrbeu.edu.cn; sunlin@hrbeu.edu.cn

X. Zhang
College of Information and Communication Engineering, Harbin Engineering University,
150001 Harbin, China
e-mail: zhangxingzhou@hrbeu.edu.cn

algorithms (like Cramer's rule, LU decomposition) any LSEs can be solved in at most cubic time (Cormen et al. 2001; Galoppo et al. 2005; Toshio and Taura 2005). The standard algorithm for solving a system of linear equations is based on Gaussian elimination with some modifications. However, for some areas current algorithms and their software implementations are insufficient or unsatisfying in practice. This is often due to the large dimension or number of LSEs that must be solved in order to accomplish a special task.

In this paper, in order to develop such hardware architecture we analyzed the suitability of various algorithms while taking possible parallelism over solving LSEs efficiently. A parallel methodology is introduced based on the modified Gaussian elimination, which turned out to be basically best suited for hardware implementation due to its simplicity and weak data dependency between each elimination element. Two parallel operators are issued in the hardware-optimized algorithm. Moreover, to be the proof-of-concept, the proposed parallel methodology is implemented as hardware structure in the cases for solving LSEs over GF(2) (primarily are bits operation) and LSEs with floating-point (IEEE-754 standard, 32-bit single precision) coefficient matrix. To the best of our knowledge the alternative of an efficient implementation in hardware has not been analyzed thoroughly yet.

The rest of the paper is organized as follows: in section "Proposed Methodology", the detail methodology is introduced with given hardware-optimized algorithm. In sections "Case Study I: Solving LSEs over GF(2)" and "Case Study II: Solving LSEs over Single Precision Floating-Point", real world cases equipped with the proposed architecture have been studied, addressing solving LSEs over GF(2) and LSEs with floating-point. The experimental results of performance are given to prove the theory analysis.

2 Proposed Methodology

The LSEs is a collection of linear equations involving the same set of variables. For example, a general system of n linear equations with m unknowns can be written as:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad\qquad \vdots \quad\ \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_n
\end{aligned}
\tag{1}
$$

The vector equation is equivalent to a matrix equation of the form "$A \cdot x = b$", while A is an $m \times n$ matrix, x is a column vector with n entries, and b is a column vector with m entries.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \; x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \; b = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix} \tag{2}$$

A general solving can be described as: firstly transform the "$A \cdot x = b$" to "$U \cdot x = b'$"; the U is an upper triangular matrix as the result of applying elementary row operations; and it can trivially be solved by backward substitution. In implementation, the column b is always added to the right hand side of matrix A to form a new matrix A' which is treated as the *adjoint matrix*. In the rest part of the paper, the coefficient matrix A' is taken as the main operated object.

$$A' = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{bmatrix} \rightarrow \begin{bmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \dots & a'_{2n} & b'_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a'_{mn} & b'_n \end{bmatrix} \tag{3}$$

The pseudo-code of Gaussian elimination is given in Algorithm 1 (Trefethen and Schreiber 1990). In Algorithm 1, each step of the outermost loop consists of pivoting phase, specifically row exchange phase and update phase. Firstly, choose a row below the diagonal with a max-value element. Selecting a large element as a pivot is important for numerical accuracy. Then, execute column elimination. The pivot element is always moved to the diagonal. Lastly, obtain the result identity matrix after the outer while-loop has been executed for every column of A.

Algorithm 1	Algorithm 2
for$k := 1$ **to** n **do**	**for** each column k $:= 1$ **to** n **do**
finds pivot element A[max, k]	**while**$A[1, 1] = 0$ **do**
exchange the max-th row & the kth row	$A := \boldsymbol{shiftup}(n - k + 1, A)$
for $i := k$ **to** n **do**	$A := \boldsymbol{eliminate}(A)$
$ratio := A[i,k] / A[k,k]$	
for $j := 1$ **to** n **do**	
$A[i,j] := A[i,j] - ratio \times A[k,j]$	

A re-written algorithm is introduced in Algorithm 2. In this algorithm, we only take care of the first column/row to find the pivot element. Therefore, $A[1,1]$ is the only pivot element according to what the whole matrix operation is performed to the next step.

In this paper, we introduce the parallel operators "*shiftup*" and "*elimination*" in our methodology which is hardware-optimized. In order to obtain a pivot element, we define a cyclic shift-up operator of all rows not yet used for elimination (due to

the shifting approach these are the first $n - k + 1$ rows) until the value of the fixed pivot position presents non-zero value (in fact, the probability of a zero element in dense matrices is so tiny that it is not a critical step). A "*shiftup*" is defined as:

$$\left(i, \ (row_1, \ldots, row_\mathrm{n})^T\right) \mapsto (row_2, \ldots, row_i, \ row_1, \ row_{i+1}, \ldots row_n)^T \quad (4)$$

Then, the elimination operation can be described as proposed mapping process combining the cyclic "*shiftup*" and the cyclic "*shiftleft*". After a normal elimination, we put the rows already used for elimination down to the bottom of the matrix. An "*eliminate*" is defined as:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}, \ r_1 = \frac{a_{21}}{a_{11}}, \ldots r_{n-1} = \frac{a_{n1}}{a_{11}}$$

$$\downarrow eliminate(A) \quad (5)$$

$$A = \begin{bmatrix} a_{22} - r_1 a_{12} & a_{23} - r_1 a_{13} & \ldots & a_{2n} - r_1 a_{1n} & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n2} - r_{n-1} a_{12} & a_{n3} - r_{n-1} a_{13} & \ldots & a_{nn} - r_{n-1} a_{1n} & 0 \\ a_{12} & a_{13} & \ldots & a_{1n} & a_{11} \end{bmatrix}$$

3 Case Study I: Solving LSEs Over GF(2)

This is the case in the cryptanalysis of asymmetric encryption schemes like RSA or ElGamal, where very large but sparse LSEs over GF(2) or GF(*p*) (*p* prime) need to be solved to obtain a secret key. Since these LSEs must be solved accurately, approximate solutions using fast iterative algorithms are simply useless.

3.1 Parallel Solution

For linear systems of equations over GF(2), solving becomes quite simple: The required row operations consist of the conditional XOR of two rows and the swapping of two rows (Parkinson 1984). Given a binary regular matrix A, Algorithm 2 can also perform Gaussian elimination and backward substitution in parallel, i.e., when the algorithm terminates the result matrix is not just an upper triangular matrix U but the identity matrix I_n, so the corresponding LSEs is already solved.

In order to illustrate Algorithm 2, Fig. 1 depicts an example application to the LSEs:
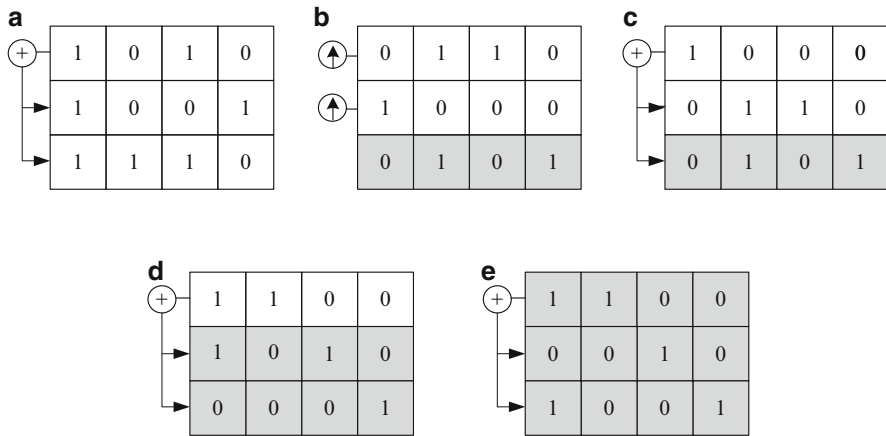
Fig. 1 Example: matrix solution with binary case. (**a**) eliminate, (**b**) shiftup, (**c**) eliminate, (**d**) eliminate, (**e**) finish

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{6}$$

For the sake of concreteness we explicitly treat the right hand side as an additional passive column to A and view the changes on this extended matrix resulting from each execution step. Rows already used for elimination are visualized by shading the corresponding elements. For convenience, a small arrow indicates the position of the passive column.

When the algorithm terminates, the coefficient matrix has been transformed into the identity matrix I_3. The passive column is the leftmost column and contains the solution:

$$x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \tag{7}$$

3.2 Mapped Architecture

In order to implement Algorithm 2 in hardware, we use a mesh structure of register cells: the whole device consists of $m \times n$ cells, which are able to execute the two basic operations defined in last section in a single clock cycle. Below, we will describe the basic constituents of the design and its interrelation.

The design at hand comprises a parallel implementation of both operations where the pivot element is used as multiplexing signal for the appropriate result. The element cells that compose the mesh structure are called basic cells (BCs). Each
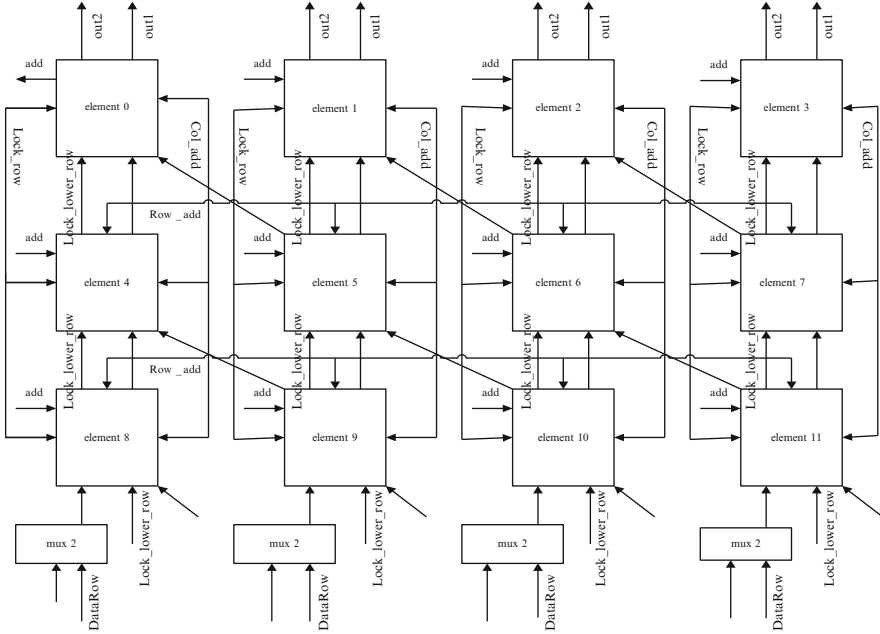
Fig. 2 Mesh network of the basic cells

of the BC stores a single bit and has local connection to four direct neighbors; it also connects to the global network like global clock (*clk*) and reset (*rst*) network. All cells are aligned in a rectangular array as shown in Fig. 2. The up-most left cell always contains the pivot element.

shiftup: in case of the actual pivot element being '0', the architecture performs a simple *shiftup* operation. All elements in the matrix with the used-flag set to '0' (lock row = 0) simply shift their value one row up. Values in the up-most row will be shifted to the lowest unused row, resulting in a cyclic shift of all unused rows.

eliminate: elimination is performed when the pivot element equals '1' (add = 1 = a_{11}). Every element except for those in the first row will compute an XOR with the up-most column entry (col-add = a_{1j}) if the first entry in the row is '1' (row add = 1 = a_{i1}). In the same step, the result is shifted to the element on the left in the upper row.

End of computation: when all rows have actively been used for pivoting, the used-flag reaches the top-most row and the architecture stops.

3.3 Performance Analysis and Experimental Result

Let's estimate the worst and average case runtime of Algorithm 1. The worst case runtime overall needed for pivoting is quadratic. Furthermore, *n* column
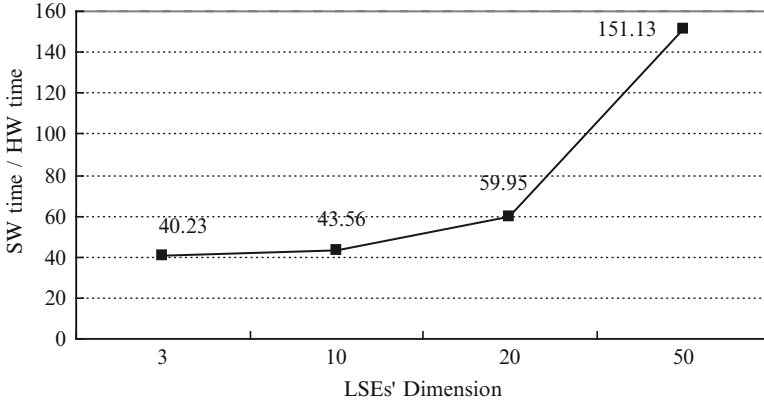
Fig. 3 Results comparison between hardware and software of solving LSEs for binary case

eliminations are necessary, each consisting of n row operations. A row operation during the k-th column elimination consists of $n - k + 1$ XOR operations. Hence, in total about n^3 XOR operations are required yielding a worst case runtime of $O(n^3)$.

In Algorithm 2, we analyze the running time of Algorithm 2. Since we assume that an application of *shiftup* and *eliminate* can be computed in a single clock cycle, we simply have to count the required total number of these primitive operations. In the best case we have no application of shiftup at all and in the worst we have exactly $n-k$ applications of *shiftup* during the k-th iteration (assuming A is uniquely solvable). Hence, we obtain the time complexity $R(n)$ of Algorithm 2 is:

$$n \leq R(n) \leq \frac{n^2 + n}{2} = O(n^2) \tag{8}$$

Hence, in the analysis we gain at least n times speedup over the software original algorithm. The accuracy of our estimations has been approved by a benchmark of an emulation of the proposed hardware architecture running at frequency of 300 MHz against the software running on a desktop (2.2 GHz, 1 GB memory), where the number of required clock cycles to solve matrices of different dimensions and densities has been counted. The results of this benchmark are depicted in Fig. 3.

4 Case Study II: Solving LSEs Over Single Precision Floating-Point

Hardware floating-point programming can hardly speed up in general evaluation since either the single, the double or the long double precision format of floating-point numbers is much time-consuming and area-demanding while the process is always suffering from long latency of floating-point computing (Abdul Gaar et al. 2002).

a

÷	1.00	-2.50	4.00	6.00
→	2.00	-5.00	1.50	4.50
→	1.00	4.00	-1.00	2.00

b

↑	0.00	-6.50	-7.50	0.00
↑	6.50	-5.00	-4.00	0.00
	-2.50	4.00	6.00	1.00

c

÷	6.50	-5.00	-4.00	0.00
→	0.00	-6.50	-7.50	0.00
→	-2.50	4.00	6.00	1.00

d

÷	-6.50	-7.50	0.00	0.00
→	2.08	4.46	1.00	0.00
→	-5.00	-4.00	0.00	6.50

e

÷	2.07	1.00	0.00	0.00
→	1.77	0.00	6.50	0.00
→	-7.50	0.00	0.00	-6.50
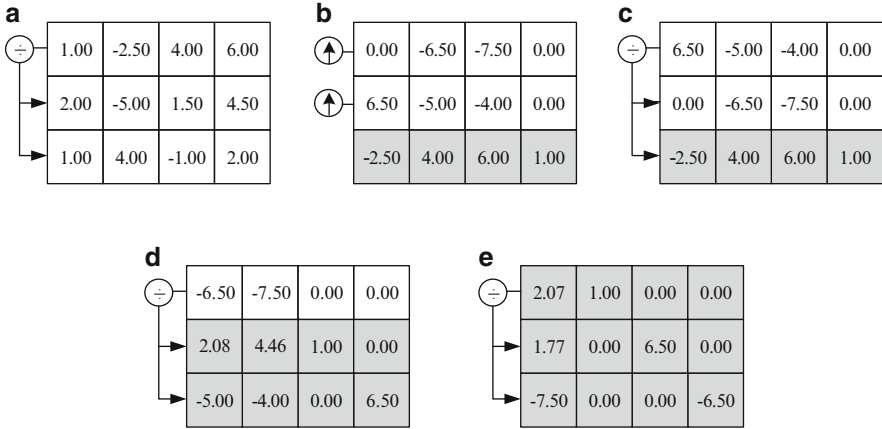
Fig. 4 Example matrix solution with floating-point. (**a**) eliminate, (**b**) shiftup, (**c**) eliminate, (**d**) eliminate, (**e**) finish

4.1 Parallel Solution

Similar to solving linear systems of equations over GF(2), the parallel methodology is suitable to solve the LSEs with complicated data type. However, unlike the bit-operations which mostly can be handled in a single clock, fine-grained controlling is required to accurately schedule the whole structure's behavior due to the arithmetic operators' different processing latency. The Fig. 4 depicts an example to solving the floating-point LSEs with our methodology:

$$\begin{pmatrix} 1.00 & -2.50 & 4.00 \\ 2.00 & -5.00 & 1.50 \\ 1.00 & 4.00 & -1.00 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6.00 \\ 4.50 \\ 2.00 \end{pmatrix} \tag{9}$$

The matrix for elimination is the *adjoint matrix A'*, when we explicitly put the right hand side as an additional passive column to original $n \times n$ *matrix A*. Rows already used for elimination are visualized by shading the corresponding elements in Fig. 4.

When eliminations terminate, the coefficient matrix has been transformed into the diagonal matrix. And the b column is the leftmost column. Thus, we obtain the result:

$$x = \begin{bmatrix} 2.07 & /1.00 \\ 1.77 & /6.50 \\ -7.50 & /-6.50 \end{bmatrix} = \begin{bmatrix} 2.07 \\ 0.27 \\ 1.15 \end{bmatrix} \tag{10}$$
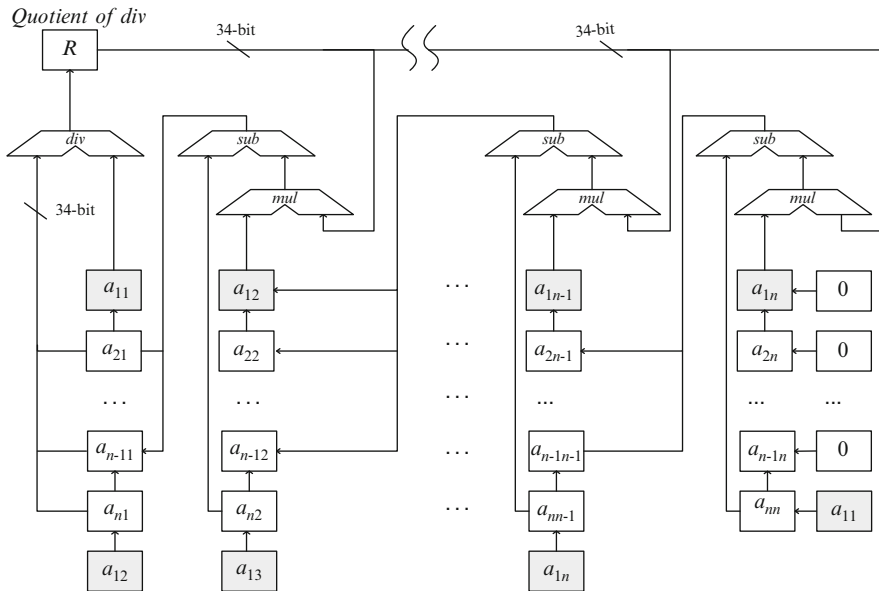
Fig. 5 Global architecture with controller and arithmetic operators

4.2 Integrated FPlibrary Operators

In the case, single precision floating-point (IEEE754 standard) data type is considered. Regarding of the complicated arithmetic operation of floating-point, an open source library namely FPLibrary Detrey and de Dinechin (2006, 2009) written in VHDL is used to synthesizable pipeline operators (FP operator with 34-bit width, delivering one output per cycle after the pipeline building time).

4.3 Mapped Architecture

An inter-link of all parts are aligned in a rectangular array as shown in Fig. 5. A single basic cell which stores 34-bit FP data has four local connections. Two of the four connections-(*Out*1, *In*1) are for local connections to two direct neighbors: above (*Out*1), below (*In*1); other two are conjunct with MUXs for serial queue transformations, respectively. The special column (the left-most column) consists of a DIV floating-point operator preserving a fix-position divisor of a_{11} and an alternative dividend switching input.

One of the MUL operator's inputs connects to 34-bit block register 'R' providing serial quotient while the other one always gets value from the first row. From the second normal column to the $(n-1)$th, they share source BCs as the destination BCs of their right hand neighbor normal columns in order to update the whole matrix with one time *eliminate.*

4.4   Performance Analysis and Experimental Result

Firstly, we analyze the worst case runtime of the Algorithm 1 with floating-point case. A row operation during the k-th column elimination consist itself of $n-k+1$ latency floating-point arithmetic MUL and SUB operation. Using c_i for the cost of line j. we sum the products of the cost and times columns:

$$R(n) = c_{10} \sum_{k=1}^{n} \sum_{i=k}^{n} \sum_{j=1}^{n} + c_8 \sum_{k=1}^{n} \sum_{i=k}^{n} + (c_3 + c_5)n \tag{11}$$

Thus, we obtain:

$$R(n) = \frac{c_{10}}{2}n^3 + \left(\frac{c_{10}}{2} + \frac{c_8}{2}\right)n^2 + \left(c_3 + c_5 + \frac{c_8}{2}\right)n \tag{12}$$

It is a cubic function of n, and depends on the statement costs c_{10}. According to the structure principle, considering limited quantity absent overhead an approximate number of spending cycles of elimination procedure can be calculated. We obtain running cycles (Δ_e) of a single elimination as:

$$\Delta_e \approx DivLatency + MulLatency + SubLatency + n \tag{13}$$

A single *eliminate* can be performed in ($DivLatency + MulLa\text{-}tency + SubLatency + n$) cycles. Therefore, taking c_{opt} indicate the *OptLatency* we can obtain as below. An $O(n^2)$ average case against $O(n^3)$ of the original algorithm is gained.

$$R(n) = n^2 + (c_{div} + c_{mul} + c_{sub})n + P_r \cdot c_2 n \tag{14}$$

Figure 6 depicts average running time curve between software (desktop, 2.2 GHz, 1 GB memory) and our hardware architecture (125 MHz), loading the benchmark with matrices in varies of dimensions. The result indicates that hardware leads an approximate linear increasing ratio outperforming over the traditional software implementation, inosculating with the theory analysis.
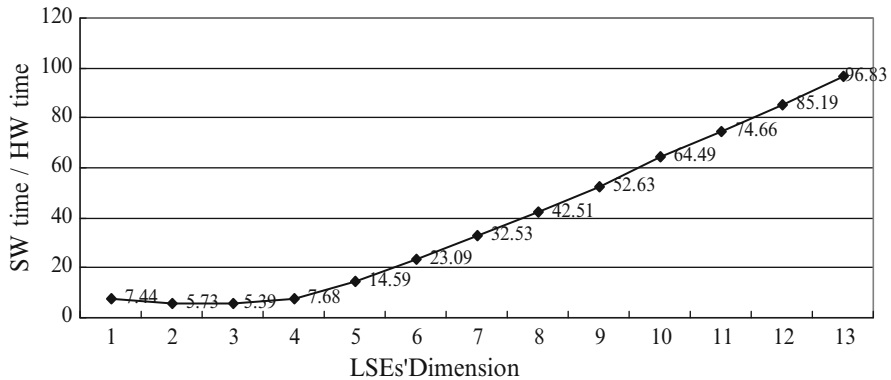
Fig. 6 Results comparison between hardware and software of solving LSEs for 32-bit floating-point case

References

Abdul Gaar A, Luk W, Cheung PY, Shirazi N, Hwang J (2002) Automating customization of floating-point designs. In: Proceedings of the 12th international workshop on field programmable logic and application (FPL 2002), LNCS 2438, Montpellier, France, 2–4 September 2002, pp 523–533

Cormen T, Leiserson C, Rivest R, Stein C (2001) Introduction to algorithms, 2nd edn. MIT Press, Cambridge

Detrey J, de Dinechin F (2006) A VHDL library of parametrisable floating-point and LNS operators for FPGA, http://www.ens-lyon.fr/LIP/Arenaire/Ware/FPLibrary/, LIP-ENS Lyon

Detrey J, de Dinechin F (2009) P.07. FPLibrary v0.91 user documentation, LIP-ENS Lyon

Endo T, Taura K (2005) Highly latency tolerant Gaussian elimination. Grid computing workshop, 13–14 November 2005, Seattle, WA, pp 91–98

Galoppo N, Govindaraju N, Henson M, Manocha D (2005) LU-GPU: Efficient algorithms for solving dense linear systems on graphics hardware. In: ACM/IEEE SC-05 conference, Seattle, WA, 12–18 November 2005

Parkinson D (1984) A compact algorithm for Gaussian elimination over GF(2) implemented on highly parallel computers. Parallel Comput 1:65–73

Trefethen LN, Schreiber RS (1990) Average case stability of Gaussian elimination. SIAM J Matrix Anal Appl 11(3):335–360

The Study of the Application of Virtual Reality Technology on the Campus Network Management

Xiao Wei and Zhao Yun

Abstract With the expansion of network, the increase of network elements and the growing complexity of the relations between these elements, the management and visualization of a great deal of complex information quickly and easily has become the bottleneck of two-dimensional user interface. The use of virtual reality technology can enhance the visualization of network information and increase the efficiency of network management comprehensively. This thesis, based on the study of the campus network management, firstly offers the solution to the crucial problems of the construction of the VR-based computer network management system, such as scene modeling, scene optimization, reality feedback and so on; and then, in the experimental test, it makes models of three virtual buildings and network equipments, and realizes simple scene walkthrough and network management function by the use of programming technology, such as VRML, X3D and so on; finally, this thesis summarizes the work done at this stage, points out the weakness, and makes clear the direction for future study.

Keywords Component • Virtual reality technology • Network management • Virtual scene • Reality feedback

1 Introduction

Nowadays, with the popularity of Internet, the network has been greatly expanded and the internet users have increased very quickly, which has caused the steady increase of amounts and types of network hardware devices and the complexity of interconnecting links so as to raise the difficulty of the network management.

X. Wei (✉) • Z. Yun
Information Technology Center, Chongqing Normal University, Chongqing, China
e-mail: ghost.xiao@yahoo.com.cn; zhaoyun@cqnu.edu.cn

However, network management software used popularly at present employs static two-dimensional pages which have greatly limited multi-dimensional display, visualization and visuality of information (Zhao Dan-yang unpublished). So, network faults tend to appear more frequently and diversely, especially in places, like campus, where network users crowd. Campus network covers a large area, and many schools have more than one campus which scatters in different places, so it is very hard for the network administrators to monitor the whole network. If Virtual Reality Technology can be applied to Campus Network Management and three-dimensional virtual campus scene is used to replace the two-dimensional pages used at present, the visualization and visuality of network management system would be greatly enhanced. Thus, the operation system can be more consistent with the norms of human behaviors and thoughts in the real world, and it would be more convenient for the network administrators to identify network faults more accurately and efficiently.

2   Research on Implementation Method

This part focuses on key problems in constructing the VR-based computer network management system, including scene modeling, scene interactive, reality feedback and so on.

2.1   Construction of Three-Dimensional Modeling

- Modeling of the terrain and its feature
  Input contour layer of the CAD topographic map of the school area collected to Sketchup software directly, and create the terrain of the new campus by use of tools of Sandbox plug-in unit for generating terrain with contour, and then, put the ground grains onto the prepared terrain model by projection; models of objects on the ground can be zoomed in proportion to the size and shape of the real objects, and intake the appearance of the objects with a digital camera and clip the photo material, and put the object models onto the prepared terrain model in Sketchup accurately; it can also be modeled by way of push and pull on the basis of the contour line of the layout (Li Ming et al. 2008; Chen Alin et al. 2007).
- Modeling of Electro-weak distribute network
  As campus network concerns only the construction of Electro-weak circuit, this thesis only discusses the modeling of Electro-weak distribute net of underground pipeline net. Firstly, remove the needless part of campus CAD map and only left Electro-weak distribute net, and save it as a .Dxf files; then, import this file into Sketchup software and adjust it to match the prepared terrain model according to the coordinate; lastly, integrate the plane model of Electro-weak distribute net with the prepared terrain model by projection. In the integrated model, modify
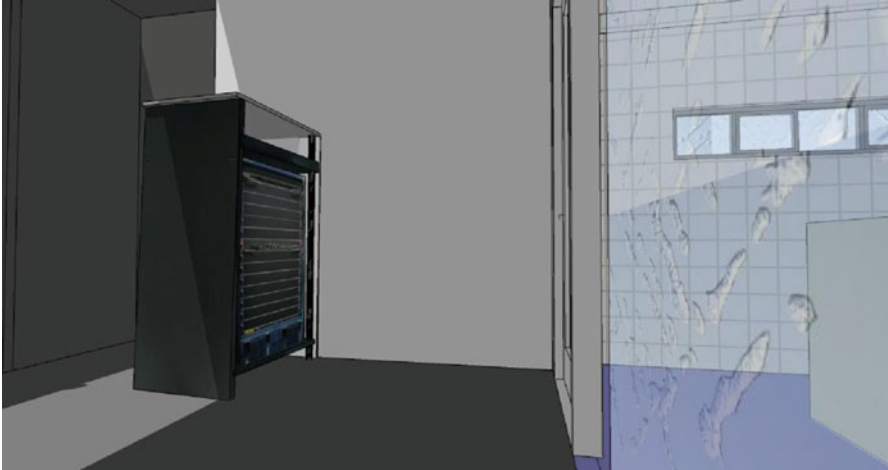
Fig. 1 The drawing of integration of campus network equipments and buildings

the inaccurate part of the position of Electro-weak circuit, and create the three-dimensional net model on the surface of Electro-weak circuit by way of cylinder modeling. As the CAD map has only main Electro-weak circuit, it is necessary to create the three-dimensional model of optical fiber circuits connecting buildings and connecting buildings and the trunk-line net on the basis of the topology of the real campus network.

- Modeling and integration of campus network equipments
  Construct models for the switches, machine cabinets, computer-room air-conditioning, UPS and so on according to their actual sizes by use of 3DMAX software, and map all sides of the models with the real digital photo materials of these equipments to create more realistic three-dimensional models, and export the models of various network equipments as .3DS files, and import these 3DS models into Sketchup software. Modify the position of the models of network equipments according to the coordinate of buildings, locate the position of the switches in the machine cupboard, group together the machine cupboard and the switches, and finally integrate these models; locate the position of the group model and put it in the building, and then, integrate all of the models. Thus, the integration of network equipments is completed. As shown the "Fig. 1".

2.2 The Realization of Scene Interactive

As for the realization of scene interactive, especially virtual scene interactive based on WEB, VRML/X3D programming technology is most popularly used at present. The Node technology of VRML/X3D supports many functions, such as collision detection, view navigation, and the event-driven function and so on.

- The realization of collision detection
  VRML/X3D provides a collision node used exclusively for collision detection (Zhang Xuan 2008). It can create a space, and when the observer's viewpoint collides with objects in it, a given event would be produced. Collision node can detect whether a collision has happened between observer viewpoint and objects, but it could not detect the collision between objects.
- The method of view navigation
  Viewpoint node designates the observing position and direction of the viewpoint of the user in the scene, while NavigationInfo node sets the way the scene is viewed and the physical features of stand-ins. Stand-ins refer to a character appearing in the scene when we choose " show alias" in shortcut menu of the scene. Different VRML browser plug-ins would provide different shapes for stand-ins (Zhang Xuan 2008; Ye Hai-zhi and Gui Wei-feng 2007).
- The realization of event-driven function
  In VRML/X3D, the event-driven function is realized mainly by use of Script node. There are two kinds of Script descriptive languages supported by Script node: Java and Javascript or ECMAscript (it is the standardized name of Javascript) (Huang Zheng-jun and Zhou Jian-zhong 2007; Pang Liao-jun et al. 2007). In Script Node, Script programming can connect to the external program files through url and be quoted, and it can also be packaged within the node by way of <![CDATA[]]>. Combined with the mechanism of Java world and VRML/X3D world, Script node can obtain fields, manipulate their values, and return the result to the scene. The network management of this system and network communication function can be realized only in the way this node invokes the script.

2.3   Research on Reality Feedback

Presently, many virtual reality systems mainly concern the simulation of the real scene, and lack the mapping of the feedback information of the real world, including the state of objects in the scene, the behavior mapping and so on. This thesis will conduct research on reality information feedback and state mapping of virtual network management system by use of the mathematical model of Finite State Machine, and point out the method of realization theoretically in order to instruct the future experimental test of this system. Take the following figure of the working state of network equipments as an example.

As can be seen from the "Fig. 2", when the device receives the order of "shutdown", it turns into the state of "power-off'", and when the device receives the order of "start", it turns into the state of "power-on". When the device is powered on, network management system will start SNMP Agent to collect the device information. If the state of link port, the state of port forwarding, CPU usage and memory usage are normal, it will turn into "working normally"; if any of the states appears abnormally, it will be turn into "working abnormally". The port link
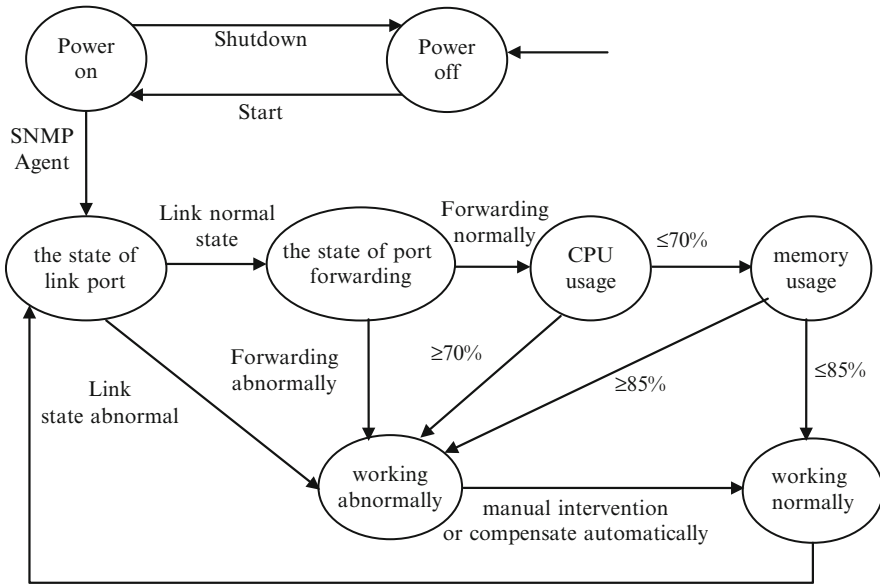
Fig. 2 The following chart of the working state of network equipments

state means whether the hardware parameters of the port are normal or damaged; the port forwarding state means whether the sending and receiving of data packages is normal, that is, whether the parameters of port traffic, throughput, response time and the like are beyond threshold. Usually, network equipments have more than one port, when it is powered on, it will enter state checking, but only one port has been taken as an example in the chart, for the description of the state of every port is the same. In addition, if the abnormal state disappears or the network administrator makes manual intervention, the device can compensate automatically and turn into "working normally".

3 Experimental Test

The hardware of the experimental test of the writer is ordinary desktop PC, so it is impossible to carry out the complicate virtual scene. However, the school scene of virtual network management system designed in this thesis includes the terrain and its features, models of various objects on the ground, the model of underground pipe net, models of the external and internal structures of the buildings, network device models, various materials and texture mappings, but under the hardware and network condition of the writer, running the browser would bring about a crash immediately, so this experimental test has only used three buildings, the fibers connecting the buildings and the model of the switch going with them to construct a simple virtual scene. As shown the "Fig. 3".
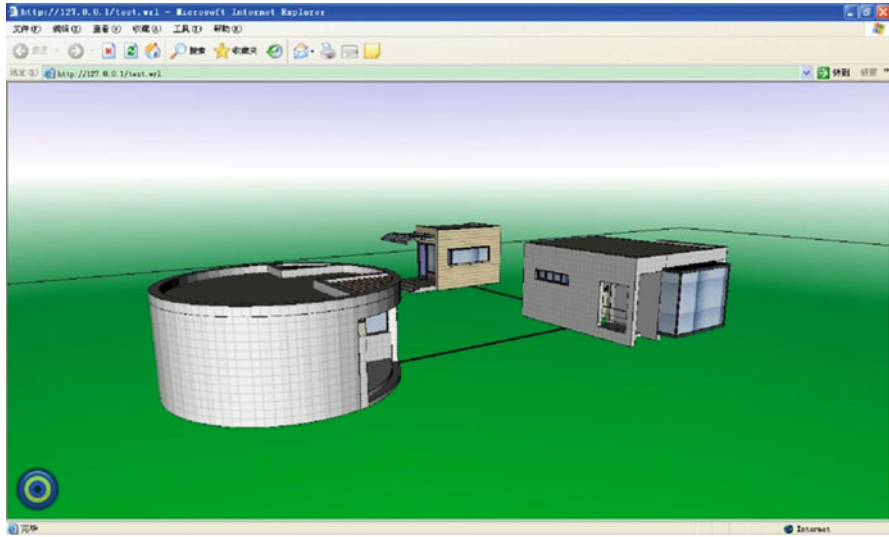
Fig. 3 Virtual scene of the experimental test
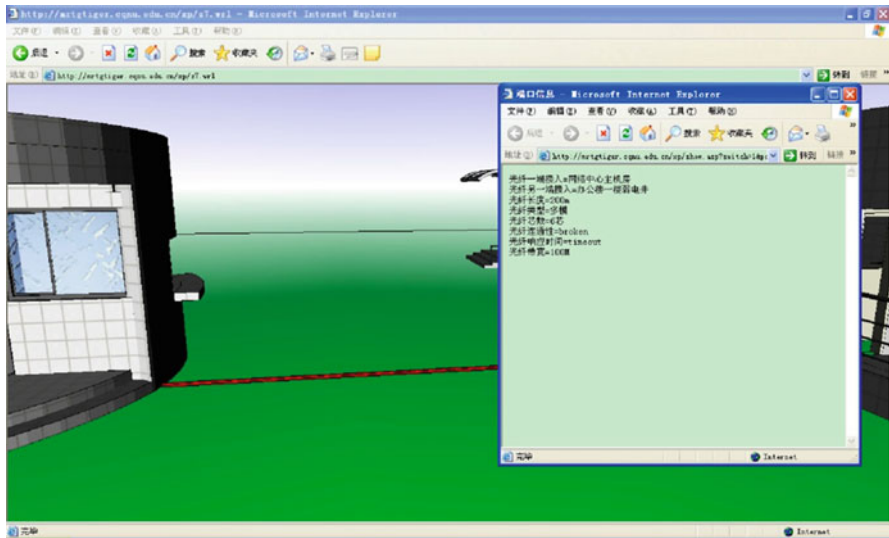
Fig. 4 View of fiber link attribute

There are three SQL Server datasheets in the test: Switches attribute table, Switch port attribute table and fiber attribute table, which are used to store all kinds of network data during the system test. Then, call up data by Java or other technologies. Finally, link and display them onto the Web page by way of Script Node of VRML/X3D. As shown the "Fig. 4".

4 Conclusion and Prospect

The wide application of network management system based on Web and the growing complexity of network have made it necessary and urgent to create visualization of network management, and the application of virtual reality technology provides us a very good resolution. This thesis has put forward a design of network management system based on virtual scene and a model of visualization of virtual network management system, made a research on various scene models, scene interactive and the method of reality feedback, and carried out an experimental test. Nevertheless, there are still much to be done in optimization of the big scene, the speed of scene roaming and the realization of network management function. The future study should care more about the interactivity and operability of models of network equipments, and simulate the real work of network administrators.

References

Chen Alin, Hu Zhao-hui, Qi Xiang-zhi (2007) The study of building virtual reality campus with three dimensional scene and modeling technology. J Chongqing Norm Univ (Nat Sci Ed) 24(4):37–40, Chongqing

Huang Zheng-jun, Zhou Jian-zhong (2007) X3D-based research on the dynamic interactive technology for virtual scenes. Comput Eng Sci 29(7):55–57, Hunan

Li Ming, Liu Gui-sheng, Li Nan (2008) The research of urban 3D visualization fast modeling and browsing system. Mod Surv Mapp 31(3):13–15, Jiangsu

Pang Liao-jun, Li Hui-xian, Wang Li, Wang Yu-min (2007) Implementation of VRML's network communication function based on java. Comput Eng 33(11):47–49, Shanghai

Ye Hai-zhi, Gui Wei-feng (2007) The space environment modeling based on the VRML. J Henan Norm Univ (Nat Sci) 35(4):177–179, Henan

Zhang Xuan (2008) Brief discussion of hypothesized scene interactive question. J Xingtai Polytech Coll 25(3):32–35, Hebei

Zhao Dan-yang. Research and realization on visualization of network management system based on VRML-Java. (Unpublished)

Learning Behavior Tracing in Online Course Based on SCORM Standard

Hu Zhengquan

Abstract In the process of E-learning, in order to grasp their learning dynamics and given real-time evaluation and feedback to learners, we need to monitor their learning behaviors. SCORM standard defines a powerful learning tracing mechanism which can trace the behaviors of learners. This paper analyzes the principles and Methods of tracing learning behaviors based on SCORM standard.

Keywords SCORM standard • Online course • Tracing • Learning behavior

1 Introduction

In distance education environment, teachers and students are separated in time and space, students learning mainly through online course. It's difficult to grasp student's learning dynamic and preference for teachers; you can not make appropriate evaluation and feedback in time, also difficult to implement individualized instruction. However, tracing and monitoring learner's learning behavior through certain technical specifications, we can record their learning progress, learning state, learning time, learning achievement, learning interaction, learning preferences and other conditions.

SCORM (Sharable Content Object Reference Model) developed by the U.S. Department of Defense's Advanced Distributed Learning (ADL) (ADL.SCORM version1.2-The SCORM content 2004), is one of the most widely used online course specifications in currently. Online course meet the SCORM specification are accessibility, adaptability, interoperability and reusability, LMS (learning management

H. Zhengquan (✉)
College of Electronics and Information Engineering, ChongQing Three
Gorges University, ChongQing 404100, China
e-mail: hzq250314@163.com

system) will be very detailed tracing and recording a variety of learning behaviors when learning online course (Yan ren-quan and Wu li-ping 2009).

This paper explained how to tracing learning behavior based on SCORM specification from the network platform and online course.

2 Tracing Theory Based on SCORM Standard

In the SCORM standard, SCO (Sharable Content Object) is the lowest level of granularity of learning resources that can be traced by LMS through the SCORM RTE (Runtime Environment) (ADL.SCORM version1.2-The SCORM runtime environment 2004), and their communication is complete by the API Adapter.

API defined eight major functions, namely LMSInitilize, LMSFinish, LMS-SetValue, LMSGetValue, LMSCommit, LMSGetLastError, LMSGetErrorString, LMSGetDiagnostic (Sun Xing-hua et al. 2010), which LMSGetValue and LMS-SetValue are the most complex, but is the most frequently used in trace, Data Model used in the communication process to record tracing information.

Learning behavior tracing is achieved through the interaction among LMS and online course, the process as follows (Fig. 1):
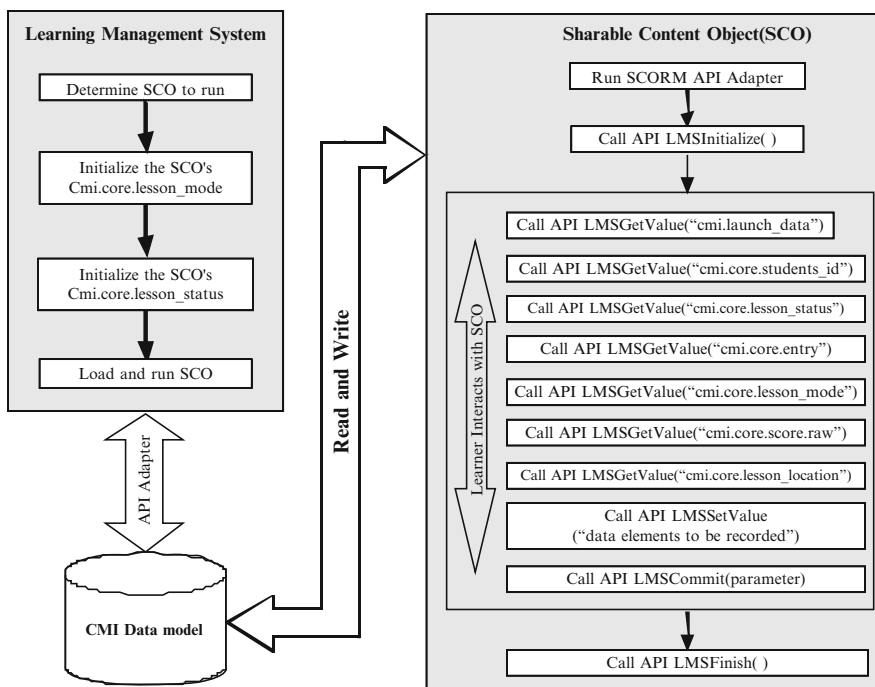
Fig. 1 Tracing theory based on SCORM

2.1 In the LMS Server

(1) First, LMS decides which SCO will be run, while loading Data Model Library; (2) Setting the operating mode of SCO, involving the preview mode, normal mode and review mode, which preview mode just run courseware, rather than tracing learning behaviors; (3)Setting learner's learning status, including three statuses: "initial" (first study), "resume" (re-learning) and "finish" (completed); (4) Loading and running SCO; (5) Completing the read and write data model by using API Adapter to achieve interaction among LMS and online course.

2.2 In the LMS Client

(1) Running SCORM API Adapter. (2) Calling the API initialization function. (3) Loading SCO initial data programs. (4) Obtaining user ID and user name in the Data Model. (5) Obtaining the value of cmi.core.lesson_status in Data Model, which is the current user's current learning status of SCO, including Completed, Incomplete, Passed, Failed and other seven statuses. (6) Obtaining the value of cmi.core.entry in data model, which is determined whether the current user learning the current SCO for the first time. If the value "initial", it means for the first time learning; if the value "resume", it means it's re-learning the current SCO. (7) Obtaining the value of cmi.core.lesson_mode in data model, which is determined the way current user access to the current SCO, including three modes: Browse (preview, does not tracing learning behaviors), Normal (tracing learning behaviors) and Review; (8) Obtaining the value of cmi.core.score.raw in data model, which is used to store each student's total scores accumulated in every learning of SCO. The total scores of current user will be recorded if the current SCO are evaluated, otherwise not record. LMS could analyze the performances of learners by the Data Model. System will initialize the value of cmi.core.score.raw be " " (empty value) if the user is learning the current SCO for the first time. (9) Obtaining the value of cmi.core.lesson_location in data model. By the Data Model, LMS could recorded which SCO each user learned finally, function like a bookmark, you can jump to the last position to continue learning when re-enter the program the next time. (10) Users began the learning of current SCO, it will record a series of data model value in the learning process for tracing the learning behaviors. cmi.core.session_time is used to record each student each learning time, including the residence time and access date of enter each SCO. When the current user finished the current SCO learning, LMS would recorded the value of cmi.core.total_time, which it means the current user spent total time for each SCO learning. (11) Executing LMSCommit (Parameter) interface function. This function is mainly used to submit the value of data model elements, LMSCommit start up before LMSfinish ( ). (12) Executing LMSFinish ( ) function. It will call LMSFinish ( ) when user after learning and leaving a SCO. LMSFinish ( ) is mainly responsible for reset the environment variables, and determined whether the SCO

has called LMSCommit ( ) to store all the records in the LMS before the end. If not stored, it will automatically calling LMSCommit ( ) to save the learning record of current user in the current study (Operating mechanism of LMS 2010).

3 Tracing Mechanisms of Learning Management System Based on SCORM Standard

If want to tracing the learning behaviors of online course, LMS RTE (Runtime Environment) should include two parts: the server and client (Chang Jia-yu 2009).

The server includes learning database management and communication data protect; the client need to provide a API Adapter for learners, which is used to implement the function of API package.

By browser the SCO running on the client to obtain the API Adapter instance, API provides a standard way for communication between SCO and LMS, while its hide the implementation details of communication for SCO, thus promoting the reusability and interoperability of SCO. Implement of tracing learning behaviors by reading and writing the Data Model in the Interaction between SCO and LMS (Sun Xing-hua et al. 2010).

In the process of tracing learners learning behaviors, the design of API Adapter is the key technology in the system. Functions of API Adapter include three categories:

1. Operating status functions: LMSInitialize( ): is responsible for initiating SCO, when the learners begin to learn a SCO, it need to call LMSInitialize( ) first. LMSFinish( ): when learners finished reading and want to leaving a SCO, SCO will be calling LMSFinish( ).
2. Status management functions, contains three API functions: LMSGetLastError ( ): This function will return an error code, every time after API called; the function value will be reset (except LMSGetErrorString and LMSGetDiagnostic). LMSGetErrorString( ): returns the error messages correspond to the error code. LMSGetDiagnostic( ): Returns detailed diagnostic information corresponds to the error code provided by LMS.
3. Data transfer functions: contains three functions: LMSSetValue (data model element, value): is responsible for storing all the relevant records of learning process. LMSGetValue(data model element): is mainly responsible for removing data from the LMS. LMSCommit(""): is responsible for restoring all temporary learning records in memory to LMS.

The interaction between SCO and LMS is through the browser, and API adapter is provided by LMS, it is needed automatically when LMS start (Sharable Content Object Reference Model 2004). Applet is a small client program of Java that used the HTML file of the Applet to download and automatically by web browser support for Java.

Therefore, Java Applet LMS can be used to achieve the API Adapter, LMS can tracing the online course of SCORM by embedded Java Applet in the network curriculum framework page, and then running the onload event to call init () function to deploy API Adapter. Here are the codes embedded API Adapter in the page.

```
<body onLoad="init( );">
<object classid = "clsid:8AD9C840-044E-11D1-B3E9-00805F499D93"
    width="0" height="0" id="APIAdapter" codebase="
http://java.sun.com/products/plugin/autodl/jinstall -1_4 -
windows-i586.cab#Version=1,4,0,0">
  <param name = "code" value = "org/adl/samplerte/
    client/APIAdapterApplet.class" >
  <param name = "codebase" value = "/elearning">
  <param name = "type" value="application/x-java-applet;jpi-version=1.4.2">
  <param name = "mayscript" value="true" >
  <param name = "scriptable" value="true" >
  <param name = "archive" value = "cmidatamodel.jar,
lmsclient.jar, debug.jar" >
    <comment>
<applet code = "org/adl/samplerte/client/
APIAdapterApplet.class"
      archive = "cmidatamodel.jar,lmsclient.jar,
debug.jar"
      codebase="/elearning"
      src="/elearning"
      height="0"
      id="APIAdapter"
      name="APIAdapter"
      width="0"
      mayscript="true">
</applet>
</comment>
</object>
```

4 Implement Tracing of Online Courses Based on SCORM Standard

In the online course, one page usually as a chapter. When production SCORM standard Online course, as a SCO, this page is the smallest unit for tracing (Yan ren-quan and Wu li-ping 2009). How to trace such a SCO?

First, we put two SCORM documents of all about packing—APIWrapper.js and SCOFunctions.js into the webpage. APIWrapper.js is a interface of online course which is used to invoke API Adapter, we used findAPI( ) to seek for API Adapter while SCO being started. We can invoke the tracing function of API Adapter to

interact with LMS after finding SCO. SCOFunctions.js is used to date analysis; we can add custom functions in it. We need to invoke the function in this file when SCO is jumping and exiting, In addition, it include another function that is used to calculating the time when the computer user is learning, those time should be saving in the date model. The code as follows when Adding SCORM package file:

<script type="text/javascript" src="SCOFunctions.js">

<script type="text/javascript" src="APIWrapper.js">

Second, add initialize event loadPage( ) and finish event unloadPage( ) of APIAdapter when the page loading and exiting, the codes fragment as follows:

<body onLoad="loadPage( )" onunload="return unloadPage( )">

Specific definition of these two files are in the package file, the loadPage( ) is defined as follows:

```
function loadPage( )
{
    var result = doLMSInitialize( );
    var status =doLMSGetValue("cmi.core.lesson_status")
    if (status == "not attempted")
      {
          doLMSSetValue("cmi.core.lesson_status", "incomplete")
      }
    exitPageStatus = false;
    startTimer( )
}
```

Third, trace current user's learning status by using functions of LMSSetValue( ) and LMSGetValue( ) to read and write data model in the page.

5 Conclusion

To achieve the learning behavior tracing is an interesting topic for learning platform developer and courseware production staff. Learning trace based on SCORM standard is a very good solution. Due to limited space, this article only describes the key technologies of learning tracing, there are many details that need to be resolved if want to trace learning behaviors. Expected this text can inspire and help e-Learning workers.

References

ADL.SCORM version1.2-The SCORM content aggregation model (2004) http://www.adlnet.org/ADLDOCS/Documents/SCORM_1.2_CAM.pdf

ADL.SCORM version1.2-The SCORM runtime environment (2004) http://www.adlnet.org/ADLDOCS/Document/SCORM1.2Run-TimeEnv.pdf

Chang Jia-yu (2009) Design and implement of sharable network course based on SCORM. Shanghai Second Polytech Univ 26(2):139–143

Operating mechanism of LMS run in client based on scorm (2010) http://www.iscorm.cn/post/623. html

Sharable Content Object Reference Model (SCORM) Content Aggregation Model (CAM) version 1.3.1 (2004) http//www.adlnet.Org/

Sun Xing-hua, Hao Li, Ye Yong-fei (2010) Research and implementation of tracing mechanism based on SCORM. E-Educ Res 65(7):65–68

Yan ren-quan, Wu li-ping (2009) Design and implement of learning management system based on SCORM specification. Comput Syst Appl 18(7):15–18

The Application of Information Technology in the Logistics

Chen Haiyan

Abstract According to the characters of logistics supports in the modern war, this paper puts forward the concept of modern logistics, analyzes the necessity for the application of information technology in the logistics, and also introduces several kinds of information technology which play an important role in the logistics system. Finally, the paper points out that we should enhance the research and application of advanced information technology in order to lay a solid technical foundation for promoting logistic support capability.

Keywords Logistics • Information • Technology • Application

1 Introduction

They say an army marches on its stomach. It is clear that logistic support has an important strategic position in the war. Under high tech conditions, modern local war has become more and more complex. The course and the result of the war has been mostly depend on the high-efficient logistical support system—logistics.

In modern wars, the type of modern materials is specially varied, the model is mixed, the technique content is high, and the amount of materials consumption is very big.

For example, Some 554,000 US troops have been dispatched in the Gulf War, the United States spent 5 months transporting about 7.7 million tons materials form America and European military bases to the Middle East in order to ensure the clothing, food, shelter and transportation for the army. This is equivalent to transporting a medium-sized city form mid-west America to Mid-East desert. The

C. Haiyan (✉)
Department of Force Management, The Chinese People's Armed Police
Force Academy, Lang fang, China
e-mail: chhaiyan@163.com

distance between the Gulf and the United States is far away, about 17,000 km from the sea, 11,000 km from the air. US must employ a large number of transport resources to complete their large-scale logistic support missions. During the Gulf War, Daily average carrying amount is 4,200 t. The massive consumption and wide range of materials set an even higher demand on logistics. Across the Gulf War, US not only won the modern high-technology war, but also successfully accomplished high difficulty multi-national logistic mission.

After the Gulf War, US thorough researched logistics, and put forward "accurate support". The so-called "accurate support" is the logistics support which provides the right number and right kind of supplies for the military at right time as well as right location (Wei Liu et al. 2008). In "Joint Military Operation Vision of 2010", the U.S. military proposed, "It is not necessary to establish a large-scale resources base for the war today. But it requires the support of a technical system with a very high standard. The desired amount of supplies should be distributed to the desired location at the desired time. But it is of no necessity to reserve military supplies in case of possible emergency" (Shaochuan Yan et al. 2008).

Modern war calls for high-efficient logistics supply system. Facing the strong pressure of the future information war, logistical service if wants to carry on the scientific deployment and supplement in the unprecedented broad battle field for millions of species, thousands of tons of materials, by no means human brain within its capacity. Realizing orderly material support relies on modern logistic command system in the modern war, and material logistics modernization must be supported by information technology. For this reason, the study on the application of information technology in logistics has become one of the current hot spots.

2 The Connotation of Modern Logistics

Logistics is the entity flow process of the commodity from provide place to the accepted place. According to the practice demand, it can carries out organic combination of the basic functions such as transportation, storing, loading and unloading, carrying, packaging, processing, distribution and information handling GB/T18354-2006 (2007).

Modern logistics is based on the information technology and the computer network; it is a comprehensive logistics activity by integrating functions of transportation, storing, loading and unloading, carrying, packaging, processing, distribution and information handling (Feng Wang et al. 2003). Its core content is cooperation of all elements, so as to achieve the efficiency optimization. Military logistics has the vital role for translating economic might into military might (Feng Wang et al. 2005).

3 The Significance of Application of Information Technology in the Logistics

3.1 *It Is Beneficial to Realize the Accurate Support*

Information war set an even higher demand on accurate support, but the traditional logistics is difficult to accurately predict and judge combat requirement. Due to the limitations of technical conditions in all previous wars, logistics have to reserve a lot of materials in various security link in order to meet the needs of war, at the same time, data of the peacetime and wartime can not be combined, data is very scattered, collection channel is not unified, logistics can only through the advantage of materials quantity scale to offset the lack of speed. After the Gulf war, countries army accelerated the pace of developing information technology to achieve logistics accurate support. In fact, it is the ideal goal of countries army since the Gulf war.

3.2 *It Is Beneficial to Realize the Materials Scientifically Management*

In the Gulf War and the Kosovo War, with the help of advanced equipment and information system, US Army has been able to successfully accomplish material support task. Among them, logistics information system marking as computer and satellite communications network played an important role.

For example, the U.S. TAV (Total Asset Visibility) System is an information system which dynamically monitors all links of materials with advanced technology. It can automatically track the information of the whole supply system and accurately display their real-time data, such as varieties, quality, location, transport tools and units, etc. The TAV System realizes the logistics "visible" and "controllable", at the same time, it provides scientific basis for the commanding decision of logistics. Therefore, by using information technology such as computer network and automatic identification technology, we can effectively control "material flow" through "information flow", improve reliability of the data's quantity and quality, reduce supplies levels and link, and greatly improve material supply speed. In modern war, the material support opportunity is fleeting. Logistics goals will be difficult to achieve without timely adjustment of advanced information system. Therefore, our military should accelerate the construction pace of the logistics information system.

4 The Application of Modern Information Technology in the Logistics

During the process of the logistics system construction, the use of large advanced technology greatly enhances logistics information level. So we should spread the advanced technology such as Radio Frequency, computer technology, network technology, modern communications technology, geographic information systems technology, global positioning system technology, electronic data interchange technology, etc.

4.1 Two-Dimensional Bar Code Technology

Bar code technology is basic means in logistics management; especially two-dimensional bar code technology is widely used in logistics field. Two-dimensional bar code records data symbol information using black-and-white figures which is distributed in the two-dimensional by specific geometrical figures according to certain rules. Using two-dimensional bar code technology, we can conveniently mark and descript personnel, goods, equipment and assets, the major marking contents include graphics, text, numbers and other information.

Thus, even in the database network interruption, you can also get basic information and achieve the basic management functions of logistics. When the database network is restored, you also gain further more item information and achieve more logistics management functions. In addition, the two-dimensional bar code also has advantages of high reliability, large information density, low production cost and great correction ability. It can also achieve a certain degree of information confidential by setting the password (Fig. 1).

These characteristics of two-dimensional bar code make it suitable for logistics applications. Therefore, it has been widely used in various military. For example, the US military requires that all military suppliers must use two-dimensional bar code shipping label based on PDF417, and issued labeling standards MILSTD197
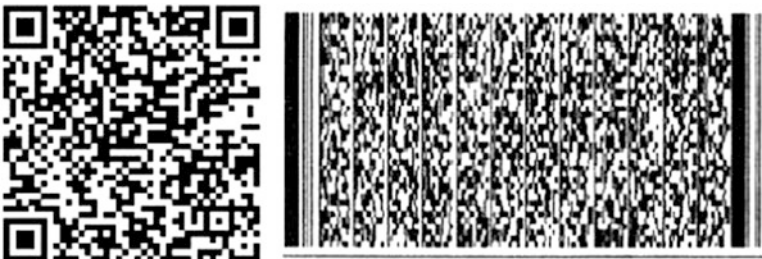
Fig. 1 Two-dimensional bar code

in 1999. Meanwhile, the two-dimensional bar code has also been applied to the US military ordnance store, ordnance maintenance, military ID, ammunition marking, material transportation and other fields. British troops also used two-dimensional bar code PDF417 which includes parts name, parts number, code designation of a military unit and other information in the weapons, equipment and facilities; it greatly improved the precision management of weapons ready. In the French army, two-dimensional bar code has also been widely used in transportation management. They use the laser-etched PDF417 bar code in the high-temperature metal plate to achieve the military personnel identification. In the military rescue, basic medical information of military personnel including identity, blood type, drug contraindications can be fast obtained from the two-dimensional barcodes under out of network environment, which accelerates the medical care process. At present, the conditions of applying two-dimensional bar code in our logistics has gradually matured, we have to give full play to two-dimensional bar code identification in logistics.

4.2 RFID Technology

Radio Frequency Identification System (RFID) is an automated data collection (ADC) system that enables businesses to wirelessly capture and move data using radio waves. RFID system consists of three parts: the Antenna, Reader and Tag. When the identified items with an RFID tag pass through its identifiable scope, the reader can automatically read the relevant information about the technical labels through the non-contact manner, and achieve functions of items automatic identification and information automatic collection. The operating principle of RFID is shown in Fig. 2.

As RFID systems can complete identification without touching between the RF card and reader, so it's the identification distance is farther away than distance in the optical system, it is not restricted line of sight, while radio frequency identification card also has the ability to read and write, carry large amounts of data, difficult to forge, it is particularly suitable for realizing automatic management in logistics system. Radio frequency identification technology can improve the efficiency and accuracy of the information collection, it is mainly used for monitoring position and condition of the equipment materials, warehousing management, searching and distributing of specific articles or goods in transit.

In the logistics process, there are hundreds, even thousands of containers in a container yard. When facing urgent searching task, the reader can propose the inquire according to the items name and code, all items in the containers with the RF card will make a response, then reader activates screaming device of radio frequency card, staff can tracks the sound to find the container. If the sound is outside the search scope, you can locate and find items according to the distance showed in the reader. In the Iraq War, US Army whole-process control and follow-up the position,
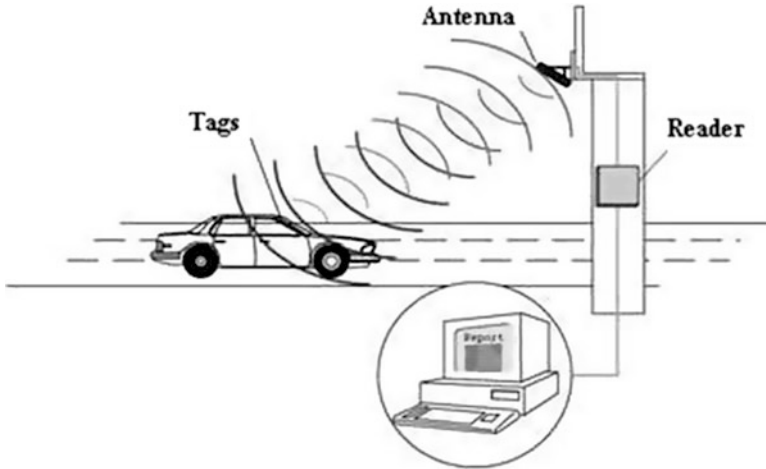
Fig. 2 The operating principle of RFID

attributes, condition of the personnel, equipment, and materials in transit by using this technology. They process 300,000 military materials container in 40 countries every day, and greatly improve ensures efficiency of military supplies. Therefore, we should also spread the RFID technology in the logistics.

4.3 GPS Technology

GPS is the abbreviation of Global Positioning System, which uses the satellite to navigate and measure location, moving targets on the ground or in the air can determine their own position.

GPS system has three main components, namely space constellation, the ground control segment and user equipment part. Space constellation means the satellites and its related tracking, telemetry, control and network control stations, the ground control segment include monitor station, master monitor station and ground antenna, user equipment part is GPS receiver. See Fig. 3.

GPS produced in the 1970s, the entire system has been improved in 1994, it is one of the most advanced information infrastructure today. The U.S. Department of Defense originally established GPS for military needs based on a global scale; it has exerted outstanding contributions for the US military for over 30 years. In the Gulf War, the United States massively used GPS to provide accurate location data for the multinational forces moving targets and realized the dynamic and real-time logistics management.
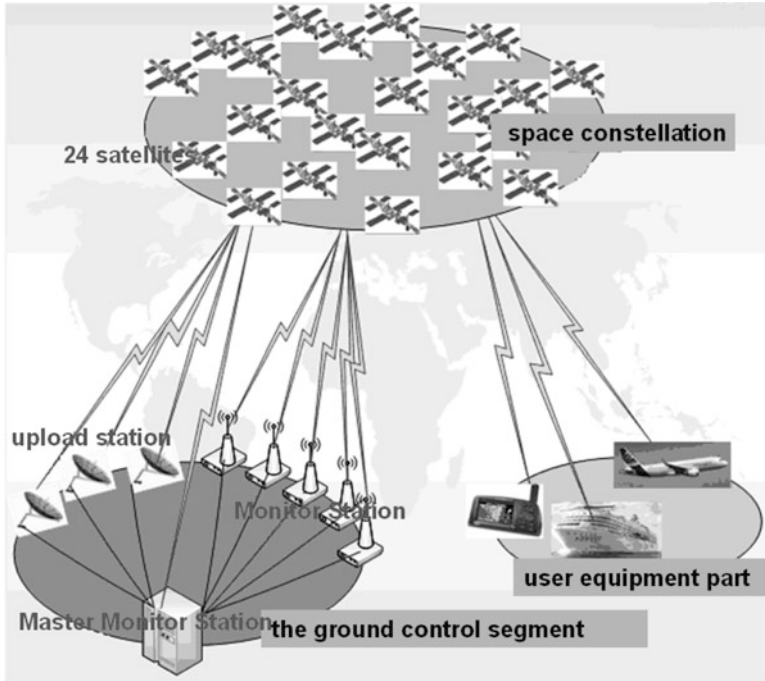
Fig. 3 Global positioning system

4.4 GIS Technology

Geographic Information System is a high technology, which combines computer technology and spatial data; it is a unique database in the world —Geodatabase. Basically, it is an extraordinarily effective technical system of comprehensive procession and analysis of space datum. GIS expressed geographic features as geometric objects with the classification property, and then organize, store, modify and display them though Layer and Layer has become essential basic features of GIS. See Fig. 4.

The Layer idea of GIS brings conveniently for the management of geographic object, using GIS systems, the resources and their situation is shown visually with rich maps form, it is easy to understand resources and decision-making. The development of visualization technology such as GIS creates favorable conditions for the close combination of logistics and geographic information. GIS technology widely used in logistics, it solves many practical problems, such as choice of transportation routes, logistics center location problem, the capacity of storage, handling strategy, transport vehicle scheduling and the choice of delivery routes and so on.
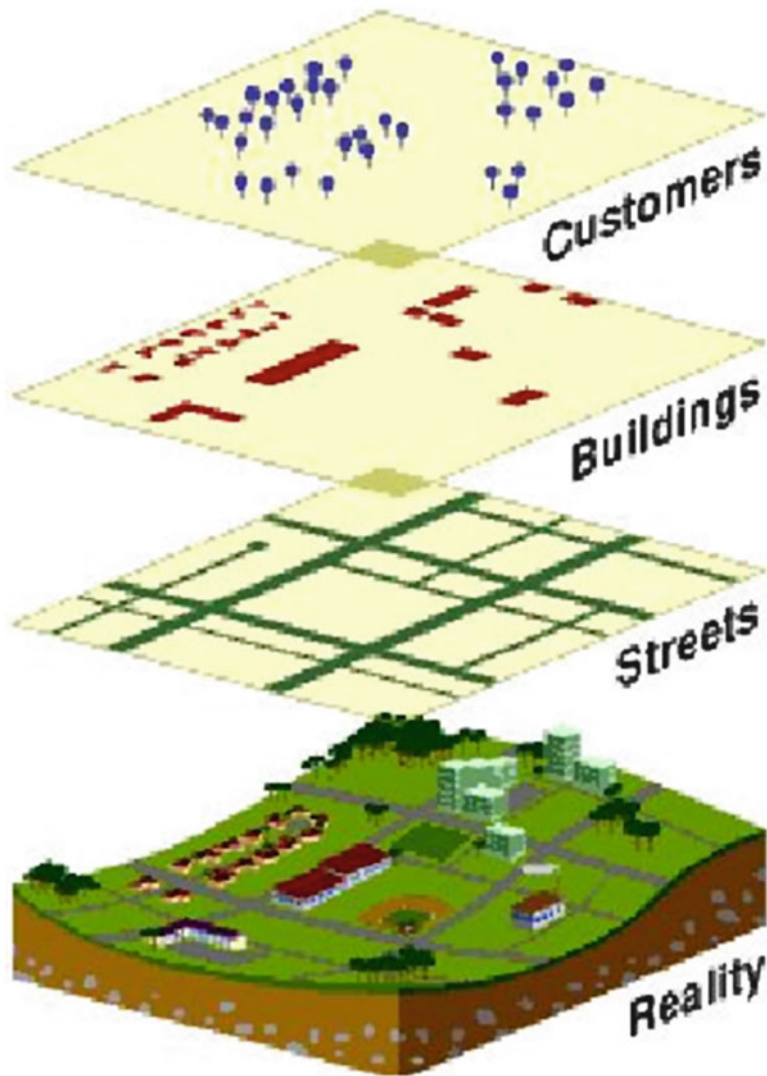
Fig. 4 GIS layer

By using the geographic data of GIS, we can improve logistics analysis technology, inquire goods in transit, solve the problems on distribution, services and query in logistics activities, automatic determine the position of warehouse and battlefield or warehouse and rescue sites for timely and accurate distribution. It can be seen that application of GIS in logistics will fundamentally change the management model of traditional logistics, and it has broad application prospects.

5 Conclusion

Currently, the world is in the information age, developed countries have increased efforts to intensify military reform. As the backbone of the military support, the logistics become the focus of information construction. In order to achieve logistics information, we must establish corresponding technical support system. For example, bar code technology provides a method for automatic identification and description of materials in logistics process; RFID technology can be used for tracking goods, vehicles, shelves and other information in the places where needs data collection and exchange with the non-contact manner; GPS which has the characteristics of real-time, rapid and high precision can be used for determining the location of items and transport scheduling management; GIS can be used to organize logistics and solves problems of the transportation route optimization options, logistics network rational distribution and other logistics problems. At present, developing information technology is an urgent requirement to our logistics; we must strengthen research and application of advanced technology to enhance the logistic support capability.

References

Feng Wang, Dali Jiang, Liang Peng (2003) Military logistics. China Material Press, Beijing

Feng Wang, Dali Jiang, Xilong Yang (2005) Modern military logistics. China Material Press, Beijing

GB/T18354-2006 (2007) S. National standards logistic sterm

Shaochuan Yan, Lei Jiang, Junjie Liu et al (2008) U.S. military logistics theory research overview. J Logist Technol 7:128–129

Wei Liu, Yajun Wang, Guowen Jia et al (2008) The study of prediction method for military logistics. J Logist Technol 4:113–115

Study on the Application of MAC-Based VLAN and Multi-domain Authentication in the Campus Network Management

Pu Heping, Li Pengxi, and Long Shiwen

Abstract Campus network as an important part of the digital campus construction is a platform that school administrators, teachers and students exchange information or deal with daily affairs. The campus network management and maintenance reflect not only the level of information technology, but also the direction and quality of its personnel training. Based on years of experiences in campus network management, we proposed the program of using MAC-based VLAN and multi-domain authentication in campus network management, and applied the program to practical management. It provided a new and effective method for campus network management, and achieved the goal that campus network management should be secure, stable and efficient, and make full use of existing resources to serve the education and teaching.

Keywords Campus network management • VLAN • MAC • Multi-domain authentication

P. Heping (✉)
The Center of Modern Educational Technology, Southwest Petroleum University, Chengdu, China
e-mail: puheping@swpu.edu.cn

L. Pengxi
Information Center, University of Electronic Science and Technology of China, Chengdu, China
e-mail: lpx@uestc.edu.cn

L. Shiwen
School of Foreign Languages, Southwest Petroleum University, Chengdu, China
e-mail: longshiwen@swpu.edu.cn

1 Introduction

With the expansion of the campus network and continuation increasing of users, the campus network operation has become the focus of campus network management. To ensure the campus network smooth operation, it must address the following two questions:

1.1 Curate Identification of Users

The whole campus network users are divided into two categories, one is free user, and the other is charging users. Taking account of the safety and effective using of resources, we need to make an accurate identification of the two categories of users by commonly using authentication technologies. At present, there are varieties of authentication techniques, such as WEB certification, 802.1x authentication and PPPOE authentication and so on. Of them, the 802.1x authentication is widely used in the campus network authentication, because it not only has the capacity of terminal management, but also controls users from the edge of the network access. It's a good solution to the internal campus network security issues, while meeting the billing needs. This paper expatiates by taking 802.1x authentications as an example. As most of the student hostel can only provide 1–2 ports and the actual residence of students is 4–6, each port usually links other equipment's bought by students such as non-management switch, Hub and so on. It's requested to distinguish multiple users under the same port when authenticating and billing.

1.2 Flexible User Charging Scheme

The coexistence of multiple billing models is the frequent situation in actual management of the campus network. Usually there are four access ways.

Firstly, the school constructs the entire network and manages users by itself with leasing network exit from Telecom Operator. In this case, the charges are generally lower, but access controlling is more stringent, especially for limiting bandwidth resources.

Secondly, network is constructed by the school, and network exit is supported by Telecom Operator. Users are still managed by the school, but the school and enterprises have to consult together about how to charge.

Thirdly, Telecom Operator constructs the entire campus network. It's responsible for managing of network users. In the case, the charge is relatively higher.

Fourthly, there are four access ways coexisting in the campus network in varying degrees. Different types of users share the same set of network. Users are unified managed by the school or it could be separately managed by the school and Telecom Operator, so that we need to take a variety of billing strategies and access control

policies. In the case of single account and a variety of billing strategies, we usually distinguish users by the ISP domain authentication. So this paper also uses the same method to distinguish users.

2 Mac-Based VLAN

VLAN (Virtual Local Area Network) is called "virtual LAN". VLAN equipment will be divided into one logical segment in order to achieve emerging data exchange of Virtual Working Group (Dongzhu and Deng 2010).

As noted earlier, a single port in the campus network has multiple services, and each service belongs to different VLAN. So the flexible configuring of VLAN under the switch port to identify different services has become a key issue of the campus network management.

In order to solve the above-mentioned problems, we propose a MAC-based VLAN solution. MAC (Media Access Control) address is burnt on a Network Interface Card (NIC, NIC), also known as the hardware address. It's composed of 48 bits long (6 bytes), 16 hex digits.

MAC-based VLAN is another way to distinguish VLAN that tag of VLAN is added to packet according to the source MAC address. This is often in combination with security technologies (such as 802.1X) to achieve the purpose of the terminal's safety and flexible access.

3 Multi-domain Authentication

Authentication is the process that net users establish a right to use the identification (E.g. user name). In the case of multi-domain authentication, one user may have a lot of identifications (Li Yu 2007). These identifications will be identified with a specific identity meaning in the default semantics. To facilitate the management of multi-domain authentication, this paper adopts "AAA" management. "AAA" is the short form for "Authentication, Authorization and Accounting". It's a network security management mechanism which provides such three security functions as authentication, authorization and accounting.

Generally, "AAA" adopts the Client/Server structure that the client runs on NAS (Network Access Server) which intensively manages the user's information. In terms of the user, NAS is a server; but in terms of servers, it's the client. The basic network structure of "AAA" is shown as Fig. 1. When the user wants to establish connection with the NAS through a certain network to gain the right to access to other networks or obtain the certain network's resources, NAS has played a role of verify users or corresponding connection. In the following figure, access user can choose to use multiple "AAA" servers, e.g. RADIUS Server (Remote Authentication Dial In User Service) or HWTACACS (Huawei Terminal Access Controller Access Control System) Server.
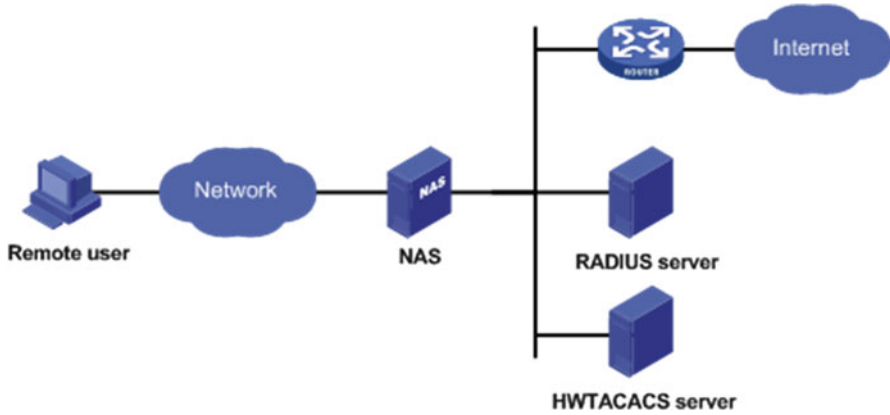
Fig. 1 AAA (*Authentication, Authorization and Accounting*) diagram of the basic network structure

4 The Model of Multi-domain Authentication Based on Dynamic VLAN

VLAN is devised by MAC based on the following ways: When the port receives an untagged packet, the equipment can, with the source MAC address of packets matching the keywords, learn the terminal binding VLAN through finding MAC VLAN entry, so that the packets are transmitted in the specified VLAN (Li Lu 2006).

MAC VLAN entries are generated in two ways: static configuration and dynamic configuration.

4.1 Static Configuration

In this way, the user binds terminal of MAC address and VLAN by the command line. The device generates corresponding MAC VLAN entries. When manually configuring, the binding is a single MAC address and VLAN if all specified MASK values are 1. If the high orders of specified MASK are continuous 1, low orders are 0, the binding is a kind of MAC address and VLAN. It's usual to divide the communication equipment of a manufacturer into the same VLAN.

- Advantages:
  This way is easily realized; it involves only the access device.
- Disadvantage:
  Disadvantage of this way is that initialization of all users must be configured, if there are hundreds or even thousands of users, the configuration is tedious. And this way has led to reduction in the efficiency of the switch, because in each switch port may exist lots of members of the VLAN group, making it

impossible to limit the broadcast packets. In addition, the notebook PC users may change their network cards frequently, so VLAN configuration must be incessantly changed too.

4.2   Dynamic Configuration

In this way, the user needs to configure both MAC VLAN and MAC-based authentication on the device (Such as MAC address or MAC-based 802.1X authentication) (Okayama et al. 2005). If the user requests authentication, server will authenticate the user's name and password. If it is certified, the VLAN information will be issued by server. Then the device generates MAC VLAN entries in the light of source MAC address and VLAN information. And MAC VLAN will be added to the untagged VLAN port list automatically. After the user offline, the device deletes MAC VLAN entries automatically. MAC VLAN will be deleted from the list.

• Advantages:
  This way is flexible and secure. It can identify the MAC address, create MAC VLAN entries and allow MAC VLAN to pass through access ports automatically. Therefore, when this way is used in large networks the configuration will be predigested greatly, and it is flexible to use. Only when the user authentication is successful can it pass through the appointed VLAN accessing network, so the security of network is increased.
• Disadvantage:
  MAC VLAN can only be enabled on Hybrid port, disabled on ACCESS interface. The server needs to check whether it conflicts with static VLAN configuration, if there is some conflict it will lead to issue failure.
    Taking various factors into account, the solution is as follows: the dynamic configuration MAC VLAN coordinates with MAC-based "AAA" authentication; "AAA" server is arrayed in the network, and VLAN is issued through the server. The campus network environment is single-user and multi-service, "AAA" server offers domain-based authentication and VLAN policy. Different businesses under the same port can be issued with different VLAN; it achieved the goal of dynamic VLAN multi-domain authentication.

5   Design and Implementation of Multi-domain Authentication Campus Network Management System

5.1   Network Model

We need to carry out authentication, authorization and billing of the network access users to improve the security and operation of campus network. Strategy is as follows:
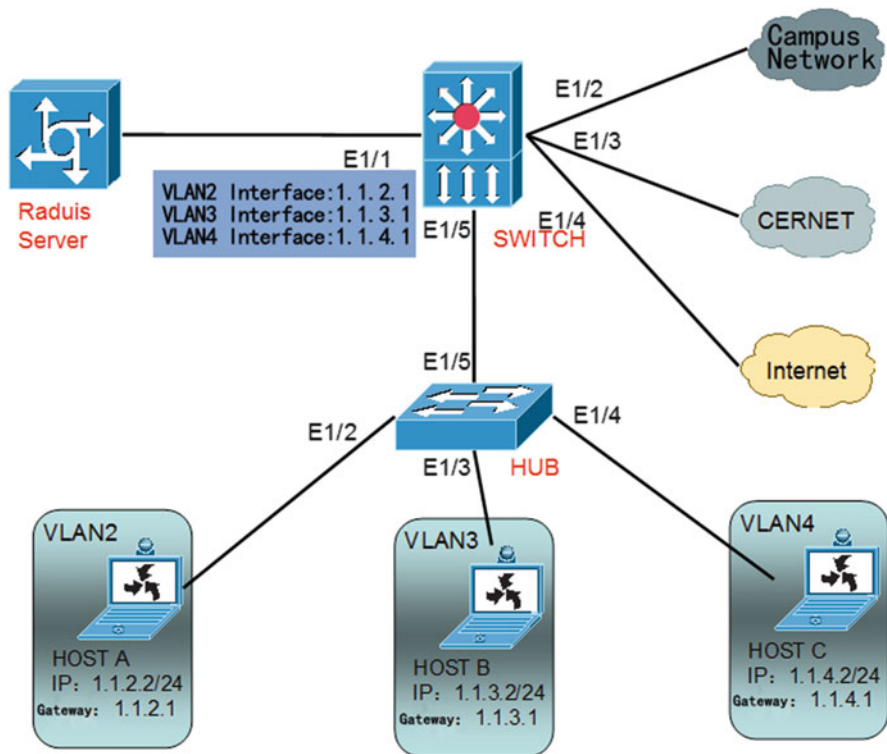
Fig. 2 Network model

1. Users who access the internal server on campus use "XXX@CAMPUS" Certification. It is certified but no charge. Server distributes the segment address 1.1.2.0/24 to them;
2. Users who access the CERNET use "XXX@CERNET" Certification. They can only access CERNET resource. We execute "A" charges. Server distributes the segment address 1.1.3.0/24 to them;
3. Users who access the Internet use "XXX@Internet" Certification. They can access Internet resource. We execute "B" charges. Server distributes the segment address 1.1.4.0/24 to them.

A student dormitory in the campus network accesses E1/5 port of switch (shown as Fig. 2). Students in the dormitory connect the network through the HUB. The Student "A" needs to access the campus server. The Student "B" needs to access CERNET. The Student "C" needs to access Internet. We are going to give student address and charging according to the overall address planning of school.

5.2   Design and Implementation

The following device configuration takes the campus network switch command of H3C as an example. We list the key configuration steps only in order to facilitate the readers understand. Other programs are similar to this.

1. Configuring "AAA" program

```
# creates a RADIUS authentication MAC VLAN, the
appointed authentication and accounting server IP
addresses are 192.168.1.15. Keys are SWPI. (This
parameter needs to be consistent with the configuration
of RADIUS Server)
<Switch> system-view
[Switch] radius scheme macvlan
New Radius scheme
[Switch-radius-macvlan]server-type extended
[Switch-radius-macvlan]primary authentication
192.168.1.15
[Switch-radius-macvlan] primary accounting
192.168.1.15
[Switch-radius-macvlan] key authentication SWPI
[Switch-radius-macvlan] key accounting SWPI
[Switch-radius-macvlan] quit
# configures the domain parameters. Creating three
domains which are CAMPUS, CERNET and Internet, taking
CAMPUS for example, so the others.
<Switch> system-view
[Switch]domain CAMPUS
[Switch-isp-CAMPUS] authentication lan-access
radius-scheme macvlan
[Switch-isp-CAMPUS] authorization lan-access
radius-scheme macvlan
[Switch-isp-CAMPUS] accounting lan-access
radius-scheme macvlan
[Switch-isp-CAMPUS] quit
```

2. Configuring 802.1X

```
#Global 802.1X function is enabled.
[Switch] undo port-security enable
[Switch] dot1x
802.1X is enabled globally.
#Students are authenticated with 802.1X on interface
Ethernet1/5.
[Switch] dot1x interface ethernet 1/5
802.1x is enabled on port Ethernet1/5.
```

3. Configuring MAC VLAN

```
# Configuring port Ethernet1 / 5, the port type is
Hybrid, opening the MAC VLAN function.
[Switch] interface ethernet 1/5
[Switch-Ethernet1/2] port link-type hybrid
[Switch-Ethernet1/2] mac-vlanenable
[Switch-Ethernet1/2] quit
```

4. Configuring RADIUS Server
   Here the ideas of configuration are briefly described since RADIUS Server configuration interface is different. The following configuration can be achieved on the Radius Server in the mainstream.

   Adding SWITCH, the shared key is SWPI; adding CAMPUS Serving, Assign VLAN2 is issued after selecting the authorization policy; no charging; adding CERNET serving, Assign VLAN3 is issued by selecting the authorization policy, configuring "A" tariff; adding Internet serving, Assign VLAN4 is issued after selecting the authorization policy, configuring "C" tariff. Creating a user Host A, the account "Host A", the password "aaa", service CAMPUS, CERNET and Internet; creating user Host B, Host C and so on and so forth, three services are created for each user.

5. Verifying results
   The connection of 802.1X was used by Host A, user's name is "Host A", and password is "aaa", connection is successful. The connection of 802.1X was used by Host B, user's name is "Host B", password is "bbb", and connection is successful. The connection of 802.1X was used by Host C, user's name is "Host C", and password is "ccc", connection is successful.

6   Conclusion

Managing campus network with the dynamic MAC-based VLAN greatly expanded the scope of application of multi-domain authentication. Especially when the dormitory in the campus network accesses, the complex campus network environment makes it difficult to carry out the security policy of static configuration and address planning; meanwhile, the concurrent access of port of varieties of users causes the authorization policies of 802.1x authentication to become more complex. The above -mentioned problems can be solved by the dynamic MAC-based VLAN. It provides an effective way for the campus network management. On the one hand, it achieves a safe and efficient management; on the other hand, it optimizes allocation of resources and saves management cost.

References

Feng Dongzhu, Yang Deng (2010) The application of VLAN-based technology in campus network construction. Microcomput Inf 6(June):133–134 (in Chinese)

Li Lu (2006) The precaution and control against the safe risk of the campus network. Netw Secur Technol Appl 8(August):72–73 (in Chinese)

Li Yu (2007) The development of multi-domain authentication and authorization technology. Libr Theory Pract 2(February):101–103 (in Chinese)

Okayama K, Yamai N, Miyashita T, Kawano K, Okamoto T (2005) A method of dynamic interconnection of VLANs for large scale VLAN environment. Information and telecommunication technologies. APSITT 2005 proceedings. 6th Asia-Pacific symposium, December 2005, pp 427–432 doi:10.1109/APSITT.2005.203697

An Empirical Study on Performance Evaluation of Construction Project Manager

Hui Cao and Changfa Xiang

Abstract The purpose of this paper is to develop a scale and an integrated assessment method of measuring construction project manager's performance. A performance scale of construction project manager is developed in this paper, which contains three dimensions of execution power, coordination power and competency and 24 indicators. The performance scale combines theories of manager performance, person-job matching and competency and is developed by semi-structured interviews and extraction of the key indicators from the perspective of objective, process and person capability. Moreover, 46 construction project managers are selected as the samples from Citic Heavy Industries Co., LTD and the methods of entropy, analytic hierarchy process (AHP) and fuzzy integrated judgment are integrated to evaluate the performance of construction project managers. The results of this paper provide a way to evaluate construction project manger's performance theoretically and methodology.

1 Introduction

The research on performance has been a significant topic in the field of management since 1990s. In the connotation sense there has been no common viewpoint of the performance. The three major perspectives proposed are result-based performance, behavior-based performance and value-based performance.

H. Cao • C. Xiang (✉)
School of Economics and Management, China University of Geosciences, Wuhan, China
e-mail: afanlpb258@126.com

Recently, project manager performance is being more and more concerned by scholars home and abroad. Most of them concerned with performance scale of project manager. However, the existing scales of project manager performance focuses on task performance, which reflects the manager's task execution results, but seldom considers the competency and coordination performance of project managers. Accordingly, there isn't a unified and common scale to evaluate the performance of project managers.

At present, the methods to evaluate project manager performance mainly focus on single evaluation method, such as management by objective (MBO), key performance indicator (KPI), balanced score card (BSC), analytic network-level process (ANP) and data envelopment analysis (DEA), etc. (Li Shenghai 2003; Sheu and Hong 2007; Cao and Hoffman 2010). Besides, some scholars also do some researches on applications of integrated evaluation methods of project manager's performance, especially, the integration of analytic hierarchy process (AHP) and fuzzy integrated judgment (Shang Mei 2004). However, the subjectivity of these methods may reduce the accuracy of project manager performance evaluation. Construction project manager is an important kind of project managers. However, its performance researches are still in infancy. Thus, in this study, we'll develop a three-dimension scale of construction project manager performance based on the theories of manager performance, project management and person-job matching. In order to overcome the subjectivity of the existing methods, we'll integrate the entropy method, AHP and fuzzy integrated judgment to evaluate the performance of construction project manager. Moreover, we take 46 construction project managers from Citic Heavy Industries Co., LTD as an example to evaluate the performance of construction project managers, which is expected to provide a way to evaluate construction project manger's performance theoretically and methodology.

2 The Scale of Construction Project Manager Performance

In this paper, we develop a scale to evaluate the performance of construction project managers, which is based on the theories of performance management, person-job matching and competency and proposed by semi-structured interviews and extraction of the key indicators from the perspective of objective, process and person capability.

The scale of construction project manager performance includes three dimensions of execution power, coordination power and competency and 24 indicators, which reflects the degree of implementation of tasks, coordination level of management and the quality of their own ability of construction project manager. (Seen in Table 1).

Table 1 The index system of construction project manager performance

Goal hierarchy	Criteria hierarchy	Indicator hierarchy
Performance of construction project manager (POCPM)	**Execution power (EP)**	u_1: The compliance rate of project quality when checking for the overall project
		u_2: The acceptance rate of the key and important process
		u_3: The saving rate of material cost
		u_4: The rate of progress ahead in key nodes
		u_5: The rate of project completion ahead
		u_6: The implement rate of security measures
		u_7: Occupational health and environmental system operation
		u_8: The gross profit margin of project
		u_9: The rate of accounts receivable on project
		u_{10}: Owner satisfaction
		u_{11}: The complete degree and situation in time of project acceptance data
	Coordination power (CP)	u_{12}: The plan of project-related personnel, equipment, materials and progress
		u_{13}: Clearing the relationship among the project-related people's responsibilities, rights and interests
		u_{14}: The cooperation of supervision unit, the owner, designer, and other construction units
		u_{15}: To evaluate and select the construction method properly
		u_{16}: Coordinating the supply of workers in construction site, material, machinery transport and kinetic energy to meet the requirements of construction schedule
		u_{17}: The coordination of cross-site operations to ensure the construction progress
		u_{18}: To encourage and create conditions for subordinates to gain new knowledge, experience and skills
		u_{19}: Have strong risk awareness and identify risks and implement countermeasures correctly
	Competency (CY)	u_{20}: Have good team spirit
		u_{21}: Work seriously, diligent, honest and trustworthy, self-discipline, sacrifice when needed
		u_{22}: Verbal communication skills, writing skills and social skills
		u_{23}: With project management expertise and experience
		u_{24}: The ability to acquire new knowledge and skills

3   Methodology

The performance scale we proposed involves three dimensions and 24 indicators. It is a multi-level and multi-indicator evaluation system. Thus, the fuzzy integrated judgment is the most appropriate method to evaluate the performance of construction project manager. The key point is to obtain the weights of execution power, coordination power, competency and 24 indicators, when we apply fuzzy integrated judgment to evaluate the performance of construction project manager. In existing literatures, AHP is mainly used to obtain weight, but the subjectivity of AHP may affect the final result. In order to overcome this shortcoming, we apply AHP and entropy method to obtain the subjective weight and objective weight.

The principle of the integrated method in our study may describe as following: Firstly, we obtain the subjective weight and objective weight by AHP and entropy method and take the average of them as the combined weight, which is treated as the weight set of fuzzy integrated judgment. Secondly, according to the fuzzy integrated judgment model, we get the final result by fuzzy synthesis of the weight set and fuzzy judgment matrix. In our study, the principle and process of the integrated method is described as Fig. 1.

4   Empirical Study

4.1   Sample and Data

We select 46 construction project managers from Citic Heavy Industries co., LTD and evaluate their performance according to the scale proposed above by 360 scores method. The reliability of data is tested by SPSS16.0. The results show that reliability of data is very good and satisfies the requirement of statistics (seen Table 2).

4.2   Constructing Factor and Judgment Sets

Two levels are identified for factor sets according to the scale. The first level includes three dimensions: execution power, coordination power and competency. The second level includes the 24 indictors:

The first level: $POCPM = \{EP, CP, CY\}$

The second level:

$$EP = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}\};$$

$$CP = \{u_{11}, u_{13}, u_{14}, u_{15}, u_{16}, u_{17}, u_{18}, u_{19}\};$$

$$CY = \{u_{20}, u_{21}, u_{22}, u_{23}, u_{24}\}$$
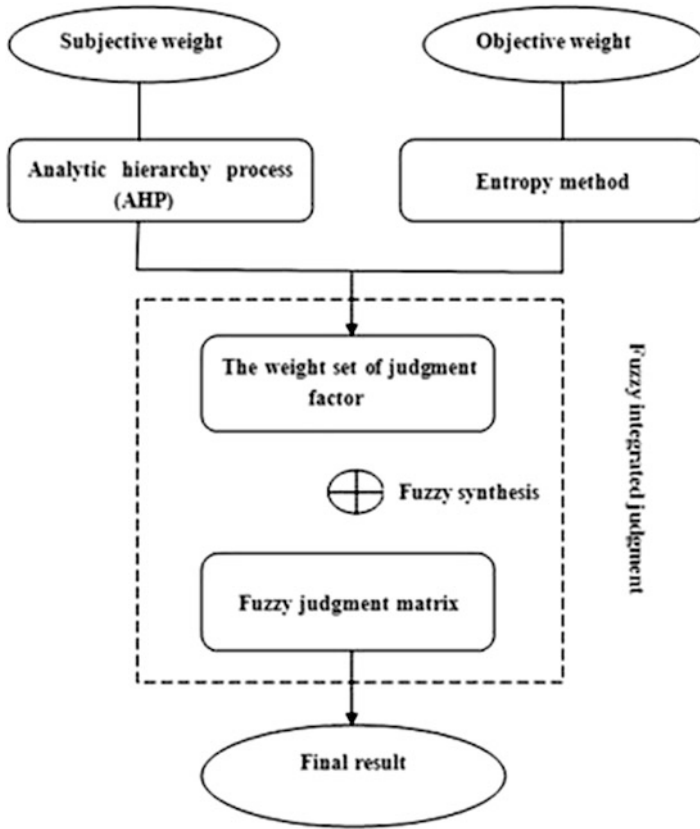
Fig. 1 The principle and process of the integrated method

Table 2 The result of reliability test

Dimensions	Item	Alpha coefficient
Execution power	11	0.892
Coordination power	8	0.941
Competency	5	0.935
Total performance	24	0.960

Let $V = \{v_1, v_2, \ldots v_j\}$ denote the judgment set. In this study, j is denoted as 5. Thus, the factor set is defined as $V = \{v_1, v_2, v_3, v_4, v_5\}$, where $v_1 = $ rank A, $v_2 = $ rank B, $v_3 = $ rank C, $v_4 = $ rank D, *and* $v_5 = $ rank E.

4.3 Constructing Judgment Matrix

We utilize 360 scores method to evaluate the performance of construction project managers, there are many scores of one indicator for every construction project

manager. Therefore, frequency method is used to construct the judgment matrix in this study. The judgment matrix obtained by frequency method is more objective than that given by experts directly. For example, the judgment matrix of execution power for the first sample is as follows:

$$R_{11} = \begin{bmatrix} 0.3750 & 0.5000 & 0.1250 & 0.0000 & 0.0000 \\ 0.7143 & 0.1429 & 0.1429 & 0.0000 & 0.0000 \\ 0.2857 & 0.1429 & 0.1429 & 0.1429 & 0.2857 \\ 0.3750 & 0.1250 & 0.2500 & 0.2500 & 0.0000 \\ 0.0000 & 0.3750 & 0.2500 & 0.3750 & 0.0000 \\ 0.3750 & 0.2500 & 0.3750 & 0.0000 & 0.0000 \\ 0.2500 & 0.7500 & 0.0000 & 0.0000 & 0.0000 \\ 0.7143 & 0.1429 & 0.0000 & 0.1429 & 0.0000 \\ 0.0000 & 0.4286 & 0.2857 & 0.0000 & 0.2857 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.1250 & 0.6250 & 0.2500 & 0.0000 & 0.0000 \end{bmatrix}$$

With the same principle, we get the judgment matrixes of coordination power and competency for the first sample and the judgment matrixes of other samples.

4.4 Indentifying the Weight Sets for Each Level

According to the integrated method identified before, we obtain the subjective weight and objective weight by AHP and entropy method and take the average of them as the combined weight sets. The subject weights are computed through the judgment matrixes which are given by experts about each indicator under execution power, coordination power and competency. The computing process is completed by MatlabR009a. The subject weight sets, w_1', w_2' and w_3', are as follows:

$$w'_1 = [0.1929 \quad 0.1929 \quad 0.1209 \quad 0.0707 \quad 0.0401 \quad 0.0707 \quad 0.0707$$
$$0.1209 \quad 0.0401 \quad 0.0401 \quad 0.0401];$$

$$w'_2 = [0.2545 \quad 0.1469 \quad 0.1469 \quad 0.0762 \quad 0.0762 \quad 0.0762 \quad 0.0762$$
$$0.1469];$$

$$w'_3 = \begin{bmatrix} 0.0986 & 0.1758 & 0.3134 & 0.3134 & 0.0986 \end{bmatrix}$$

The objective weight sets, w_1, w_2, and w_3, are computed by the empirical data according to entropy method. They are as follows:

$$w_1 = [0.0753 \quad 0.0810 \quad 0.1374 \quad 0.0812 \quad 0.1040 \quad 0.0984 \quad 0.0763$$
$$0.0585 \quad 0.0817 \quad 0.1254 \quad 0.0808];$$

$$w_2 = [0.1083 \quad 0.1272 \quad 0.0809 \quad 0.1081 \quad 0.1773 \quad 0.1192 \quad 0.1474$$
$$0.1316];$$

$$w_3 = \begin{bmatrix} 0.1906 & 0.1508 & 0.2209 & 0.2882 & 0.1496 \end{bmatrix}$$

The combined weight sets, $\overline{w}_1$, $\overline{w}_2$ and $\overline{w}_3$ are computed by the average of subjective weight sets and objective weight sets. They are as follows:

$$\overline{w_1} = [0.1341 \quad 0.1370 \quad 0.1291 \quad 0.0721 \quad 0.0846 \quad 0.0735 \quad 0.0897$$
$$0.0609 \quad 0.0827 \quad 0.0605]$$

$$\overline{w_2} = [0.1814 \quad 0.1370 \quad 0.1139 \quad 0.0921 \quad 0.1267 \quad 0.0977 \quad 0.1118$$
$$0.1393]$$

$$\overline{w_3} = \begin{bmatrix} 0.1446 & 0.1633 & 0.2671 & 0.3008 & 0.1241 \end{bmatrix}$$

The rest may be deduced by analogy, the combined weight of execution power, coordination power and competency is computed as follows:

$$\overline{w} = \begin{bmatrix} 0.5102 & 0.2755 & 0.2143 \end{bmatrix}$$

4.5 Fuzzy Integrated Judgment of Construction Project Managers' Performance

In this paper, we take the first sample as an example to illustrate the process of fuzzy integrated judgment of construction project managers' performance. Based on the fuzzy integrated matrixes and the weight sets obtained before, the first-order fuzzy integrated judgment for the execution power of the first sample is obtained as follows:

$$B_1 = \overline{w_1} \cdot R_{11}$$
$$= [0.1341 \quad 0.1370 \quad 0.1291 \quad 0.0759 \quad 0.0721 \quad 0.0846$$
$$0.0735 \quad 0.0897 \quad 0.0609 \quad 0.0827 \quad 0.0605]$$

$$
\cdot \begin{bmatrix}
0.3750 & 0.5000 & 0.1250 & 0.0000 & 0.0000 \\
0.7143 & 0.1429 & 0.1429 & 0.0000 & 0.0000 \\
0.2857 & 0.1429 & 0.1429 & 0.1429 & 0.2857 \\
0.3750 & 0.1250 & 0.2500 & 0.2500 & 0.0000 \\
0.0000 & 0.3750 & 0.2500 & 0.3750 & 0.0000 \\
0.3750 & 0.2500 & 0.3750 & 0.0000 & 0.0000 \\
0.2500 & 0.7500 & 0.0000 & 0.0000 & 0.0000 \\
0.7143 & 0.1429 & 0.0000 & 0.1429 & 0.0000 \\
0.0000 & 0.4286 & 0.2857 & 0.0000 & 0.2857 \\
0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.1250 & 0.6250 & 0.2500 & 0.0000 & 0.0000
\end{bmatrix}
$$

$$
= \begin{bmatrix} 0.3352 & 0.3773 & 0.1560 & 0.0773 & 0.0543 \end{bmatrix}
$$

Thus, the result of execution power for the first sample which stands for the distribution in the five ranks is as follows:

$$
B_1 = \begin{bmatrix} 0.3352 \ 0.3773 \ 0.1560 \ 0.0773 \ 0.0543 \end{bmatrix}
$$

One score is given for every rank, where rank $A = 95$, rank $B = 84.5$, rank $C = 74.5$, rank $D = 64.5$ and rank $E = 29.5$. The score of execution power is obtained by the score for each rank multiplying by the result of fuzzy integrated judgment for the execution power. The formula is as follow:

$$
\begin{aligned}
EP_{(score)} &= b_1 * 95 + b_2 * 84.5 + b_3 * 74.5 + b_4 * 64.5 + b_5 * 29.5 \\
&= 0.3352 \times 95 + 0.3773 \times 84.5 + 0.1560 \times 74.5 \\
&\quad + 0.0773 \times 64.5 + 0.0543 \times 29.5 \\
&= 81.94
\end{aligned}
$$

Similarly, the results (B_2 and B_3) can also be obtained by the fuzzy integrated judgment for the coordination power and competency of the first sample.

Therefore, we can obtain the fuzzy judgment matrix R of the first level.

$$
R = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix}
0.3352 & 0.3773 & 0.1560 & 0.0773 & 0.0543 \\
0.4991 & 0.4834 & 0.0174 & 0.0000 & 0.0000 \\
0.5988 & 0.3856 & 0.0155 & 0.0000 & 0.0000
\end{bmatrix}
$$

The second-order integrated judgment of the construction project managers' performance is made as follows:

Table 3 The scores of EP, CP, CY and POCM for 46 construction project managers

N	EP	CP	CY	POCPM	N	EP	CP	CY	POCPM
1	81.9	89.6	90.6	85.9	24	87.3	89.2	90.9	88.6
2	83.6	88.8	91.7	86.8	25	82.0	84.5	83.2	83.0
3	79.4	86.7	89.6	83.6	26	85.1	88.2	88.9	86.7
4	83.5	90.0	92.5	87.3	27	83.0	87.8	85.9	85.0
5	87.4	90.7	93.6	89.6	28	79.5	84.0	82.4	81.4
6	87.4	90.9	92.2	89.4	29	80.9	85.4	85.6	83.2
7	88.2	90.5	91.0	89.4	30	83.9	86.3	87.0	85.2
8	88.4	90.9	92.4	89.9	31	82.5	85.2	86.3	84.1
9	88.4	90.1	91.7	89.6	32	77.1	82.3	81.7	79.5
10	86.5	89.9	91.5	88.5	33	78.1	82.7	82.3	80.3
11	85.6	89.1	90.8	87.7	34	81.7	85.7	86.9	83.9
12	85.7	89.2	90.5	87.7	35	82.2	85.0	85.2	83.6
13	86.2	89.2	90.7	88.0	36	86.7	89.7	91.0	88.4
14	86.1	88.2	89.1	87.3	37	84.7	85.9	87.1	85.6
15	85.9	88.4	89.8	87.4	38	76.8	82.8	81.6	79.5
16	87.2	91.3	93.5	89.7	39	84.4	87.3	88.4	86.1
17	88.9	89.1	90.1	89.2	40	79.3	84.5	83.2	81.6
18	89.7	90.9	91.3	90.4	41	86.1	90.2	89.0	87.8
19	88.9	91.3	92.0	90.2	42	86.4	89.8	89.1	87.9
20	88.5	90.7	91.7	89.8	43	86.1	89.7	89.0	87.7
21	84.7	88.7	88.8	86.7	44	86.0	90.1	92.2	88.4
22	88.0	89.1	90.7	88.9	45	84.3	89.2	89.9	86.9
23	85.8	88.6	91.3	87.7	46	88.7	94.1	93.1	91.1

$$B = w \cdot R$$

$$= \begin{bmatrix} 0.5102 & 0.2755 & 0.2143 \end{bmatrix} \cdot \begin{bmatrix} 0.3352 & 0.3773 & 0.1560 & 0.0773 & 0.0543 \\ 0.4991 & 0.4834 & 0.0174 & 0.0000 & 0.0000 \\ 0.5988 & 0.3856 & 0.0155 & 0.0000 & 0.0000 \end{bmatrix}$$

$$= \begin{bmatrix} 0.4369 & 0.4083 & 0.0877 & 0.0394 & 0.0277 \end{bmatrix}$$

Similarly, the score of construction project managers' performance can be computed as follows:

$$POCPM_{(score)} = 0.4369 \times 95 + 0.4083 \times 84.5 + 0.0877 \times 74.5$$
$$+ 0.0394 \times 64.5 + 0.0277 \times 29.5$$
$$= 85.90$$

With the same principle, we can compute the scores of execution power, coordination power, competency and construction project managers' performance of other samples. The results can be seen in Table 3.

5 Conclusion

On the basis of manager performance theory, person-job matching theory and competency theory, a scale is proposed in this study to evaluate the performance of construction project manager, which involves three dimensions and 24 indicators. It provides a theoretic framework to evaluate the performance of construction project manager in management practice. An integrated method containing entropy method, analytic hierarchy process (AHP) and fuzzy integrated judgment is developed, which is used to evaluate the performance of 46 construction project managers in Citic Heavy Industries Co., LTD. The result shows that the construction project managers of this company have good performance, especially the performance of coordination power and competency.

References

Cao Q, Hoffman JJ (2010) A case study approach for developing a project performance evaluation system. Int J Proj Manage 2(18):68–74

Li Shenghai (2003) Performance evaluation of the middle managers and project manager in Jiangsu SanXing Co., Ltd. Xi'an University of Technology 9 (in Chinese)

Shang Mei (2004) Employment and assessment of project manager in the construction enterprise based on fuzzy comprehensive evaluation. Xi'an Univ Sci 24(1):133–135 (in Chinese)

Sheu HC, Hong TL (2007) Performance evaluation model for project managers using managerial practices. Int J Proj Manage 25:543–551

The Design of Ontology-Based Intelligent Answering System Model in Network Education

Yao Zhen and Zhang Zheng-wan

Abstract This article describes the importance of Answering System in network education, and analyzes the current common problems of Answering System. Based on the intelligent, interactive and resource sharing needs of Answering System, the article proposes an ontology-based intelligent Answering System model in network education and suggests related technologies.

Keywords Ontology • Intelligent answering • Semantic web • Network education

1 Introduction

Distance Answering System's main function is for students to answer the problems in the learning process, tracking the effect of the students. Relatively isolated the teachers and students in distance education environment, it plays an indispensable role in strengthening the exchanges between teachers and students, helping students with learning disabilities and accessing the correct answer information. In the current practice of online education, they use e-mail, online discussions, keyword queries, etc. to display answering (Androutsopoulos et al. 1995).

Y. Zhen (✉)
College of Electronic and Information Engineering, Chongqing Three Gorges University, Wanzhou, Chongqing, China
e-mail: yzhen43@126.com

Z. Zheng-wan
Online Education College of Southwest University, Southwest University, Beibei, Chongqing, China
e-mail: James129@swu.edu.cn

1.1 The Problems of Remote Answering System

1. The intelligence of Answering System is not strong. It requires students to have the ability to extract certain key words, human-computer interaction unfriendly.
2. The accuracy of Answering System is not high. For example, in some database based on keywords of Answering System, will often return a lot of problems may be related to the information or may not, students also need to identify themselves from a large number of materials. This is not consistent of the actual needs of students.
3. The answering is difficult to achieve a wide range of knowledge sharing. This is particularly serious in our network education resource's sharing, often different network college have different Answering System, and the different Answering Systems are often strong coupled in different network learning platforms, so that knowledge can not be shared. As a result, most teaching resources are serious overlapped (Studer and Fensel 1998).

Thus, in-depth researching of Answering System in network education, designing a network education Intelligent Answering System which feedback timely, answering intelligence, knowledge sharing and answering accurately, to solve the problems raised in practice, has become an important issue which need to be resolved in network education.

2 The Theory of Ontology

2.1 The Definition of Ontology

"Ontology" is a term from philosophy, it is a science which to study the various entities of the world and how they related. In the area of computer artificial intelligence, the most recognized Definition given is Gruber Ontology in 1993: "Ontology is a clear conceptual model specification" (Copestake and Sparck Jones 1990). On this basis, Studer et al. in 1998, a more detailed definition is given: a clear conceptual model of shared formal specification. He pointed out that the body of the four meanings:

1. Conceptual model: abstract of the objective world by some phenomena are related to the concept of the model, meaning the independent of their specific environmental condition.
2. Explicit: the use of the concept and these concept's constraints have a clear definition.
3. Formal: Ontology is a computer-readable.
4. Sharing: Ontology is reflected in the common recognition of knowledge, reflects the generally accepted concept in the relevant fields set, it is aimed at groups rather than individuals.

In summary, ontology is a kind of knowledge can be formalized, but also a knowledge which can be managed. The character is expression of group knowledge sharing, knowledge reuse, knowledge exchange to support and so on. Ontology is a set of protocols, a language; it also can be a standard, a framework (Huang and Yao 2003).

2.2 The Modeling Primitives of Ontology

Perez, who believes that ontology, can be organized according to classification, and he concludes five basic modeling primitives of ontology.

- Classes or Concepts
  It means any business, such as job description, function, behavior, strategy and reasoning process. Semantically speaking, it represents a collection of objects; the definition commonly used frame structure, including the name of the concept, the collection with other concepts relationship and use of natural language description of the concept.
- Relations
  The interaction between the concepts in the field of ontology is defined as the n-dimensional form a subset of Cartesian product: R: $C1 \times C2 \times \ldots \times Cn$, such as subclass-of. Semantically, relations are corresponding to the set of tuples.
- Functions
  It's a group with special relationship. The first $n-1$ elements of the relationship can be the only decision the n elements. Formally defined as: F: $C1 \times C2 \times \ldots \times Cn-1 \rightarrow Cn$. Such as the Mother-of is a function, mother-of (x, y) that y is x's mother.
- Axioms
  Behalf never wrong, such as the concept B is in the range of concept A.
- Instances
  Represent element. Semantically speaking, instance is the object.

Semantically speaking, there are four kinds of basic relations: part-of express the concept of the relationship between the part and whole; kind-of express of the inheritance relationship between concept, similar to the relationship between object-oriented parent classes and subclasses; instance-of express the relationship between instance and concept, similar to the relationship between object-oriented objects and classes; attribute-of express the concept is a property of another concept, such as "price" is an attribute of the table. In the actual modeling process, the relationship between the concepts are not limited the four basic relationship listed above, we can define the corresponding relationship according to specific circumstances of the field (Huang and Yao 2004).

2.3 Ontology Modeling Tool

Currently, ontology development tools are lot. Corroding to the ontology description language which these tools support by, we can divide these into two categories. The first category includes ontolingua, ontosaurus, Webonto and so on. The second category includes Protégé, Webbed, ontoEdit, oilEd. The more famous is the Protégé, in this study we using this tool to develop a prototype system, it is an open source ontology editor which developed by the Stanford Medical Informatics at Stanford University, and its main features are: By the cross-platform written in java language, and without specific operating system restrictions; It is an extensible knowledge model, the user can redefine the system that used the original language; support the concept hierarchy, concept attributes, and the definition of principles and constraints; support for RDF (S), OIL, XML, and OWL input and output formats; The background support database storage (Neches et al. 1991), you can use JDBC and JDBC-ODBC bridge to access multiple databases (Oracle, MySQL, Microsoft SQL Server, etc.)

2.4 Ontology-Based Knowledge Representation

Intelligent answering system inevitably requires knowledge base, of course, involves the formal representation of knowledge. In recent years, ontology-based knowledge representation has become a hotspot research. In the Ontology-based knowledge representation, people think: any complex knowledge constituted by the most basic concepts, which called ontology; ontology is a detailed description of the basic concepts. The importance of ontology is reflected important role in the knowledge reusability and sharing. Ontology-based knowledge is different from frame or other representation. Ontology expressed the structure of the concept, the relationship between the concepts, which is "shared conceptualization", while other knowledge representation methods such as semantic networks, to express an individual's knowledge of entities in the field, not necessarily the inherent characteristics of the entity. The difference of ontology knowledge representation is an important consideration by we introduced the ontology-based knowledge representation technology in this study. It's ability to meet our needs for knowledge sharing of educational resources.

3 Requirements Analysis of the System Model

3.1 The Design Goals of Intelligent Answering System

The author summarizes the existing network education answering system in the full survey, and based on the related theory technology, I think that the answering system meets contemporary network education should do the following:

1. Answer timely: learning disabilities solved in time can keep learning enthusiastic and enhance their confidence in learning.
2. The share of answering knowledge easy: It avoids duplication of resources and is conducive to a wide range of knowledge sharing.
3. Answer intelligence: the system is more intelligent, so that the learners have better experience. This is also conducive to resolving the status of relative shortage of teacher resources, and liberates them to do more efficient teaching work (Chen 1999).
4. Answer accurate: Integrated variety of answering ways, artificial and systematic way to improve the accuracy and completeness of answering system.

3.2 The Overall Design of the Model

3.2.1 System Model Overview

System knowledge base mainly starts from two aspects, one is the domain ontology, and another is the formal semantics of educational resources. Domain ontology defines a related field (a subject) of the domain model and determines the area of things works. The formal semantics of educational resources is an instance of the field mode. In short, the domain ontology defines a model (Schema), and the semantic educational resources defines the date which in line with the model specification. In view of the current fact standards and level of support of API consideration, the actual paper used in the prototype system is the OWL ontology, RDF resource description framework. The overall operation of the system is shown in Fig. 1.

System design is divided into two parts, one is the asked portal of answering system, the second is answering system management portal. Asked portal include: Part of the traditional manual answering (BBS, Chat room, voice and video answering system), FAQ answering module, intelligent answering module. Management portal including: FAQ management module, ontology management module, semantic knowledge management module, rules management module. Answering portal mainly provides comprehensive answering services, and management portal mainly co-manage and maintain knowledge base through experts and knowledge engineers work together. Two portals, under the use of FAQ database, domain ontology library, and semantic knowledge base, complete the work of online education answering.

The users of Network Education Answering System are students, and the administrators are teachers, experts, knowledge engineers. Under the three modules' common services, students settle questions. Three modules in the system carry out their duties and promote each other (Li Shuang and Chen Li 2003). The FAQ answering module in the system mainly used to answer frequently question, and for the general simple question, it not only ensure the speed of feedback, but also ensure the accuracy of the answer. If the student gets answer, the process is ended. Otherwise, enter the intelligent question answering module, it mainly through the questions
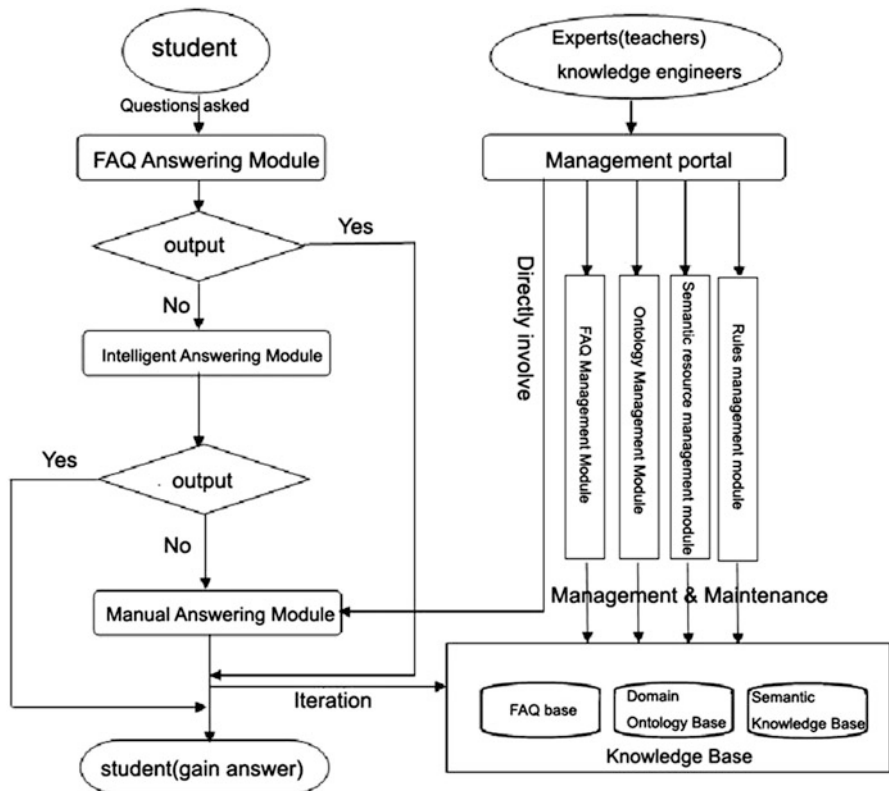
Fig. 1 Overall system flow chart

analysis component, queries produce components, knowledge query components and the interaction of systems persistent domain ontology and semantic knowledge database, from which the extraction of the answer back to the students. Similarly, if there is no satisfactory answer, the questions are entered manually answering module. Manually answering module mainly handle questions which the above two modules can not handle, it mainly rely on teachers (experts) manual answer. While the manual answered, it iteration manage the FQA base and domain knowledge base through the management portal, and continuously enhance the system's answering levels. On the other hand, the corresponding answering system management portal including: FAQ management module, ontology management module, semantic knowledge management module and the rule management module. Four modules are mainly through experts, teachers and knowledge engineers together to protect the knowledge base.

In summary, the structure of the system and process operation which author proposes are clear, convenient and practical, should be able to meet our current request of network educational answering system.
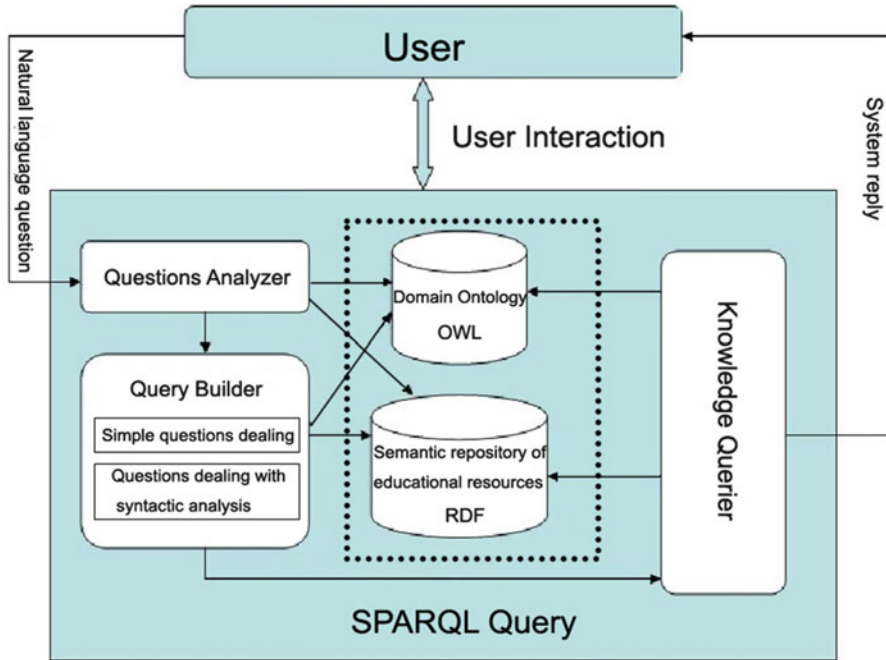
Fig. 2 Intelligent modular system frame

3.2.2 The Detailed Design of Intelligent Question Answering Module

The model design of intelligent in the answering system, the part of the domain ontology and semantic educational resources, through various steps of intelligent answering, played a key supporting role to the operation of the system. The part of the system framework as shown in Fig. 2.

The main processes are as follows:

When a user enters a natural language query, the system through Questions Analyzer, first do tag to natural language query segmentation and part of speech, and then do semantic mapping through the semantic base and the knowledge base element. Finally through the two levels questions module in the Query Builder (simple questions processing module, syntactic analysis processing module), gain the SPARQL query language which are natural language questions converted. Simple questions processing module is responsible for simple, straightforward questions, the effect is obvious. The syntactic analysis processing module is responsible for complex query, and through NLP tools to get the syntax tree information, then in combination with common domain ontology to get complex SPARQL query. Finally, knowledge query grab the user's query goals. To the query problems which the system can not be converted to a SPARQL, the system by way of human-computer interaction to re-determine the query target, and then removed answer from the ontology knowledge base.

3.2.3   Technology Implementation of System Model

- Ontology: Ontology Description Language use Web standard ontology language OWL, we can make use of a large number of application tools to do building ontology, reasoning ontology, validation ontology (e.g: Protégé, OntoEdit, Jena, etc.).
- Pre-questions: word segmentation, lexical, POS tools by the system ICTCLAS by the institute of Chinese Academy of Sciences Computing College, it uses a multi-layer Hidden Markov Model, the model is applied to the atom splitting, respectively, are not simple and complex Log in word recognition and class-based sub-word Hidden Markov and other levels.
- Curriculum knowledge representation: with established ontology vocabularies, RDF triple model used to express the curriculum knowledge, finally, establish a RDF/XML course knowledge base, in order to extract the answer with the Jena semantic inference.

References

Androutsopoulos I, Ritchie GD, Thanishch P (1995) Natural language interfaces to databases-an introduction. Nat Lang Eng 1(1):29–81

Chen Y-L (1999) The modeling analysis and design method of IDFF. Qinghua University Press 1(1):25–26

Copestake A, Sparck Jones K (1990) Natural language interface to database. Knowl Eng Rev 5(4):225–249

Huang Z, Yao H-H (2003) A system for Chinese question answering. IEEE/WIC WI-2003 international conference 10

Huang G-T, Yao H-H (2004) Chinese question-answering system. J Comp Sci Technol 19(4):479–488

Li Shuang, Chen Li (2003) The comparison study of on-line intelligent answering system between China and Abroad. China Audio-V Educ 1(1):80–83

Neches R, Fikes RE, Gruber TR (1991) Enabling Technology for Knowledge Sharing. AIMagazine 12(3):36–56

Studer R, Fensel D (1998) Knowledge engineering principles and methods. Data Knowl Eng 25(122):161–197

Applications Based on the Principal Component Analysis of Informatization Evaluation Model in University Informatization Evaluation

Xiao Danyan and Xi Guangwen

Abstract According to the reality of the Chinese college education informatization development and the relying on the national informatization index system. This paper designs the university informatization evaluation index system and establish mathematical evaluation model by using principal component analysis method. This model quantitatively measure university informatization development level in order to provide necessary information and decision for establishing university education development strategy and policy. Finally, it shows that the university informatization evaluation index system has got a good result when it is used in practice of College Education Informatization Evaluation through questionnaire investigation and expert consultation method.

Keywords Educational informationization • The evaluation indicator system of college • Educational informationization • Principal component analysis

1 Introduction

In today's society, the informationization becomes a macro demand for the present state and social development. The "National education reform and development of long-term planning programs" (2010–2020) have made the education informatization into the national informatization development strategy of whole and clear-cut "accelerating education informatization process forward." Professor Fu DeRong

X. Danyan (✉)
Information Technology Center, Chongqing Normal University, Chongqing, China
e-mail: xdy@cqnu.edu.cn

X. Guangwen
Network Center, Chongqing Three Gorges University, Wanzhou, China
e-mail: irains@vip.qq.com

from Central China Normal University defined: "education informatization means that information will become an element of our education system on the basis of analysis. The information technology will be widely used in all aspects of education, to improve education information environment, cultivation of teachers and students information ability. It will also promote the reform of education and teaching under the informatization environment and will accelerate the realization of education modernization." Thus, educational information has great significance on promoting the comprehensive reform of education and teaching.

After the efforts of the "Ninth Five-Year" "Tenth Five-Year" and "Eleventh Five-Year", the higher educational informatization has made great process. However there are still some shortages such as the inadaptation about infrastructure construction to the requirements of network education, the lack of funding input, the inadequate educational software development, the lack of technical personnel and high quality resource, the imbalance of informatization development and the low level of information standardization. Because the lack of scientific evaluation system and method some universities and colleges can't make effective evaluation about the level of informatizaiton construction. This leads to some high schools can't recognize their informatization construction stage correctively and waste a lot of manpower and material resources and influence the next step informatization construction severely. At present, the researches of evaluation indicator system of higher educational informatization and theoretical methods are very limited. So we must accelerate the development of scientific evaluation indicator system of higher educational (Bonham and Beichner 2004) and set up evaluational technical method which can make quantitative reflection and evaluation about the level and development progress and existed issues about higher educational informatization. Thus we can know the present situation of higher educational informatization more clearly. According to the feedback of evaluation, we can recognize the issues in the development of informatization and establish corresponding measures and impulse to guide the high quality development of high educational informatization.

2   Construct the University Informatization Evaluation Index System

Through the statistical analysis of the university informatization index and quantitatively measurement of regional informatization development degree as well as the statistical rules of the university informatization development perspectives we can improve the scientificalness and accuracy of government's decision-making to promote education informatization construction. It can also effectively guide and promote the informatization construction work.

2.1 The Guidelines and Principles of Evaluation Index System

In order to promote national informatization construction the government establishes a unified national informatization index system (NIEC 2006) to scientifically evaluate the informationization level of different countries and regions in order to guide the information development. The national informatization system contains six factors: the information network, information resource, information technology industry, information talent, information technology application, informatization policies and regulations and standards. The University informatization system is a sub-system of education informatization and a special structural and functional part. Based on the national informatization index system (Yang Jingying et al. 2007) and other academic research results the author (Zhang Chenghong 2006) redesigned the evaluation index system for university. In this process we completed with the relevant state higher education informatization (Roach 2008) policies and Chongqing colleges and universities "Digital Campus" evaluation index system in Chongqing municipal education commission.

2.2 Choose the Evaluation Index System

In Chongqing colleges and universities "digital campus" evaluation index system there are six primary index, 22 secondary indexes and 61 level 3 index. This system is suitable for examination evaluate college. Directly use this system to analyze the schools' informationization level failed to provide scientific judgment due to no clear quantitative index. This paper focuses on reducing qualitative indexes, increasing the quantitative indices, merging secondary indexes, and deleting some indexes which were difficult to quantify or obtain. The university informatization level evaluation index system of this paper is shown in Table 1.

3 The Mathematical Evaluation Model Based on Principal Component Analysis

3.1 The Mathematical Principle and Processing Method of Principal Component Analysis

The principal component analysis (Yu Xiulin and Ren Xuesong 1999) uses dimensionality reduction thoughts to reduce multiple indexes to a few independence comprehensive index. At the same time it makes those indexes to reflect the original index group medium information as much as possible. These integrated indexes can

Table 1 University informatization evaluation index system

Assessment index	Evaluation content
Information organization setup M1	School leader team, operation management institution, resource development and technical support department for informatization construction
Informatization short-and long-term planning M2	Planning time, feasibility, rationality and advanced of informatization
Informatization system construction M3	The system of informatization safety management, students political ideology and propaganda work system on internet, training system of informatization
Informatization funds guarantee M4	Informatization investment funds percentage, informatization funds increase total of this year
Informatization personnel M5	According to 1500–2000:1 ratio on student number, equipped with high-quality digital campus construction specialty management maintenance personnel
Informatization network infrastructure M6	Export bandwidth, wireless coverage, teacher and student of computer internet connectivity rates, teacher and student computer Ratio, classroom and multi-media classroom, the quantity of server, network failure rate
Network management application software M7	Good network teaching platform, network management system, server management, perfect MIS software, library management software, digital learning environment (Wacher and Gupta 1997)
Informatization safety construction M8	Provide network security and service
High-performance computing equipment or server cluster M9	High-performance computing equipment or server cluster of internet connectivity can satisfy large computing needs
Library's digital resource construction M10	More than 20TB literature database, retrieval own literature database and outside the library information of internet connectivity, access to domestic and foreign library collection material, library integrated management system, university library management system in the same city can interconnect each other
Teaching source base M11	Multimedia material database, Subject teaching database, web course database (Chia-I Chang 2002)
Management database construction M12	Student enrollment and performance management, teachers and teaching situation system, equipment and school property system, personnel and industrial relations system, financial system, researching system, office automation database
Informatization education application level M13	Web portals, "one-card" system to cover campus management, teaching, life, scientific research, etc. Scientific research personnel information technology application method
Education and literacy training M14	To cultivate students' information literacy and information literacy for the purpose of informatization public courses (The student guide to evaluating information technology on campus), students political ideology and propaganda work construction on internet
Electronic school education and management informationization level M15	Office automation management system ratio to 90%, administrators can skillfully use business management information system for handling business

better reflect the differences between each sample, and they are independent of each other in statistical sense. The model realizes the principal component analysis as follow:

1. Collect primitive data from matrix A according to the index system, X_{ij} is the values of university i and informatization index

$$A = \{X_{ij}(i = 1 \dots n; j = 1 \dots m), \ (n = 5, m = 15)\} \tag{1}$$

2. First, normalize the primitive data matrix and get Correlation Matrix R. Since the evaluation index is composed by many indexes we need to normalize the data to non-dimensional data for avoiding the influence of dimensional and orders of magnitude.
3. Compute the matrix R of eigenvalue λ and eigenvector β.
4. Compute the cumulative contribution of variance, choose the principal components number P and then compute the corresponding eigenvector.
5. Make the comparative analysis by computing the principal component score of each university.
6. According to the principal component score, get the total principal component index F of each university.

3.2 Use the Evaluation Model to Analyze Data that be Collected by According the University Informatization Evaluation Index System

In 2009, based on Chongqing statistical yearbook, by questionnaire investigation (Kobulnicky and Rudy 2002) and expert consultation method, the author obtained education informatization and corresponding indicators of informatization data from Chongqing university, Chongqing university of posts and telecommunications, Chongqing normal university, Chongqing medical university, Chongqing technology and business university. By using principal component analysis to analyze the data obtain in every school's main indexes of informatization at different development situation we give the quantitative objective evaluation and make improvement and suggestions in the insufficient places.

This paper adopts the objective way that most scholars used to determine each index of concrete numerical value cause university informatization evaluation indexes contain both quantitative indices and contains qualitative indexes. We selected index variables such as M1- M15 and process the primitive data according to the above steps using SPSS17.0 statistics software. Specific steps (Ganesalingam and Kumar 2001) are:

1. Normalize the 15 primitive index data and get Correlation Matrix R.
2. Compute the matrix R of eigenvalue λ and the cumulative contribution of variance as shown in Table 2.

Table 2 Total variance explained

Component	eigenvalue λ	% of variance	Cumulative %
1	3.685	19.25	19.25
2	3.142	24.53	43.78
3	2.538	17.15	60.93
4	2.265	11.98	72.91
5	1.956	12.13	85.04

X1-X15 is M1-M15 normalization

3. From Table 2, the first five the principal components of the cumulative contribution rate was 85.04%. So we can use the change of these five components after dimension-reduction to express the change of the primitive indexes. At this moment it cannot get the principal component expression yet. We need to Compute the matrix R of eigenvector,

$$\beta : \beta_{ij} = Y_{ij}/SQR(\lambda_i) \quad (i, j = 1, 2, \dots 15). \tag{2}$$

Here Y_{ij} is the load quantity of the principal component matrix and λ_i is the eigenvalues of the principal components F_i.

4. The first five principal components can be expressed as a linear combination like:

$$F_i = \sum_{i=1,j=1}^{15} \beta_{ij} \times X_{ij} \tag{3}$$

X_{ij} is the index's normalization. Compute the principal component score of each university according to the data had be calculated. Then Compare and analyze through the score. We finally get the total principal component index F of each university according F_i.

$$F = \sum_{i=1}^{n} A_i \times F_i \ (i = 1, 2, \dots 15) \left(A_i = \frac{\lambda_i}{\sum\limits_{i=1}^{p} \lambda_i} \right) \tag{4}$$

5. Through our calculation we get principal components and total principal component index. Then we sort all five university informatization data in Chongqing.

Through the empirical analysis we can see that Chongqing technology and business university and Chongqing university belong to advanced level in education informatization from Table 3. Chongqing university of posts and telecommunications and Chongqing normal university belong to the medium level in those five universities. It shows that each university's informatization network infrastructure are developed relatively perfect and balance. However there are certain gaps in information organization setup, high-performance computing equipment or server cluster, informatization education application level, informatization basic database and resources and network management application software. In the future we will focus on building these targets in the development of informatization.

Table 3 The university in Chongqing informatization evaluation index sort table in 2009

School	F1	Sort	F2	Sort	F3	Sort	F4	Sort	F5	Sort	F	Sort
Chongqing Technology and Business University	5.15	1	3.975	1	4.231	1	3.659	1	3.321	1	4.0676	1
Chongqing University	5.09	2	3.753	3	4.202	2	3.562	2	3.198	2	3.9608	2
Chongqing University of Posts and Telecommunications	4.93	3	3.786	2	4.009	3	2.892	3	3.096	3	3.7426	3
Chongqing Normal University	4.86	4	3.694	4	3.986	4	2.642	4	2.984	4	3.6332	4
Chongqing Medical University	3.94	5	3.682	5	3.897	5	2.599	5	2.962	5	3.4161	5

4 Conclusion

Through the empirical analysis and by using the quantitative analysis method to calculate the development level of university informatization we can reveal different areas or different universities informatization development situation. We can also the university informatization level at the same region in different periods of development in order to provide necessary information and decision for establishing university education development strategy and policy (Soudien).

5 Prospect

There is a certain limitation of this university informatization index evaluation system because it was established by using principal component analysis according to Chongqing university informatization constructions. In the future work, we will optimize it to make the university informatization index evaluation system (Li Gang et al. 2008) face the nationwide usage. Our university informatization construction are concentrating developing. Since different regions and different schools have different levels of development the university informatization construction can be divided into different grades by clustering methodology. It will help them to judge what type and which level they belong to in the university informatization development. After find a reasonable position it will help people make a better informatization development plan.

References

Bonham SW, Beichner RJ (2004) Education riser using web-based assessment systems. J Reser Comput Educ 33(1):101–104

Chia-I Chang Flor (2002) Intelligent assessment of distance learning. Inf Sci 140:1–2

Ganesalingam S, Kumar K (2001) Detection of financial distress via multivariate statistical analysis. Manage Finance 27(4):45–56

Kobulnicky P, Rudy JA (2002) The third annual EDUCAUSE survey identifies current IT issues. EDUCAUSE Quarterly 25(2):8–21

Li Gang, Chen Ping, Qi Er-shi (2008) Design of index system for comprehensive evaluation & its application in enterprise informationization. Comp Integr Manuf Syst 14(1, January): 96-101-104

NIEC (2006) The structural scheme and description of nation informatization index. http://www. niec.org.cn/gjxxh/zbfa01.htm

Roach R (2008) Meeting new challenges: diverse issues in higher education 25:120–126

Soudien C. Education in the network age: globalization, development and the World Bank. Int J Educ Dev 22:35–40

The student guide to evaluating information technology on campus (developed by EDUCAUSE in cooperation with the American Association of Collegiate Registrars and Admissions Officers (AACRAO) and the National Association for College Admission Counseling (NACAC)). http:// www.educause.edu/studentguide:23

Wacher RM, Gupta JND (1997) Distance education and the use of computers as instructional tools for systems development projects: a case study of the construction of expert systems. Comput Educ 29(1, August):13

Yang Jingying, Lv Haiqi, Yang Hongjun (2007) Measure informatization development index. China Stat 2(February):15–18

Yu Xiulin, Ren Xuesong (1999) Multivariate statistical analysis. M. Beijing, China Statistics publishing house

Zhang Chenghong (2006) Plan and evaluate the university informatization. M. Fudan University publishing house, pp 163–164

An Application of Classification Models in Credit Risk Analysis

Ruan Ling-ying

Abstract A default risk is defined as the possibility that a borrower will not be able to pay back the principle or interest associated with a lending. Credit card business has high risk of delinquency as there is no collateral required before borrowing the money. Lenders usually collect a lot of information to learn the consumer risks. A conventional method to this problem is to examine combinations of the information variables that are likely to have influence. However, hunch can leave out important variables without being noticed. In this article, we introduce statistical models to conveniently predict the default risk based on an application to a real data of credit card business. Several potential improvements are also discussed.

Keywords Credit risk • Classification models • Logistic regression • Boosting • Random forests

1 Introduction

Credit risk analysis (finance risk analysis, loan default risk analysis) and credit risk management is important to financial institutions which provide loans to businesses and individuals. Credit loans and finances have risk of being defaulted. A default risk is defined as the possibility that a borrower will not be able to pay back the principle or interest. The biggest default risk involves unsecured lines of credit, such as credit cards. With an unsecured line of credit, it may be impossible for the lender to get back much of the investment, in case of a default. Credit cards and other unsecured lines of credit therefore carry a bigger risk when a default does occur. For this reason, credit providers must understand risk levels of credit.

R. Ling-ying (✉)
Chongqing Three Gorges University, Wanzhou, Chongqing, China
e-mail: ruanlingying@163.com

Personal credit scores are normally the most critical information lenders need to obtain. They are provided by external credit bureaus and ratings agencies. Credit scores may indicate personal financial history and current situation. However, it does not tell you exactly what constitutes a "good" score from a "bad" score. More specifically, it does not tell you the level of risk for the lending you may be considering.

Except credit scores, credit providers often collect a vast amount of information on credit users. So credit risk profiling (finance risk profiling) is very important. The Pareto principle suggests that 80–90% of the credit defaults may come from 10 to 20% of the lending segments. Profiling the segments can reveal useful information for credit risk management. Information on credit users (or borrowers) often consists of dozens or even hundreds of variables, involving both categorical and numerical data with noisy information. Profiling is to identify factors or variables that best summarize the segments. Analyzing such vast information is an extremely difficult and challenging task. As the total number of variables increases, the number of combinations to be examined in this way grows exponentially. When a large number of variables involved, the number of combinations is too large to be examined manually. Thorough systematic accurate analysis is all but impossible.

Fortunately, this problem can be overcome with the methodology described here. Predictive modeling is an excellent technique for credit risk management. Predictive models are developed from past historical records of credit and consumer behavior. From the past information, predictive models can learn patterns of different credit default ratios, and can be used to predict risk levels of future credit loans. In this paper, we present several methods of predictive modeling and illustrate their performance based on a real data task. In Sect. 2, we briefly introduce the models and we apply them in the real task of credit card default risk classification in Sect. 3. And we conclude the findings in Sect. 4.

2 Methodology

Logistic regression (Hastie et al. 2009) is a popular linear model in classification to model the posterior probabilities of the K classes based on linear functions in the features x. The model has the form

$$\log \frac{P(G = j | X = x)}{P(G = K | X = x)} = \beta_{j0} + \beta_j^T x, \quad j = 1, \cdots, K - 1 \tag{1}$$

It is specified in K-1 log-odds or logit transformations to satisfy the constraint that the probabilities sum to one. The choice of the denominator is arbitrary and the estimates are equivalent under the choice. Here the K-th class is chosen as the baseline. It is solved by iteratively reweighted least square (IRLS). For typical two-class problems, it is simplified and the posterior probabilities have the form

$$P\left(G = 1 | X = x\right) = \frac{\exp\left(\beta_{10} + \beta_1^T x\right)}{1 + \exp\left(\beta_{10} + \beta_1^T x\right)},$$

$$P\left(G = 2 | X = x\right) = \frac{1}{1 + \exp\left(\beta_{10} + \beta_1^T x\right)} \tag{2}$$

Random forests (Breiman 2001) select a subset of variables for each node and choosing one best split from the subset to build one tree based on bootstrap samples (Efron 1979). Many trees are grown iteratively and it returns the probability by voting. One promising property of random forest is that it can handle high dimensional data and it does not over-fit. It is robust to outliers and it's easy to implement because there are not many parameters to be tuned. The only tuning parameters for random forest are the total number of trees to grow and the size of subset to split at each node. However, random forest is not sensitive to the size of subset in a wide range actually. Several choices of the parameter is tried to get the best random forest model.

Boosting was first proposed in Schapirer (1990) and was received much attention since then (Freund 1995; Freund and Schapirer 1997; Schapire et al. 1998; Friedman et al. 2000). The idea of boosting is to generate weak learners iteratively and combine them as a committee to do the classification. Each weak learner is trained sequentially to perform better than random guess. The iteration goes with down weighting the correctly classified samples and increasing the weight for those misclassified samples to get the next weak learner. Each weak learner has a weight to vote which depends on its performance, i.e., the less error rate, the more weight. Classification is the weighted vote of these weak learners. There are many boosting algorithms. AdaBoostM1 is most commonly used. In this study, we tried different boosting algorithms including AdaBoostM1, LogitBoost and BlackBoost. Also, it is flexible to choose the type of weak learner. Decision stump is used as the base learner when using AdaBoostM1 and LogitBoost. Decision tree is the base learner when using BlackBoost. See Schapire et al. (1998) and Friedman et al. (2000) for more discussion of boosting.

3 Application

Credit card provides provide as many variables as possible to understand the risk. Starting with the raw variables, we discussed several ways to extract features (Ruan 2010). With all the possible features ready, stepwise logistic regression chooses ten variables. Compared with the logistic regression model with the raw variables only, i.e., out feature extraction, the AUC of the new logistic regression is lifted by 10.1% on in-sample validation and 9.7% on out-of-time validation. The kolmogorov smirnov (KS) statistic is lifted by 12.3% on in-sample validation and 11.8% on out-of-time validation. Figure 1a is the ROC on the out-of-time validation data set.
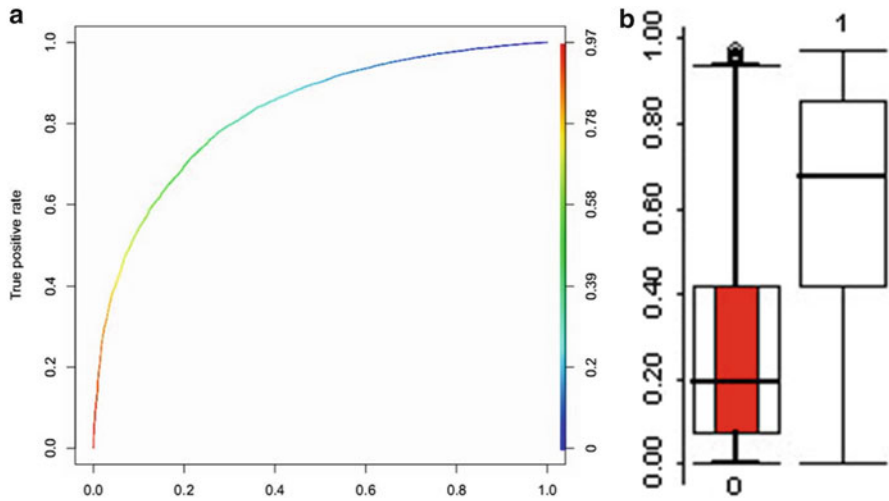
Fig. 1 (**a**) ROC on the out-of-time validation data set with AUC 92.6%; (**b**) *box plot* for the probability of good among the two classes on out-of-time validation, where 0 (*red*) represents bad and 1 (*white*) represents good

And Fig. 1b is the box plot for the probability of good among the two classes on out-of-time validation, where 0 (red) represents bad and 1 (white) represents good.

In random forest, all variables are first put into the model with subset size of 14 and 500 trees, which is decided by cross validation. By bagging (Breiman 1996), the importance of each variable available and is assessed by mean decrease in accuracy and mean decrease in node impurity. The previous one measures how one variable helps others and the latter one measures the individual contribution it has. To get a sparse model, we select part of the variables according to the importance order. For example, the geometric mean of the two metrics can be used as a criterion to select variables. Figure 2 is the importance picture.

For boosting, 500 iterations of AdaBoostM1 returns the best AdaBoostM1 model and 110 iterations of LogitBoost is the best one. To look into the importance of each variable, we check the weight of each decision stump. Table 1 lists the first ten weak classifiers in AdaBoostM1 and the corresponding weight and error rate. The higher the weight, the more important the variable is. And the more times one variable is selected as the split variable, the more important it is. For example, U_alpha is firstly selected as the split variable and the corresponding error rate and weight is 0.2809 and 0.94 respectively. U_alpha is the intercept from the regression, which means it is similar and highly correlated to current utilization. It is more than current utilization because it contains the information with the current utilization and the utilization trend over time. The performance of the two types of Boosting is similar to logistic regression.
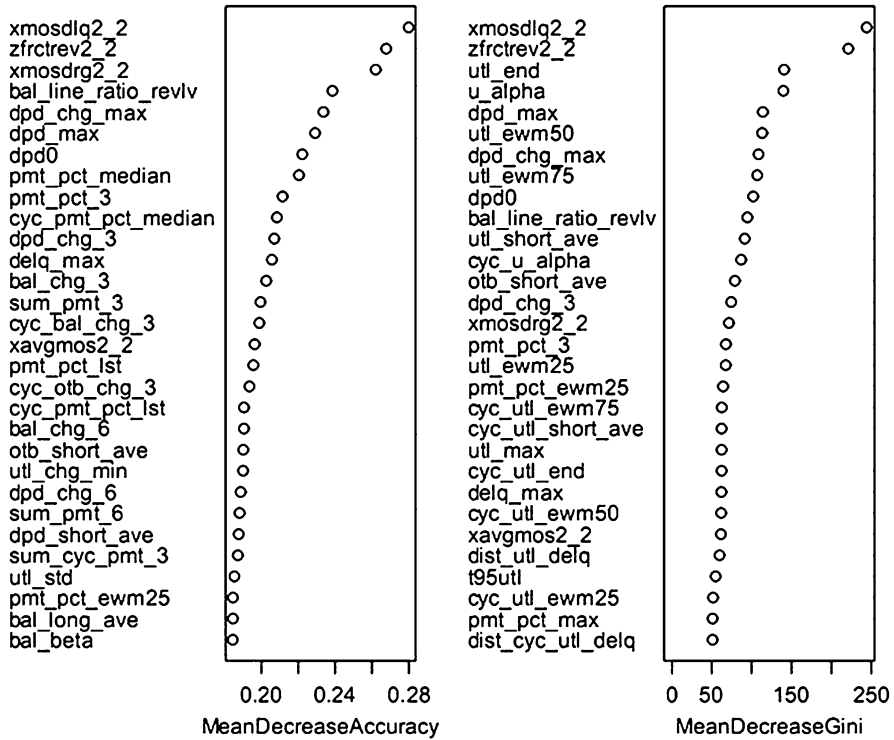
Fig. 2 Importance measures of variables in random forests. The *left one* is mean decrease in accuracy and the *firth one* is the mean decrease in node impurity

Table 1 The first ten weak classifiers in ADABOOSTM1

	Variable	Split	<=Split	>Split	Weight	Error rate
1	U_alpha	0.8232	Good	Bad	0.94	0.2809
2	Xmosdlq	1.5	Bad	Good	0.9	0.28905
3	Cyc_pmt_pct_3	0.114	Bad	Good	0.45	0.389361
4	zfrctrev	58.5	Good	Bad	0.29	0.428004
5	Dep_tot_am	832.74	Bad	Good	0.37	0.408541
6	U_alpha	0.8215	Good	Bad	0.29	0.428004
7	Cyc_ca_ewm25	7.9248	Good	Bad	0.2	0.450166
8	xavgmos	65.5	Bad	Good	0.23	0.442752
9	xmosdrg	8.5	Bad	Good	0.21	0.447692
10	Dpd_short_ave	2.166	Good	Bad	0.15	0.46257

4   Conclusion

This paper has empirically investigated the application of three classification models in the application of credit risk analysis. The data is based on the feature extraction illustrated in Ruan 2010. The Logistic regression shows better separation than the old one without feature extraction, which shows great advantage of doing feature extraction. The logistic regression model is a simple parsimonious model and it is easy to interpret. Random forests and boosting are suitable to handle a large set of variables.

References

Breiman L (1996) Bagging predictors. Mach Learn 24:123–140
Breiman L (2001) Random forests. Mach Learn 45:5–32
Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7(1):1–26
Freund Y (1995) Boosting a weak learning algorithm by majority. Inf Comput 121:256–285
Freund Y, Schapirer E (1997) A decision-theoretic generalization of online learning and an application to boosting. J Comput Syst Sci 55:119–139
Friedman J, Hastie T, Tibshirani R (2000) Additive logistic regression: a statistical view of boosting (with discussion). Ann Stat 28:337–407
Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference, and prediction. Springer, New York
Ruan L (2010) An empirical study of feature extraction in the analysis of credit card risk. Technical report. Chongqing Three Gorges University
Schapire R, Freund Y, Bartlett P, Lee WS (1998) Boosting the margin: a new explanation for the effectiveness of voting methods. Ann Stat 26(5):1651–1686
Schapirer E (1990) The strength of weak learn ability. Mach Learn 5:197–227

Exploration of the Examination Reform in Advanced Language Program Design Course

Xiaoling Tan and Xing Zou

Abstract There are still many problems in the content and methods of current college course examination, thus, quality education to be the center of the test has been put forward. On the basis of combining computer-based courses with teaching experience and analyzing the Advanced Language Program Design curriculum features, a special implementation program of the course examination reform is designed, which has received good results.

Keywords Course examinations • Examination reform • Talent training • Practical ability

1 Introduction

Advanced Language Program Design Course is not only the core curriculum of computer science major, but also one of the fundamental courses in many science and engineering majors for the undergraduate and junior college students. It is of significance to foster the students' abilities to deal with the practical problems by using computers and to understand the program design ideas. To strengthen the construction of the course, some necessary researches and improvements of test program reform have to be done in addition to the teaching content, teaching methods and teaching tools, so that the students' learning motivation and initiative can be promoted, and the quality of teaching can be improved through a flexible examination and evaluation mechanism.

X. Tan (✉) • X. Zou
School of Electronic and Information Engineering, Chongqing Three Gorges University, Chongqing, China
e-mail: fenglingcao86@163.com; zouxin1981@126.com

2 The Present Problems Existed in Teaching and Exam

2.1 Test Mode Is Single

In the assessment of Advanced Language Program Design Course, the students' theoretical knowledge as well as the doing ability should be both examined. The current testing method, which is based on a written examination, is single. This test has primarily examined the basics of program design, but the practice-and applicable- topics account for less weight. It is not profound and comprehensive. It is also less conductive to arouse the students' study enthusiasm and to improve their hands-on operational capability.

2.2 The Test Content Focus More on Knowledge than Capability

Studies have shown that learners' learning purposes and test methods are much related to each other. Scientific and rational test mode is aimed at enabling learners to truly understand the learning content; Unreasonable test mode may lead to the learners' passive reception of knowledge, and the mechanical reappearance of parts of learning content in the examination. The current test content is mainly based on knowledge and is too attached to material. In such examination, the memory assessment is in a primary position, and the operation ability, analytical ability and critical thinking ability are less important. It will cause an inertial thought for the students in this test mode. So this kind of examination cannot measure the students' true quality, and also it hoodwinks their own concerns on capacity and development, and limits the student's growth of creativity and personality as well (Chen Li-xin et al. 2007).

2.3 The Exam Results Is More Important than the Processs

The final exam is the chief test mode, there are little other exams in the teaching process. The Students' achievement cannot be well reflected. The students would not be engaged in review until the final exams come. The purpose of their learning is just to get high scores in the exams but not to improve their abilities, which causes a sudden review and rote learning. However, they would forget most of the knowledge they got from preparing the exams soon. This greatly prevents the students' intellectual development, and is not conducive to the cultivation of their innovation ability.

Course Examination is like the inspection process of product quality in an enterprise. The products must undergo a rigorous process monitoring before get out of the factory. Examination is true of this; teachers should strengthen the monitoring in the students' learning process and know the students' learning state and learning effects. The traditional one-time test mode should be replaced by getting the test run through the whole process of teaching. So that the test can truly play its evaluation and guidance function (Xu Yuqiong 2008).

2.4   There Exists Test Paper Management Problems

At first, because of the restraints of manpower and space, the test paper cannot be conserved intensively and long-timely, and they are easy dispersed and easy to damage. All these will lead much inconvenience in the later reference, statistics, checking and so on. Then the management would become even difficult after the increase of the students' numbers. Meanwhile, the printing of examination papers will also consume a large amount of paper, which is neither conductive to economics nor environmental protection.

3   Taking Effective Measures to Promote the Examination Reform

The basic principle of selecting test ways is that this measure can be able to promote students' learning motivation and initiative, can be able to collaborate with teachers' high-quality teaching task, and can be beneficial to the cultivation of students' innovative ability (Craig Larman 2006).

As a basic course for computer science major, Advanced Language Program Design Course is very practical. Through a practical learning of this course, the students can strengthen their in-depth understanding about syntax, grammar and semantics, and cultivate their programming ability. Assessment program is divided into four parts: (1) Normal operations, accounting for 10%, mainly to urge the students to learn in daily time; (2) Experiment and reports, accounted for 20%, aims at strengthening the students' practical ability; (3) Procedural examination, accounting for 20%, to consolidate their learning outcomes; (4) final exam, accounting for 50%, mainly to study the basic theory of knowledge and the comprehensive examinations of the term. The main idea of the program is to highlight the process assessment with the supplement of the theoretical examination. The assessment program will to some extent limit some students who are weak in learning desire, so that they would not be too lax in the learning process, and play a certain role in the requirements of the students' attending class. The specific implementations must be done to check

whether the examination plan is reasonable and whether it can reach the initial supposition. In order to make sure that the examine plan is taken effectively, the following two keys should be emphasized (Chen Li-xin et al. 2006; Ramez Elmasri et al. 2005).

3.1 Strengthening the Procedural Examination

Strengthening the procedural examination is to test after each stage of a teaching mission is completed in order to consolidate the related knowledge learned in daily time, which requires students to do more in daily time, so that the students' reviewing teaching content becomes normal. There is another effect is that the students who try to pass the examination in one time through cheating will not succeed. So it is beneficial for correcting their study attitudes.

3.2 Request the Students Strictly in the Experiment, and Standardized the Reports After the Experiment

Experiments and reports, accounting for 20%, clearly direct the students to pay high attention to each experiment. In general, teachers require the students to prepare related experimental content in advance so that students can focus on debugging and analyzing algorithms and procedures, and discuss difficulties in the practice of experimental subjects with teachers. Students should not only prepare for the experiment and complete it, but also finish the experimental report.

In this way, the students' writing skills and analytical capabilities are sufficiently trained, which is helpful to foster a rigorous research attitude.

The procedural examination increases the teachers' workload, but the traditional written exam, which needs questioning, making test papers and evaluating papers, statistics of scores and a series of efforts, increases the workload of educational administration. In order to ensure that the procedural examinations and final examinations are carried out smoothly, a research of Web-based online examination system have been undergoing since 2008, and a network examination system which suit the actual situation in our school has also been designed. Students can login to participate in the examination at an appointed time and in a designated examination room through the campus network. That will achieve a good result in actual use of the system, and it will improve the automation and science of the examination process, which can provide better services for the teaching management in our school (MacDonald 2005; Lu Jingping and He Yulin 2003).

4   The Network Test System Developed by Us Has the Following Advantage

4.1   Ensuring the Fairness in the Test

Due to the randomness of pumping problems, we should ensure the same questions in different examination papers with different order. That will reduce the possibility of cheating and ensure the fairness in examination while also be helpful to maintain the examination rules.

4.2   Saving Manpower and Resources, and Be Conducive to Examination Administration

Adopting unified management by computer will save a great deal of manpower and resources. Selecting papers is fully automated. Examinations are carried out on the computers, which need no printing papers. To a large extent, that simplifies the examination process, and it has not only saved manpower and resources and reduced the cost, but also is more conducive to exam management and environmental protection.

4.3   It Is More Convenient Evaluate and Conserve Papers

The testing papers will be evaluated and statistics automatically by the computer, which will reduce the teachers' work. And after being conserved by the computer, the statistics and inquiry of scores are also very convenient, so that the work efficiency can be improved.

5   The Practical Results

Through the examination reform of this course, it has pre-achieved the following results: First, it has added some stage assessment in the examination, guiding the students to value their usual learning, reducing their psychological pressure in the final examination, and mobilizing and improving students' learning motivation. That will enable students to pay high attention to every aspect of teaching and improve learning efficiency, which has corrected some students' adverse learning attitudes of wearying of studying. Second, the assessment of practice-oriented

operational capability is paid attention to, and their level of applied skills is improved by enhancing practical links. Third, through reforming the more reasonable curriculum assessment and evaluation programs, the process of examination will be more standardized and scientific. To give prominence to the objectives of cultivating applied talents who have social responsibilities by combining assessment capacity with evaluation capacity, in order to maintain the main role of the examination (Wang Zhi-yue and Yang Hai-ming 2008). Fourth, the network test system will become a strong complement and development to the current test methods. Network online test has not only reduced manpower and resources, improved the efficiency of scoring and reporting, but also improved the objectivity and impartiality of test scores, speeded up the construction of teaching information technology in colleges (Ningrong Liu 2008; http://tecenter.buaa.edu.cn/kechen/gaige/9.html).

References

Chen Li-xin et al (2007) Exploration and practice on curriculum test reform in colleges and universities. J Tonghua Teach Coll 28(2):94–96

Craig Larman (2006) Applying UML and patterns: an introduction to object-oriented analysis and design and iterative development, 3rd edn. Prentice Hall PTR, Upper Saddle River

Lu Jingping, He Yulin (2003) A web product data management system based on Simple Object Access Protocol. J Chongqing Univ (Eng Ed) 2(1):57–61

MacDonald M (2005) The book of visual basic 2005 – NET insight for classic VB developers. No Starch Press Inc., San Francisco

Ningrong Liu (2008) Restructuring China's adult higher education: an examination of the driving forces behind the reform. Res Post-Compul Educ 13(1):107–121, ISSN: 1359–6748

Ramez Elmasri, Qing Li, Jack Fu, Yu-Chi Wu, Babak Hojabri, Swathi Ande (2005) Conceptual modeling for customized XML schemas. Data Knowl Eng 54(1):57–76

Shyh-Kwei Chen, Ming-Ling Lo, Kun-Lung Wu, Jih-Shyr Yih, Colleen Viehrig (2006) A practical approach to extracting DTD-conforming XML documents from heterogeneous data sources. Inf Sci 176(7):820–844

Wang Zhi-yue, Yang Hai-ming (2008) Exploring the reform of specialization course examinations. J Yangzhou Univ (Higher Education Study Edition) 12(6):84–86

Xu Yuqiong (2008) On the examination reform in application university courses. Cont Educ Res 3:123–125

A Recommendation Method in E-Commerce Based on Product Taxonomy Graph

Qian Liu, Hongzhi Wang, Hong Gao, Qi Lv, and Jianyu Fu

Abstract The data of e-commerce is growing at a rapid speed. As a result, customers are no longer able to achieve what they want to buy in a relatively short time. Collaborative Filtering (CF) is the most acceptable method about recommendation. However it has two limitations. One is sparsity, the other is scalability. In this paper we give a methodology to solve the problems based on product taxonomy graph. Data mining on product taxonomy graph helps make the transaction data in more aggregated way which is expected to solve the sparsity and scalability problem in CF.

Keywords CF • Product taxonomy graph • Recommendation • Top-k

1 Introduction

With the rapid growth in size and number of databases in e-commerce, it is difficult for customers to discover what to buy. The first reason is that overloading products leads to the difficulty for customers to choose the products they are interested in. The second reason is that recommendation methods do not work efficiently as expected. Therefore it is crucial to improve the quality of recommendation methods. Collaborative Filtering (CF) is the most successful technique (Basu et al. 1998; Terveen et al. 1997; Adomavicius and Tuzhilin 2001), and is widely used in many ways like recommending movies, article, products, etc. However, there remains important research questions in overcoming two fundamental challenges for existing collaborative filtering recommendation systems.

Q. Liu (✉) • H. Wang • H. Gao • Q. Lv • J. Fu

School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
e-mail: prideliuqq@gmail.com; wangzh@hit.edu.cn; honggao@hit.edu.cn; seven0329@gmail.com; 7lemonsoul@gmail.com

The first challenge is sparsity. In large e-commerce sites, there are millions of products. Customers may rate only a very small share of them and a product may be rated by few customers. Therefore it is difficult to make proper recommendation with a lack of information. The second challenge is related to scalability. The CF algorithms may confront problems when facing millions of customers and products. Sometimes the algorithm is time-consuming and works poorly in practice. The methods for solving the problems include dimensionality reduction techniques and model-based approaches (Sarwar et al. 2001). Latent Semantic Index (LSI) (Basu et al. 1998) is a widely used dimensionality reduction technique. Model-based approaches use data mining methods such as Bayesian network, clustering, and association rule mining (Lin et al. 2000). However, the model is expensive to build, and usually focused on specific problems.

CF based on tree-structured product taxonomy has been proposed in (Choa and Kim 2004). However, tree-structured taxonomy has difficulties in giving suitable recommendation for the following two reasons.

- It is common that a product or a class can be sorted into two or more classes. As product taxonomy has too many products like that, if the product taxonomy is presented as a tree, it will consume too much space.
- In a tree, it is more difficult to find obvious relationships among products than in a graph.

To overcome the shortcoming of tree-structured taxonomy, in this paper we use product taxonomy graph to find similar product groups to build customer profiles and search for neighbors in the reduced dimensional space. These help to solve the scalability and sparsity problem of CF.

The contribution of this paper includes:

A taxonomy-graph-based CF method is proposed. To our knowledge, this is the first paper where graph-model taxonomy is used for recommendation. We propose a novel algorithm based on taxonomy graph, and it is efficient in performance and help solve the sparsity and scalability problem in CF.

2 Related Work

2.1 Data Mining of Customers' Behavior

The overall process is generally divided into two main tasks: data preprocessing and pattern discovery. Data cleansing, user identification, session identification and path completion are used in the preprocessing tasks. The pattern discovery tasks contain the discovery of association rules, sequential patterns, usage clusters, page clusters, user classifications or any other pattern discovery method (Mobasher et al. 2000).

2.2 Using Product Taxonomy for Recommendation

In most Web retailers, product taxonomy is available. Traditional product taxonomy is practically represented as a tree that classifies a set of product at a low level into a more general product at a higher level. However, nowadays the product taxonomy must be presented as a graph. In different customers' views, the products can be classified in different ways. And so many different methods of taxonomy may cause a product being involved in different categories. Our recommendation methodology is based on product taxonomy graph.

2.3 Data Mining of Concept Hierarchies

As is mentioned in Han and Fu (1994), the study of dynamic generation and refinement of concept hierarchies shows we can propose some algorithms based on automatic generation of concept hierarchies.

The concept of "Top-down big nodes promotion and bottom-up small nodes merging" can be used to discover the potential information of large database.

Recently, the usage of product taxonomy in data mining has been emphasized by many researchers (Han and Fu 1995, 1999; Berry and Linoff 1997). Han and Fu (1995, 1999) discuss that we can discover many interesting and useful relations by data mining in different levels of product taxonomy. Berry and Linoff (1997) stresses the importance of mining the appropriate level of the product taxonomy.

3 PTG-CF Recommendation Methodology

3.1 Dynamic Adjustment of Product Taxonomy Graph

3.1.1 Definition of Product Taxonomy Graph

Graph $H = (V, E)V = \{0, 1, \ldots, n-1\}$, each node v represents a category cat (v) and each edge (u, v) represents that cat(u) is a sub-class of cat(v). If these is an edge (u, v), v is called u's father. Since one sub-class may belong to multiple classes, in the graph, each node may have more than one father. If a node belongs to more than one father F1, F2, …, Fn, then the information of the node contains the percentage it belongs to each father.

In the adjacency matrix presentation of the graph, assuming all the nodes which represent the product classes are numbered in a way as 1, 2, …, |V|, and then the adjacency matrix of graph H is a matrix $A = (a_{ij})$ is $|V| \times |V|$, it satisfies:

$$a_{ij} = \begin{cases} X \ if \ (i, j) \in E, X \text{ is the propotion that i belongs to j} \\ 0 \ else \end{cases}$$
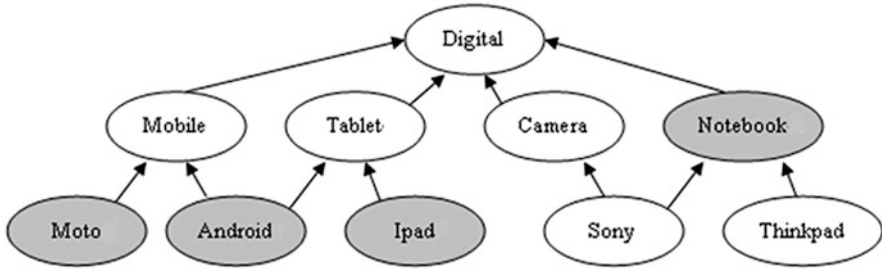
Fig. 1 Level specified product taxonomy

The product taxonomy graph can be transformed into hierarchy graph, and it makes convenience for processing of the algorithm. In the hierarchy, top level is "All", level 1 contains relatively big categories, and the bottom level contains certain product that customers purchase.

3.1.2 Definition of Galaxy and Star

Definition 1 (Galaxy) In a taxonomy graph, a Galaxy is a subset of all the nodes except the meeting point, $G \subseteq V - \{a0\}$, $a0$ is the meeting point.

Definition 2 (Star) Each element in set Galaxy is called Star. For any source point x, its corresponding galaxy product class $class_G(x)$ is defined as follows:

$$class_G(x) = \begin{cases} x & if & x \in G \\ class_G(M(x)), & otherwise \end{cases}$$

M(x) is a set of the categories which x belongs to.

For example, Fig. 1 shows a product taxonomy graph, and then Fig. 1 $G = \{Moto, Android, ipad, Sony, Thinkpad\}$ is a possible result of corresponding galaxy.

Specifying the star at different levels would give different recommendation results, however the proper leveled stars would reflect the transaction data in a more aggregated way and the popularity of products. Hence, the recommendation based on the right star is expected to solve the sparsity and scalability problem in CF.

3.1.3 Algorithms of Finding Galaxy

In this section, the algorithm of building galaxy is presented. As mentioned above, data mining algorithms usually produce the best results when product-related transaction distributions among nodes are even (Han and Fu 1994; Berry and Linoff 1997). They are suitable for making better customer profiles.

In this paper, the algorithm of finding galaxy is based on the idea mentioned in Han and Fu (1994). Algorithm 1 finds the proper leveled stars dynamically.

The pseudo-code of our algorithm is shown in Algorithm GDA (Graph Dynamic Adjustment).

Algorithm 1 (GDA)

Galaxy is the set of stars. T is the number of stars. Use a struct v[] to keep the information of all nodes, including father[], father_num, index_children[] (the name of its children), children_num, count (each node's popularity), mark (classify big or small nodes). Father[] is also a struct which contains index (the name of its father) and the ratio that node v belongs to node[index].

1. **Initialization:** The initialization information of v[i] is presented by scanning the matrix. Buffer is a set used to store nodes being processed. Initial Buffer stores all the source nodes in I. Total is the sum of popularity of all the nodes in Buffer. Set Galaxy initializes as Φ. t = total / T, t is used to make comparison with nodes in Buffer.

2. **Top-down promotion:**

   Call CalculateWeight(0);
   Call ExtendGalaxy;
   if |Galaxy| +|Buffer| <= T {
       move all nodes in Buffer to Galaxy
       output Galaxy;}
   else do refresh{
       T <- T - |Star|, total <- sum of v[i].count in Buffer and t <- 1 / T.
   Call CalculateWeight(1);
   Call ExtendGalaxy;
   until no change in Buffer}

3. **Bottom-up merging:**
   if Buffer is not Φ then{
       for all nodes v[i] in Buffer {
         create father_num new v[];
         delete v[i] in Buffer, add new v[] to Buffer;}
       find v[index].children_index
         do merge two of the nodes into a new node v[m];
             if v[m].weight >= t
         move v[m] to Galaxy;
         else merge another node with m;
         until there's no children_index;}

   if |Galaxy| + |Buffer|<= T {

       move all the nodes in Buffer to Galaxy;
       output Galaxy;}

(continued)

(continued)

---

else {

T <- T - |Galaxy|;
total <- the sum of counts of nodes in Buffer; t <- total / T;
call CalculateWeight(2);}}

output Galaxy.

4. **Two procedures:**
   procedure CalculateWeight(F):
   if F =2 then return;
   else

   compare v[].count to t
   if v[].count >= t, v[].mark <- B;
   else v[].mark <-S;}

   procedure ExtendGalaxy:
   {move all the nodes which v[].mark are B to Galaxy.}

The complexity of Initialization is $O(n^2)$, for the number of nodes is n and the matrix of all the nodes is $n \times n$. The complexity of Top-down promotion and two procedures is O(n). The complexity of Bottom-up merging is $O(n^2)$. This is because dividing each node in Buffer cost O(n), therefore it costs $O(n^2)$ to divide all the nodes. Finding the father of a node in buffer cost O(n), and the father found takes O(1) to find all its children. All in all, the complexity of the algorithm is $O(n^2)$.

3.2 Creating Customer Profiles

In e-commerce, customer profiles are often extracted based on the transaction data. With detail, the preferences of customers can be detected in following ways: click-through, put into collection, place into shopping basket, purchase and comment. These behaviors have different proportion of weight in creating a customer's profile. A customer's profile is defined as the matrix of ratings.

Matrix C_{ij}, $i \in [1, m]$, m is the number of customers. $j \in [1, n]$, n is the number of galaxy product classes. $C_{ij}^1, C_{ij}^2, C_{ij}^3, C_{ij}^4, C_{ij}^5$, respectively means the number of occurrences of customer's click-through, collection, shopping basket, purchase and comment on galaxy product class j.

$$P_1, P_2, P_3, P_4, P_5(P_1 \leq P_2 \leq P_3 \leq P_4 \leq P_5, P_k \in [0,1], \sum_{1 \leq k \leq 5, k \in Z} P_k = 1)$$

means the proportion each behavior has in the weight calculation of a customer, respectively. Following formula is used to compute the preference of user i to product class j. The calculated result *Cij* means the ratings of user i to product class j. $l \in [1, 5]$, means the class of the behavior of customers. Different features of a customer's behavior are involved and the result of the formula varies in different customers.

$$C_{ij} = \sum_{1 \leq l \leq 5} P_{ij}^l \frac{C_{ij}^l - \min_{1 \leq j \leq |G|} C_{ij}^l}{\max_{1 \leq j \leq |G|} C_{ij}^l - \min_{1 \leq j \leq |G|} C_{ij}^l} * 100 \qquad C_{ij} \in [0, 100] \qquad (1)$$

3.3 Neighborhood Formation

CF is used to help a customer to find people with similar interests. The process is called neighborhood formation. Those people are called neighbors.

Correlation is willing to give better result when data are not normalized (Breese et al. 1998). Therefore in PTG-CF Recommendation Methodology, we choose this approach.

The similarity between two customers a and b is measured by calculating the Pearson-r correlation in formula (2) and (3). Previous research (Breese et al. 1998) has shown its superiority in performance over the others.

$$sim(a, b) = corr_{ab} = \frac{\sum_k (C_{ak} - \overline{C}_a)(C_{bk} - \overline{C}_b)}{\sqrt{\sum_k (C_{ak} - \overline{C}_a)^2 \sum_k (C_{bk} - \overline{C}_b)^2}} \qquad (2)$$

$$\overline{C}_a = \frac{\sum_{k=1}^{|G|} C_{ak}}{|G|} \qquad (3)$$

Where C_{ak} and C_{bk} are customer a and b's ratings on certain product class k. $\overline{C}_a$ and $\overline{C}_b$ are customer a and b's average ratings on all product classes, respectively.

3.4 Recommendation

This algorithm is used to calculate top-N most suitable items for a certain user.

Algorithm 2

//using the neighborhood ratings' weighted average to make recommendation
   Totals = {},simSums = {}
   //Prefs[person] is products which have been rated by the customer
   If item $\notin$Prefs[person] or Prefs[person][item] = 0
Totals[item] + = prefs[other][item] * sim
simSums.setdefault(item,0)
simSums[item] + = sim
   //All the products of neighbors are calculated and ranked
   Rankings = [(total/simSums[item], item) for item, total in totals.items()]
   Return Rankings

---

4 Experimental Evaluation

4.1 Data Sets

In the experiments, the transaction data we used are based on a web survey in Harbin Institute of Technology. One hundred and twenty-three students and teachers are involved in the survey. They are asked to provide their shopping behaviors like purchase, comment, collect, etc. in e-commerce web site like taobao.com, amazon.com.

4.1.1 E-Commerce Data

In total, the database contains 5,738 records with 3,873 click-through records, 345 put into collection records, 634 places into shopping basket records, 671 purchase records, and 215 comment records.

4.1.2 Product Data

The product taxonomy contains 1,401 product classes, and the adjacent matrix is 1401×1401.

4.2 Evaluation Metrics

The main objective of this research is to develop an effective and efficient recommendation methodology with better quality and less computation time compared to other methodologies. In this paper, two methods are involved in to evaluate our methodology.

4.2.1 Quality Evaluation Metric

To evaluate top-N recommendation, two metrics are used namely, recall and precision.

N(a) = the number of products in both the test set and the recommendation set
N(b) = the number of products in the test set
N(c) = the number of products in the recommendation set

$$\mathrm{Re}\,call = \frac{N(a)}{N(b)} \qquad \mathrm{Pr}\,ecise = \frac{N(a)}{N(c)}$$

Recall is the portion of the correct recommendation in products in the actual customer purchase list. Precision is the portion of the recommended products belonging to the actual customer purchase list.

These measures are simple to compute and intuitively appealing, but they are in conflict since increasing the size of the recommendation set leads to an increase in recall but at the same time a decrease in precision (Sarwar et al. 2000). Hence, a widely used combination metric called F1 metric (Sarwar et al. 2000) that gives equal weight to both recall and precision was employed for our evaluation. It is computed as follows:

$$F1 = \frac{2 \times recall \times precision}{recall + precison}$$

4.2.2 Performation Evaluation Metric

To evaluate the scalability issue, we used a performance evaluation metric in addition to the quality evaluation metric. The response time was employed to measure the system performance. The response time defines the amount of time required to compute all the recommendations for the training set.

4.2.3 Experiment Environment

The system to perform our experiments was implemented using Apache server. Apache Version: Apache/2.2.17 (Win32) mod_ssl/2.2.17 OpenSSL/0.9.8o PHP/5.3.4 mod_perl/2.0.4 Perl/v5.10.1. The experiment is running on Windows NT 6.1 build 7600 i586 with Intel(R) Core(TM) 2 CPU having a speed of 2.26 GHz and 3.00 GB RAM. The result of our methodology is compared with a benchmark recommender system. Therefore benchmark algorithm is also implemented.
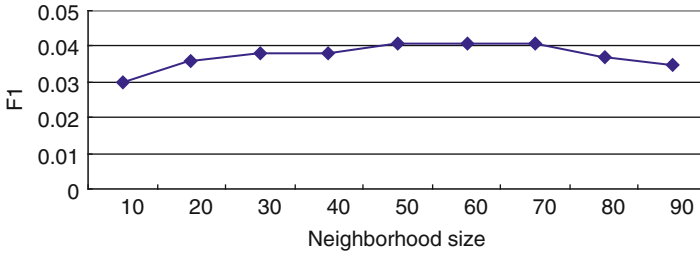
Fig. 2 Impact of neighborhood size

Fig. 3 Impact of different
galaxy specification
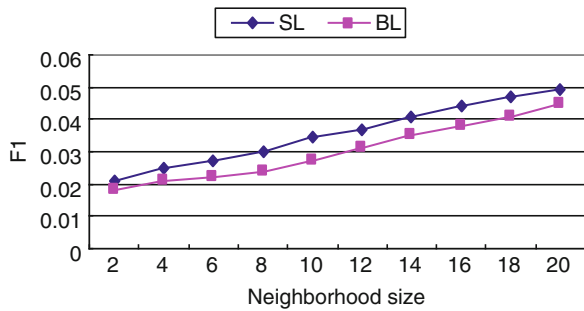
4.3 Experiment Results and Discussions

4.3.1 Impact of Neighborhood Size

To discover the sensitivity of neighborhood size, different neighborhood size is provided to compute the corresponding F1 metric.

The result in Fig. 2 shows that the recommendation quality increases with the increasing of the neighborhood size at first. Then, the quality reaches a peak point. After that, the quality does not grow with the size of neighborhood, for the noise grows with the neighborhood size. This leads to the recommendation quality fluctuate. From the experimental results, the neighbor size has significant effect on recommendation quality.

4.3.2 Impact of Galaxy Specification

As mentioned in above algorithm, the set of galaxy contains product classes in different levels. In the experiment, the quality provided by the product classes in the bottom level (BL) compares with stars in galaxy (SL) in based on F1 metric. In conclusion, according to Fig. 3, PTG-CF methodology provides better quality as the product classes have relatively evenly distributed weight.

Fig. 4 Quality comparison between benchmark CF and PTG-CF
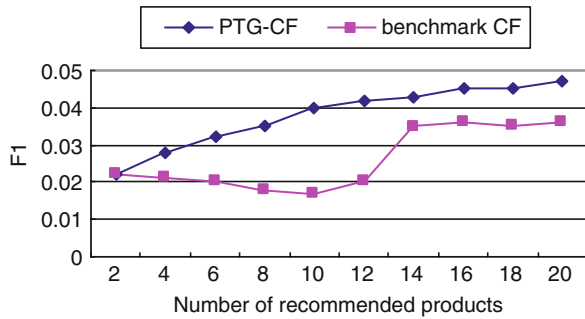
4.3.3 Quality Comparison with the Benchmark CF Algorithm

In Fig. 4, the recommendation quality of PTG-CF is better than the baseline CF method. And the average improvement is 45.8%.

5 Conclusion

The research work presented in this paper makes several contributions to the recommender systems based on CF to solve the sparsity and scalability problem. Firstly, we propose an algorithm using product taxonomy to find product classes. Secondly, the algorithm is based on dynamic adjustment of graph instead of traditional tree. Thirdly, a new method of calculating neighborhood is proposed. The algorithm performs better than benchmark CF method in both quality and efficiency.

References

Adomavicius G, Tuzhilin A (2001) Expert-driven validation of rule-based user models in person-alization applications. Data Min Knowl Disc 5(1–2):33–58, Springer

Basu C, Hirsh H, Cohen W (1998) Recommendation as classification: using social and content-based information in recommendation. In: Proceedings of the 15th national conference on artificial intelligence, pp 714–720, AAAI Press

Berry JA, Linoff G (1997) Data mining techniques: for marketing, sales, and customer support. Wiley, New York

Breese, JS, Heckerman D, Kadie C (1998) Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the 14th conference on uncertainty in artificial intelligence, pp 43–52, AAAI Press

Choa YH, Kim JK (2004) Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce. Expert Syst Appl 26(2):233–246, Elsevier

Han J, Fu Y (1994) Dynamic generation and refinement of concept hierarchies for knowledge discovery in databases. In: Proceedings of the AAAI'94 workshop on knowledge discovery in databases, pp 157–168, AAAI Press

Han J, Fu Y (1995) Discovery of multiple-level association rules from large databases. In: Proceedings of the international conference on very large data bases, pp 420–431, Morgan Kaufmann

Han J, Fu Y (1999) Mining multiple-level association rules in large databases. IEEE Trans Knowl Data Eng 11(5):798–804, IEEE Press

Lin W, Alvarez SA, Ruiz C (2000) Collaborative recommendation via adaptive association rule mining. In: Proceedings of the WEBKDD, ACM Press

Mobasher B, Cooley R, Srivastava J (2000) Automatic personalization based on web usage mining. Commun ACM 43(8):142–151, ACM Press

Sarwar B, Karypis G, Konstan JA, Riedl J (2000) Analysis of recommendation algorithms for e-commerce. In: Proceedings of the ACM E-commerce, pp 158–167, ACM Press

Sarwar B, Karypis G, Konstan JA, Riedl J (2001) Item-based collaborative filtering recommendation algorithm. In: Proceedings of the 10th international world wide web conference, pp 285–295, ACM Press

Terveen L, Hill W, Amento B, McDonald DW, Creter J (1997) PHOAKS: a system for sharing recommendations. Commun ACM 40(3):59–62, ACM Press

From Network Cultural Adaptability Perspective to See Web-Based Learning

Diao Xiuli and Song Zhengguo

Abstract There are many barriers in web-based learning and the factors of its impact at present. Learner-adaptive network culture is the one of its important influence factors. This paper explores the impact of learning on the network from the perspective of learners' network culture adaptability. It bases on the elaborate network of study and research the status and network culture adaptability.

Keywords Network culture • Network culture adaptability • Web-based learning • Learning disability

1 Introduction

With the increasing popularity of computer network technology, web-based learning has become a new important learning style of learners. With "web-based learning" for keyword, there are 448 related papers which are searched in China Knowledge Resource Integrated Database from 2008 to 2009. With "web-based Education" for keyword, there are 562 related papers which are searched in China Knowledge Resource Integrated Database from 2008 to 2009. Thus, from the number of the relevant papers retrieved nearly 2 years to see, web-based learning is still an

---

DIAO Xiuli, femail, (was born in September, 1978 in Liaocheng), lecturer, major: education technology, research direction:online education and digital media technology.

D. Xiuli (✉) • S. Zhengguo
College of Information Science and Engineering, Shandong University of Science and Technology, Shandong Qingdao 266510, China
e-mail: dxlszg@126.com

important research field where domestic scholars have concerned widely. Through to analysis of the retrieval 448 papers for "web-based learning" as keywords, we have found that there are many research papers about "learning disabilities" "emotional loss", "cognitive barriers", "psychological factors", "information trek", and "interactive shortage". Thus, in fact inadaptability and incomplete success which are displayed by distance learners on online learning have showed soft ribs that are restricted network education play its unique efficacy (Zuo Jiming 2003). Among them network culture adaptability of network learners is one of the important crux.

2 Learner Factors in Web-Based Learning

Online learning is the self construction as the center a way of learning. Learner uses all kinds of information resources, communication tools that are offered by computer networks to develop self-identify, self-selected, self-motivation, self-monitoring, online communication cooperation of constituting self-centered on. Its specificity is: firstly, learners can set their own learning objective, study progress and learning strategies on web-based learning, etc. This has reflected learning personalization and autonomy. Secondly, web-based learning is a way of learning which need to resort to the network tools under network culture background. This learning is immersed in network culture. Lastly, web-based learning is also the study of network culture and in this study learners can upgrade their own information literacy. Thus, learners themselves are one of important influence factors of web-based learning effect. At present the researches about network learners mainly focus on the following aspects: the learners' learning disabilities analysis, autonomous learning ability research, about information literacy cultivation of network learners, and the learners' cognitive ability.

From the above concern and research of network learners to see, we have found that the research content and problems of web-based learning disabilities, autonomous learning ability, information literacy and metacognitive ability reflect web-based learning effect of learners themselves. All kinds of inadaptability phenomena of web-based learning which are displayed by distance learners under network background are also inadaptability phenomena of network culture of learners. Learners don't accommodate network cultural environment and lack of network culture adaptability, they are the disadvantages for developing web-based learning. To ensure the healthy development of web-based education and response to the cultural conflict in it, the need is to enhance cultural consciousness of web-based education, improve conformity ability of network culture and promote the popularization of network culture resources (Zuo Jiming 2003). In web-based learning, if learners' network culture adaptability is better, they can develop web-based learning smoothly. Instead, if learners' network cultural adaptability is weaker, web-based learning will goes badly.

3 Network Culture and Network Culture Adaptability

Culture is based on the practice, and what kind of practice mode will produce what kind of culture. At present, computer network are forming and shaping a new virtual practice mode, and a new cultural adapting to the practical form – the network culture has formed.

3.1 The Definition of Network Culture

About the concept of the network culture, academic circles have not unified recognition. The view, content and its connotation emphasis of network culture research are different, but we can find in the definition of network culture: network culture is a subculture formed under the network environment; Network culture is the culture took the network as the carrier; Network culture is a kind of activities (act) way; Network culture includes content of mentality, morals, values and knowledge and so on. Thus, network culture through a variety of activities and carriers influences person's knowledge structure and imagination and emotion, will, etc., which affects people ideological of thinking process, and, eventually, to change the world outlook, the outlook on life, the values and overall quality. Under the network culture environment, only to adapt to network culture atmosphere, can learners study, work and life better in this environment.

3.2 Network Culture Adaptability

The emergence of any kind of culture, people always have an adjustment process firstly, then practice and survive in this cultural background. The cultural adaptability process involving the changes of individual psychology and the behavior ability is the comprehensive process including many factors such as cognition, emotion, behavior and etc. Cognitive adaptability is reflected in ability enhancement of distinguish and integration, capability improvement of cognitive complexity and accuracy, and the ability of obtaining information. Emotional adapt is reflected in emotional positioning, aesthetic, motivation, attitude, etc. However, behavior adaptability includes two aspects of technical ability and social ability (Ren Yuhai 2003). Similarly, learners in network environment also face the problem that "how to adapt to network culture". It is required that learners have web-based learning style cognitive level, learning motivation, learning attitude, communication and cooperation skills, learners can learn effectively in a free, open network virtual environment. Through the cultural adaptability, learners can obtain network culture content and skills, symbols and means of expression, and it makes them form the digital learning and living action ability.

Adaptability is a kind of psychological adaptable ability, is balanced mental ability individual acquired through the behavior of reacting to surroundings adversely and actively in the process of individual interact with the surrounding environment, and contact with the surrounding people, which has the individuality and sociality double attributes (Zhao Shijun 2005). Network cultural adaptability is the dynamic process in which individual cognitive, behavior, psychology, emotions, attitudes, values and so on, is changing and achieving the harmonious relationship finally in network environment, the network culture adaptability is the tendency of individual surpassing the obstacles in the network situation or the adaptability of individual forming in individual adapting to network environment. Network cultural adaptability contains several dimensions of cognitive and metacognition adaptability, values adaptability, social and communication adaptability, psychology and emotion adaptability, morality adaptability and so on (Huang Lifeng and Wen Xue 2008).

4 The Influence of Network Cultural Adaptability to Online Learning

Network culture adaptability is the survival, development ability of people in the network environment, and the so-called survival of the fittest is this meaning. Online learning is a kind of important study way under network culture environment, if learners don't adapt to network culture, what about web-based learning? Therefore, the network cultural adaptability is the premise and foundation of online learning. Regarding the above analysis, network learners' learning disabilities mainly come from the supported web-based learning activities cognitive, values, attitudes, moral, emotion, self-confidence, and communication, namely network learners' inadaptability to the network culture. The following mainly discusses the influence of network cultural adaptability dimensions to online learning.

4.1 The Influence of Network Cognitive and Metacognition Adaptability to Online Learning

In the network learning, learners need to judge and obtain information accurately, handle information creatively and collect, sift, classify and summarize the information and knowledge that they need, and then carry on the processing of these knowledge integration. Therefore, network learners should have network cognitive and metacognition adaptability under the network environment.

Cognitive adaptability refers to the process of acquire knowledge or application knowledge in cyberspace for learners, including the relevant facts, knowledge and belief, etc. Network cognitive adaptability mainly embodies in adapting to the basic knowledge and mastery the basic skills of network, including the network application ability, familiar with the situation of network environment and network

operating level. In online learning, because of the restriction of many factors, parts of the network learners' original computer operating level are not high, which can lead to direct learning difficulties and psychological barriers. Plus the complex environment, on the one hand, learners will be surrounded by trek into information, on the other hand, also information can't be get accurately, efficiently because of the operation obstacle, thus leading to information loneliness (Wu Yingyin 2009). The main reason is the learners' bad cognitive adaptability, which influences web-based learning conducting, thus further affects the learning effect.

Network metacognition adaptability is learners' self-awareness on cognitive activities of their network, learners' self-monitoring, guidance and adjustment on the network learning activities in the network environment. Network learners need network metacognition adaptability, such ability makes learners firmly grasp learning goal and determine next plan of study; aware of their system position in study, effective screening information, control learning activities, from irrelevant information interference off learning objectives, reduce the learning efficiency; in light of their own cognitive style, choose proper media present study content; introspection learning effect, and constantly control of and improve their learning methods and means.

Learners of high network metacognition adaptability are able to proper use of self adjusting and control strategy, evaluation, supervision, examination and modify his web-based learning activities, and can show higher divergent thinking and the ability to locate information, thus improving the quality of online learning.

4.2 Values Adaptability on the Influence of Web-Based Learning

Network cultural values adaptability is the adapting and evaluation about cultural values of learners. Its core includes equality, sharing, and opening up values. Meanwhile, learners build the consciousness about network social development and the sense of responsibility, and the sense of mission of construction network society. Learners need effectively supported web-based learning activities values, so learners can establish equality mutual virtual learning environment, building and sharing network learning resources, be glad to dedicate their knowledge in the network, pay attention to the harmonious development of online learning environment, so as to promote the development of online learning.

4.3 The Influence of Social and Communication Adaptability to Web-Based Learning

The social adaptability in the network is people's adaptability to community activities and sharing resources in the network activities, in order to actively construct personal social relationships, and people have social existence in the

network. Communication is the person's social need, interpersonal relationship shaping in network communication is often networking integrated embodiment of social adaptability ability. Network relationships harmony will be conducive to social adaptability and development, promoting information exchanges and enhance feelings. Positive network social adaptability has certain sense of community, with enthusiasm, rational, tolerance and cooperation skills, then learners are able to actively participate in positive network community activities.

Teacher and student are separated on space and time in web-based learning, unlike traditional classroom teaching that has very good learning environment and more directly interpersonal communication and exchange. Therefore, in the web-based learning process, learners lack the communication and motivation from teachers and students, and are not easily motivated learning motivation, and the learning effect is affected. At present, researches have showed that network learners generally feel lack of exchange and communication in the learning process (Wang Aimei and Yang Xiaopin et al. 2006). The main reason is that online learning environments interact function has not been fully used. Also the sociality and communication adaptability networks are bad, learners can't achieve the construction of knowledge significance through the communication, cooperation in certain contexts. Therefore, only having good sociality and communication adaptability, can learners study effectively in online learning activities.

4.4 The Influence of Emotions Adaptability to Online Learning

As the network teaching emotional loss problem increasingly, emotional loss has become one of research topics in twenty-first century (O'Regan 2003). At present, learners are faced with cold machine throughout the learning process, teachers and students' separation in time and space leads to the serious lack of the emotional exchange in between teachers and students, and in between students and students, which leads to students accumulate some problems which couldn't be solved on the time effectively, so they will lose interest in learning; when they having confusion in psychological could not get timely help, they will produce psychological shadows. So students produce loneliness easily, and facing setbacks and failures, produce helplessness, frustration, without promptly company or teachers' help, thus influences learning effectiveness and efficiency of students. If learners have network psychology and emotion adaptability, which always has the strong series absorbs force on learners' attention, emotion, can touch the learners' spiritual needs and inspire their positive web-based learning, avoid the ill effects of emotional loss to online learning.

4.5 Moral Adaptability of the Influence of Online Learning

Network moral refers to codes of conduct combined showed by network activity subject, in the network activities, who respect science, starting with respecting for

humanity, the social harmony and network technology features, and criticize with good and evil ideas through the social public opinion, traditional habits and inner sustain, adjusting relationships between individuals and the society (Tang Huirong 2007).

Good network morality is very important to promote web-based learning. Learners will form respecting knowledge, honesty and trustworthiness, and resist in bad information ethics corresponding network safe use. This is the fundamental conditions of network normal operation and the foundation for web-based learning smoothly. At present, moral questions which most learners have are: "hackers" and "virus", intellectual property rights, personal privacy issues, social responsibility issues, the credibility and moral emotion indifferent questions in virtual world. The main reasons triggering university students' network morals are the harmful information overflowing from the openness of network, college students' split personality caused by network concealment, without binding and virtuality (Bai Juan and Ge Shufang 2004). Accordingly, we can eliminate network moral problems, evaluate good and evil of network behaviors, increase the network moral cognition, standard ethical behaviors and achieve network moral self-discipline, thus we will have good information immune ability and avoid direct learning failure caused by network trek and intaking error message under the background where learners have different degree of barriers because of the influx of information when they screening information.

5   Conclusion

Online learning is a very complex learning system and process under network culture background. Only to have the network cultural adaptability, learners can successfully and efficiently develop web-based learning under network environment.

How to adapt to this new culture form to become effective network learners, and developing the network culture adaptable education are what we should explore and the survival of information society needs.

References

Bai Juan, Ge Shufang (2004) The analysis of network learners learning disabilities. Mod Educ Technol 6:25–27

Huang Lifeng, Wen Xue (2008) Network cultural adaptability research paper. Open Educ Res 8:54–58

O'Regan K (2003) Emotion and e-learning. J Asynchronous Learn Netw 7(3):78–92

Ren Yuhai (2003) The possibility and the intrinsic mechanism of cross-cultural adaptability. J Anhui Univ (Phil Soc Sci) 1:106–107

Sang Xinmin (2001) In network education research deepen the cultural background. Sixth world of computer education application meeting of global Chinese proceedings (CD-ROM Version), Chongqing Xinhua Multimedia Development Co., Ltd. Produces

Tang Huirong (2007) Establishing and perfecting our country's education mechanism of college students' network moral. Theory Pract Educ 3:42–43

Wang Aimei, Yang Xiaopin et al (2006) The study of online learning environment social interaction characteristics. Technol Educ 7:56–58

Weng Hansong (1999) We need what kind of network culture. Workers Daily, August

Wu Yingyin (2009) China's online learning effect factors of the current research situation analysis [J]. Softw Guide. Educ Technol 3:34–36

Zhao Shijun (2005) The analysis and countermeasures of psychological barriers of network learners. Educ Vocat 33:54–55

Zuo Jiming (2003) About network cultural collisions, explore network education countermeasure. J Liaoning Educ Coll 5:64–72

A Survey of Computer Network Management Technology

Tan Li

Abstract With the development of computer technology and the internet, the administration and maintenance of computer network have acquired crucial importance, especially so in large computer network. Being a key technique in computer network, network management has an immediate impact on the efficiency of network operation. As a result, to ensure its stability and reliability is fundamental to an efficient network operation. This paper, from the perspective of the concept and development of computer management, takes an analysis of its current problems and proposes several solutions with wide application.

Keywords Computer • Network management • Technical • Network management technology

1 The Concept of Computer Network Management Technology

It refers to the centralized management of the resources on the internet through some network management programs, including the management of configuration, function, accounting, problem, operating process and variable data.

T. Li (✉)
Network Center, Chongqing Three Gorges University, Wanzhou Chongqing 404000, China
e-mail: 375355845@qq.com

2 The Present Technologies of the Computer Network Management

Network management, as a science subject has been highlighted since 1980s. Many international organizations of standards, forums and research institutes have issued some kinds of standards, or agreement to conduct the network management and designs. Each one has its own properties, but no one has a universally accepted standards. Up to now, there are only three main network management technologies widely used. The first one is SNMP (Chen Xingyuan 2008). Invented by Internet Company, SNMP is specially designed for internet management. It is easy to run and practical to use, so SNMP has become the practical standard for the internet world. However, due to its short time development, it also has some insufficiencies, which makes it hard be used in complex network management, only in the TCP/IP management. Furthermore, SNMP is not powerful enough to guarantee the security of the computer network. The second is CMIP. It can provide an all-round support for a complete network management, which is more advanced in technology and better completed in criteria (Li Fuliang 2007). The best privilege is that the variable data in the contract is designed for the data on the terminals, but also for finishing some tasks of the network management. Anyhow, it is designed mainly to make up the insufficiencies of SNMP, and it is so complicated and expensive that it can't be accepted widely. The third is distributed object network management technology, which is generated from the applying process of CORBA on the internet. All the management application and elements are viewed as the distributed objects. And the interaction of the distributed objects is the task of the internet management. The distinguished properties are that it screened the differences of the program languages, internet contracts and operating systems. It is more incisive and easily see-through so that it can be researched and developed more easily and used more widely. SNMP and CMIP with their own properties appeal to different clients, so they can not be replaced by either of them for a long period. It is also impossible for the system of CORBA to replace them, because the time, cost and human resources for the research and development are so mass. Consequently, the research for the combination of the CORBA, SNMP and CMIP is suggested for the network management of the third technology (Niu Yan 2010).

3 Application of the Network Management Technology

The aim of network management technology is to guarantee the internet and the network system to operate to some extent stably, reliably and efficiently, namely to make sure all the network resources are running smoothly and to satisfy the needs of the clients. In the past, there were some simple devices to help the network administrator to manage the resources. But now with the development of

the internet, a powerful and easy-to-use network management technology is needed urgently to cope with the enormous and complicated internet resources. Only this kind of technology can fulfill the aim (Qin Zheng 2007).

3.1 Introduction of WBM Technology

Since the internet business is increasing, and the intranet technology is becoming popular, some major network companies attempt to use MIS technology in a new way, hence the Web-Based Management technology. As a complete new network management mode (based on Web network mode), its strong vitality, special flexibility and good serviceability are welcomed by a large number of technicians and clients, which was commended as "a final solution to the network management".

WBM technology based on the functions of the Web and its management technology, has greater capability than the traditional devices for the network administrators. WBM allows the administrator to use any kinds of Web browser, to allocate quickly at any point of the network, to control, store and withdraw any items on it. As a result, the administrators work is no longer confined to their offices; at the same time many operating problems caused by the different computer systems can also be resolved. Furthermore, WBM provides quicker and easier user interface for the internet log-in than the traditional driver. The browser and Web surface are familiar to the WWW clients, so the cost for the training of operation is much reduced and the network operation information is possibly well used by more and more internet clients (Zhang Baosheng 2007). Therefore, we can say it is "a final solution to the network management".

3.2 Introduction of WBM Technology

3.2.1 Design Objective

The objective of the design is to develop Chinese network management system software in the "Web" pattern with independent copyright on the basis of campus network. By employing the advanced WBM technology and efficient algorithm, the software is expected to reach the level of the same products abroad.

This network management system will provide a complete set of solution based on WEB. Its effective resource management of distributed IP network enables the user to carry out strain type management and control to network, equipment as well as relevant system and service via WEB browser from anywhere. In this condition, the best operating status of the network resources, availability and reliability of the network can be guaranteed (Tan Xianhai 2006).

3.2.2   System Architecture

The system is designed in the current technical conditions and with a view to technology development trend while taking advanced products abroad of the same kind as references. The three-tier architecture used by us, which has incorporated advanced WBM technology, offers the administrator a simple and convenient way of management (Tian Yuan 2009).

3.2.3   The Characteristics of the Three-Tier Architecture

The software responsible for management task works in background process way as middle tier to carry out polling of network equipment and collection of fault information.

The management middleware resides in between network equipment and browser, thus users just need to manage the homepage of middle tier in order to access the equipment.

The management middleware relay forwards management information and makes protocol conversion between SNMP and HTTP. No change is necessary for the equipment in three-tier architecture (Ye Shuhua 2006).

3.3   The Design of Network Topology Discovery Algorithm

To effectively manage the network, the administrator needs an intuitive and friendly user interface, of which the fundamental role is unfold the topological relation of network equipment graphically before the users' eyes, namely, topology discovery. At present, the topology discovery algorithm in widespread adoption is based on SNMP. It is very effective and fast to some extent, but there exists a defect (Liu Ruiling 2010). That is, in a particular field, all information on subnetwork depend on that the equipment posses the characteristic of SNMP. Therefore, the algorithm will become invalid in a system incompatible with SNMP. Moreover, the SNMP function may be closed due to lack of attention to network management or for security reasons, as a result, imperfection of topology will be caused by the difficulty in obtaining MIB values of equipment. With the purpose of resolving this problem, an improvement of the algorithm above will be discussed in the following – topology discovery based on ICMP.

3.4   PING and the Establishment of Routing

PING is mainly used to send messages, and to wait for answers. PING got its name because it's a simple echoing agreement, using ICMP request and response to send

messages. PING is mainly used by the system programmers to diagnose and debug, and it can be realized through the following processes. First, send an ICMP message to the target machine, and then wait for the response from it, until overtime. If the response messages can be received, it proves that the target machine works well. Then the program withdraws.

The function of Routing is to use TTL domain in IP. At first, the source defines the TTL value of IP as zero, and sends messages to the informationsink. After finding out that the TTL value is zero, the first gateway rejects the messages and sends an overtime-typed ICMP message to the source, which then analyzes the message (Zhang Pusheng 2010). Therefore the first gateway address is created. Then the source sends the message with TTL 1 to the informationsink. The first gateway reduces its TTL value to zero, and then sends it to the second gateway, which rejects the message, and sends the overtime-typed ICMP message to the source, after finding out that the TTL value is zero. Therefore the second gateway address is created. When the message reaches the informationsink, the Routing to the informationsink is created.

3.5 Processes to Realize Network Topology Discovery Algorithm

1. In the confined IP space, take turns to detect each IP address, by using PING, and record the detected IP addresses in the list of IP addresses.
2. To trace route every IP address, to record the routing of these IP addresses, and to add the gateway addresses into the IP list.
3. For every IP address in the list, find out its corresponding subnet mask.
4. According to the submask, find out the corresponding subnet website, and its network, adding each subnet to the list.
5. Try to get the corresponding Domain Name of each IP address. The same Domain Name means that the same network facility has many IP addresses, i.e. many network interfaces.
6. Create the connection profile according to the routing got in the 2nd step and the subnet got in the 4th step.

4 Conclusions

At present, the application of computer network is in an explosive growth period. With the rapidly expanding network, the network complexity is also aggravating. With the development of network technology and the network business application, it becomes more and more important to manage and maintain the computer network. We must pay more attention and strive to manage the network well, to let the highly efficient and safe network promote the economic and social development.

References

Chen Xingyuan (2008) An agreement for security of internet communication. Higher Education Press, Beijing, pp 50–55

Li Fuliang (2007) An intensive study of network application. Electronics Industry Press, Beijing, pp 62–63

Liu Ruiling (2010) Computer netword technology and Application. Zhejiang University press, Zhejiang, pp 179–185

Niu Yan (2010) Computer network technology and the application. Tsinghua University Press, Beijing, pp 185–185

Qin Zheng (2007) Network application psychology. Science Press, Beijing, pp 132–133

Tan Xianhai (2006) Network program technology and application. Tsinghua University Press, Beijing

Tian Yuan (2009) Security on internet. People's Telecommunication Press, Beijing, pp 98–99

Ye Shuhua (2006) A course book for network program technology. People's Telecommunication Press, Beijing, pp 121–122

Zhang Baosheng (2007) Innovations and competitiveness of network technology. Business Management Press, Beijing, pp 89–89

Zhang Pusheng (2010) Computer network technology. Science Press, Beijing, pp 125–125

The Internal Control of Computerized Accounting in Information Era

Xiaoqin Tang, Chunlan Yan, and Yanyu Kang

Abstract In the present, computer and modern information technology were widely used in accounting. Therefore, a good surrounding of university's financial management was created by computerized accounting. Since 1980s, computerized accounting was developing so fast, It was used in over 90% of firm's financial management. This essay analysed the risks and problems appeared in Computerized Accounting, some advices were developed based on working experience, expect to offer little promotion to computerized accounting of China in information era.

Keywords Information era • Computerized accounting • Internal control

As computer and modern information technology were widely used in accounting, accounting's working mode was gradual developing from manual mode to computerization. The changes what computerized accounting made were not only changed the dealing and storage ways of accounting information, but also impact the theory and application of it. Today, computerized accounting are applied successfully in various professions.

1 Features of Computerized Accounting

Computerized accounting is defined as accounting work which used electronic digital computer and modern data processing technology (Du Dechun and Ai Yali 2007). In computerized accounting, computer process substituted manual work in book-keeping, account, reimbursement and also the analysis, forecast and decision

X. Tang (✉) • C. Yan • Y. Kang
Chongqing Three Gorges University, Wanzhou, Chongqing 404000, China
e-mail: 23944473@qq.com

about accounting information, this kind of substitution aimed to promote the financial management and economical benefit and modernize the accounting work. After study, there are four major features comes out: (1) Data processing substitute the traditional manual work, meanwhile, computer is as a major working tool; (2) Magnetic media is the major storage of accounting data; (3) The centralized and automatized data processing way, it makes form making easier, data analysis more accurate and faster the inquiry; (4) Fast computing, massive storage, highly information sharing (Campus Internal Control Program Manual of Stony Brook University 2012).

The follow advantages were taken by the computerized accounting:

(a) The accounting software released accountants from manual accounting work, promoted working efficiency, also the quality of accounting information is improved.
(b) The functions (transfer, check etc.) of accounting software reduced error rate as opposed to accuracy rising.
(c) A multitude of inter-disciplinary talents who master both computer and accounting were trained up, it made domestic computerized accounting work towards standardization, commercialization, generalization and specialization.
(d) Provided information construction experience to the firms, promoted the development of enterprise management software market.

2 The Risks Computerized Accounting Faces

Computer and accounting software are the most important material basis of computerized accounting. But everything is not perfect, the defect of computer design takes some risks to the computerized accounting at the same time. The computer hardware ensures the processing of software and data, all the data store in the storage equipment which can access from computer. Any error of hardware system can impact the accounting work negatively, such as data lost and no recovery of data. The major factor influence the safety of hardware: The quality of machine components, hardware damage caused by human or natural disaster etc (Zhang Jincheng 2002).

Accounting software is the fundamental of computerized accounting, the accounting software choosing is the major problem in computerized accounting. The compatibility of software is so important, bad compatibility will make data sharing not so easy, it cause hard communication in the firm, influence the whole work negatively. Meanwhile virus prevalence, hacker attack, software bugs threaten the running of computer devices.

Computerized accounting is a connection between human and computer compare to traditional accounting. In computerized accounting, account, cash, data storage, file transfer are digital files, it makes data be deleted and changed without clue possible. The database of accounting data connecting to Internet, it's possible for

those people who are not authorized access the database and view the data, further copy, forge and delete important accounting data. So compare to traditional accounting, it will be harder to find cheating and forgery in computerized accounting, so the computer cheating and forgery will cause more financial lost. Otherwise, because of potential physic damage of storage medium, the data will lost at that time, so automatic backup of accounting data is so important, it can avoid potential data lost.

3 Problems in Computerized Accounting's Internal Control

In the present, most of accounting data are processed by computer with network support. As the most important component of accounting information system, in computerized accounting working environment, the contribution was enhanced. The diversification is the trend of internal control, the control will be much stricter than before. The accounting information system of computerized accounting developing in upper degree, this kind of change bring the enterprises both financial benefit and challenge of the control itself.

3.1 Paperless Cuts Traditional Internal Control System

Computerized accounting system changes the form of accounting vouchers, business processing methods and procedures of accounting and financing undergone great changes in the Computerized accounting system. In traditional manual accounting system, accounting voucher, bookkeeping voucher, detail account, general ledger and so on are recorded and kept by different accountants. Responsibilities are divided according to the nature of economic businesses, different people in charge, and form a network with internal mutual restraints and less fraudulent. Once problem arises, ascertain where the responsibility, at the same time any unauthorized person can not browse all the accounting data.

3.2 Defects of Internal Control System Construction

The internal control system under computerized accounting environment consists of system control and management control. System control is mainly the software integration; management control sets and manages enterprise itself. However, many enterprises are lack of awareness, both exist shortcomings. In computer system control, there haven't established such systems which prevent fraud, effective division of responsibilities system and the separation of incompatible duties. But in management control, there haven't establish effective responsibility of com-

puterized accounting, operation management system of computerized accounting, computer hardware and data management system, accounting archives management system. And internal control system is lack of operability according to the realities of company. A good internal control system is established on the basis of thorough investigation on the nature of business, management requirement, staff, software and hardware and so on. Every business has its own features, computerized accounting internal control system should be established according to the actual situations of business, but some enterprises are lack of progress, internal control management model are backward, and can not adapt to a new age.

3.3   The Efficiency of Financial Software Should Be Improved

With the rapid development of information technology, computer and software update quickly, a wide range of financial software put into use, their functions become more perfect and powerful. The application of financial software aims not only to improve efficiency, but more importantly, to provide relevant departments with useful and effective data to make a decision, and accounting plays its role in the participation of enterprise management. However, computerized accounting level of many enterprises are low, the function open and utilities are not enough, poor generality of accounting software and low degree of integration. Software of different companies can not exchange data, share information and control management, and can not integrate other subsystem in the company, so it is difficult to construct enterprise-wide management information system.

4   Measures of Computerized Accounting Internal Control

In order to ensure the authenticity and reliability of accounting information, realize computerized accounting system operate normal, safe and effective, eliminate possible hazards, and enhance the overall management of enterprise, achieve management goal, we must establish and improve internal control system and measures of computerized accounting.

4.1   Establish Computer Hardware, Software and Data Management Systems

The normal operation of computer is the precondition of computerized accounting implementation, so we should regularly carry out maintenance on the equipment to prevent accidents; ensure the security of accounting data and accounting software, prevent unauthorized modification and delete of data and software; the data stored

on magnetic should be preserved and double backup; ensure the continuity and safety of actual accounting data in the process of software modification, upgrading and hardware replacement, and should be supervised by relevant staff; improve management measures when computer hardware and software fail and the elimination of fault, ensure the integrity of data; enhance computer virus protection, increase security, improve the management of accounting files.

4.2 Establish Operation and Management System for Computerized Accounting

Establish post responsibility of computerized accounting system, every job has a clear scope of official duty, and make sure everything was tubes, everybody is expertise, work is required, work is inspected. Clearly define the operation contents and authorities of operators, prevent accounting data such as original voucher and bookkeeping voucher registered the computer before them are audited. When the operators leave, they should quit accounting software; and according to actual situation, the operation records should be kept by special person, and recording contents including operator, operation time, operation contents, fault conditions and so on.

4.3 Strengthen the Training of Computerized Accounting Staff

Computerized accounting has become a comprehensive subject integrating accounting, management science, computer technology and information technology. Accountants not only should be proficient in accounting business, but also proficient in computer and data processing technologies. Accountants should be systematic in-service training, there should be plans to train staff, promote enterprises establishing high efficiency accounting information system, make standardization of computerized accounting, commercial, universal and specialized. At the same time, enterprises should strengthen the ethical standards of accountants, be honest, observe disciplines and obey laws, ensure the smooth progress of work.

5 Conclusions

Perfect computerized accounting internal controls can effectively guarantee the reliability of accounting information and all the accounting processing can be deal with in the line with national accounting regulations, to improve the efficiency of accounting, and provides management decisions for the enterprises with high quality accounting information.

References

Du De-chun, Ai Ya-li (2007) Internal control of computerize accounting system. Mod Mark 2:32
Office of management Analysis and Audit, Campus Internal Control Program Manual of Stony Brook University 2012 from http://www.sunysb.edu/audit/icmanualnew.shtml
Zhang Jincheng (2002) Computer information system control and audit, vol 4. Peking University Press, Beijing

Study of Personalized Service About WAP Based on Recursive Learning

Jia-Liang Xu, Jiao Guo, Chun-Tao Zhang, and Xiao-Qin Ye

Abstract The development of WAP business allows users to obtain fast, convenient resources. With the increase of WAP services, the traditional way in the WAP terminal services can not meet the needs of users. In the paper, the traditional vector space model method is improved, Introductions thinking of recursive learning, Implements a recursive type of user interest for personalized service, It is through the collection and analysis of user access to information to get the user's interest, in order to achieve the purpose of the recommended initiatives.

Keywords WAP • Recursive learning • Personalized service

1 Introduction

WAP is a mobile terminals to provide Internet content and advanced value-added services to the global unity of open standard protocols, it will make mobile network and Internet and corporate LAN closely linked with the network, provided a type, operators and terminals are independent mobile value-added services (Mladenic 2000; Bollacker et al. 2000). WAP has become an important way to get information. With WAP information growing, people have to spend a lot of time to search for browsing the information they need. Personalized service is proposed for this problem, providing different service strategy and services for different character-istics and interests of users (Kurki et al. 1999). By collecting and analyzing user information Personalized service learn the users' interests and behavior in order to achieve the purpose of the recommended initiatives.

J.-L. Xu (✉) • J. Guo • C.-T. Zhang • X.-Q. Ye
Chongqing Three Gorges University, Chongqing 404000, People's Republic of China
e-mail: ywing123@163.com; legend_j@163.com; zhangchuntao@163.com; yxq_leaf@163.com

2 Personalized Service Model

Personalized service that is in line with visitor interest, identity and needs in the form of information and applications provides special treatment. Personalized service by collecting and storing the user's information, analysis and study the information of the users' interests and behavior, and then learn the users' interest, at the right time to send each user the right information, in order to achieve the recommended active Objective (Chen et al. 2001).

2.1 Vector Space Model

A widely used model is based on content filtering system or named vector space model (LI Xue-lei and Zhang Dong Mo 2003; Yang and Ding 2003). The model system is mainly to filter for the following models: the user submits a standard document that mostly reflects the users' interests and hobbies. Others need to filter the document and the document comparison, the more similar the document with the standard documentation of their chances to obtain recommended. The advantage is simple and effective; disadvantage is difficult to distinguish resource quality and style, and can't discover the users' newly interested resources, and users can find the resources have similar interests. In order to compensate the deficiency of Vector space model, introduced the recursive learning model.

2.2 Recursive Learning Model

Establishment a user Interest dictionary is to reflect the user's interests. Calculated in accordance with certain rules in the user interest in each term the weight of the dictionary, and then adjust the users' interests in the dictionary: Appearing in the dictionary entries for the weight increase is not in the dictionary entries will be added to the dictionary. If interests in the dictionary is full, then the dictionary entries will delete the low weight, allowing users to maintain interest in the dictionary at a fixed capacity. Entries in the dictionary user interest in the learning process is gradually added. This allows the user to get the interest in time, so as to provide a higher quality of service. Recursive learning reduces learning time, improving the learning efficiency, it's good to do a dynamic update knowledge base, to get the user's interest in time (Schwab et al. 2000).

3 Recursive Learning Personalized Service

The introduction of recursive learning personalized services, tracking and learning the users' interests and behavior, and design a suitable expression, so you can follow the user's interest in providing personalized service (Mobasher et al. 2000). In this

regard, first, to collect user's information, treatment information recorded in the tables in a user; and then give a learning algorithm for the tables, run the algorithm of interest by users dictionary, and then use the interest to keep a dictionary Update user profile, enabling the user interested in obtaining recursive personalized service.

3.1 Collect User Information

For personalized service, the most important is the user's participation, in order to get the user's interest and behavior. First, the user must track the collection of information, user information is divided into explicit and implicit messages of information are two kinds of explicit information obtained by the user's registration information, registration information stored in the user table, implicit information can be collected and analyzed by the user WAP access logs are kept in the user access log table. The main table is defined as follows:

Definition 3.1. UserReg (Uid, Attr, Val) of the table records the basic characteristics of the user information, Uid that the user phone number, Atrr feature attributes that the user name, Val said that the corresponding attribute values, such as name, Hobbies and other non Required information.

Definition 3.2. Userlogin (Uid, LTime, LAmount), the table records the user's login time and total number of landing.

Definition 3.3. UserCost (Uid, DownSort, DownCost, Time), the table records the user's consumer content, consumer resource type and consumption time.

Definition 3.4. ResInfo (RID, Rname, Weight), Weight for the system in which the value of the recommended priorities, in the content rating system used on the representation of forests, as a whole does not exceed three. The first level navigation pages in a database using two-digit ID, a collection of Ravi = {00,01, ..., mn}, m, n ∈ (0,9); second level for the category pages, numbered from Left to right, set RSort = {1,2, ..., n}, n ∈ (0, ∞); the third level for a content page, in the representation similar to the second level classification, collection RCont = {1,2, ..., N}, n ∈ (0, ∞).

Definition 3.5. UserLog (Uid, Attr, Val) of the table records the user's online behavior. Each record is a record of a user session Uid conditions contained in Attr k1, k2, ..., km, which is a key that the session is continuously accessed page belongs to the category of resources, Val contains access Times, access times and other information.

The table in this deal, the result stored in the user interest table (UserInter ()) and the user description table (UserProfile ()), and the user description table is essentially a multi-decision making table.

Definition 3.6. UserInter (Uid, RID, BrowAmount, AuthValue, Time) table each record corresponds to a period of time the user Uid to access information. BrowAmount that users of resources in a period of time for the RID class visits, and AuthValue said that the period of time the user's interest bias.

Definition 3.7. UserProfile (Uid, RID, Weight, Time) recorded in the corresponding table is the user UID interests and hobbies, which Weight category of resources that the user RID of the weights.

Described in the user table, the user's weight will love interest as the user records in the table is changed, each over a period of time will be interest in favor of user interest table to replace the maximum value of the right of users to describe the table records the lowest.

3.2 Learning User Interest

Users interested in learning a machine learning process is based on the user's choice of browsing information, to take a clear step by step learning method to a process of user interest through a variety of machine learning methods.

Collected and collated the information for the user then to study the user's interest. Through the analysis of user behavior and actions (Mobasher et al. 2000), the recent interest in favor of the user, the end of the user profile. Defined as follows (Fan and Gauch 2000):

Algorithm 3.1 to obtain the user profile

a) for manual classification of resources as a collection of training resources;
b) created user profile;
c) Statistics of the entries in the user profile in the frequency;
d) the frequency of the various entries, calculated the corresponding weights of the classification of resources, draw the user's interests.

3.3 Recursive Updating User Interest

Algorithms 3.1 through the user's continued to access the amount of new information of interest to the user preference, and then tend to go with the new update the user's interest profile. Defined as follows:

Algorithm 3.2 (for users interested in the dictionary)

a) for manual classification of resources as a collection of training resources;
b) create user interest in a dictionary;

c) Statistics of the entries in the user dictionary in the frequency of interest;
d) the frequency of the various entries, calculate the corresponding weights of the classification of resources, draw the user's interests.

Algorithm 3.3 (Updated user profile)

a) taking out the dictionary from the user's interest in the highest three power, if it contains user profile in non-classified, then take the first;
b) using the data obtained to update the user profile, if the user profile blank, it will have to take the data directly into; If you add new data which is beyond the capacity of the user profile, then the user profile in the weights of the lowest three Item deleted;
c) every other time, repeat step1 and step2.

Users interested in obtaining recursive process: building a user interested in documents and user profile, through the algorithm to get the user interested in documents 3.2 by Algorithm 3.3 to get the user profile. Repeating the above process, the user can get a real hobby that you can also prompt the user interest in the current bias.

Thus, for the time the new information collected, recursive get the user interested in has been completed, the user profile has been dynamically updated.

4   Simulation Results

4.1   Experimental Test

To verify the recursive personalized service for users' interests of the effectiveness in the program, in the Windows platform, with the JAVA to achieve a non-recursive algorithms for user interest and recursive algorithms for user interest, and on a single machine Testing of the experimental data. The test data was coming from the Web site user registration information, and user access log information for half a year.
Experimental steps are as follows:

• First 6 months, respectively, the former co-exist in the database organize information, statistical information in the database as the original training data, user interest generated by algorithm 3.2 dictionary, and then use the users' profile construction algorithm 3.3.
• Every other time statistics the user information, the following two treatments. The first method is the input training data, run the algorithm 3.1, the user's profile; The second method is to input the training data, the use of algorithms 3.2 and 3.3, the user's profile.
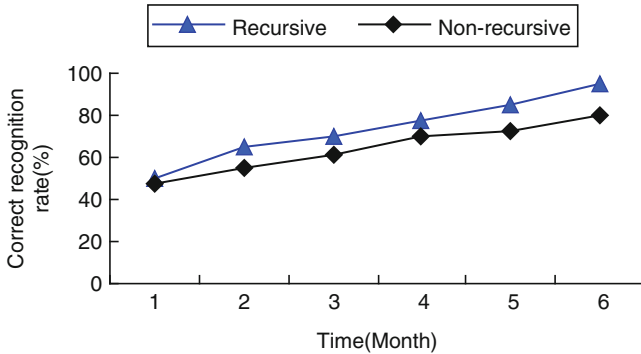
Fig. 1 Correct recognition rate comparison

4.2   Results Analysis

The results shown in Fig. 1. Figure 1 can see that the recursive method to get user interest than non-recursive difference methods the results are satisfactory recognition rate of more than closer to the user's real interest.

- Non-recursive method to increase the time used very fast and recursive methods of the time growth is not obvious;
- Recurrence than non-recursive greatly reduce the time spent. This shows that the recursive algorithm to get the user interested in the time and space costs than non-recursive low, and as time goes by, more and more information processed, recursively increasingly used in a clearly advantageous. Therefore, this recursive method to get user interest more effectively update the knowledge base.

5   Using the Template

As the users continue to access, the information collected more and more which will increase the analysis of the reliability of user interests. In order to provide personalized services better. The paper presents a recursive user interest for personalized service system, truly tailored for the user "system", it is a good way to do the personalized recommendation service, in order to achieve the "useful information for the most suitable people".

References

Bollacker KD, Lawrence S, Giles CL (2000) Discovering relevant scientific literature on the Web [J]. IEEE Intell Syst 15(2):42–47

Chen CC, Chen MC, Sun YS (2001) PVA: a self-adaptive personal view agent system [A]. In: Schkolnick M, Provost F, Srikant R (eds) Proceedings of the 7th ACM SIGKDD international conference on knowledge discovery and data mining [C]. ACM Press, New York, pp 257–262

Fan Y, Gauch S (2000) Adaptive agents for information gathering from multiple, distributed information sources [A]. In: Murugesan S, Macarthur S, O'Leary DE (eds) Proceedings of the AAAI symposium on intelligent agents in cyberspace [C]. AAAI Press, Menlo Park, CA, pp 40–46

Kurki T, Jokela S, Sulonen R et al (1999) Agents in delivering personalized content based on semantic metadata [A]. In: Murugesan S, Macarthur S, O'Leary DE (eds) Proceedings of the AAAI spring symposium workshop on intelligent agents in cyberspace [C]. AAAI Press, Menlo Park, CA, pp 84–93

Li Xue-lei, Zhang Dong Mo (2003) A text classification method based on vector space model [J]. Comput Eng (17):17–21

Mladenic D (2000) Machine learning for better Web browsing [A]. In: Rogers S, Iba W (eds) AAAI 2000 spring symposium technical reports on adaptive user interfaces [C]. AAAI Press, Menlo Park, CA, pp 82–84

Mobasher B, Cooley R, Srivastava J (2000) Automatic personalization based on Web usage mining [J]. Commun ACM 43(8):142–151

Schwab I, Pohl W, Koychev I (2000) Learning to recommend from positive evidence [A]. In: Riecken D, Benyon D, Lieberman H (eds) Proceedings of the international conference on intelligent user interfaces [C]. ACM Press, New York, pp 241–247

Yang XH, Ding J (2003) Chinese information retrieval technology based on vector space model [J]. Comput Eng Appl 282–286

An Efficient Design of Publication and Subscription Model Based on WSN

Xianli Li

Abstract The purpose of the Web Services Notification (WSN) is to define a set of specifications that standardize the way Web services interact using "Notifications" or "Events". They form the foundation for Event Driven Architectures built using Web services. These specifications provide a standardized way for a Web service, or other entity, to disseminate information to a set of other Web services, without having to have prior knowledge of these other Web Services. They can be thought of as defining "Publish/Subscribe for Web services". We provide an overview of the WS-Notification specification and describe a modified Publish and Subscribe model based on WS-Notification. The model is an adaptive policy-driven notification framework that helps enterprises to meet the flexibility and responsiveness requirements of the enterprise. With the modified publish/subscribe model, information consumers can dynamically and declaratively create and configure entities on their behalf to manage their distribution requirements.

Keywords Publish/subscribe • Notification broker service • Notification consumer proxy service

1 Introduction

In a Service Oriented Architecture (SOA), there is often a need to monitor situations. This occurs from a computer management perspective or a much more broad scope of keeping up to date on real world events such as weather, financial transactions, etc. To monitor these events, a client can poll for status changes or subscribe for status changes. In polling case, the client is configured to actively ask the resource

X. Li (✉)
Computer Network Information Center, Yangtze Normal University, Chongqing, China
e-mail: lisansanchqi@163.com

for its latest state. The more often the client polls for state, the more likely the client has an accurate resource representation. However, frequent polling requires bandwidth and resources on both sides to handle the connection. Thus polling is useful when monitoring timeliness is not an issue or network and hardware resources are abundant. But in the SOAP world, polling is less common as a client typically receives requests from the producer of events. With this peer to peer style, a publish/subscribe system can be created. In this system, the client requests that notifications be sent when they occur. This reduces the latency between the event occurring and the client processing it.

WS-Notification has been standardized by OASIS and is a standard that solves this business problem of event distribution in heterogeneous complex event processing systems. It specifies an interface for a consumer to subscribe, filter notifications, and manage subscriptions and an interface for publishers to send notifications. Further, it describes a notification broker to allow for scaling of the system (Chumbley and Eisinger 2009).

2 WS-Notification

Associated with the WS-Resource Framework, IBM, Sonic, and other companies introduced a family of related specifications called WS-Notification. The basic idea behind WS-Notification is to standardize the way that a Web service can notify interested parties (other Web services) that something of interest has happened. It is not meant to replace all messaging infrastructure such as low latency buses, industry standards, or existing infrastructure like JMS. However, WS-Notification systems should be able to integrate with these systems through simple adapters.

Obviously, the key value to WS-Notification is its ability as a standard to allow for greater interoperability with a greater number of vendors and, thus, a lower cost for implementation. The key features of WS-Notifications allow for it to be used as well in general purposes publish/subscribe (pub/sub) situations. It defines a unified message format to achieve interoperability between kinds of systems, procedures and components in different platforms and systems, and it defines a set of a standard Web services approach to notification using a topic-based publish/subscribe notification pattern.

WS-Notification family is made up of the following three components (Sams 2006): WS-Topics, WS-BaseNotification and WS-BrokeredNotification. Figure 1 shows the relationship between them.

Based on the WS-Notification, the publish/subscribe model is needed to handle the real-world information integration scenario by allowing a subscriber to specify on behalf of an information consumer filtering rules and policy constraints, not only to select what types of messages or content the subscriber wants the consumer to receive, but also to specify transformation, scheduling, distribution, and other constraints to be applied to selected messages before they reach the consumer. The architecture must enable the generation (on behalf of a consumer) of a proxy service,
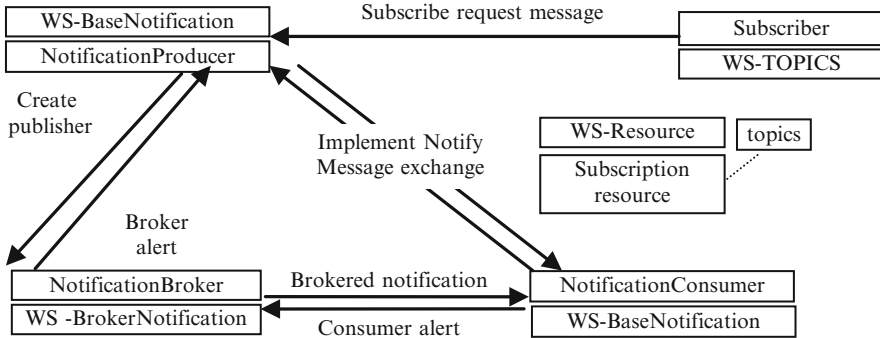
Fig. 1 Relationship between WS-Topics, WS-BaseNotification and WS-BrokeredNotification

or agent, that includes the engines to enforce and manage these constraints at runtime. The proxy service can receive the messages on behalf of the information consumer and apply the specified constraints to the message before delivering it to its consumer (Czajkowski et al. 2004).

2.1 WS-Topics

The WS-Topics specification (Vambenepe et al. 2006a) defines a mechanism to organize and categorize items of interest for subscription known as "topics." This is achieved by associating each notification with a topic, and means that subscribers can define the specific category of event that they are interested in hearing about. A web service can publish a set of topics used to organize and categorize a set of notification messages that clients can subscribe to, and receive a notification whenever the topic changes.

WS-Topics are very versatile, as they even allow us to create topic trees, where a topic can have a set of child topics. By subscribing to a topic, a client automatically receives notifications from all the descendant topics (without having to manually subscribe to each of them). As part of the publication of a Notification-Message, the Publisher associates it with one or more Topics.

WS-Topics also provide a coarse-grained filtering mechanism that allows large sets of uninteresting notifications to be excluded quickly. For example, in a sport results scenario a subscriber can indicate that he or she is only interested in receiving notifications about football, which excludes any events about baseball or hockey. More fine-grained control of filtering can be achieved using other filtering mechanisms, such as the message content filter defined in WS-BaseNotification. For example, by selecting only those football games in which the home team beat the away team. In many situations, the topic does not actually appear in the body of the notification message itself since the classification of the notification is made at a higher level than the generation of the notification content.
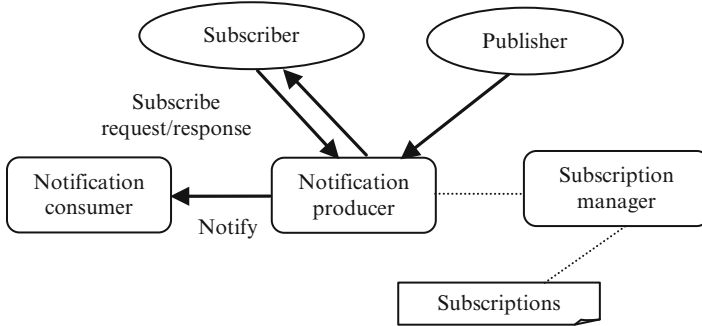
Fig. 2 A typical WS-notification interaction

In order to avoid naming collisions, and to facilitate interoperation between independently developed Notification Producers and Subscribers, every WS-Notification Topic is assigned to an XML Namespace. The set of Topics associated with a given XML Namespace is termed a Topic Namespace.

2.2 WS-BaseNotification

The WS-BaseNotification specification defines the standard Web services interfaces for Notification Producers and Notification Consumers. It includes standard message exchanges to be implemented by service providers that wish to act in these roles, along with operational requirements expected of them. Notification producers have to expose a subscribe operation that notification consumers can use to request a subscription. Consumers, in turn, have to expose a notify operation that producers can use to deliver the notification (Sotomayor 2007; Vambenepe et al. 2006b). Figure 2, "A typical WS-Notification interaction" shows how the five primary entities work together to pass data through the WS-BaseNotification. Initially, the Subscriber is responsible for setting up a subscription between the Notification-Producer Web service and a NotificationConsumer Web service. This subscription is managed by the SubscriptionManager Web service working on behalf of the producer. Subsequently, when a Situation is observed by the Publisher, the Publisher creates a notification message and passes it to the NotificationProducer. It is the responsibility of the producer to establish whether the notification message matches the subscription that has been registered, and, if so, to send the notification message to the consumer.

2.3 WS-BrokeredNotification

In even the most simple publish/subscribe environments, the amount of connections and boot strapping information can grow very quickly. If there are only two

publishers and two consumers, and each consumer wants to be notified from each publisher, four connections need to be set up. Add one more consumer and you now have six connections. The number of connections starts to grow very quickly as more distributors and consumers are added; the required number of connections for m publishers and n consumers is m × n connections. To simplify this topology, WS-Notification standardized a notification broker that acts as an intermediary between publishers and consumers. Here, publishers and consumers are decoupled from each other and instead only require boot strap information to the broker. So, in the scenario of m publishers and n consumers, the required number of connections is m + n.

The WS-BrokeredNotification specification extends the definitions made in the WS-BaseNotification specification to define the concept of a NotificationBroker, which is an intermediary service to which producers and consumers can connect in order to pass notifications. Critically, the NotificationBroker is capable of accepting subscription requests from consumers, as well as receiving notification messages from producers. The broker is then responsible for matching the notifications with the subscriptions and sending them to the consumer. In this way, the broker takes on some of the more painstaking functions of the producer, freeing developers of producer applications to concentrate on the business task of observing situations and generating the appropriate notifications without having to worry about the challenging but mechanical task of managing subscriptions and matching them to notifications. Advantages of the brokered notification pattern are as follows:

- Relieves the publisher of having to implement message exchanges associated with the notification producer; for example, managing subscriptions (SubscriptionManager) and distributing notifications (NotificationProducer);
- Avoids the need for synchronous communications between the producer and the consumer;
- Can reduce the number of inter-service connections and references;
- Acts as a finder service; for example, if a new publisher is added that publishes notification x, a consumer does not have to issue a new subscription if it is already subscribed to the broker with x.
- Provides anonymous notification, which means that publishers and consumers need not be aware of each other's identity.

In many scenarios, the NotificationBroker service is implemented by a middleware provider, ensuring that the brokering facilities are written to enterprise quality expectations and often providing additional value-add services over and above the basic definition of the service, for example logging, transformation, or quality of service enhancements above those required by the specification (Vambenepe et al. 2006c). As shown in Fig. 3, "A typical brokered WS-Notification interaction", the producer must register with the broker and publish its topics there. The subscriber must also subscribe through the broker, not directly with the producer. Finally, when a notification is produced, it is delivered to the consumer through the broker.
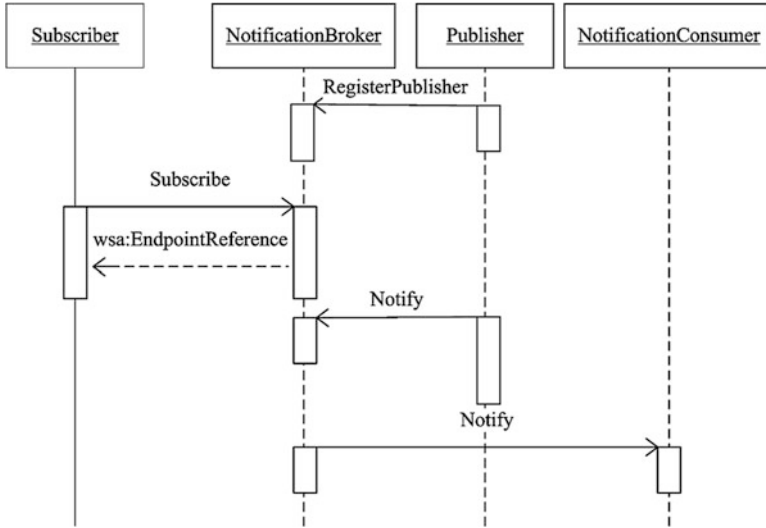
Fig. 3 A typical brokered WS-notification interaction

3 Publish/Subscribe Model Based on WS-Notification

The modified publish/subscribe architectural model is as shown in Fig. 4. It extends the basic publish and subscribe pattern by extending the subscription capabilities to include the specification of transformation, distribution, and scheduling constraints as part of the publish and subscribe subscription (Bou-Ghannam and Roberts Matt 2007).

Additionally, this architecture enables non-pub/sub-enabled systems (that is, information consumers that are not able to consume notification messages of the pub/sub system) to participate in the pub/sub pattern by allowing the model to dynamically create a proxy service to receive pub/sub notifications on behalf of the consumer. This is the Notification Consumer Proxy Service (NCPS) shown in Fig. 4, which also manages the distribution of notifications to the consumer based on the transformation, distribution, and scheduling constraints specified by the consumer upon subscribing.

As shown, the model highlights the following components:

Notification producer: Contains information of interest to a consumer. Good examples of information producers are systems that manage business information for an enterprise and include master data stores for customer, product, order information, and so on, in addition to enterprise operational data stores.

Notification consumer: Depends on and must consume information from an information producer. For example, many enterprise business applications like order fulfillment systems depend on data from the business information sources.

Subscriber: Requests creation of a subscription. It sends a subscribe request message to a notification broker (pub/sub broker). The subscribe request message
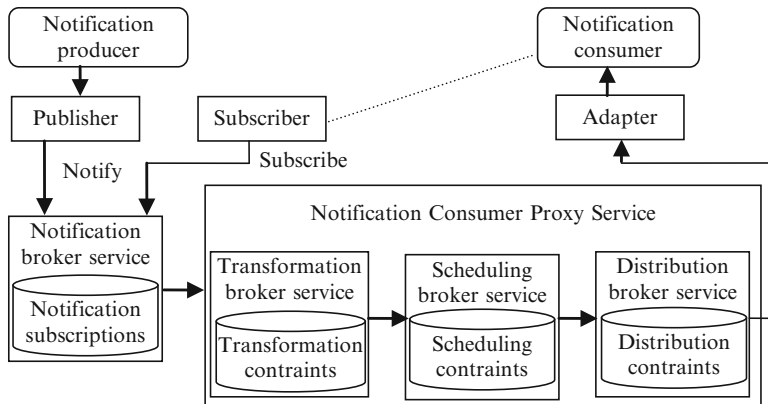
Fig. 4 The publish/subscribe architectural model

identifies a notification consumer. A subscription is an entity that represents the relationship between an information consumer and an information producer. It records the fact that the consumer is interested in some or all of the notifications that the producer can provide. It can contain filter expressions, and may be long-running or have a limited lifetime.

Publisher: Creates notification message instances. A publisher receives information from entities in the information producer that monitor and detect a situation. A situation is an occurrence that is noted by one party and is of interest to other parties. A notification is a one-way message that conveys information about a situation to other services.

Notification broker service: Performs a notification broker function between notification consumers and notification producers, and is responsible for sending notifications to the appropriate consumers. It also acts as a subscription manager and manages requests to query, delete, or renew subscriptions.

Notification Consumer Proxy Service (NCPS): Receives notifications from the notification broker on behalf of the information consumer. Typically, the consumer is not able to receive notification messages, hence the need for this service to act on its behalf, collect the notifications, perform some business logic (if the scenario calls for it), enforce the transformation, scheduling, and distribution constraints for the consumer, and then send the results to the consumer.

Adapter: An entity that enables the interaction with an information consumer.

4 Conclusions

This paper discussed how the WS-Notification bundle of standards, WS-BaseNotification, WS-Topics, and WS-BrokeredNotification, can be used as a general purpose publish/subscribe interface for a Service Oriented Architecture.

We described an adaptive, policy-driven notification architectural model to support a generalized publish/subscribe interaction pattern. This model is based on the WS-Notification standards, a set of reusable integration services. We introduced the teacher-student interactive scenario to help demonstrate the WS-Notification features and explained how the publish/subscribe model is the standard of choice for event distribution and processing.

References

Bou-Ghannam A, Roberts M (2007) GPASS: a generalized publish and subscribe solution using WS-Notification standards. 2007-8. Online available at http://www.ibm.com/developerworks/websphere/library/techarticles/0708_boughannam/0708_boughannam.html. Accessed 27 Aug 2009

Chumbley RB, Eisinger JD (2009) Leveraging key WS-notification features in your business applications, 2009-04. On line available at http://download.boulder.ibm.com/ibmdl/pub/software/dw/webservices/ws-wsnotificationWAS7/ws-wsnotificationWAS7-pdf.pdf. Accessed 25 Aug 2009

Czajkowski K, Ferguson D, Foster I et al (2004) WS-resource framework. 2004-06-9. Online available at http://www.globus.org/wsrf/specs/ws-wsrf.pdf. Accessed 9 May 2009

Sams Publishing (2006) WS Notification and WS Topics in the WS resources framework, 2006-07. Online available at http://www.devarticles.com/c/a/Web-Services/WS-Notification-and-WS-Topics-in-the-WS-Resources-Framework/. Accessed 9 May 2009

Sotomayor B (2007) The globus toolkit 4 programmer's tutorial, 2007-08-19. Online available at http://gdp.globus.org/gt4-tutorial/multiplehtml/ch08s02.html. Accessed 9 June 2009

Vambenepe W, Graham S, Niblett P (2006a) wsn-ws_topics-1.3-spec-c, 2006-10-09. Online available at http://docs.oasis-open.org/wsn/wsn-ws_topics-1.3-spec-cs. Accessed 10 May 2009

Vambenepe W, Graham S, Niblett (2006b) P wsn-ws_base_notification-1.3-spec-cs.pdf, 2006-08-09. Online available at http://docs.oasis-open.org/wsn/wsn-ws_base_notification-1.3-spec-cs-01.pdf. Accessed 9 June 2009

Vambenepe W, Graham S, Niblett P (2006c) wsn-ws_brokered_notification-1.3-spec-cs, 2006-08-09. Online available at http://docs.oasis-open.org/wsn/wsn-ws_brokered_notification-1.3-spec-cs-01.pdf. Accessed 9 June 2009

Part IV
The Application of Information Technology

On the Study of Information and Ultra-Wide Band RADAR Based Information Retrieval

Abdul Waheed and Yang Li

Abstract Different information has different importance in respective sectors. Information can be about security, about monitoring, about observations or about detection. Similarly depending on the information sensitivity different methods of Information retrieval and algorithms are used with the help of suitable retrieval systems. This paper first discuss about Information retrieval systems based on retrieval methods and their use in different areas and then starting from the birth of radar, the evolution and advancement in radar systems as information retrieval system is studied to retrieve the required and suitable information for the purpose of security, monitoring, observation and tracking. Finally based on the technologies, electromagnetic consumption and the FCC standards, radar based time a line of the information retrieval systems is studied.

Keywords Wireless information retrieval methods • UWB • RADAR

1 Introduction

Diversity of information retrieval systems has a variety of retrieval methods in order to retrieve the required information for example information like Electro-cardiograph, Magnetic resonance image, sonograph, X-Ray image are important in medical sectors while the information retrieval systems in electronics and electrical Industry start from voltage, current, frequency, resistance and capacitance measure meters and then move to oscilloscopes and spectrum analyzers and the advancements reaches to the involvement of computers. Thus based on information type the information retrieval system are divided into different categories

A. Waheed (✉) • Y. Li
School of Information and Electronics, Beijing Institute of Technology, Beijing, China
e-mail: laserspc_tech@yahoo.com

- Everyday information systems.
- Industrial Information systems.
- Scientific Information system.
- Security and defense based Information systems.

These categories partially emerge into each other and the information retrieved by the system become more sensitive as we move to the system of security and defense category therefore a suitable information retrieval system is required in order to avoid information thievery. Security and defense based information is a major category and therefore we will focus on the fourth category only. From personal security to the national defense security, the band of type of information and information retrieval system has a large variety of retrieval methods.

Wireless Information retrieval systems are discussed in Sect. 2 while the UWB RADARs as information retrieval systems in different sectors are discussed in Sect. 3.

2 Wireless Information and Information Retrieval Systems

Depending on the required information, the security and defense based wireless information retrieval system can be divided into different categories.

2.1 Personal and Industrial Information and Retrieval System

A variety of security systems for personal assets fall in this category so the information retrieval system can be limited to the boundary of assets. Indoor security systems, home security and car security systems are the examples of such systems.

Industrial information systems are the combination of personal information, equipment information and information retrieved by the equipments. Thus the security of information is the combination of employee surveillance, equipment surveillance, production information and equipment working observation.

2.2 Scientific Information and Retrieval Systems

These kinds of systems cover the area of scientific research. Scientific research contains the use of machinery in order to perform the research experiment. Most of scientific equipments contain two types of information. First one is the useful scientific data either in image form or in statistical form while the second type covers the security alarms of the equipment.

2.3   Defense Related Information and Retrieval System

Defense related information and related system has its own importance. It covers the indoor as well as outdoor activities relevant to the security.

RADAR and detectors are two major information retrieval systems used to generate and retrieve the exact information in order to maintain the security. Depending on the range parameter, Radar systems are short range as well as long. Since radar systems cover a large variety of information acquisition in a vast area of fields therefore we will focus on the evolution, history and future of radar systems based on different parameters and standards.

3   History of Radar with Respect to Different Countries

The history of radar with respect to different countries has the following time lines

1865—Dynamical theory of electromagnetic field published by Scottish physicist and mathematician James Clerk Maxwell.

1887—Phenomenon of transmission and reflection through different materials was found by German physicist Heinrich Hertz.

1899—An inventor of Italy Guglielmo Marconi Performed reported his radio based Experiments.

1904—Radio waves were first used by a German Christian Huelsmeyer to detect the presence of distant metallic objects.

1917—Serbian-American Nikola Tesla published primitive radar units.

1922—A. Hoyt Taylor and Leo C. Young, researchers in U.S. Navy, found the possibility of detection of range and bearing of nearby ships in the Potomac river with radio waves broadcasted at 60 MHz.

1934—the French man Émile Girardeau built a radio based apparatus for hindrance-locating by following the principles stated by Tesla. Same year Soviet military engineer P.K. Oschepkov with Leningrad Electrophysical institute introduced RAPID to detect an aircraft within the range of 3 km. Both systems were based on continuous-wave operation. In late 1934 an American Robert M. Page, working at the Naval Research Laboratory demonstrated the improved form of RADAR with pulsed system which was followed by Rudolf Kühnhold and the firm GEMA in Germany and by an Air Ministry team led by Robert A. Watson Watt in Great Britain.

1939—Eight countries secretly and independently developed their own RADAR systems.

1940— U.S Navy used the term RADAR.

1943— Robert M. Page, improved the monopulse technique.

3.1 Frequency Ranges and UWB RADAR

From HF range to UWB range, depending on the application, RADAR exists in every frequency band of electromagnetic pattern but UWB radar is unique due to its frequency range i.e. 1.6–10.5 GHz whereas the ranges like 1–2, 2–4, 4–8 and 8–12 GHz are L-Band, S-Band, C-Band and a part of X-Band respectively.

UWB radar is an information retrieval system to retrieve the information of objects across the wall, hurdles and obstacles and therefore also named as through wall radar.

Due to the efficiency of even low signal to noise ratio the UWB signals are much better for information retrieval in different applications.

According to Federal Communications Commission (2002) 3.1–10.6 GHz is the license free range for UWB but power spectral density emission for radiators is limited to −41.3 dB/MHz.

UWB radar has a vast range of applications for the purpose of information retrieval in an indoor environment and short range boundaries. The next section is based on the applications and usage of UWB information retrieval system.

3.2 Applications of UWB RADAR as Information Retrieval Systems

According to Dr. Robert J. Fontana in (2000), late 1960s is the starting time for UWB technology but with different name and approximately 1989 was the year to recognize the technology with the name of UWB.

The current applications consisting of UWB radar are spread over approximately every area of life but with different purposes of information retrieval.

In the medical field the most important thing is the effect of UWB waves on human body. Klemm and Troester (2006) gives the effect of UWB in human tissues while (Staderini 2002; Tan and Chia 2004; Chia et al. 2005; Li et al. 2010; Venkatesh et al. 2005) gives the study of application of UWB radar as information retrieval system like medical imaging, remote sensing of heart beat breathing rate, detection of presence of water and infection into human body, chest movements assessment.

The indoor applications along with the security and defense related application of UWB radar based information retrieval systems contains human body detection also called vital sign detection, moving target detection, direction estimation of moving target, speed estimation of moving target (Zito et al. 2008; Leib et al. 2010; Ossberger et al. 2004; Yang and Fathy 2009; Ahmad et al. 2007; Earp et al. 1996; Tao et al. 2008).

4   Conclusion

There was a time when doctors use to observe the temperature of patients by observing the heartbeat and by physically touching the patient on wrist or the neck time. Then a time came when use of mercury based thermometers started for same kind of observations, then those mercury based thermometers welcomed the strip based thermometers and digital thermometers. These thermometers helped the doctors to avoid the physical touching but the physical interaction between patient and the temperature measuring device increased. Now a day's all those kinds of thermometers are replaced by infra red based device which helped the patient and doctor both to avoid the physical interactions.

This mean that the temperature is a kind of information in very basic doctor-patient system while the device used to measure the temperature is the information retrieval device and the way to use the temperature measuring meter is called the information retrieval method.

Secure Information retrieval method is necessary for the security of system. Ultra-wide band RADAR based information retrieval system provides more security and accuracy for short range environment, either for through wall imaging, patient or asset monitoring or target detection and tracking while provides the required isolation.

References

Ahmad F, Amin MG, Mandapati G (2007) Autofocusing of through- the-wall radar imagery under unknown wall characteristics. IEEE Trans Image Process 16(7):1785–1795

Chia M, Leong S, Sim C, Chan K (2005) Through-wall UWB radar operating within FCC's mask for sensing heart beat and breathing rate. Europ Radar Conf EURAD, Paris, pp 267–270

Earp SL, Hughes ES, Elkins TJ, Vickers R (1996) Ultra-wideband ground-penetrating radar for the detection of buried metallic mines. IEEE AES Systems Magazine, September 1996, pp 30–39

Federal Communications Commission (2002) Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems. First report and order, ET Docket 98–153, FCC 02–48, April 2002. http://www.fcc.gov

Fontana RJ (2000) A brief history of UWB communications. Multispectral Solutions, Inc., Germantown, MD, United States. www.multispectral.com/history.html

Klemm M, Troester G (2006) EM energy absorption in the human body tissues due to UWB antennas. Prog Electromagnetics Res PIER 62:261–280

Leib M, Menzel W, Schleicher B, Schumacher H (2010) Vital signs monitoring with a UWB radar based on a correlation receiver. In: Proceedings of the EuCAP 2010, Barcelona, Spain, April 2010, pp 3425–3429

Li X, Pancera E, Zwirello L, Wu H, Zwick T (2010) Ultra wideband radar for water detection in the human body. GeMiC 2010, Berlin, Germany, March 2010

Ossberger G, Buchegger T, Schimback, B Stelzer A, Weigel R (2004) Non-invasive respiratory movement detection and monitoring of hidden humans using ultra wideband pulse radar. In: Proceedings of the joint UWBST & IWUWBS, Kyoto, Japan, May 2004, pp 395–439

Staderini EM (2002) UWB radars in medicine. IEEE Aerospace and Electronic Systems Magazine, January 2002, pp 13–18

Tan A, Chia M (2004) UWB radar transceiver and measurement for medical imaging. In: 2004
    IEEE international workshop on biomedical circuits and systems, Singapore, 1–3 December
    2004, pp S3.1–9
Tao R, Zhang F, Wang Y (2008) Fractional power spectrum. IEEE Trans Signal Process
    56(9):4199–4206
Venkatesh S, Anderson C, Rivera NV, Buehrer RM (2005) Implementation and analysis of
    respiration- rate estimation using impulse-based UWB. In: IEEE Military Communications
    conference (IEEE Milcom'05), Atlantic City, NJ, October 2005, pp 395–399
Yang Y, Fathy AE (2009) Development and implementation of a real-time see- through-wall radar
    system based on FPGA. IEEE Trans Geosci Remote Sensing 47(5):1270–1280
Zito D, Pepe D Mincica M et al (2008) Wearable system-on-a-chip UWB radar for contact-less
    cardiopulmonary monitoring: present status. In: Proceedings of 30th annual conference of the
    IEEE Engineering in medical and biological society, Vancouver, British Columbia, Canada,
    August 2008

The Algorithm of Data Preprocessing in Web Log Mining Based on Cloud Computing

Guanglu Zhang and Mingxin Zhang

Abstract In the structure of distributed cluster server, web log data mining model based on data warehouse has the defects of bottlenecks in the network and computing, transmission errors caused by the large data transmission, the paper makes use of the advantages of cloud computing, distributed processing and virtualization technology, designs a type of Web log analysis platform based on cloud computing Hadoop cluster framework, finally, a new hybrid algorithm of distributed procession in the cloud computing environment is proposed. To further verify the efficiency of the platform, we use the improved data pretreatment algorithm on the platform of processing large number of Web logs, experimental results show that it can improve the efficiency of Web data mining.

Keywords Web log mining • Expected page • Website structure • Navigated path

G. Zhang
College of Mathematics and Information Science, Northwest Normal University, Lanzhou, China
e-mail: gluzhang@163.com

M. Zhang (✉)
School of Computer Science and Engineering, Changshu Institute of Technology, Suzhou, China
e-mail: mxzhang163@163.com

1   Introduction

Web Log Mining is an important research content in Web mining: By mining server log, publishers can analysis and study the laws of web logs, obtain the user's access schemas, finally prove the site's organizational structure and services. Generally, web log mining contains three main steps: data preprocessing, pattern discovery and pattern analysis. Moreover, many large websites use distributed web servers, which are distributed to multiple different geographic locations, server cluster structure is presented in Fig. 1 (Ni et al. 2009).

$P_1, P_2, : \ldots, P_n$ denote the documents on the site. At present, domestic and foreign large sites are using the web cluster structure shown in Fig. 1, which is characterized by the page file on the site of the whole image. Based on cluster structure in Fig. 1, there are two models of distributed web log mining: one is presented by Mehmet Sayal and Peter Scheuermann, which is based on data warehouse technology; another is based on Synthesis of distributed data mining results. If using the first technique, the key is to build data warehouse out of the web server log files distributed in each server, and then in data mining. The distributed technology based on data warehouse has the advantages of heterogeneous data processing, but because of large web log files, building a data warehouse need to invest significant human and material resources, and so this technology have a big disadvantage. Such as transferring large amounts of data to the central site on the distributed server, it is likely to cause network bottlenecks, transmission errors and other defects. If using the second technique, merging the mining results of each distributed log file, but because of the logs of a distributed single server geographically has one-sidedness, it doesn't reflect the information contained in the overall log, which leads to the result of digging out has limitations. Cloud computing can get a powerful network computing resources. It distributes complex calculation consumes a large amount of computing resources to multiple nodes process by network, which is the current effective solution (Ni et al. 2009; Jun Ji 2009; Xianghe Zhu 2010).

In this article, we use highly progressive Hadoop distributed computing platform to research and analyze the logs of some typical books reading web, design and implement a novel web log pre-processing method based on Hadoop, which can quicken the later work of web log mining and promote the overall efficiency of web log mining.
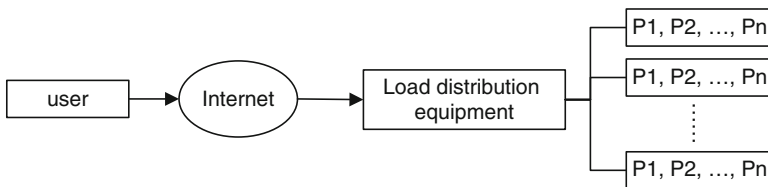

Fig. 1 Web cluster structure

2 Description of Relative Question

2.1 Web Logs

Web server log, when user request pages through the browser to the server, web server records each request in the file to form log file, it records users' access and interactive information. According to that different web server has different log format, there are two mainly formats: Common Log Format and Extended Log Format, in this article we use Log in ECLM mode.

2.2 Hadoop Technology

Hadoop, a distributed system infrastructure, is developed by Apache Foundation. Using it (Benjun Guo et al. 2009; Miao Cheng and Huapin Chen 2011), even if the users don't know the distributed low-level details, they can develop distributed program, and use clusters' power to super-speed computing and memory. Briefly, Hadoop is a software platform, which can more easily run and analyze large-scale data. The core idea of platform is Map/Reduce. Map/Reduce is a programming model used to process large data. meanwhile, it also is an efficient job scheduling. It divides a task into lots of more fine-grained subtasks which can be dispatched between free processing nodes, and so, the faster processing speed the nodes have, the more tasks they can handle, also because of such, it can avoid prolonging the task completion time due to the nodes with slow processing speed. Performing a MapReduce operation requires five steps: input file, divide file into many segments and assign to multiple work nodes to execute parallel, write locally middle files, merger middle files, output final files (Mdamet Sayal 2009).

3 Implement of Distributed Preprocessing Model

3.1 Design a Web Log Analysis Platform Based on Hadoop

The overall framework of web log analysis platform based on Hadoop cluster framework shown in Fig. 2 (Zhu Zhu 2008):

First, user submits mining tasks by the browser. After receiving the tasks, scheduling nodes split them into small task units, and check nodes' idle lists, then assign the task space to idle compute nodes in the cluster. According to data source location information in the task units, the compute nodes access to data from data storage nodes and for the corresponding computational work. Last, it returns the
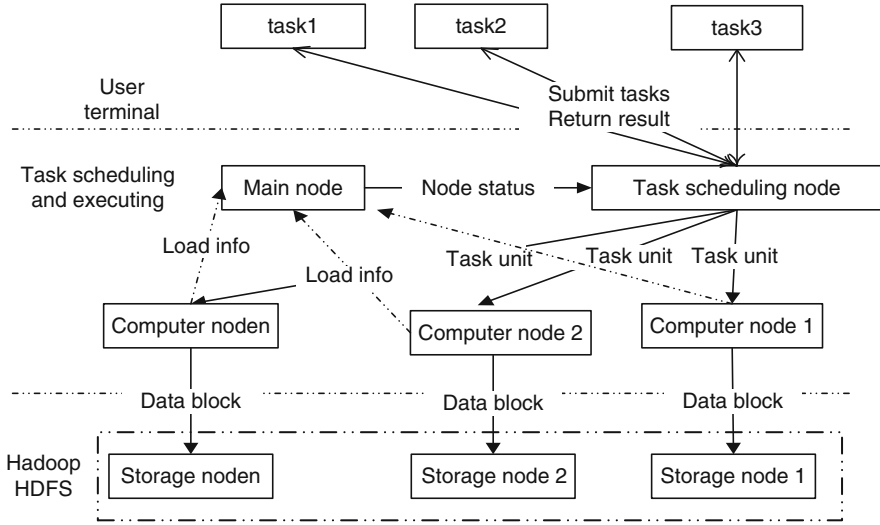
Fig. 2 the overall framework of web log data preprocessing based on Hadoop

result to task scheduling nodes combine, and the task scheduling nodes return the final result to user. All compute nodes need to send a signal to the master node at regular intervals to prove them work normally, not in an idle state. Meanwhile, the master node need to real time return the compute nodes' state to the task scheduling nodes (Ni et al. 2009; Peng Liu 2010; Peng Wang 2010).

3.2　The Web Log Preprocessing Using Map/Reduce

Web log preprocessing need to deal with all of its access for each user, so it is very suitable for the <key,value> model of MapReduce. In this model, you can easily get together to the same user's web log by selecting the user ID as the key. Mapper will break up the input record, make up the <key,value> keys' pair in accordance with user requirements, then sent the group of the same key to the Combiner, the Combiner aggregates the same keys, these data are sent to the Reducer, finally, Reducer gets the final result by calculating the value of each key.

3.2.1　Mapper Operation

The input of Mapper operation are the records of Web log, set the key userid, value is <URL|ReferURL|Time>, and then sent to Combiner.

```
Function: clean each record, form the pairs of
<URL|ReferURL|Time>, and sent it to Combiner.
input[] divide the input record using '' ''
if input[5] = ''GET'' and input[6] 's end is not
{gif, jpg, jpeg, wav, rm, css, cgi,js}
and 200 <=input[8] <= 299
    mapKey ← input[0]
    mapValue ← input[5]|input[6]| input[8]
collect {mapKey,mapValue}
end if
```

3.2.2   Combiner Operation

In order to optimize the MapReduce operation, an optional Combiner operation is added, which is running on the map task running node after Mapper before Reducer. Make Mapper output on the receiving node as input, aggregate the output of the Mapper, at the same time the aggregated results will be further input as a Combiner data aggregation, the aggregation is not stop until as each key on node is aggregated to a record. Combiner output will be sent to the reducer.

3.2.3   Reducer Operation

We can define USID as the key of Combiner's output, define URLIt as iterator which contains Combiner value of output, define US as the triple which format for <URL,ReferURL,Time>. Define USIt as the iterator which contains the US, defined USL as queue which contains the US. The Reducer receiving series of <USID,URLIt>. Sort the collection of records for each user to access the page in chronological order, and output all user's sessions before the session identification, path added and session split method.

```
Function: Combiner output access to the record
  aggregate for a user session
for each urls URLIt
urlArray[] ← divide the urls input record using \ "
for each url in urlArray
USIt ← USIt + {url}
end for each
end for each
sort the USIt in ascending order by Time field
us1 ← USIt.next()
USL ← {us1}
reduceValue ← null
while USIt.hasnext() do
```

```
us2 ← USIt.next()
if (us2 = ''index.jsp'') or (us2.referURL =
  ''index.jsp'' and us1.URL ≠ ''index.jsp'') or
  (us2.referURL = null and us2.Time us1.Time <
  10s) or (us2.Time us1.Time > δus1.URL) then
reduceKey ← URL field in all USL join one after
  another by ''|''
collect {reduceKey,reduceValue}
us1 ← us2
USL ← {us1}
else if us2.referURL = us1.URL then
USL ← USL + {us2}
us1 ← us2
else if us2.referURL ≠ us1.URL then
reduceKey ← URL field in all USL join one after
   another by ''|''
collect {reduceKey,reduceValue}
us1 ← us2
scan USL in reverse order until finding usk, usk.
URL = us2.referURL
  USL ← USL - {us1...usk} + {us1}
  if none usk.URL = us2.referURL
  USL ← {us1}
  end if
 end if
end while
```

4 Experiment

The experiments run on a Hadoop clusters which are form ten servers. One of the servers are considered as NameNode of the HDFS in a distributed file system and the JobTracker in the MapReduce running process, thus, the server is called the master node. The other nine servers considered as the TaskTracker of the HDFS DataNode and MapReduce running process, these nodes are called node. In order to verify the effectiveness and efficiency of Hadoop to mine Web log, we use java that has realized the same web log preprocessing method in the stand-alone and Hadoop cluster servers. On both platforms, the file size of Web log files were processed to calculate the execution time, the result is shown in Fig. 3:

Test for a lot more data on a Hadoop cluster, the results shown in Fig. 4:

From the test mentioned above, we can see that it's an effective way to handle web log using the Hadoop platform. When dealing with small amount of data, the Hadoop total parallel computing time is greater than the stand-alone execution, due to startup time-consuming and time-consuming intermediate file transfers. The
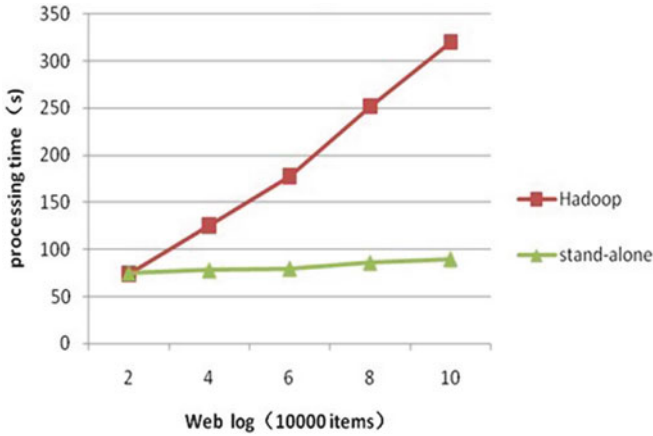
Fig. 3 Stand-alone environment and cloud computing parallel data processing environment
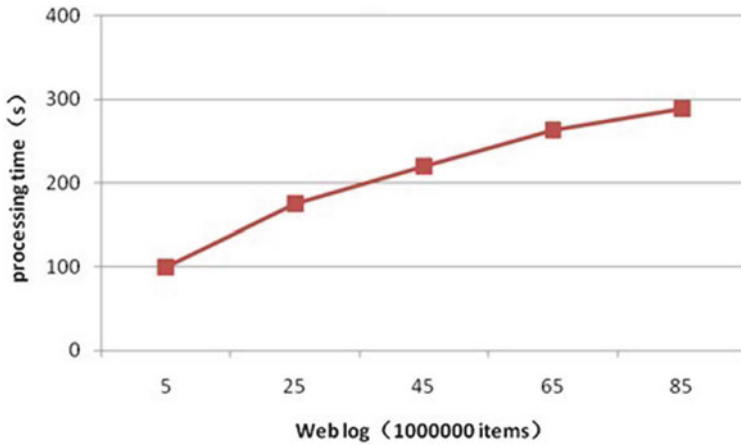
Fig. 4 E execution time large amounts of data implied in Hadoop cluster

Hadoop will divide data and assign partition data to multiple computing nodes to parallel processing. With the increases of the amount of data, parallel computing speed grows slowly, whereas stand-alone processing time basically showing the linear growth trend. When the amount of input data grows, the efficiency gap is also growing. When the number of the Web log is about 106, because of compute nodes of the cluster is already full, the time growth shows slightly less than linear growth. In that conditions, memory overflow or other mistakes may possibly happen when using stand-alone computing, by reason of the limitations of memory, processor power, etc. And splitting of the data is necessary.

5 Conclusion

The accuracy of data is an important precondition and basis for web log mining, high-quality web log mining must rely on high-quality data (Peng Wang 2010). In this paper, an improved method based on existence of session identification and path complement and session split is proposed according to the mass of web log, and the is implied on the Hadoop platform, at the same time, compare it's processing efficiency with stand-along machine, confirm the effective and efficient of Hadoop platform in processing web log.

References

Benjun Guo, Pen Wang, Gaoyun Chen (2009) Cloud computing model based on MPI. Comput Eng 35(24):84–85, 96

Jun Ji (2009) Architecture design and implementation of a cloud computing-based data mining platform. Qingdao University, Qingdao

Mdamet Sayal (2009) Peter Scheuemann Distributed web logs mining using M&xi

Miao Cheng, Huapin Chen (2011) Web log mining based on Hadoop. Comput Eng 06(11):37

Ni P, Liao JX, Wang C, Ren KYC (2009) Web information recommendation based on user behaviors. In: 2009 WRI world congress on computer science and information engineering, 31 March 2009–2 April 2009, p 426

Peng Liu (2010) Key technology and application examples of cloud computing. The Press of People Posts and Telecom, Beijing

Peng Wang (2010) Cloud calculative key technology and application example. The People's Posts and Telecommunications Press, Beijing

Xianghe Zhu (2010) The study of data preprocessing algorithm of web log mining. Jiao tong University, Dalian:11

Zhu Zhu (2008) Model of research and application of Hadoop-based mass data processing. Beijing University of Posts and Telecommunications, Beijing

A Simple Attempt to See if Artificial Intelligence Tool Is Helpful in Long Term Earthquake Prediction

Xiaxin Tao and Zhengru Tao

Abstract Artificial Neuron Network tool is adopted in an attempt to long term earthquake prediction for the Japan Trench subduction zone where an shock with magnitude 9 occurred last year and the probability model failed in forecast even after the shock. The preliminary result shows that the AI tool is helpful in such difficult a prediction, it can recognize some kind of rhythm of seismicity fluctuation that people can also find in the time series, but cannot be clear described.

1 Introduction

As well known, earthquake prediction is a very difficult problem worldwide at present. Long term prediction is a basis of taking engineering countermeasures to save loss of life and property, and is mainly from statistics. Some probability models are developed for estimating occurrence probability of earthquakes and are adopted in seismic hazard assessment, such as Poisson model, renewal model (National Research Institute for Earth Science and Disaster Prevention 2009; Working Group on California Earthquake Probabilities 2003; Earthquake Research

X. Tao (✉)
Institute of Engineering Mechanics, China Earthquake Administration, 29 Xuefu Road, Harbin 150080, China

Harbin Institute of Technology, 202 Haihe Road, Harbin 150090, China
e-mail: taoxiaxin@yahoo.com.cn

Z. Tao
Harbin Institute of Technology, 202 Haihe Road, Harbin 150090, China
e-mail: taozr@foxmail.com

Committee 2000; Tao and Tao 2012). In cases, the assessments failed, for example for the area of 311 earthquake with magnitude 9.0 occurred last year in eastern Japan Trench (National Research Institute for Earth Science and Disaster Prevention 2009). The shock and a very bad tsunami followed up made a serious disaster. After the earthquake, researchers re-evaluated the occurrence probability of an earthquake with magnitude more than 8.5 in the next 30 in that region (Earthquake Research Committee 2009), and failed again (Earthquake Research Committee 2011; Tao and Tao 2012). Seismologists and earthquake engineers want to know if the failure comes from the model unreliability, or the lack of data, or problem in data process. A simple test is presented in this paper, based on the historical data in the eastern Japan subduction zone, by means an Artificial Intelligence toll, two layer neuron network (Tao and Zheng 1991, 1998; Tao and Du 1992). The preliminary result shows a large earthquake 8.5 or more can be predicted.

2   Seismicity in Eastern Japan Subduction Zone

There has been a historical record for the earthquake in the subduction zone, more than 500 years. There were 82 shocks with magnitude than 7.0, in which 15 from 7.5 to 8.0, and 18 more than 8.0, the maximum magnitude is 8.6 in 1707, as listed in Table 1.

A time series can be constituted from the data in Table 1. A fluctuation is quite obvious in the plot of the series with vertical axis for magnitude and horizontal axis for time, as shown in Fig. 1. Seismologists tried to recognize rhythm in the fluctuation with the physical nature of energy release in the crust. One could think from the fluctuation that an earthquake with magnitude higher than 8.5 is possible at the time period after 2008. However the rhythm structure cannot be understood and described well.

3   Sampling Seismicity Tendency Data for ANN Training

The authors of this paper believe that rhythm structure must imply in the time series. The earthquake energy released in a time period may relate with those released in the past decades, even 100 years. A moving widow train is designed to sample data from the series. There are four windows in the train, each time window with width 30 years and the beginning of the next is exactly the end of the previous, the total time period covered by the train is 120 years. In each window, four values are acquired for four magnitude intervals ($7.0 \leq M_j < 7.5$, $7.5 \leq M_j < 8.0$, $8.0 \leq M_j < 8.5$ and $8.5 \leq M_j$), as 1 if there is at least one quake occurred with

Table 1 Seismicity data in the subduction zone

Serial number	Occurrence date	Magnitude	Serial number	Occurrence date	Magnitude
1	1498.9.20	8.3	42	1933.3.3	8.1
2	1605.2.3	7.9	43	1935.10.18	7.1
3	1611.12.2	8.1	44	1936.11.3	7.4
4	1633.3.1	7	45	1938.11.5	7.5
5	1649.3.17	7	46	1938.11.5	7.3
6	1662.10.31	7.6	47	1938.11.6	7.4
7	1677.4.13	7.9	48	1940.8.2	7.5
8	1677.11.4	8	49	1941.11.19	7.2
9	1686.1.4	7.2	50	1943.6.13	7.1
10	1703.12.31	8.1	51	1944.12.7	7.9
11	1707.10.28	8.6	52	1945.2.10	7.1
12	1763.1.29	7.4	53	1946.12.21	8
13	1782.8.23	7	54	1952.3.4	8.2
14	1793.2.17	8.2	55	1958.11.7	8.1
15	1833.12.7	7.7	56	1960.3.21	7.2
16	1835.7.20	7.3	57	1961.2.27	7
17	1843.4.25	8	58	1961.8.12	7.2
18	1854.12.23	8.4	59	1962.4.23	7.1
19	1854.12.24	8.4	60	1963.10.13	8.1
20	1854.12.26	7.4	61	1964.6.16	7.5
21	1856.8.23	7.5	62	1966.3.13	7.8
22	1857.10.12	7.3	63	1968.4.1	7.5
23	1861.10.21	7.4	64	1968.5.16	7
24	1893.6.4	7.7	65	1969.8.12	7.8
25	1894.3.22	7.9	66	1973.6.17	7.4
26	1894.6.20	7	67	1978.3.23	7
27	1895.1.18	7.2	68	1978.3.25	7.3
28	1896.6.15	8.2	69	1978.6.12	7.4
29	1897.2.20	7.4	70	1978.12.6	7.2
30	1897.8.5	7.7	71	1982.7.23	7
31	1900.12.25	7.1	72	1983.5.26	7.7
32	1905.6.2	7.2	73	1984.8.7	7.1
33	1915.3.18	7	74	1989.11.2	7.1
34	1918.9.8	8	75	1993.1.15	7.5
35	1921.12.8	7	76	1993.7.12	7.8
36	1923.6.2	7.1	77	1994.10.4	8.2
37	1923.9.1	7.9	78	1994.12.28	7.6
38	1924.7.1	7.6	79	1995.12.4	7.3
39	1924.12.29	7	80	2003.9.26	8
40	1931.3.9	7.2	81	2004.11.29	7.1
41	1931.11.2	7.1	82	2008.5.8	7
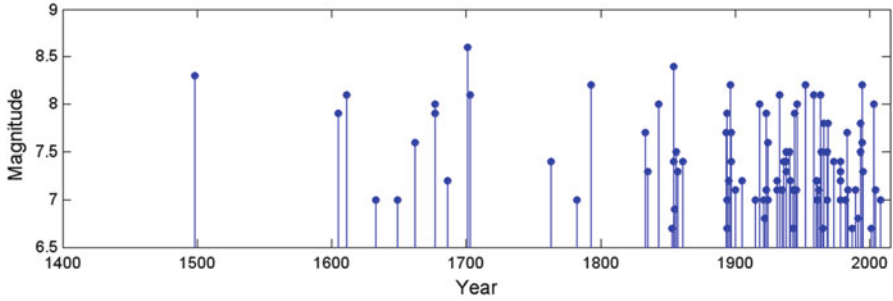
Fig. 1 M-T plot of earthquake series in the subduction zone

magnitude in the interval, but 0 if there is no such a quake. By this way, 16 values from the 4 windows form a training set, 12 of previous 3 windows for input and the 4 of the last one for output. The total 46 sample sets are acquired by moving the train along time axis with a step of 10 years from the historical data (1498–2008) in the subduction-zone.

Of course, the number of the windows in a train, the width of the window, the moving set and the values sampled in a window are all adopted without systematical determination for this simple attempt. Other number such as 5 or 6 windows, 50 year, 5 or 20 years can also be taken into account in the further deal, especially the sapling value need more detail trails to optimize.

4 Design of Two Layer Network and Training

An error back propagation (BP) network is designed with two neuron layers. There are 8 and 4 neurons in the middle and output layers respectively. In the additional input layer with 12 nodes and in the middle layer, there is an additional node with constant output 1. Each neuron links with all nodes in the next layer bellow one by one, as shown in Fig. 2.

One neuron is represented as

$$X_{is}^k = \frac{1}{1 + e^{-\sum_{j=0}^{N_{k-1}} w_{ij}^k \cdot x_{js}^{k-1}}} \tag{1}$$

where, x_{is}^k is the output of ith neuron in kth layer to the sample s; x_{js}^{k-1} is the output of jth neuron in k-1th layer to the sample s, and it is the output of the bias unit on k-1th layer when j equals to 0; w_{ij}^k is the weight between ith neuron on kth layer and jth neuron on k-1th layer, which is assigned as a random between 0 and 1 at the beginning.
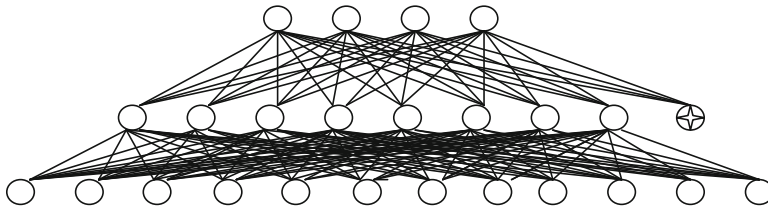
Fig. 2 Structure of the neuron network

Table 2 The result of seismicity in 2008–2038 in the subduction zone

Magnitude	$7.0 \leq M_j < 7.5$	$7.5 \leq M_j < 8.0$	$8.0 \leq M_j < 8.5$	$8.5 \leq M_j$
Possibility	0.958	0.679	0.755	0.323

The training is to fit the expected output for given input by modify the all weights in Eq. 1 systematically for all neurons. The modification is carried out by each input and output set of samples by the Error Back Propagation (B-P) Learning Approach, and the learning rate is selected by Davis-Swan-Campey (DSC) searching approach (Tao and Zheng 1998, 1991; Tao and Du 1992). The network training is completed by 46 modifications with all sample sets one by one.

5   The Result of the Test

After the network trained, the structure of seismicity fluctuation rhythm or called as the knowledge of the fluctuation rhythm is memorized in the weights of the network as a whole, without clear physical meaning for each weight or some part of weights. A forecast can be obtained by input the samples of the last three windows up to present. The preliminary result is listed in Table 2.

If the output value can be considered as possibility from the definition of 1 for at least one occurring and 0 for no one, one can see an earthquake with magnitude more than 8.5 may occur with possibility 32.3%, while the quakes in other magnitude interval may occur more possibly.

6   Conclusions

A simple attempt is deal with in this paper by Artificial Neuron Network tool for long term earthquake prediction in the Japan Trench subduction zone. There was a quite large earthquake occurred last year, the test is assumed for the time period from the data before the shock. The preliminary result shows that the Artificial Intelligence toll is helpful in such difficult a prediction, while statistical model

failed. The AI tool can recognize some kind of rhythm of seismicity fluctuation that people can also find in the time series. The structure of the network and the B-P learning approach are old ones at least 20 years. The forecast ability of AI tool will further improved by new developed method after this international conference.

References

Earthquake Research Committee, Japan (2000) Long-term evaluation of the Miyagi-ken-Oki earthquake (in Japanese)

Earthquake Research Committee, Japan (2009) Partial revision of long-term evaluation of seismic activity from Sanriku-Oki to Boso-Oki (in Japanese)

Earthquake Research Committee, Japan (2011) Long-term evaluation of seismic activity from Sanriku-Oki to Boso-Oki, 2nd edn (in Japanese)

National Research Institute for Earth Science and Disaster Prevention, Japan (2009) Technical reports on National Seismic Hazard Maps for Japan

Tao X, Du W (1992) Seismicity trend along the Southeastern coastal seismic zone assessed by artificial intelligence. Earthquake Research in China 6(1):1–6

Tao Z, Tao X (2012) A review on long-term evaluation of occurrence probability for subduction-zone earthquakes in Eastern Japan. In: Proceedings of the CEABM, Yantai, China

Tao X, Zheng G (1991) An intelligence aided system for assessing seismicity parameters. In: Proceedings of the China-USSR bilateral workshop on the earth dynamics and seismic hazard assessment, Beijing

Tao X, Zheng G (1998) An intelligence aided system for estimating seismicity tendency. Earthquake prognostics strategy, Germany

Working Group on California Earthquake Probabilities, USA (2003) Earthquake probabilities in the San Francisco bay region: 2002–2031

E-Learning Through Cloud Computing: Shaping the Future of Learning for Learner of Tomorrow

Priyanka Soni and Pooja Gupta

Abstract Innovation is necessary to ride the inevitable tide of change. Indeed, the success of the transformation of traditional education system to E-learning education system depends on driving the right balance of Accessibility, Manageability, Reusability, Interoperability, collaboration, and innovation to achieve enhanced knowledge and performance. – And bottom line high-quality, innovative and need-based programmes. Cloud computing takes E-learning to a new level and allows an Education Entity to further reduce costs through improved utilization, reduced administration and infrastructure costs, and faster deployment cycles. This paper tries to clear confusion and hesitation about cloud computing and the value of cloud computing in higher education, This paper also performs SWOT analysis to show main Strengths, Weaknesses, Opportunities and Threats of Using Cloud Computing in Higher Education.

Keywords Cloud • Higher education • Virtualization • SWOT

1 Introduction

No generation is more tech-savvy than today's young people—"Net generation", who have grown up on internet. They come to classroom armed with smart phones, laptops and iPods. Traditional teaching methods are not suitable for this net generation. Today we foresee that internet driven teaching-learning solutions can change the future of higher education. Higher education demands scalable solutions like those provided by cloud computing. The cloud is a next Generation platform that provides dynamic resource pools, virtualization, and high availability. Cloud

P. Soni • P. Gupta (✉)
Computer Application Department, UIM, Dehradun, Uttarakhand, India
e-mail: contactpriya_2007@rediffmail.com; poojagupta010@gmail.com

Table 1 Average fee for cloud computing services

Provider	Average CPU cost (per hour)	Average bandwidth cost (per GB/month)	Average storage cost (per GB/month)
AMAZON	0.11$	0.12$	0.15$
GOOGLE	0.10$	0.11$	0.15$
MICROSOFT	0.12$	0.125$	0.15$

Computing is a way to provide computer applications to users without the need for those users to purchase, installs, or support software on their local computers and/or servers. According to NIST "Cloud Computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction". This cloud model is composed of five essential characteristics, three service models, and four deployment models (http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf). Cloud Computing can be classified as a new paradigm for the dynamic provisioning of computing services supported by state-of-the-art data centers that usually employ Virtual Machine (VM) technologies for consolidation and environment isolation purposes (Barham et al. 2003). There is several cloud computing services providers that offer support for educational systems. Among them are Amazon, Google, Yahoo, Microsoft etc. It is not an easy decision for any higher education varsity to host all their operations and activities online, as even today a large volume of data processed in higher education is still done manually in paper and the staff at most of the colleges is yet to get tech savvy. With more and more institutions realizing the long-term benefits and cost savings of using cloud computing – the transition to cloud based learning has taken a jump-start. At present, the buzzword in the $60 billion Indian software industry is 'cloud computing', and our home-grown technical institutions realize that. While many have already introduced specialized courses or full semesters on cloud computing, some are on their way (http://www.indianexpress.com/news/technical-colleges-cheer-cloud-education/746841/1, Accessed on 16 Feb 2012) (Table 1).

2 Cloud on the Rise

Cloud computing has evolved through a number of phases which include grid and utility computing, application service provision (ASP), and Software as a Service (SaaS) (http://www.wikinvest.com/concept/Cloud_Computing; http://www.computerweekly.com/feature/A-history-of-cloud-computing). The big question is: what is different about the cloud? The first key difference is technical: the maturity of standards throughout the stack, the availability of high performance network capacity, and the emergence of virtualization technologies are combining to enrich the sourcing options at our disposal. Some educators have wrongly assumed that cloud computing refers to any services provided over the Internet which are not

hosted by their institution. In addition there is confusion between the terms Web 2.0 and cloud computing. Web 2.0 can be regarded as a particular type of application while cloud computing is a method by which applications and data are hosted and delivered (http://net.educause.edu/section_params/conf/CCW10/highered.pdf). The generation that has grown up with information technology referred as generation Y or the Net generation reflects what makes them really different. The aptitudes, attitudes, expectations, and learning styles of Net Gen students reflect the environment in which they were raised—one that is decidedly different from that which existed when faculty and administrators were growing up (http://www.educause.edu/educatingthenetgen). If the Net Generation values experiential learning, working in teams, and social networking, what are the implications for classrooms and the overall learning environment? They are more likely to embrace the idea of cloud-based services in their enterprise roles, just as they embrace them in their private lives. Such users—who are driving the rising sales of netbooks—are likely to fuel the drive toward lower cost and lightweight computing clients and applications (http://net.educause.edu/section_params/conf/CCW10/highered.pdf). Why Is Cloud Computing Important? Cloud computing is becoming very important in present scenario due to on demand infrastructure, applications and its support services. Basically it is next generation of internet computing. McKinsey suggests that "using clouds for computing tasks promises a revolution in IT similar to the birth of the web and e-commerce." Burton Group concludes that "IT is finally catching up with the Internet by extending the enterprise outside of the traditional data center walls."

- Driving Down total costs by facilitating the on demand or pay per use services.
- Fostering further IT standardization.
- Save time by reducing IT supply bottlenecks.
- Creating a pathway to $24 \times 7 \times 365$ environment and Offers universal document access.
- Instant software updates and Latest version availability.
- Helps to use application without installation.
- Effective computation by centralizing storage, memory processing and bandwidth.

3 Adoption of Cloud Computing in Higher Education

There biggest challenge for Higher Education is Traditional lectures are not fulfilling the learning potential of typical students today. So what do Net Generation want from learning technology? Interactivity—whether it is with a computer, a professor, or a classmate. They want it; they crave it. Distance education and online courses don't work well with Net Generation—the social component of learning is required. As technology in the classroom progresses, more and more students are going to demand it be included. This will pose challenges, though (http://net.educause.edu/ir/library/pdf/pub7101.pdf). As a solution to these challenges and to achieve faster and hassle-free campus-based activities, many colleges are now
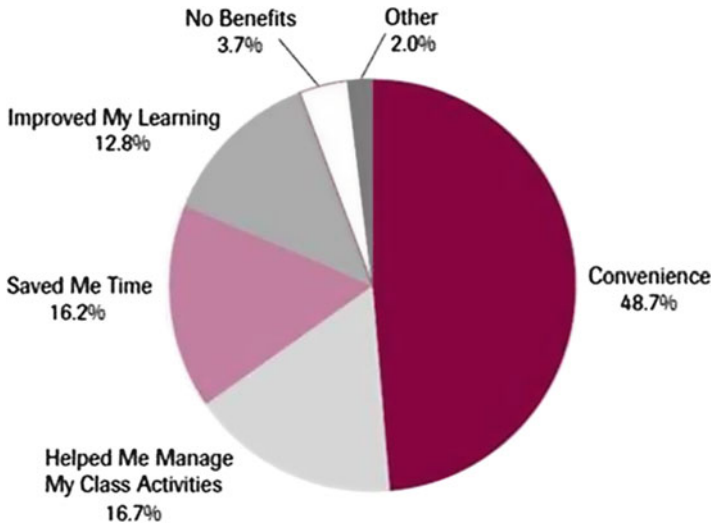
Fig. 1 Depicts benefits from the student perceptions when IT is used in the classroom (http://net. educause.edu/ir/library/pdf/pub7101.pdf)

increasing the use of IT in classrooms and campus management solution. This allows students to access all the campus-related information easily (Fig. 1).

Cloud based teaching-learning solutions playing a key role in the future of higher education. The reasons for their success would be their reach, efficiency, scalability and affordability. Cloud computing allows institutions to information from anywhere in the world in a matter of seconds. In the field of education, this is pretty important as it gives the teachers and the learners to constantly update their knowledge. Renowned names like IITs, IIIT-Hyderabad, and IIIT-Delhi comprehend the importance of cloud computing in their educational curriculum and research projects, along with the infrastructural support that cloud can provide them (http://www.indianexpress.com/news/technical-colleges-cheer-cloud-education/746841/1, Accessed on 16 Feb 2012). Practically it is difficult to move traditional education system to cloud but educational institutions seeking long term benefits using cloud computing. As competition is rising amongst the educational institutes to attract the best students and offer world-class campus environment, cloud computing offers the tech savvy students an e-friendly learning environment.

4 Main Benefits and Limitations of Using Cloud Computing in Higher Education: SWOT Analysis

Higher education in the future will need to come to the grips with globalization and will be subject to worldwide competition. It is not an easy decision for any higher education varsity to host all their operations and activities online.

Strength	Weakness
Support for teaching and learning	Not all applications run in cloud
Software free or pay per use	Risks related to data protection and security
24 h access to infrastructure and content	Dissemination politics, intellectual property
Access to application anywhere, anytime	Lack of confidence and Trust
Echo-friendly by using green technologies	Standards adherence
Increased openness to students to new technologies	Speed/lack of internet can affect work
Increased functional capabilities	

Opportunities	Threats
Way to business environment and advanced research	Data security is the major threat
Better learning environment for students	Unsolicited email or advertising
Offline usage with further synchronization opportunities	Lock- in companies such as Google and Microsoft allow institutions to co brand their cloud products
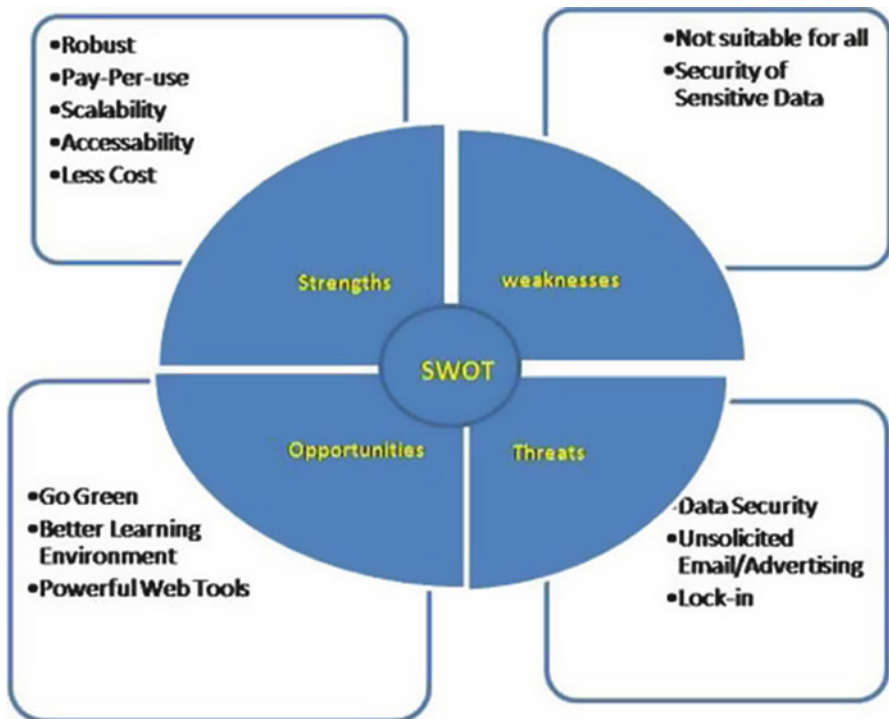
5 Conclusion

Although cloud computing promises a great change in Higher Education but education community needs guidance to fully adopt cloud computing in academic Arena. To use solutions provided by cloud computing, the educators, administrator, and students need to be trained. Different clouds have different implications, so choosing the right cloud service that goes best with the need of your organization is the concerned area. Cloud Computing is undoubtedly a growing revolution within IT. However, as with any paradigm shift, there are strengths and weakness which need to be finely balanced. To take full advantage of Cloud Computing the students can go for on demand web pages. Students are benefitted by removing, license cost, hardware cost and maintenance cost. It also facilitates flexibility to educational institute and universities and offers new features and tools.

Reference

Barham P, Dragovic B, Fraser K, Hand S, Harris T, Ho A, Neugebauer R, Pratt I, Warfield A (2003) Xen and the art of virtualization. In: Proceedings of the 19th ACM Symposium on Operating Systems Principles, SOSP 2003, Bolton Landing, NY, USA, p 177

The Research on Image Extraction and Segmentation Algorithm in License Plate Recognition

Fang Weijian and Xin Zhou

Abstract As a vital part of Intelligent Transportation System, License Plates Recognition System is meaningful in Vehicle Positioning and Traffic Monitoring. It consists of Vehicle License Plate Location, Character Segmentation and Character Recognition. In the paper we put forward a new method of segmentation algorithm on the basis of comparison and analysis some common methods of License Plate Recognition, which firstly chooses the self-adaptive thresholds for license images and executes the binarization; then arranges with projection Method and fixed-edge methods on the basis of connected-area methods. The experiment's result shows this new method has better recognition performance.

Keywords License plate recognition • Binarization • Image segmentation • Segmentation algorithm

1 Introduction

The character segmentation method in license plate image is the premise of next step, character recognition, and it determines the character recognition rate. So, the step of the character segmentation attracts great attention around the world. Researchers promote kinds of solutions on the basis of the existed printed character segmentation technology (Hong Zhao et al. 2004). Nowadays, one popular way of segmentation is utilizing the fixed aspect ratio of character and the character

F. Weijian (✉)
Network Center, Chongqing Three Gorges University, Chongqing, China
e-mail: wanzhoufwj@yahoo.cn

X. Zhou
College of Physics and Electronic Engineering, Chongqing Three
Gorges University, Chongqing, China
e-mail: wanzhouzx@163.com

space as the Priori Knowledge (Lu Yi 2003); The threshold (T) is determined by Mathematical Expectation (E) and Variance (D): T = E−D. E and D got from the Projection-histogram of the last step. The home position is determined when the background pixel satisfies the inequality piexl[i] < T < piexl[i + 1]. The ending position is determined when the background pixel satisfies the inequality piexl[i] < T< piexl[i − 1]. In this way, all the characters can be determined; As it shows above, the threshold is important for this method's segmentation performance. So this method needs researches on every segmentation results and the readjustment if there's a specified amount of deviation (Casey and Lecolinet 1996). Firstly, the gray histogram on the basis of vertical and horizontal direction's projection must be concluded (Haijiao Wang et al. 2008). Secondly, choose the suitable threshold to segment the license plate image and character according to the gray histogram. Then the character segmentation combined with aspect ratio and the nor malized transact can be done. So the next step of feature extraction can be easily achieved. This paper will do the binarization of license image, and then project the binary image at the vertical direction. At the same time, I take the advantage of characters' aspect ratio and the character space along with contour analysis to cut the character of license plate image.

2 Choosing the Suitable Threshold and Doing the Binarization

The license plate character segmentation performance is determined by the binarization. The key factor is choosing the threshold. The background and the object are distinguished by the threshold. The Binarization Formula is:

$$g\left(x,y\right) = \begin{cases} 1 & f(x,y) \geq T \\ 0 & f(x,y) < T \end{cases} \tag{1}$$

f(x,y) is the Original Gray Images; g(x,y) is the Transformed Image.

There are many kinds of traditional ways to choose the threshold, but most methods need a prior probability distribution of destination and background. Then the optimal threshold which satisfies this distribution can be achieved by utilizing the Bayesian formula (Jin Zhang et al. 2008). But it is an impossible way for the real object for the statistic of license plate target points is a difficult problem to deal with. What's more, the license plate background values can affect the threshold.

Different ways to choose threshold according to thread:

1. Probability and Statistics, if g is threshold, the object points are consisted of points < g. P is the ratio of object points and all the points. The gray value is 0, 1, 2, . . . , g, till the ratio closing to P. This threshold is image threshold (Ni Fuyin 2011).

2. The average value of maximum gray value and minimum gray value can be simply regarded as segmentation threshold of the image when dealing with high quality image whose gray level-differential value is large (Yu Zhang et al. 2006).
3. If g is threshold, there are more target points and background points while fewer points whose gray value is g. In this case, the gray value of minimum points can be determined as threshold. This is the trough principle of gray histogram. The object and background take most of the points, so the two peaks of image can be achieved. On the other hand, the points which are close to threshold are fewer to consist of the trough of the image. If the image is complicated changing, looking for the trough is the most important part of this method.
4. When transiting from background to objects, edge points are the points with greatest change of gray value. In this case, the edge gray value can be determined as image threshold. This method is used when doing the local threshold segmentation.

This essay focuses on gray value histogram iterative method to look for trough on the basis of enhanced treatment of image. The middle point of average gray value of background and average gray value of character is threshold (T). The formula is (Lifeng Zhang and Jun Dai 2008):

$$T = \frac{1}{2} \left(\frac{\left(\sum_{i=1}^{T-1} i \times H(i) \right)}{\sum_{i=1}^{T-1} H(i)} + \frac{\left(\sum_{i=T}^{N} i \times H(i) \right)}{\sum_{i=T}^{N} H(i)} \right) \quad (2)$$

H is gray value histogram. N is maximum gray value (as 255). The formula above can conclude that the threshold is in trough. Then do the iteration till TK = TK + 1. TK is threshold. The formula is:

$$T_{k+1} = \frac{1}{2} \left(\frac{\left(\sum_{i=1}^{T_k-1} i \times H(i) \right)}{\sum_{i=1}^{T_k-1} H(i)} + \frac{\left(\sum_{i=T_k}^{N} i \times H(i) \right)}{\sum_{i=T_k}^{N} H(i)} \right) \quad (3)$$

The binarization treatment is after the threshold determined. The procedure is easy. The gray value which is less than threshold value is gray region. The gray value which is more than threshold is objection. The license plate gray value and binarization image are as follows (Figs. 1 and 2).

As above images show, some pictures are black character with white background while other pictures are white characters with black background after binarization treatment. To explain this phenomenon, we notice that there are four kinds of color: white character with blue background, black character with yellow background,

Fig. 1 Car plate gray

Fig. 2 Binarization image

black or red character with white background and white or red character with black background. When the license plates are black character with yellow background and black or red character with white background, after turning them to gray image, the gray value of objection is less than background. In this case, the technology of image determination is needed to transform the situation of black character with white background.

Choosing threshold is the key step of Binarization treatment. To find better ways to choose threshold, I analyze the histogram through Laplace Histogram, Difference Histogram and Gradient-histogram, but the result is barely satisfactory. So I still use the traditional histogram algorithm method.

3 Character Segmentation Algorithm

There are many ways of character segmentation Algorithm such as the License Plate Segmentation algorithm by Rough Transition and Priori Knowledge; They are complex. It is important to find a better algorithm of faster speed and higher segmentation rate for the system as license plate recognition with high timeliness requirement (Zhao Peng and Zhao Zhong-meng 2011).

3.1 Fixed Bindery Segmentation

The size (aspect ratio) of license plate is certain, so as the size of character, script, number and distribution. The segmentation can be done according to the location of character in the image under the situation of location accurate, image capture fully and variant little.

Fixed bindery segmentation is easy, not affected by noise and independent on the condition of image. However, the requirement is high on the size accurate. Or the character is not entire and make bad influence on the character recognition.

3.2 Projection Segmentation

The principle of projection segmentation firstly turn the license plate image to Binarization image (black is 0, white is 1). Then the vertical projection can be achieved through accumulating the pure gray value from row's horizontal direction. The gray value between characters is always 0. The projection image can form to trough or form to a special wave crest. Find out the valley points between two peaks. Use them as the position of character segmentation (Jun Tao 2008). Then complete the character segmentation. The process is as follows:

Assuming binary image is M × N matrix, its vertical projection:

$$D(j) = \sum_{i=1}^{M} f(i, j) \tag{4}$$

f(i, j) is pixels. (i, j) is gray value. Then do the differential operators to the projection vectors.

$$\begin{cases} \mathit{Diff}(j) = 1 \ldots G(j+1) - G(j) < 0 \\ \mathit{Diff}(j) = 0 \ldots G(j+1) - G(j) = 0 \\ \mathit{Diff}(j) = -1 \ldots G(j+1) - G(j) > 0 \end{cases} \tag{5}$$

Search the graphic coordinates matching recognition in the vectors Diff.

$$D(j) = \sum_{s=1}^{S} \sum_{t=1}^{T} [\mathit{Diff}(j-t) - M(s, t)] \tag{6}$$

D(j) is similarity. S is template number. T is factor number in template.

The trough between characters will be greatly affected because of the noises. The result is the bottom is not obvious. Even seriously, bottom will disappear resulting in character segmentation error or mistakes. Ultimately affect the character recognition. Therefore, the method of the plate is relatively high image quality requirements. Even already removed the border and rivets, there still have noises when interception of the general license plate, and sometimes the inaccurate judgment will affect the trough.

3.3   Connective Area Segmentation

Connective area method is based on the principle of mathematical morphology and utilizes the structuring elements to collect image information in order to understand the structure of the image.

The basic morphological operations are: corrosion and intumescent. The corrosion makes the objection reduced, the inter-bore increased and eliminates the external noise isolation. The intumescent makes the objection increased and the inter-bore reduced to form connected area. B is structural elements. A is working range. The formulas are as follows:

$$\text{corrosion } A \ominus B = \cap \{A - b : b \in B\} \tag{7}$$

$$\text{intumescent } A \oplus B = \cap \{A + b : b \in B\} \tag{8}$$

Another two important methods are opening and close operation. Corrosion before intumescent is opening while intumescent before corrosion is close operation. The formulas are as follows:

$$\text{opening AOB} = \{A \ominus B\} \oplus B \tag{9}$$

$$\text{close operation } A \bullet B = \{A \oplus B\} \ominus B \tag{10}$$

This essay combines corrosion and opening to do the license plate character segmentation.

The formula is:

$$X \oplus S = \left\{(x, y) | (S^{\wedge}_{(x,y)}) \cap X \neq \varnothing\right\} \tag{11}$$

Find the point, then recursive using the formula above till the set of elements of the character pixels are no longer increasing. Its advantage is independent on tilted image. What's more, the positioning is accurate. The drawback is that it requires high quality images and converts a good binary image. The two characters are not linked together in the binary image.

3.4   Integrated Algorithm of Character Segmentation

Connective area method, comparing with the projection and fixed boundary method, is more accurate but more sensitive to noise, when one character is next to another, it will fail in segmentation. However, projection and fixed boundary method are not sensitive to this. Character segmentation are compared as below (Table 1):

Table 1 Character segmentation comparation

	Fixed boundary method	Projection	Connective area method
Extra boundary	√×	√	√
Inaccurate threshold value	√	×	×
Noise	√	√	×

Table 2 The contrast of two algorithm

	Average recognition rate of traditional standard template		Average recognition rate of improved template	
Template matching	Standard images of license plate (%)	On-the spot images (%)	Standard images of license plate (%)	On-the spot images (%)
Chinese characters	96.24	94.45	96.56	96.36
English characters	97.13	95.63	97.46	97.27
Numbers	97.74	96.05	97.92	97.20

To achieve the optimal segmentation effect, This paper has proposed an integrated method to enhance it. The figure above has shown the best ways to assemble the three methods to achieve the optimum.

Procedure:

1. To transfer Gray Images to Binarization Images using dynamic threshold approach.
2. Scan to the right at one seventh of width of the license plate image by connective area method, to see whether six connective components could be found.
3. If failed, to determine the position of characters using projection.
4. If failed again, to break up characters using fixed frontier approach.

Connective area method acts as the main method of detection for its success rate of detection is particularly high, when the input image is of low noise or without black area, which is similar to the characters. Static boundary method, as the last method, is based on the horizontally-arranged characters. Because of the first character of the license plate is Chinese and of no connective characteristics, it is detected using projection.

This paper stated all the improved template matching method that has been utilized in the License Plate Recognition. The experiment was conducted using Matlab emulator. The original statistics are 100 original standard images of Vehicle Administration Office and 100 on-the-spot images of Toll Station. The results of the experiment are as below (Table 2):

We can see from the above figure that the algorithm we proposed in This paper has improved self-stability and timeliness, especially the on-the-spot images with high timeliness.

4 Prepare Your Paper Before Styling

This paper introduces the main methods of image extraction and separation, proposing an integrated method of character separation. We popularize the traditional connective area method, combining the merits of projection and fixed boundary method, to make mutual complementation in calculating speed and separating quality。 Experiments have shown that this method, based on the enhanced effect of image separation, accelerate the image processing.

References

Casey RG, Lecolinet E (1996) A survey of methods and strategies in character segmentation. IEEE Trans Pattern Anal Mach Intell 18(7):69–706

Haijiao Wang, Wenju Li, Xinnian Wang, Danxiao Jia (2008) Edge color-clustering and neural network-based license plate style identification. Comput Eng Appl 44(26):196–199

Hong Zhao, Liming Wang, Gongyi Wang (2004) Research on binarization method of license plate recognition in automatic recognition. Appl Tech Sci 31(3):15–16

Jin Zhang, Zhiming Zhang, Fan Huang (2008) An advanced adaptive template matching method. Microcomput Inf 24(27):166–168

Jun Tao (2008) Exploration and implementation of vehicle license plate automatic recognition. Master Degree Dissertations, Wuhan University

Lifeng Zhang, Jun Dai (2008) Research of the licence plate recognition technology in the intelligent traffic control system. Microcomput Inf 24(33):312–314

Lu Yi (2003) Machine printed character segmentation-an overview. Pattern Recognit 28(1):67–80

Ni Fuyin (2011) Design and realization of one imbedded vehicle license plate recognition system. J Jiangsu Teach Univ Technol 17(2):18–22

Yu Zhang, Siliang Ma, Xiao Hang, Zhongbo Zhang (2006) Image extraction and segment arithmetic of license plates recognition. J Jilin Univ (Sci Edn) 44(3):406–410

Zhao Peng, Zhao Zhong-meng (2011) License plate recognition method base on character dendity. Microelectron Comput 28(1):154–156

Analysis and Applications of Campus Network Flow Control

Hu Zhonghua and Fang Weijian

Abstract Assisting the teaching and administration greatly, campus network is becoming more and more popular. However, more and more problems appear. Especially the flow control becomes a key problem for the campus network management. The paper analyzed the principles of flow control and its security management, and proposed the solution of broadening and regulating.

Keywords Digital campus • Flow control • Flow strategy

1 Introduction

With the booms and popularizing of network technology and applications, the scale of campus networks grows to a level of what the developed countries' have. As the most important career of digital information, the campus network is being occupied by a lot of non-critical applications, such as P2P applications. At present, campus network is the one of the network which contains the most P2P applications. It should be something right for student to think actively and try something new. However, it brings troubles to the network management. What's more, the problems of undesirable websites, cyber-attack, etc., cannot be evaded during the management of campus network (Fig. 1).

H. Zhonghua (✉)
Financial Department, Chongqing Three Gorges University, Chongqing, China
e-mail: huzhonghua@21cn.com

F. Weijian
Network Center, Chongqing Three Gorges University, Chongqing, China
e-mail: wanzhoufwj@yahoo.cn

flv (30.4%)
http-mirror (18.28%)
others_tcp (9.27%)
qvod (6.55%)
http-browse (6.12%)
super cyclone (5.64%)
xunlei (5.59%)
qq (5%)
ppstream (3.72%)
http-others (3.5%)
http-part (3.23%)
http-download (3.02%)
bittorrent (2.33%)
others_udp (2.15%)
pplive (0.9%)
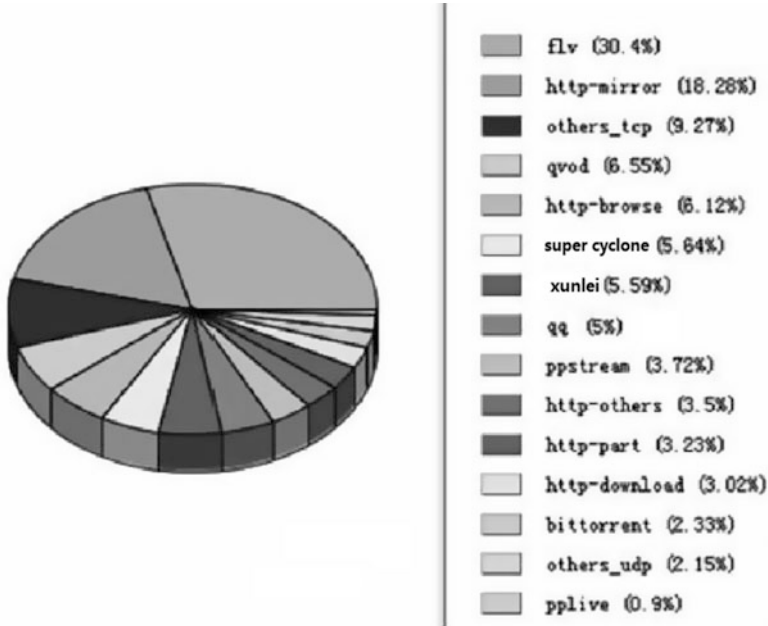
Fig. 1 Analysis of outlet flow

Digital campus bases on a suitable network speed. Although the speed grows to hundreds of megabytes, some even to gigabit, applications are updating as well. Without the proper distribution and management of bandwidth, the bandwidth resources would be wasted, and the running of regular applications might be slowed down. The flow grows quickly by the applications of VOIP, Video conference and OA. Meanwhile, the uncontrolled downloading, P2P application and worm occupy the bandwidth; even might cause meltdown of network. It could not be solved by just broadening the bandwidth. Without optimizing the controlling measure, it would only lead to a higher operating cost (Zhang Weidong et al. 2005).

2 Major Causes of the Problem

2.1 *Weak Protection from Virus*

The restore card with a virus killer is a common method to protect the campus network from malicious software and virus. It is easy to set up and operate. However, its virus killer could not be updated. When the computer turns on, the worm in the network may reproduce and produce a lot of packet to block the network.

2.2 *Excessive Downloading and Video Streaming*

It is popular to download music and movies by some downloading applications. The P2P downloading applications require a lot of bandwidth, while some students play computer games and watching videos online, the bandwidth for teaching applications is fewer.

2.3 *Illicit Links*

The number of cyber attacks of consuming bandwidth resources grows quickly. Campus network, with weak protection, the illicit links consumes the bandwidth and make the server out of service. What's worse, most illicit links are attached with Trojan applications, which also could infect computers and lead to the block of network.

Hence, we should control the outlet flow by broadening and regulating to realize a high efficient consumption of outlet bandwidth.

3 Broadening

It is common to downloading by P2P on the internet. The campus network also has an amazing number of outlet flows by downloading. The campus network should contain more downloading servers to enrich the in-campus resources (Ning Tianqiao1 and Luo Jie2 2010):

3.1 *Enhancing the Protection*

Regular and continuous updating of operation system is necessary. Users often download system patches outside the campus network, which take users more time and also bring pressure to the outlet bandwidth. In order to solve the problem, CTGU Network Center bought Kingsoft online virus killer to help the users save time and bandwidth.

3.2 *Available Downloading of Applications*

The Network Center also set up a software downloading server, providing the downloading and updating of small-scale software in common use. The users could

download the software in campus net safely, which reduce the infection of virus and release the pressure of outlet bandwidth.

3.3 Setting Up VOD System

The popularization of BT, Thunder and HDV, the bandwidth of campus network was almost occupied by stream media and video downloading. The campus network was slowed down. In order to meet with the users' requirements of entertainment and less the downloading, the Network Center of Chongqing Three Gorges University sets up the system of Boful Truran VOD and uploads the popular TV programs to it. Some of them are HDVs. It shares the videos and the downloading reduces, while the bandwidth resource is saved.

3.4 Providing Online Broadcasting

Live broadcasting is another critical factor influence the bandwidth. Especially at the time of great events or sports match, it occupies a lot of band width. The Network Center opened online live broad casting on the net, providing 8 channels in common use to solve the problem.

However, compared with the regulating, the broadening could only release the outlet pressure in some scale.

4 Regulating

4.1 Set Up a RT Analysis System to Observe the Abnormal Flow

Cyber attack and virus often appear in a network. It is difficult to analyze or prevent technologically. However, it is much easier to observe by statistical significance. The number of cyber attacks of consuming bandwidth resources grows quickly. Campus network, with weak protection, the illicit links consumes the bandwidth and make the server out of service. What's worse, most illicit links are attached with Trojan applications, which also could infect computers and lead to the block of network. The system could be used to manage the outlet information, observe the flow, and display the flow by chart or HTML. Hence the users could see the flow directly. The function of alarm in the system could provide the alarm service of abnormal flow (Ren Fengyuan et al. 2003).

4.2   Distribute Different Applications with Different Bandwidth

Web is the main application, so it gets a priority on bandwidth distribution. For example, it's minimum bandwidth is 100K, standard of 2M, and maximum of 10M. When the network was congested, 2M could be provided; when it was free, 10M; when WEB flow was few, some other applications could use it's bandwidth. The maximum for Thunder is 5M, while the other P2P applications are limited to 1M. Even more strict limitation could be proposed, such as the maximum uploading of Thunder is 0, which means to change the P2P style to a single-track flow; or limits the P2P flow to an amount of which is less than 30% of the total until the network is free (Bin Cheng et al. 2010).

4.3   Upper Limit According to IP

It is used to solve the problem of some users occupy too much bandwidth. The users could be divided into different groups, such as teacher group, student group, network management group. Different groups have different upper limits. The teacher group and network management group are prior than the student group. The bandwidth can also be distributed by different times. The student group may have more bandwidth in the evening. The linkage limitation is similar to the IP limitation. Different IP has different linkage limitations. It should be noticed that the linkage should be more in the computer center, for each IP in the computer center contains a lot of links. Although applications of P2P and streaming media are controlled and limited, some user's daily flow is over 10G (Zhou Zhongwei 2009). A blockade system developed by the Network Center was in use to warn and blacklist the user. The blockade would be over in 24 h. At last, the bandwidth could be distributed by different times. During the working time, office group and teaching group owes more; while the other time for student group and teachers dormitory (Zhao Shuguang 2010).

Besides the controlling and limiting, ACL could be used at the outlet devices to limits the outside users linking to then in-campus users. It could help to keep away from the virus and garbage flows (Fan Peiying 2010).

5   Conclude

The controlling strategy determines the bandwidth distribution. The proper strategy could help to release the problem of bandwidth occupying by BT, Thunder and video streaming. The bandwidth is distributed properly. The critical applications and service are guaranteed while the teaching, administrating, researching could be in good condition.

References

Bin Cheng, Guoqiang Wei, Guangying He (2010) Campus network traffic monitoring and analysis based on application layer. J Shanghai Univ Electr Power 26(1):75–78

Fan Peiying (2010) Flow control of campus network double outlets. Value Engineering 44(21):78–91

Ning Tian-qiao1, Luo Jie (2010) Deploys the flow control equipment to build high-effective campus network. Sci Technol Inf 89(11):250–252

Ren Fengyuan, Lin Chuang, Liu Weidong (2003) Congestion control in IP network. Chin J Comput 26(9):1025–1034

Zhang Weidong, Wang Wei, Han Weihuan (2005) Design and completion on network flow measurement and monitor system. Comput Eng Appl (32):160–163

Zhao Shuguang (2010) A method of flow control based on RTCP. Inf Res (5):45–47

Zhou Zhongwei (2009) Applicaiton research of flow control in campus network. J Huan Ind Ploytech 9(5):26–27

A Knowledge Base System Model for Dam Safety Monitoring

Tiesheng Wang and Pengli Cheng

Abstract In the base of the detailed analysis of dam safety monitoring expert system of knowledge, it is introduced that the method of expert system structure of knowledge and the characteristics of knowledge's representation. Maxing the method and ANN to indicate dam safety and the physical structure of the knowledge base is studied. The reasoning strategy based on hybrid knowledge base is also discussed and the technical's problems which forming mixing knowledge base should be paid attention to is analysed.

Keywords ANN • Dam • Knowledge base • Safety monitoring

1 Introduction

Our expert system for the dam safety monitoring work are studied in the 1990s. So far, it has achieved remarkable social and economic benefits. In the development, many problems is encountered and the knowledge acquisition and maintenance is the most prominent. Therefore, the construction method of knowledge base for further research is particularly urgent.

T. Wang (✉) • P. Cheng
North China University of Water Resources and Electric Power,
Zhengzhou 450000, Henan Province, China
e-mail: 707413719@qq.com

2 Knowledge and Knowledge Representation

2.1 *Knowledge Representation*

Knowledge representation indicate that the knowledge is instead of some appropriate data structure (Cios et al. 2005). In order to facilitate knowledge management. Knowledge expression need be simple, clear, easy to understand and flexible to expand. There are several methods which are different in character and ability to express knowledge. In the following chapter, the production rule and ANN is introduced.

2.1.1 Production Rules

The general form of production-type rules are as follows: if <premise> then <action or conclusion> with CF (Kheifits). There are two main parts of the rules: premise and conclusion. Premise is condition. If the premise is true, the rules are operational and the conclusion is got. Activation of the rules influence the knowledge base by adding or modifying the facts. The rule is influenced by its certainty (CF value). It define the new thing's CF value that the premise's CF value and the rule's CF value.

The advantage of the production-type is close to the way of people thinking. Knowledge representation is intuitive, natural, and easy reasoning. But production-type is weak in the level of expression power. It can't omit the relation which have been identified in inferential process. Thereby, it reduces the efficiency of reasoning.

2.1.2 Neural Network Representation

In the study of neural network, the representative is "Parallel Distributed Processing System". This model suppose that information processing through a large number of the interactions of simple processing elements called "unit" to conduct. Each unit sent excitable or restrained signal to the top of the unit. Parallel means that all the targets are processed at the same time. Distribution means that each node and the connection is expressed only partial information, rather than a complete concept.

Neural network knowledge has the following advantages. It express information by its distributed manner. Any knowledge of the rules can transform into number forms. This facilitate the organization and management of the knowledge base. Automatic acquisition of knowledge and adaptation to environmental change is also easy to achieve. It implements the knowledge representation, memory and reasoning carried out by the same neural network. Neural network knowledge representation drawback is that the reasoning process is not clear for a given input. Therefore it is difficult to explain (Panizzo et al. 2009).

3 The Knowledge of Dam Safety Monitoring and Its Classification

Dam safety monitoring expert system relates to structure, safety monitoring, artificial intelligence and other disciplines. It mainly include the following aspects:

1. The knowledge about evaluation of design standard. Such knowledge is mainly used for review of the overall hub design flood standards and seismic design standards. Such knowledge usually use less. When hub structure or the hub regional environment changes, it will have a systematic analysis according to the specific engineering.
2. The knowledge about evaluation of structural safety. the knowledge is mainly used for building strength and stability of safety evaluation. When the building's structure materials properties basic state operating environment and other significant changes occurred, such knowledge is applied on the part or whole building safety evaluation. but such knowledge usually use less.
3. The knowledge about evaluation of monitoring data. Knowledge which is diagnosis and analysis of monitoring data is the dam safety monitoring expert system using one of the most frequent. Such knowledge is mainly used for monitoring the reliability of physical examination. It can help to find abnormal monitoring data timely and to analysis of what cause abnormal monitoring data. At present, it mainly include: the criterion of space-time evaluation, the criterion of monitoring model evaluation, the criterion of monitoring index evaluation, the criterion of mechanical rules evaluation, etc.
4. The knowledge about evaluation of daily inspections. Daily inspections usually carry out by the person responsible. According to the level and characteristics of the building, model of conduct have the clear checking item and a dedicated inspection checklist and need be filled by the inspection personnel at the scene.
5. The knowledge about physical causes of abnormal data. Currently, the causes of abnormal values is divided into three cases. Namely: monitoring system of abnormal, external causes of abnormal and structural changes caused by the abnormal state. The case that monitoring system cause the measured value is frequently encountered in dam safety monitoring. It is mainly caused by the measuring instruments' damage and calculation errors. It does not affect the safety of the structure. This kind of mistake can be solve with correlation analysis or spot check of the monitoring system. External exception which is caused by the great change of abnormal water level, temperature, rainfall, earthquakes external factors affect the dam of the work of state. At present, the primary means of the external analysis is finite element analysis. Structural changes caused by the modal analysis of the abnormal values is internal analysis. It is a further analysis after the system get rid of the first two abnormal values. If the test was abnormal, the inference should inferre what changes in the structure and provide a reasonable explanation. Such knowledge is the most difficult to collect and arrange in the expert system.

6. The knowledge about evaluation of state building. Such knowledge mainly assess the work of state building. Since the specific requirements of each project and characteristics of each building is different greatly. Therefore, the judgement should be specific for each project.

4 Knowledge Construction of Dam Safety Monitoring

4.1 Knowledge of Dam Safety Monitoring

For the six dimensions of knowledge described above, The first two categories of knowledge is rarely used. The next two is frequently used. The first two parts of the fifth class knowledge is also used more. This knowledge is essentially a deep knowledge and the rules can easily express it. Because the traditional expert system based on the rules structure technology has quite mature, Therefore its constructor is no longer tired in this state.

For internal analysis and comprehensive evaluation of knowledge. Because it is difficult to collect and collate and knowledge of different projects requirement are quite different, using the rules to describe such knowledge have some difficulties. But if using the neural network to describe the relatively is relatively easier.

Neural network expert system knowledge base include two processes: knowledge of acquisition and knowledge of storage. Knowledge acquisition refers to the gain and choice of training sample. The training samples are from characteristic parameters of the same type of object in the normal operation and operation with fault. The training sample need to be compatible and ergodic. Neural network expert system store the knowledge in the connectible weights and threshold of the neural network. In the training process of neural network expert system, the output values based on the teacher sample compare with the target output, then calculation error is output. Last the connection weights and threshold of the neural network is adjusted accorded the calculation error and until it meet the specified requirements. After the adjustment of the connection weights and threshold matrix, the expert system is developed (Tianqi Yang 2005).

In summary, the knowledge base use two kinds of methods to indicate, one is using rules to indicate design standards, safety evaluation, judgment of abnormal values, external analysis, etc. Other is using the neural network to express the internal analysis, the workability modal analysis, etc.

4.2 Reasoning Mechanism

Comprehensive analysis of the reasoning mechanism is divided into two parts. The first part is that it accords the knowledge base of rules based on knowledge of the original monitoring data (including Monitoring instruments measuring data, Manual inspection data and Safety evaluation results data) to check abnormalities,
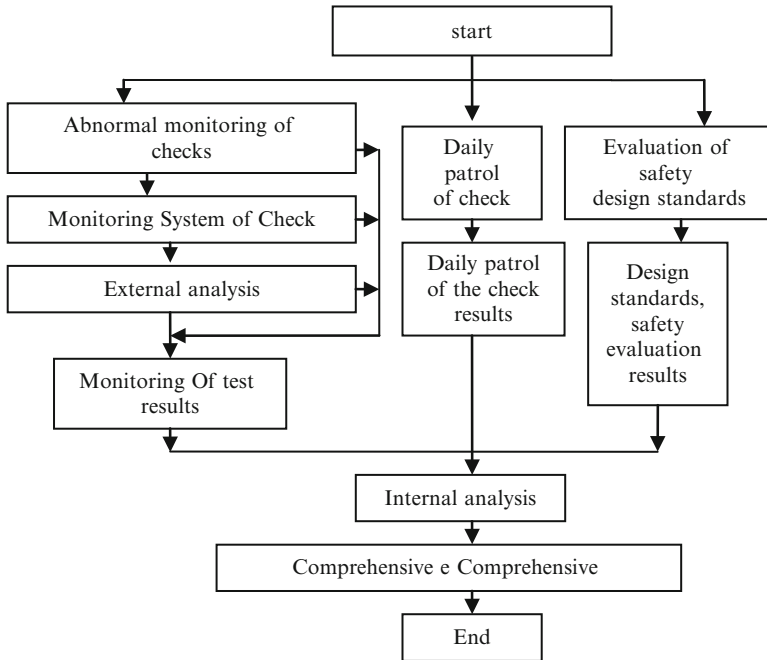
Fig. 1 Comprehensive evaluation of reasoning flowchart

search and inference. The measured value of abnormal data which is caused by monitoring systems and the change of the external environment is analysed and excluded and we strive for the results of judgement from the error of measurement data. In the second part, it accord reasoning results of the first part and the neural network knowledge to analyse abnormal measured values which is causes of the internal structure and to infer the location and extent of structural abnormalities. In addition, it base on the results obtained to evaluate the whole building workability state further.

Comprehensive process of reasoning inference engine is shown in Fig. 1.

5 The Formation of Neural Networks

5.1 Principle

In the neural network expert system formation process, there are several key technical issues to be resolved. (1) The determination of input layer neurons; (2) The output layer data form; (3) The collection and collation of training samples. These issues should be based on the specific circumstances of the case project. Here to make a general discussion (Xu Ri-qing 2000).

The neural networks have many forms. Among them, forward multilayer neural network based on the BP algorithm is the most common. It constitute of one input layer, one hidden layer, an output layer. The Layer connect with different weights. After the training, this neural network can learn the input—output mapping relationship from the training sample and approach any function with arbitrary precision.

In expert system based on neural network. It generally regard evaluation results of test points as symptom data. That is to say the system see monitoring the amount of test results, the results of manual inspection checks, safety test results as the input data. Values of the input data range $[+1, -1]$. The general value of the output data also range $[+1, -1]$. As buildings measuring points is to many, therefore it is not appropriate to all measuring points as a network for training, so we must accord the project site to select a representative measurement point.

5.2 *Example*

The main monitoring sections of a dam has laid displacement, seepage, seepage pressure, stress and other monitoring equipment. Accordance with the actual situation of project, we select 4 displacement measuring point, a seepage measurement points, two osmotic pressure measuring points, as well as the daily safety inspections and test results as input data. The principle of input data values is that: the value of measuring points measured normal take $+1$; equipment damage or data measurement error taken 0; non-measured factors cause the measuring point's anomalies take -1; labor inspections and safety checks normal take $+1$; fundamental normal take 0; behavior disorder (dangerous) take -1.

Taking many years measurable data and considering the requirements of ergodicity, 150 sets of data is selected as training samples. Then passing iterative calculation, we get neural network knowledge base. For the future's Measured data, after evaluation by the rule knowledge, we can get its normal or abnormal characteristics. According to characteristics of the measured point, it constitute the input data set, then using the trained neural network to get the output data. If the input data is $(1,1,0,1,1,1,0,1,1)$, the output is $+0.96$ that indicate that the work of the project site is normal. Using of the network to judge the 3 years measured data, the accuracy rate is 92% that show that the method will have better prospects after it was further refined.

6   Conclusion

Through the above analysis and discussion, following conclusions can be obtained. First, although the research work of the dam safety monitoring expert system is not long in our country, several dams which have been built the expert system have

obtained a significant socio-economic effectiveness. So strengthening the research work is necessary. Second, that using of neural networks and traditional rules form knowledge base bring the traditional knowledge representation and the neural networks to represent knowledge into full play. It make the knowledge representation to be simple and clear. Third, the capability of the current neural network expert system interpret the results is weaker than the traditional interpretation. Therefore, it need continue to study in this regard.

References

Cios KJ, Kurgan LA (2005) Trends in data mining and knowledge discovery. Advanced techniques in knowledge discovery and data mining: 8–12

Kheifits VZ. Routine instrumental monitoring – a guarantee of dam safety. Power Technol Eng 29(5):34–36

Panizzo A, Petaccia A (2009) Analysis of monitoring data for the safety control of dams using neural networks. New Trends Fluid Mech Res Part 5:107–109

Tianqi Yang (2005) Neural networks for solving on-line outlier detection problems. Adv Neural Netw, ISNN 2005:78–79

Xu Ri-qing (2000) The GA-ANN method for determining calculation parameters for deep excavation. J Zhejiang Univ Sci A 1(4):408–413

A TPM-Based Protection Mechanism for Remote Attestation Evidence

Shi Guang-yuan, Gong Bei, and Zhu zhen-shan

Abstract Remote Attestation is very importance in trusted computing, and the attestation evidence is foundation of the Remote Attestation. However, the existing technologies are not pay close attention on the protection of attestation evidence. In this paper, through research the main technologies of Remote Attestation, and analyze the attestation evidence in each method, providing a TPM-based protection mechanism for remote attestation evidence. The mechanism for Remote Attestation to ensure the secure transmission of the attestation's information, and designing a common communication protocol to guarantee the confidentiality and integrity of remote attestation's information in the course of transmission.

Keywords Remote attestation • Protection • Communication protocol • Trusted computing

1 Introduction

With the rapid popularization of Internet applications, people are more closely associated with them in everyday life. At the same time, since the Internet is open and interconnected, there are many insecurity factors, such as virus, worms and Trojans attacked the terminals destructively by means of vulnerabilities, which threat computer security seriously.

S. Guang-yuan (✉) • G. Bei
College of Computer and Science, Beijing University of Technology, Beijing, China
e-mail: 118sgy@163.com; HyperionUl@gmail.com

Z. zhen-shan
College of Nuclear Science and Technology, Harbin Engineering University, Harbin, China
e-mail: Heu321@yahoo.com.cn

In this case, the program used to prove the terminal communicating in a network environment is safe and trustworthy becomes a research hotspot in the field of information security. In order to ensure the trustworthiness of the entire computer system, trusted computing conducts comprehensively research such as remote attestation, a very important part of it, which is the technology that attestator proves the trustworthiness of the local environment offering the trustworthy credentials to the remote challenger. At present, the major research work is as follows: TCG first introduced the concept of Remote Attestation (T.C. Group 2003). Sadeghi and Stuble (2004) proved the trustworthiness of the platform by computing the security properties of the platform. Vivek Haldar et al. (2004) proved it through the program semantics analysis by the virtual machine; R. Sailer et al. (2004) proposed the integrity measurement system based on the TCG specifications which proved it by measuring the integrity of entity; Xiao-yong Li et al. (2006) determine the trustworthiness of platform by analyzing the behaviors of the system. The programs above mainly concerned the framework and methodologies of remote attestation, but the research on protection mechanism for remote attestation evidence is not deep enough. First, if the attestation evidence is not protected and the attacker tempers the evidence, the result of attestation is not trustworthy. Second, there needs a common trusted network framework and communication protocol to ensure the transmission of information is safe and to protect the credentials from tampering as the remote attestation is carried out between multiple platforms.

In this paper, we analyzed the main methods of remote attestation and trusted computing technologies; we proposed a TPM-based protection mechanism for remote attestation evidence. The mechanism add a Attestation Evidence Transfer Agent, AETA, into Trusted Network Connection (Trusted Computing Group 2007; Trusted Computing Group. Federated TNC Version 1.0), and the AETA transfers the attestation evidence into a general information, and the general information will be encrypted by TPM, then the AETA will transmit the general information by the AGP (Attestation Generic Protocol) to ensure the confidentiality and integrity during the transmission of information. The second section of the paper analyzes the existing attestation mechanisms and gets the evidence content in every method. The third section introduced some trusted computing technologies. The fourth section proposed the protection mechanism and gave the Attestation Generic Protocol.

2 Mechanisms of Remote Attestation in Trusted Computing

Remote Attestation is the technology that attestator proves the trustworthiness of the local environment offering the trustworthy credentials to the remote challenger. Specifically, when the user on the platform A (Attestator) requests the services from platform B (Challenger), apart from the identify authentication and privilege verification, the challenger also verify that platform A is in the state of trustworthiness. A typical scenario of remote attestation is shown in Fig. 1.

TCG's remote attestation mechanism is that the attestator is asked to submit information in which encapsulated the information about platform state by the
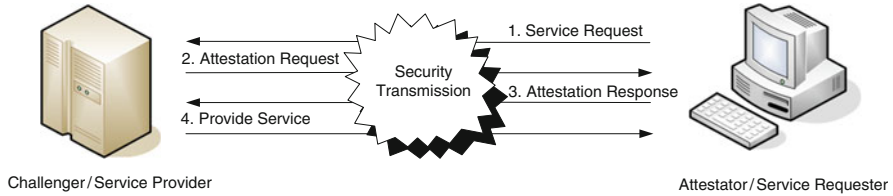
Fig. 1 A typical scenario of remote attestation

SHA-1(N.S. Laboratory 1995) algorithm including the event log and relevant PCR (Platform Configuration Register) values stored in the TPM to the challenger. PCR can be computed into one-to-one correspondence with certain state of specific events in the computing platform. The attestator signs the PCR first, and then sends the signature value, the event log, platform certificates and other relevant information to the challenger. Then the challenger authenticate the reports, determine the identity of the remote computing platform and judge whether the attestator is in the state of trustworthiness. The property-based attestation maps system configuration information to the system security properties first, and then the challenger determine whether the attestator is trustworthiness by verifying the security properties of it. In the bibliography Liqun Chen et al. (2006), it performs the functions including the anonymous attestation and revocation of binding between the configuration information and security properties. The mechanism of system behavior based attestation is that the challenger measures and verifies the system behaviors associated with the trustworthy state of the attestator's platform, and then determines whether the attestator is trustworthiness by the analysis of system behaviors based on security policies.

The methods above have their own advantages and shortcomings. TCG's method runs efficiently and develops simply. However, it only checks the PCR values and the platform certificates in which the amount of information contained is limited. It will be able to meet the requirements of users who have the lower security level. The other two methods need either to map system configuration information to the system security properties or to analyze system behaviors. However, they can provide more information and check the system security more comprehensive and accurate and are more appropriate to the users who have the higher security level. Thus, a common model compatible with a variety of attestation methods is required as different methods have their applicable targets.

3   The Trusted Computing

First, The TNC (Trusted Computing Group. Federated TNC Version 1.0) is an open solution architecture that enables network operators to enforce policies regarding the security state of endpoints in order to determine whether to grant access to

a requested network infrastructure. This security assessment of each endpoint is performed using a set of asserted integrity measurements covering aspects of the operational environment of the endpoint. The TNC is a new conceptual model proposed by TCG, while it is an open and common architecture, and the TNC does not depend on the specific technology or pattern.

The TNC architecture uses the Client/Server pattern. The three columns in this figure depict the three roles in the TNC architecture: the Access Requestor (AR), the Policy Enforcement Point (PEP) and the Policy Decision Point (PDP). The role of the AR is to seek access to a protected network in order to conduct activities on the network. The role of the PDP is to perform the decision-making regarding the AR's network access request, in light of the access policies. The PEP is the element which is connected to the AR; the role of the PEP is to enforce the decisions of the PDP regarding network access. Three horizontal layers of the architecture are indentified: the network access layer, the integrity evaluation layer and the integrity measurement layer. The TNC supports the platform integrity attestation: TCG-based attestation. It did not consider the subsequent remote attestations such as the property-based attestation and the system behavior based attestation. However, it can be the basic function module as it has a good scalability and the trusted connection can be established using the services provided by it.

The core specification of Trusted Computing Group (TCG) (Trusted Computing Group 2003) concerns the Trusted Platform Module (T.C. Group 2003), a component that provides certain cryptographic functions. The assumption is that this party is fully trusted. The current implementation of the TPM is a tamper-evident hardware chip. A TPM provides a secure random number generator, non-volatile tamper resistant storage, key generation algorithms, and cryptographic functions for encryption/decryption, digital signatures (RSA) and a cryptographic hash function (SHA-1). Moreover, the TPM includes a set of registers called Platform Configuration Registers (PCR), which can be used to store hash values.

$$PCR\_Extend(m) : PC\,R_i^{new} \leftarrow SHA - 1(PCR_i^{old}||m),$$

m is the message for measuring.

Trusted Software Stack (TSS) (Trusted Computing Group. TCG Software Stack) performs various functions like communicating with the rest of the platform or with other platforms.

4 Protection Mechanism for Evidence

The mechanism is separated into two parties: first is the evidence in the endpoints which is in attestation process need to be protected; second is the evidence should be protected in the transmission. Here are some stages:
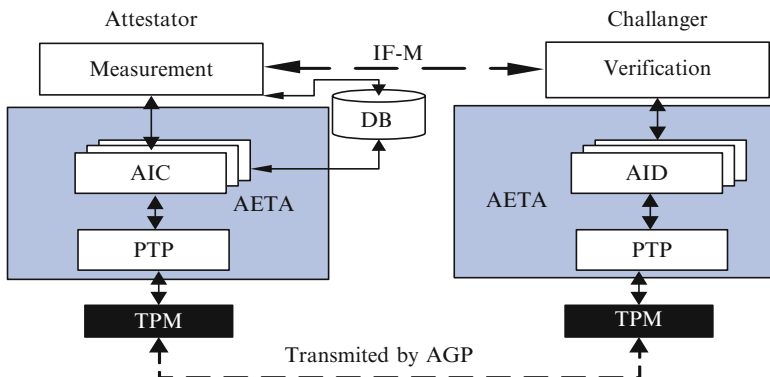
Fig. 2 Architecture of AETA

1. Collect and transfer the evidence. The evidence collected from result of measuring should be transferring into a form of general information.
2. Encrypt the general information by TPM and make sure the general information is not tampered.
3. Transmit general information by AGP.
4. Decrypt the general information, and resolve the general information.

4.1 Attestation Evidence Transfer Agent

Because of protecting the transmission of the evidence and extend the useless of the mechanism, we added a Attestation Evidence Transfer Agent, AETA, into TNC, show in Fig. 2. In the attestator, AETA is composed of AIC (Attestation Information Collector) and PTP (Protocol Transfer Proxy). AIC is responsible for collecting the evidence, which is the result of measurement. PTP is responsible for normalization of evidence and transmitting it. In the challenger, AETA is composed of AID (Attestation information Deliver) and PTP. PTP is responsible for transferring the general information into evidence, and AID is responsible for classifying the evidence and delivering the evidence to challenger. The design of the measurement module and verification module can reference the Sadeghi and Stuble (2004), Haldar et al. (2004), Sailer Reiner et al. (2004), Xiao-Yong Li et al. (2006).

PTP module is mainly responsible for normalizing the evidence files that provided by several Attestation methods, and the standard files will be used by Attestation Generic Protocol. PTP has two key functions: (1) trans- ferring the different evidence files into generic attestation information in the attestator. (2) transferring the generic attestation information into specific evidence files in the challenger.
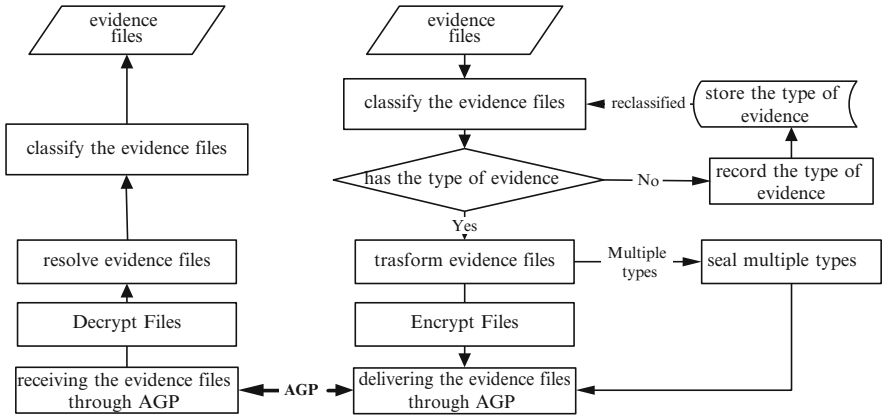
Fig. 3 Attestation
information
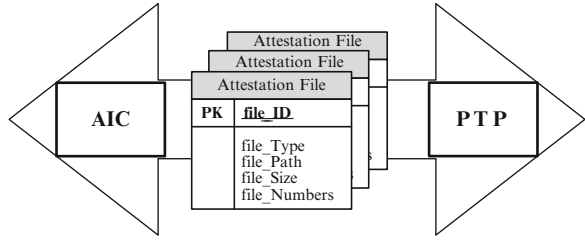
Fig. 4 Flow diagram of PTP

First, AIC collects the attestation information. The attestation information (see Fig. 3) involves the evidences files, the type of evidence, the quantity of evidence, the size of evidence and so on.

When the PTP module receives the evidence files, it uses the dual method to map the evidence files into generic information used by AGP. The flow diagram of PTP is shown in Fig. 4.

Definition 1 Set T is the set of attestation information, T = {A = {a1, a2, a3 . . . }, B = {b1, b2, b3 . . . }}, Set A is the subset of T and is the set of attestation type, each element of A represents a type of evidence files. For example, a1 represents TCG type, a2 represents property-based type and so on. The subset B is evidence files, each element represents the evidence files, for example, property certificate, kinds of measurement results and security logs.

Definition 2 Assuming ξ represents the generic information, and it is comprised of attestation type (AT), Attestation Information (AI), and some other information.

The formal description of PTP transformation mechanism: PTP uses dual method for transferring the different evidence files into attestation generic information.

Assuming the set of evidence files is Σpoof, the set of generic information is Σm, giving a information dual set K,

Fig. 5 Protocol transfer proxy module
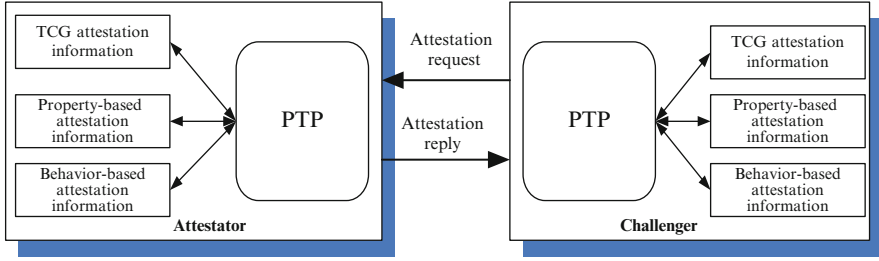
$$K = \left\{ < a, \ m > |ai \in \sum poof, \quad mi \in \sum m, \quad i = 1, \ 2, \ldots \right\},$$

$$P = \{a1, a2, \ldots, ak\} \subseteq \sum poof, \quad M = \{m1, \ m2, \ldots, mk\} \subseteq \sum m$$

K defines a mapping function φ, which maps the set P to set M, $\varphi(ai) = mi$.

$$\varphi(a_i) = \begin{cases} m_i = AT, \ a_i \in A \\ m_i = AI, \ a_i \in B \\ \ldots \end{cases}$$

each element of the dual set represents one pair of primitive information and generic information. If ai belongs to set of attestation type A, it will be mapped into AT, AT $\in \xi$, also, if ai belongs to set of attestation file B, it will be mapped into AI, AI $\in \xi$ (Fig. 5).

When the attestation platform will send the attestation information to the challenger, first, PTP should transfer attestation information $t \in T$ into generic information ξ. During the transformation, it will add some information like attestation type a_i, $a_i \in A$, attestation type is used for deciding which type the attestation information is. Then evidence files b_i, $b_i \in B$ will be mapped, for example, TCG attestation method will sign the measurement value of components PCR, Sign(PCR), and then evidence files involve the Sign(PCR), event log file EventLog, and platform certificate Cert, $b_i = \{Sign(PCR), EventLog, Cert\}$. Than a_i will be mapped into AT in ξ, b_i will be mapped into AI. Finally, ξ will be encapsulated in the AGP, and sent to challenger.

4.2 Attestation Generic Protocol

PTP module is response for transferring different evidence files into standard attestation information, therefore, a novel communication protocol for remote attestation needs to be designed. When attestation is required, the attestator firstly
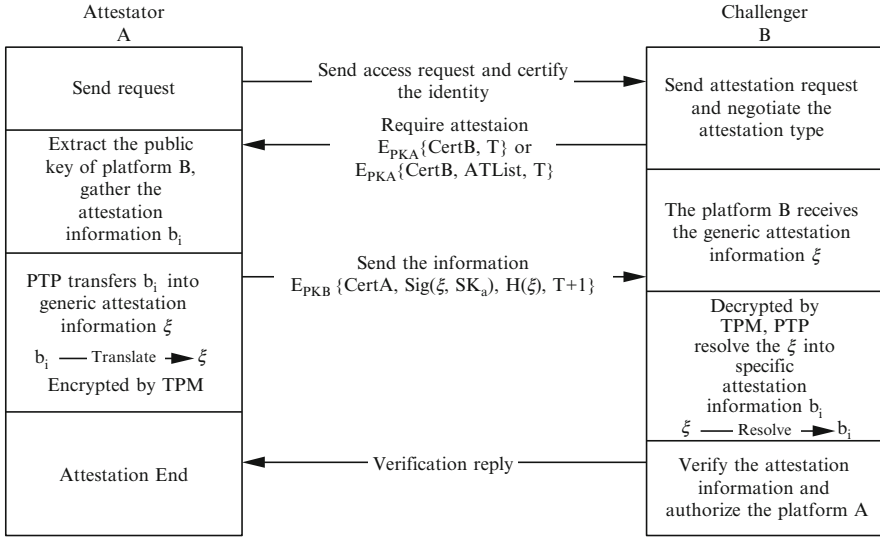
Fig. 6 Communication protocol flow

collects the attestation information, because of diversity of attestation methods, there are several kinds of attestation information. Then the attestator certifies its identity to a challenger, and attestator will send the attestation information after the certification is successful, the challenger will verify the information and authorize attestator depend on the verification results. Communication protocol flow is shown in Fig. 6.

We will give a detailed explanation of the commutation protocol flow. The platform A is attestator and platform B is challenger. The platform A firstly needs to certify its identity when A sends the serve request. Selecting a prime group G which rank is p, and an integer ring Z_p, G is a multiplication cyclic group, and g is a multiplication cyclic group, define two hash functions:

$H_{id} : \{0,1\}^{lid} \rightarrow G$, $H_m : \{0,1\}^{lm} \rightarrow Z_p$ (H_{id} and H_m have no collision probability, they are both strong one-way functions), selecting $g_1 = g^x$, $x \in Z_p$, $g_2 \in G$, define $e : G \times G \rightarrow G_T$ is the bilinear mapping from G to G_T, $e(g,g) = I$, the parameters of the whole system are $(G, G_T, e, p, g, g_1, g_2, H_{id}, H_m, I)$.

4.2.1 Certify the Identity of TPM

Considering the importance of TPM, TPM can be used to complete platform authentication. Because of the limitation of the TPM computing power, so the way of authentication is following:

TPM selects $x \in {}_R Z_P^*$, and computes $PK = g^x$, and PK is the public key of platform, then apply for Certificate Cert from CA, Cert contains identity of platform

and PK. First, attestator sends the PK of TPM to CA, then CA selects $x \in {}_R Z_P^*$, and sends it to attestator, then attestator computes $R = g^{\frac{1}{r+x}}$ and sends the result to CA, CA is certifies validity of $e(R, g^r PK) = I$, and determines whether the identity of TPM is legal, CA encrypts $v' \in Z_p$ by the attestator's PK and sends it to attestator.

4.2.2 Certify the Identity of Platform

(a) After verifying the identity of TPM, the attestator selects $r \in Z_p$, and computes $v = v' + r$ after it receiving v'.
(b) sign the message M, compute $m = Hm(M)$, attestator computes its private key $di = (g_2^x \gamma_i^v, g^v)$.
(c) select $r_1, r_2 \ldots r_n \in Z_p$, compute

$$\sigma = \left(\left(g_2^x \gamma_i^{v+m} \prod_{i=1}^{n} \gamma_i^n \right), g^{r1}, g^{r2} \ldots g^{v+m} g^{ri}, \ldots g^{rn} \right)$$

$\sigma = (S, f_1, f_2 \ldots f_n)$ as signature of platform i.
(d) verifier computes $m = H_m(M)$ according to $\sigma = (S, f_1, f_2 \ldots f_n)$, then verifier verifies the equation:

$$e(S, g) = e(g_1, g_2) \prod_{i=1}^{n} e(\gamma_i, f_i)$$

if the verification is successful, the attestator is permitted, otherwise, the attestator is rejected.

4.2.3 Attestation Information Transmission

After finishing certification, and receiving the attestation request of platform B, the CertA and CertB contains the EK of TPM respectively, the platform A firstly decrypts the request packets with private key of platform A SKA, and extracts the public key PKb from public key certificate of Platform B CertB, AIC will gather the attestation information b_i, $b_i \in B$, and deliver to PTP. PTP transfers attestation information into generic attestation information ξ, then and computes the hash value of ξ H(ξ) by using one-way hash function. Platform A send ξ, public key certificate of platform A to platform B. logical expression is:

$$PA \rightarrow PB : E_{PkB}\{CertA, Sig(\xi, SK_a), H(\xi), T + 1\}$$

and $Sig(\xi, SK_a)$ is attestation information signed by private key SK_a. Platform B unpacks the packets with its private key SKb after receiving the attestation reply,

and extracts the public key certificate CertA and signing message. Then B extracts the attestation information using public key of platform A, and verifies then integrity of the attestation information by $H(\xi)$. Then B delivers the generic attestation information ξ to PTP, PTP resolves the ξ into specific attestation information b_i, $b_i \in B$, and platform B will verify b_i.

4.3 Security Analysis of the Communication Protocol

4.3.1 Correctness

According to the quality of Bilinear group, we can deduce following equations:

$$e(S, g) = e\left(g_2^x \gamma_i^{v+m} \prod_{i=1}^n \gamma_i^{ri}, g\right) = e(g_2^x, g) \cdot e(\gamma_1, g)^{r1} \cdot \ldots$$

$$.e(\gamma_i, g)^{v+m+ri} \ldots e(\gamma_n, g)^{rn},$$

$$e(R, g^r PK) = e(g^{\frac{1}{r+x}}, g^r g^x),$$

and based on the mapping quality of Bilinear group, we can obtain:

$$e(g^{\frac{1}{r+x}}, g^r g^x) = e(g, g)^{\frac{r+x}{r+x}} = e(g, g) = I,$$

$$e(g_2^x, g) \cdot e(\gamma_1, g)^{r1} \cdot \ldots e(\gamma i, g)^{v+m+ri} \ldots e(\gamma_n, g)^{rn}$$

$$= e(g_1, g_2) \prod_{i=1}^n e(\gamma_i, f_i),$$

so we can arrive at the conclusion is that the signature and verification scheme is correct.

4.3.2 Security Analysis of Verification of Identity

Function simulate the honest attestator $P_i = \{P_1, P_2 \ldots P_n\}$, attacker A interact with P_i simultaneously, all of Pi have the same public key $PK = g^x$, attacker A pretends as verifier V, and transmits $r_i \in {}_R Z_p^*, i \in 1, 2 \ldots n$ to P_i, P_i answers the R_i or $\perp$ for $r_i \in {}_R Z_p^*$, if $r_i \in \{h_1, h_2 \ldots h_k\}$, then P_i sends $g^{\frac{1}{hi+x}}$ to attacker A, if not, gives $\perp$ to A. Because A knows the PK of attestator P_i, so $e(R_i, g^{ri} PK) = I$, A can't distinguish the true situation and idea situation.

Verification: A pretends as attestator, function F is honest verifier, F sends $r \notin \{r_1, r_2.....r_n\}$ by reference (PK, r_i, R_i), if A gets the result r and makes $e(R_i, g^{ri} PK) = I$ is true, then A finds answer of K-CCA, this is impossible in the computing ability of polynomial, so the protocol is security.

4.3.3 No Forgery of Signature

Because of only the certified member of ring can pass the verification and get v'. Computing $v = v' + x_1 + x_2$, and then generate the signature key $di = (g_2^x \gamma_i^y, g^v)$, because the signature scheme is based on DDH assumption, so it can't get any information of signature key from $\sigma = (S, f_1, f_2 \ldots f_n)$ in the Polynomial-time, so from an engineering perspective, the signature can't be forgery in the duration of validity.

5 Conclusion

Trusted remote attestation is an important research field in the trusted computing area, there has been many attestation methods to propose, in order to assure the security transmission of the attestation information between many computing platforms, this paper provides a TPM-based protection mechanism, and designs a generic communication protocol. We added A Attestation Evidence Transfer Agent in endpoint, and made use of Protocol Transfer Proxy module to transfer different attestation information into generic information. We also designed a generic communication protocol; this protocol can guarantee the integrity and confidentiality of attestation information by using the technologies of authentication, digital Signature and integrity verification. This paper is supported by the National High-Tech Research and Development Plan of China under Grant No 2006AA01Z440, 2009AA012437 and National Basic Research Program of China (973 Program 2007CB31110 0), and the open project of key state laboratory of Harbin Engineering University "the research on critical security technologies of the third generation nuclear power information system", No.HEUFN0801.

References

Haldar V, Chandra D, Franz M (2004) Semantic remote attestation-a virtual machine directed approach to trusted computing. In: Proceedings of the third virtual machine research and technology symposium USENIX 2004

Liqun Chen, Rainer Landfermann, Hans Lohr, Markus Rohe, Ahmad-Reza Sadeghi, Christian Stüuble (2006) A protocol for property-based attestation. In: STC'06: proceedings of the first ACM workshop on scalable trusted computing, ACM Press, New York, NY, USA, pp 7–16

N.S. Laboratory (1995) Secure hash standard. Federal Information Processing Standards Publication 180–1, April 1995

Sailer Reiner, Zhang Xiao-Lan, Jaeger Trent, van Doorn Leendert (2004) Design and implementation of a TCG-based integrity measurement architecture. In: SSYM'04: proceedings of the 13th conference on USENIX security symposium, USENIX Association, Berkeley, CA, USA, pp 16–16

Sadeghi A-R, Stuble C (2004) Property-based attestation for computing platforms: caring about properties, not mechanisms In: NSPW'04: proceedings of the 2004 workshop on new security paradigms, ACM Press, New York, pp 67–77

T.C. Group (2003) TPM main specification, November 2003, Version 1.2

Trusted Computing Group (2003) http://www.trustedcomputinggroup.org

Trusted Computing Group (2007) TCG specification architecture overview specification revision 1.4[Z]. TCG Published, August 2007

Trusted Computing Group. TCG Software Stack (TSS) specification, Version1.2. http://www.trustedcomputinggroup.org/resources/tcg_software_stack_tss_specification

Trusted Computing Group. Federated TNC Version 1.0, Revision 26.http://www.trustedcomputinggroup.org/resources/federated_tnc_version_10_revision_26

Xiao-Yong Li, Chang xiang Shen, Xiao-Dong Zuo (2006) An efficient attestation for trustworthiness of computing platform. In: IIH-MSP, Pasadena, California, USA, pp 625–630

Iterative Algorithm for a Family Multivalued Mapping in Banach Spaces

Zhanfei Zuo and Feixiang Chen

Abstract In this paper, we consider the convergence of iterative algorithm for a family of multivalued nonexpansive mapping, under some different conditions, the Noor multivalued algorithm converge to the common fixed point of the family of multivalued nonexpansive mapping.

Keywords Iterative algorithm • Multivalued nonexpansive mapping • Convergence theorems • Fixed points

1 Introduction

Let X be a Banach space and K a nonempty subset of X. We shall denote by 2^X the family of all subsets of X, $CB(X)$ the family of all nonempty closed bounded subsets of X and denote $C(X)$ by the family of nonempty compact subsets of X. A multivalued mapping $T : \mathrm{K} \to 2^X$ is said to be nonexpansive if

$$H(Tx, Ty) \leq ||x - y||, x, y \in K$$

where $H(Tx, Ty)$ denotes the Hausdorff metric on $CB(X)$ defined by

$$H(A, B) := \max\{\sup_{x \in A} \inf_{y \in B} ||x - y||, \sup_{y \in B} \inf_{x \in A} ||x - y||\}$$

Z. Zuo (✉) • F. Chen
Department of Mathematics and Statistics, Chongqing Three
Gorges University, Wanzhou 404000, P.R. China
e-mail: zuozhanfei@139.com; cfx2002@126.com

$A, B \in CB(X)$. A point x is called a fixed point of T if $x \in Tx$. Since Banach's Contraction Mapping Principle was extended nicely to multivalued mappings by Nadler in 1969 (see Nadler and Nadler 1969), many authors have studied the fixed point theory for multivalued mappings. In 2000, Noor (2000) suggested and analyze three-step iterative algorithm for finding the approximate the fixed point of the single-valued nonexpansive mapping. It is well known that the three-step iterative algorithm perform better numerically than the two-step (Ishikawa) and one-step (Mann) iterations. Three-step iterations algorithm are also called Noor iteration algorithm and this has initiated a quite a new direction of research in functional analysis. It is a very natural question whether the strongly convergent results of Noor iteration algorithm for a single-valued nonexpansive mapping T can be extended to the multivalued case.

In this paper, we consider the following iteration for a family of multivalued nonexpansive mapping T_n. Let K be a nonempty closed convex subset of Banach space X and $T_n : \mathrm{K} \to CB(K)$ be a family of multivalued nonexpansive mappings. For a given $x_1 \in K$ and $s_1 \in T_1 x_1$. Let $z_1 = (1 - a_1)x_1 + a_1 s_1$. From Nadler Theorem (Nadler and Nadler 1969), there exists $t_1 \in T_1 z_1$ such that $||t_1 - s_1|| \leq H(\mathrm{T}_1 z_1, \mathrm{T}_1 x_1)$.

Let $y_1 = (1 - b_1 - c_1)x_1 + b_1 t_1 + c_1 s_1$. There exists $r_1 \in T_1 y_1$ such that $||r_1 - t_1|| \leq H(\mathrm{T}_1 y_1, \mathrm{T}_1 z_1)$ and $||r_1 - s_1|| \leq H(\mathrm{T}_1 y_1, \mathrm{T}_1 x_1)$. Let $x_2 = (1 - \alpha_1 - \beta_1 - \gamma_1)x_1 + \alpha_1 r_1 + \beta_1 t_1 + \gamma_1 s_1$. There exists $s_2 \in T_2 x_2$ such that $||s_2 - r_1|| \leq H(\mathrm{T}_2 x_2, \mathrm{T}_1 y_1)$, $||s_2 - t_1|| \leq H(\mathrm{T}_2 x_2, \mathrm{T}_1 z_1)$ and $||s_2 - s_1|| \leq H(\mathrm{T}_2 x_2, \mathrm{T}_1 x_1)$. Inductively, we can get the sequence $\{x_n\}$ as follows:

$$z_n = (1 - a_n)x_n + a_n s_n$$
$$y_n = (1 - b_n - c_n)x_n + b_n t_n + c_n s_n$$
$$x_{n+1} = (1 - \alpha_n - \beta_n - \gamma_n)x_n + \alpha_n r_n + \beta_n t_n + \gamma_n s_n \qquad (1)$$

Where $\{a_n\}, \{b_n\}, \{c_n\}, \{b_n + c_n\}, \{\alpha_n\}\{\beta_n\}, \{\gamma_n\}$ and $\{\alpha_n + \beta_n + \gamma_n\}$ are appropriate sequence in $[0, 1]$, furthermore $s_n \in T_n x_n, t_n \in T_n z_n, r_n \in T_n y_n$ such that $||t_n - s_n|| \leq H(\mathrm{T}_n z_n, \mathrm{T}_n x_n), ||r_n - t_n|| \leq H(\mathrm{T}_n y_n, \mathrm{T}_n z_n)||r_n - s_n|| \leq H(\mathrm{T}_n y_n, \mathrm{T}_n x_n), ||s_{n+1} - r_n|| \leq H(\mathrm{T}_{n+1} x_{n+1}, \mathrm{T}_n y_n)$, and $||s_{n+1} - t_n|| \leq H(\mathrm{T}_{n+1} x_{n+1}, \mathrm{T}_n z_n), ||s_{n+1} - s_n|| \leq H(\mathrm{T}_{n+1} x_{n+1}, \mathrm{T}_n x_n)$.

The iterative scheme (1) is called Noor multivalued iterative algorithm. If $a_n = c_n = \beta_n = \gamma_n \equiv 0$ or let $a_n = b_n = c_n = \beta_n = \gamma_n \equiv 0$, we get the Ishikawa iteration and Mann iteration for a family of multivalued nonexpansive mappings. In fact let $\gamma_n \equiv 0$ or $c_n = \beta_n = \gamma_n \equiv 0$ or $b_n = c_n = \alpha_n = \gamma_n \equiv 0$, we also have the other three Algorithms.

2 Preliminaries

Definition 2.1 A family of multivalued mapping $T_n : K \to CB(K)$ is said to satisfy Condition (A$'$) if there is a nondecreasing function $f : [0, \infty) \to [0, \infty)$ with $f(0) = 0, f(x) > 0$ for $x \in (0, \infty)$ such that

$$d(x, T_n x) \geq f(d(x, F(\cap_n T_n))) \text{ for all } x \in K.$$

Where $F(\cap_n T_n) \neq \emptyset$ is the common fixed point set of the family of multivalued mapping $\{T_n\}$. From now on, $F(\cap_n T_n)$ stands for the common fixed point set of the family of multivalued mapping $\{T_n\}$.

Lemma 2.2 *see Nilsrakoo and Saejung (2007) Let X be a uniformly convex Banach space and $B_r := \{x \in X : ||x|| \leq r\}, r > 0$. Then there exists a continuous strictly increasing convex function $g : [0, \infty) \to [0, \infty)$ with $g(0) = 0$ such that*

$$||\lambda x + \mu y + \xi z + \upsilon \omega||^2 \leq \lambda ||x||^2 + \mu ||y||^2 + \xi ||z||^2 + \upsilon ||\omega||^2$$

$$\leq \frac{1}{3} \upsilon (\lambda g(||x - \omega||) + \mu g(||y - \omega||) + \xi g(||z - \omega||))$$

for all $x, y, z, \omega \in B_r$ and $\lambda, \mu, \xi, \upsilon \in [0, 1]$ with $\lambda + \mu + \xi + \upsilon = 1$.

3 Main Results

Lemma 3.1 *Let X be a uniformly convex Banach space and K be a nonempty convex subset of X. Let$\}$ $T_n : K \to CB(K)$ be a multivalued nonexpansive mapping for which $F(\cap_n T_n) \neq \emptyset$ and for which $T_n(p) = \{p\}$ for any fixed point $p \in F(\cap_n T_n)$. Let $\{x_n\}$ be a sequence in K defined by (1), if the coefficient satisfy one of the following control conditions:*

(i) $0 < \liminf_n \beta_n \leq \limsup_n (\alpha_n + \beta_n + \gamma_n) < 1$ *and* $\limsup_n a_n < 1$;
(ii) $0 < \liminf_n \gamma_n \leq \limsup_n (\alpha_n + \beta_n + \gamma_n) < 1$;

then we have $\lim_n d(x_n, T_n x_n) = 0$.

Proof. Firstly, we will prove that $\lim_n ||x_n - p||$ exists for any $p \in F(\cap_n T_n)$. Let $p \in F(\cap_n T_n)$, from iterative scheme (1), note that $T_n(p) = \{p\}$ for any fixed point $p \in F(\cap_n T_n)$, we have

$$||z_n - p|| \leq (1 - a_n)||x_n - p|| + a_n ||s_n - p||$$
$$\leq (1 - a_n)||x_n - p|| + a_n H(T_n x_n, T_n p)$$
$$\leq ||x_n - p||,$$

similarly $||y_n - p|| \le ||x_n - p||$, and so we have

$$\begin{aligned}
||x_{n+1} - p|| &\le (1 - \alpha_n - \beta_n - \gamma_n)||x_n - p|| + \alpha_n ||r_n - p|| \\
&\quad + \beta_n ||t_n - p|| + \gamma_n ||s_n - p|| \\
&\le (1 - \alpha_n - \beta_n - \gamma_n)||x_n - p|| + \alpha_n d(T_n y_n, T_n p) \\
&\quad + \beta_n H(T_n z_n, T_n p) + \gamma_n H(T_n x_n, T_n p) \\
&\le ||x_n - p||.
\end{aligned}$$

Then $\{x_n - p\}$ is a decreasing sequence and hence $\lim_n ||x_n - p||$ exists for any $p \in F(\cap_n T_n)$. Then it follows that $\{s_n - p\}$, $\{t_n - p\}$ and $\{r_n - p\}$ are all bounded. We may assume that these sequences belong to B_r where $r > 0$. Note that $T_n(p) = \{p\}$ for any fixed point $p \in F(\cap_n T_n)$. By Lemma 2.2, we get

$$\begin{aligned}
||z_n - p||^2 &\le (1 - a_n)||x_n - p||^2 + a_n ||s_n - p||^2 \\
&\le (1 - a_n)||x_n - p||^2 + a_n d(s_n, T_n p)^2 \\
&\le (1 - a_n)||x_n - p||^2 + a_n H(T_n x_n, T_n p)^2 \\
&\le |x_n - p||^2,
\end{aligned}$$

$$\begin{aligned}
||y_n - p||^2 &\le (1 - b_n - c_n)||x_n - p||^2 + b_n ||t_n - p||^2 + c_n ||s_n - p||^2 \\
&\quad - \frac{1}{3}(1 - b_n - c_n)(b_n g(||t_n - x_n||) + c_n g(||s_n - x_n||)) \\
&\le (1 - b_n - c_n)||x_n - p||^2 + b_n H(T_n z_n, T_n p)^2 + c_n H(T_n x_n, T_n p)^2 \\
&\quad - \frac{1}{3}(1 - b_n - c_n) b_n g(||t_n - x_n||) \\
&\le ||x_n - p||^2 - \frac{1}{3}(1 - b_n - c_n) b_n g(||t_n - x_n||),
\end{aligned}$$

therefore we have

$$\begin{aligned}
||x_{n+1} - p||^2 &\le (1 - \alpha_n - \beta_n - \gamma_n)||x_n - p||^2 + \alpha_n ||r_n - p||^2 + \beta_n ||t_n - p||^2 \\
&\quad + \gamma_n ||s_n - p||^2 - \frac{1}{3}(1 - \alpha_n - \beta_n - \gamma_n)[(\alpha_n g(||x_n - r_n||) \\
&\quad + \beta_n g(||x_n - t_n|| + \gamma_n g(||x_n - s_n||)] \\
&\le (1 - \alpha_n - \beta_n - \gamma_n)||x_n - p||^2 + \alpha_n H(T_n y_n, T_n p)^2 + \beta_n (T_n z_n, T_n p)^2 \\
&\quad + \gamma_n H(T_n x_n, T_n p)^2 - \frac{1}{3}(1 - \alpha_n - \beta_n - \gamma_n)[(\alpha_n g(||x_n - r_n||) \\
&\quad + \beta_n g(||x_n - t_n|| + \gamma_n g(||x_n - s_n||)]
\end{aligned}$$

$$\leq ||x_n - p||^2 - \frac{\alpha_n}{3}(1 - b_n - c_n)(b_n g(||t_n - x_n||) - \frac{1}{3}(1 - \alpha_n - \beta_n - \gamma_n)$$

$$[(\alpha_n g(||x_n - r_n||) + \beta_n g(||x_n - t_n||) + \gamma_n g(||x_n - s_n||)].$$

Then

$$(1 - \alpha_n - \beta_n - \gamma_n)\alpha_n g(||x_n - r_n||) \leq 3(||x_n - p||^2 - ||x_{n+1} - p||^2), \quad (2)$$

$$(1 - \alpha_n - \beta_n - \gamma_n)\beta_n g(||x_n - t_n||) \leq 3(||x_n - p||^2 - ||x_{n+1} - p||^2), \quad (3)$$

$$(1 - \alpha_n - \beta_n - \gamma_n)\gamma_n g(||x_n - s_n||) \leq 3(||x_n - p||^2 - ||x_{n+1} - p||^2), \quad (4)$$

and

$$\alpha_n(1 - b_n - c_n)b_n g(||t_n - x_n||) \leq 3(||x_n - p||^2 - ||x_{n+1} - p||^2). \quad (5)$$

Since $\lim_n ||x_n - p||$ exists for any $p \in F(\cap_n T_n)$, it follows from (2) that $\lim_n (1 - \alpha_n - \beta_n - \gamma_n)\alpha_n g(||x_n - r_n||) = 0$. From g is continuous strictly increasing with $g(0) = 0$ and $0 < \liminf_n \alpha_n \leq \limsup_n (\alpha_n + \beta_n + \gamma_n) < 1$,
then

$$\lim_n ||x_n - r_n|| = 0. \quad (6)$$

Using a similarly method together with inequalities (3) and $0 < \liminf_n \beta_n \leq \limsup_n (\alpha_n + \beta_n + \gamma_n) < 1$, then

$$\lim_n ||x_n - t_n|| = 0. \quad (7)$$

Similarly, from (4) and $0 < \liminf_n \gamma_n \leq \limsup_n (\alpha_n + \beta_n + \gamma_n) < 1$, we have $\lim_n ||x_n - s_n|| = 0$, since $s_n \in T_n x_n$, then

$$0 \leq \lim_n d(x_n, T_n x_n) \leq \lim_n ||x_n - s_n|| = 0$$

thus we get the (ii). If $0 < \liminf_n \beta_n \leq \limsup_n (\alpha_n + \beta_n + \gamma_n) < 1$ and $\limsup_n a_n < 1$; we will prove (i).

$$\begin{aligned} ||s_n - x_n|| &\leq ||s_n - t_n|| + ||t_n - x_n|| \\ &\leq H(T_n x_n, T_n z_n) + ||t_n - x_n|| \\ &\leq ||x_n - z_n|| + ||t_n - x_n|| \\ &\leq a_n ||x_n - s_n|| + ||t_n - x_n||. \end{aligned} \quad (8)$$

Since $\limsup_n a_n < 1$, then $\liminf_n(1-a_n) = 1 - \limsup_n a_n > 0$. This together with (7), (8), we obtain the result.

Theorem 3.2 *Let X, K, T_n and $\{x_n\}$ be the same as in Lemma 3.1, if T_n satisfies Condition A′ with respect to the sequence $\{x_n\}$, then $\{x_n\}$ converges strongly to a common fixed point of T_n.*

Proof. By Lemma 3.1, we have $\lim_n d(x_n, T_n x_n) = 0$. Since T_n satisfies Condition A′ with respect to $\{x_n\}$. Thus we get $d(x_n, F(\cap_n T_n)) = 0$. The remainder of the proof is similar as Theorem 2.4 in (Song and Wang 2009), we omit it.

References

Nadler SB Jr, Nadler SB Jr (1969) Multi-valued contraction mappings. Pac J Math 30:475–487

Nilsrakoo W, Saejung S (2007) A new three-step fixed point iteration scheme for asymptotically nonexpansive mappings. Appl Math Comput 181:1026–1034

Nilsrakoo W, Saejung S (2007) A reconsideration on convergence of three-step iterations for asymptotically nonexpansive mappings. Appl Math Comput 190:1472–1478

Noor MA (2000) New approximation schemes for general variational inequalities. J Math Anal Appl 152:217–229

Song YS, Wang HJ (2009) Convergence of iterative algorithms for multivalued mappings in Banach spaces. Nonlinear Anal 70:1547–1556

Weak-Linked Document in Search Engine

Wei Yanjun and Zheng Qingbi

Abstract The current search engine cannot rank well weak-linked documents such as PowerPoint files and AJAX applications. Current search engines return therefore either completely irrelevant results or poorly ranked documents when searching for these files. Roc, a novel framework is proposed for correctly retrieving and ranking weak-linked documents based on Clustering. The experiments show that our approach considerably improves the result quality of current search engines and that of latent semantic indexing.

Keywords Search engine • Clustering technology • Weak-linked document

1 Introduction

Currently, mainstream search engines like Google, Yahoo! On the web page can produce excellent sorting results, in which, the web page rank algorithm represented by PageRank (Page et al. 1998), HITS (Kleinberg 1999) based on link analysis makes use of the link information between web pages to calculate the ranking values of the web pages. However, some documents on the network do not have enough hyperlink information. For example, PowerPoint documents, PDF files and Word documents only have a few incoming links and not outer links. In this paper, this kind of document is called weak-linked document. Such interactive applications like AJAX and Flash are also weak-linked document, because the page does not link to their internal content on the Web. When the weak-link document is searched, the current search engines can not sort the results better. This paper presented a framework supporting search engine of the weak-linked document, and

W. Yanjun (✉) • Z. Qingbi
Computer Department Shangqiu Vocational and Technical College Shangqiu,
Shangqiu, P. R. China
e-mail: weiyjhs@163.com

527

emphatically describes a weak-linked document's ranking algorithm RoC (Ranking weak-linked files based on Clusters of Web pages) based on Web search results.

2 Weak-Linked Document Search

In order to be able to sort the weak-linked document, a new Search Engine ESE (Enhanced Search Engine) (Duda et al. 2008, 2009) must regard the state of weak-linked document as the smallest retrieval unit. Like the traditional search engines, ESE also is divided into online and offline. Offline includes grab and sorting of weak-linked document, and online includes retrieval and sorting of the document.

2.1 Grab the Weak-Linked Document

According to the type of weak-linked document, there are two different crawler capture methods. Some weak-linked documents such as PowerPoint files, Flash movies and so on can be downloaded once to overall, and crawler was decomposed into multiple states according to the user view after being whole downloaded. For example, a slide page is a state in the PowerPoint files, and in the Flash movie, a scene is a state. The other is some dynamic applications shaped like AJAX, Flash and so on, which must request content by internal links or client-side script. Given the initial URL, crawlers triggered some events which each user may triggered to continuously crawled the entire contents of the dynamic applications by simulating user behavior, where each triggered event, the application crawler would get a new state. ESE uses a special AJAX crawler, which analyzes script of AJAX to obtain the call graph, and then monitors and caches the parameters of Hot function node connected on a network to detect successfully duplicate state, thus greatly improved the crawling efficiency of crawlers.

2.2 Query Processing

ESE indexes the weak-linked document, the index file provides weak-linked document and state information for sorter, so the traditional inverted index file need to be extended to index which regarded the state as the basic unit of indexing. Given a query q user submits, ESE will submit q to a public Web search engines, and cluster web search results using a clustering algorithm, each cluster achieves their corresponding weight according to page rank. Cluster contains more pages of high ranking, the greater its weight is. ESE retrieves the relevant state by the index, and then ranks clusters using the weight, on the basis of which for further calculate the sorting value of the weak-linked document.

3 Sorting the Weak-Linked Document by Status

This paper takes a two-step strategy to solve the sorting problem of the weak-linked document. Firstly, using clustering technology obtains the subject contained given query keywords from the web search results; secondly, calculating the sorting value of the document according to the relevance of weak-linked document and theme. Based on the above ideas, this paper provides a weak-linked document ranking algorithm RoC (Ranking weak-linked files based on Clusters of Web pages) based on web search results clustering. RoC algorithm is as follows:

Algorithm: RoC weak-linked document sorting algorithm
Input: query q, a public search engine S, a document clustering algorithm CA.
Output: A sorted by relevance weak-linked document sequence N(q).

The experiment steps as follows:

- The query q was submitted to the search engine S, the first n pages returned were written as $R(q) = (P_1, P_2, \ldots, P_n)$;
- The query was submitted to the collection which was constituted by all states that contained the set of q obtained by ES, the collection was denoted by $N(q) = (S_1, S_2, \ldots, S_m)$;
- R(q) was clustered By CA, and the results recorded as $C(q) = (C_1, C_2, \ldots C_t)$;
- D. Each theme cluster $Ci \in C(q)$ was calculated the weight $W(C_i, q)$;
- Each weak-linked document $di \in N(q)$ was calculated the sorting value;
- N(q) was Sorted in descending order based on the sorting values.

3.1 Identification of the Subject of Query q

The third step in the RoC, taking into account the need to calculate the sorting values of weak-linked document online, clustering algorithm of the selected web pages should have higher performance and be able to produce good theme description. In this paper, a graph-based text clustering algorithm was adopted: Star clustering algorithm (Aslam et al. 2004). One advantage of the algorithm is its ability to produce high quality clustering results, and the number of clusters is not fixed. The advantage of Star is very suitable for algorithm RoC, because for an arbitrary given query q, its subject implied on the Internet, that the number of theme which the page search results covered could not be pre-determined.

Using the standard vector space model to represent each Web search result $p_i \in$ R(q), that is to say pi is expressed a vector which should be indexing entries (term) for the dimension, in which the weight set is same to TFIDF. For any two pages $p_i, p_j \in$ R(q), their similarity Sim (p_i, p_j) is defined as the cosine distance between their vectors $\mathbf{v}_i$ and $\mathbf{v}_j$, namely: **v**

$$Sim(p_i, p_j) = \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|v\| \times \|\mathbf{v}_j\|}$$

Based on the web pages similarity definition, building a website similar graph G (V, E), where each page in R(q) is a node of V, E = ((v_i, v_j) | Sim (p_i, p_j) > δ), δ is the threshold value specified minimum similarity.

Star algorithm divided G into several star clusters using a heuristic iterative strategy continuously, the so-called star clusters is constituted by a document as the center and its adjacent documents. The specific steps are as follows:

- The C(q) is initialized to empty set
- In the nods in G which have not yet been classified the Star cluster, that is to say in those nodes which has not been covered by C(q), finding the node with the maximum degree, denoted them by C.
- Building a Star cluster C centering on node c, and the neighboring nodes of C which are not classified Star cluster are classified C. Then the C is added to C(q).
- The star clusters C = (p_1, p_2, ... p_k) obtained in the last step, its calculation vector is as follows:

$$\overrightarrow{v_c} = \frac{1}{k} \sum_{i=1}^{k} \frac{\overrightarrow{v_c^i}}{\left\| \overrightarrow{v_c^i} \right\|}$$

Repeating steps 2, 3, 4 until all the nodes are classified into a certain Star cluster.

In the algorithm the only parameter which is the minimum similarity threshold δ determines the structure of G, and after the graph G is determined, the final clustering result C(q) is determined. Therefore, δ is the main factors influencing final clustering results quality and the final sorting results of weak-linked document. The larger δ is, the fewer edge G has, the greater the similarity is between web pages in Star cluster, accordingly, the less page clusters contains. Noted that, because the n pages are not repeated, while δ = 1, G will not contain any edge, and thus clustering algorithm is degraded to the state a web page form a cluster. In the experiment, the influence of sorting results of the last weak-linked document caused by δ is verified.

3.2 Confirming the Weight of the Theme

After obtained the theme cluster set C(q) of a given query q, it need to give a weight to the object in C(q) to distinguish the importance of each theme. Because in the search engine results, the more front page ranking is, the more important, the more be able to reflect the performance of search engines. According to the report which comes from the search engines market research institution iProspect, 62% of users only click the first page of search results, and only 10% of users will see the third page of search results (iProspect Search Engine User Behavior Study). Therefore,

a page with higher ranking should have greater impact on the theme of cluster compared with the page with lower ranking. For each theme cluster Ci in C(q), its weight is defined as:

$$w(c_i, q) = \sum_{j=1}^{k} \log_b^{\frac{n}{r(d_j)}}$$

In which, $p_j \in C_i$, $r(p_j)$ is a web p_j's ranking in the search results, n is the number of pages to be clustered, b is the number of results which each page contains in the returning results of the search engine. Generally, in the current search engines, the value of b is 10, that are each page contains ten search results. In the experiment, the value of b is also set to 10.

3.3 Calculating the Sorting Value of the Weak-Linked Document's State

The sorting values of weak-link document should be based on the identified weight subject. As well as the cluster web search results, each weak-linked document in N(q) is represented by the standard vector space model, and the indexing weight in vector is expressed by tf*idf, in which the value of idf come from the statistics of a collection R(q) of Web search results, rather than from the N(q). C(q) is the theme collection of a given query q, any weak-linked document's state is $s_i \in N(q)$, the sorting values are defined as follows:

$$Score(s_j, q) = \sum_{c_i \in c(q)} w(c_i, q) \times Sim(s_j, c_i)$$

In which, Sim (d_j, c_i) is the cosine distance between d_j and c_i.

4 Sorting the Weak-Linked Document

Although some applications like AJAX return the associated status to the users, which can meet the demand of users, some weak-linked document like PowerPoint files should return PowerPoint document rather than one of its pages. Therefore, after sorting the state, it is need to sort the entire weak-linked document. The sorting value of weak-linked document should be based on the sorting values of its state.

Given a weak-linked document $A = \{s_1, s_2, \ldots s_m\}$, its calculation of the sorting value is as follows:

$$Score(A, q) = \frac{1}{m} \sum_{i=1}^{m} w_i \times Score(s_i, q)$$

In which, $s_i \in A$, wi is the weights of weak-linked document. For the sake of simplicity, in the experiment, the weight is set to 1. It should be noted that ordinary web page can be regarded as the document of containing one state, that is, $m = 1$. Thus, the sorting values of ordinary web page and weak-linked document could be calculated separately using the same method, and the results list containing a weak-linked document and ordinary web page is returned to the user to achieve the uniform query of both document.

5   Experimental Evaluation

In order to verify the performance of RoC, the PowerPoint files are selected as a weak-linked document, and Google is used as a public search engine which could obtain the web search results. In addition to Google's search services, in the experimental evaluation the following two methods are also considered for comparison:

(a) *False rely on feedback (PRF)*: In fact, PRF is a special case of RoC, which is when clustering Web search results, all the results of the first n pages are classified as a theme clusters, and each page result has the equal treatment.
(b) *Latent semantic indexing (LSI)*: LSI is a successful sorting model based on the content. In The experiment a open source is obtained from The Semantic Indexing from using Project (http://knowledgesearch.org).

To make the results more convincing, we constructed 30 different queries. Each query was submitted to Google, and the top 100 PowerPoint documents which were returned by Google were grabbed. The documents downloaded failed were removed, the final total is 2,675 PowerPoint documents, which were noted the document collection D. RoC, PRF, and LSI acted on D respectively, and for each query, we determined whether the results of all methods related to a given query.

In the experiment, considering that the users are most concerned about the result of top-ranking, so the evaluation criteria P@N which is widely used in the information retrieval area is regarded as a measure of order. P@N represented the accuracy of search results in the first N position, and it is defined as follows:

$$P@N = \frac{|T \cap R|}{|R|}$$

Where R is the collection constructed by first N results, T is the collection constituted by manually label and inquiries related documents. In the experiment three evaluation criteria P@5, P@10 and P@20 were used.

The accuracy of four methods was shown in Fig. 1. Google's results were significantly worse than the other three methods, which indicate that the link-based sorting algorithm was determined by the importance of weak-linked document in the link map is not sufficient for weak-linked document; RoC has achieved the
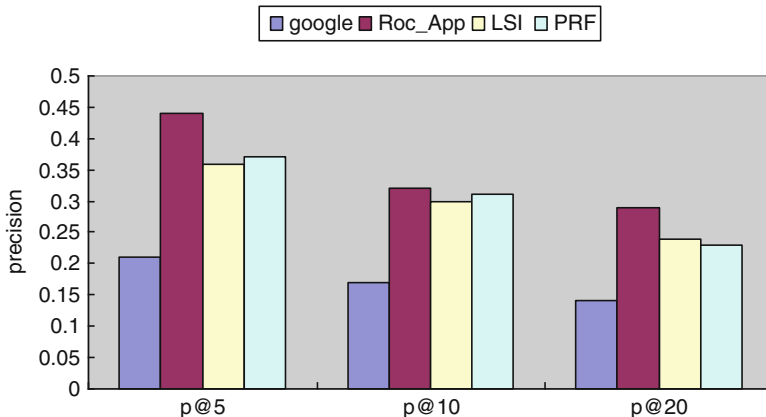
Fig. 1 Performance comparison of Google, Roc, LSI, PRF

best performance. Particularly, the accuracy of RoC in the P @ 5 is better than the standard PRF method. This is because the RoC could make the weak-linked document related to query get the high ranking in the final results by highlighting the top-ranking pages' role in the theme cluster. On the other hand, RoC in P @ 20 have also improved significantly compared with the other methods, which indicate that the clustering technique successfully identified the theme of a given query, and eliminates the noise in the page search results, which could produce more better sorting results than PRF. Experimental results show that RoC could get idea results for weak-linked document compared to existing search engines.

References

Aslam JA, Pelekhov E, Rus D (2004) The star clustering algorithm for static and dynamic information organization. J Gr Algorithm Appl 8:95–129

Duda C, Frey G, Kossmann D et al (2008) AJAXSearch: crawling, indexing and searching web 2.0 applications. Proc VLDB Endow 1(2):1440–1443

Duda et al (2009) AJAX crawl: making AJAX applications searchable. In: Proceedings of the 25th International Conference on Data Engineering (ICDE'09), Shanghai, China, pp 78–89

iProspect Search Engine User Behavior Study. http://www.iprospect.com

Kleinberg JM (1999) Authoritative sources in a hyperlinked environment. J ACM 46(5):604–632

Page L, Brin S, Motwani R et al (1998) The PageRank citation ranking: bringing order to the web. Technical report, Stanford digital library technologies project, Stanford University, Stanford, CA, USA, November 1998

Concurrent-Secure Blind Signature Scheme Without Random Oracles

Liu Xin

Abstract In the context of blind signature schemes, how to devise efficient and concurrent-secure blind signature scheme without random oracles has been a very active research topic. Recently, Gjøsteen et al. proposed a round-optimal blind signature scheme in the registered public key model. Unfortunately, this scheme has the drawback that the costs of computation and communication of its signing protocol are linear in the size of the message to be signed. The main contribution of this paper is a revised scheme which is built on the Zhou-Lin signature scheme, the compilation technique of Damgård et al., and Arita's straight-line extractable commitment scheme. The salient features of the new scheme are that the costs of computation and communication of its signing protocol do not depend on the size of the message to be signed. In addition, the resultant signature is very short and has an efficient verification process. Moreover, thanks to the technique of straight-line extraction, the security reduction algorithm is efficient and run in strictly polynomial-time.

Keywords Blind signatures • The registered public-key model • Concurrent • Non-interactive zero-knowledge

1 Introduction

Blind signature schemes allow a user to obtain a signature from an authority on message m, such that the authority cannot learn any useful information about m

L. Xin (✉)
School of Information Engineering, Shandong Youth University of Political Science, Jinan, China

Key Laboratory of Information Security and Intelligent Control in Universities of Shandong
(Shandong Youth University of Political Science), Jinan, China
e-mail: lxonne@163.com

from the signing protocol. So far, blind signatures have been intensively used in the construction of e-cash schemes, e-voting schemes and group signature schemes, etc. Typically, blind signature schemes share two main security properties, i.e., blindness and unforgeability. In Hazay et al. (2007) pointed out that it is necessary to distinguish whether different instances of the signing protocol can be carried out concurrently. There exist several efficient and concurrent-secure constructions (such as Abe 2001; Bellare et al. 2003), however these schemes can only be provably secure in the random oracle model.

In CRYPTO 1997, Juels et al. (1997) presented the first scheme in the standard model and claimed that their scheme was secure in the concurrent setting. However, this construction is based on general results about two party computation and is thus extremely inefficient in contrast with those schemes in the random oracle model. Recently, Hazay et al. (2007) argued that the solution in Juels et al. (1997) can only be considered secure in the sequential setting, because how to build a concurrently-secure protocol for two party computation without random oracles is still an open question. Camenisch et al. (2005), proposed the first concrete provably-secure blind signature scheme in the standard model. The main drawbacks of Camenisch et al.'s scheme is that it can only be proven unforgeable under the case of sequential attacks. Furthermore, its signing protocol is 8-move and the user has to execute three complicated proofs of knowledge, thus this scheme is also inefficient. Kiayias and Zhou (2006) in later work devised a 4-round efficient and concurrently-secure blind signature scheme without random oracles in the common reference string model. Later, Okamoto (2006) proposed an (partially) blind signature scheme which is secure for a concurrent run of users. Okamoto's scheme uses the similar technique to Kiayias and Zhou (2006), but it requires a general trapdoor commitment scheme and its security reduction algorithm is not very efficient, i.e., to extract the secret values of the malicious users, the security reduction algorithm has to repeatedly check whether a certain equation holds or not, and the number of iterations of this step is bounded by a polynomial. Recently, several round-optimal blind signature schemes were respectively put forward in Hazay et al. (2007), Fischlin (2006), Au et al. (2010), Gjøsteen and Kråkmo (2008). The main advantage of such schemes is that the signing protocol is 2-round which is automatically secure in a concurrent setting. However, the schemes in Hazay et al. (2007), Fischlin (2006) are only inefficient generic constructions. The scheme in Fischlin (2006) requires generic NIZK (non-interactive zero-knowledge) proofs while the solution in Hazay et al. (2007) resorts to ZAPs (2-round public coin witness-indistinguishable protocol). Au et al. (2010) very recently presented a concrete instantiation of round-optimal blind signature, which is secure in the random oracle model. But the resultant signature is too long and verifying signatures is an expensive task. The Gjøsteen-Kråkmo scheme (Gjøsteen and Kråkmo 2008) was built on Waters signature (Waters 2005) and the Σ-compiler by Damgård et al. (2006). However, this scheme requires rather large computational complexity, because the costs of computation and communication of its signing protocol are linear in n, i.e., the size of the message to be signed. Furthermore, the security reduction algorithm in Gjøsteen and Kråkmo (2008) is

not efficient. In other words, in order to reduce the unforgeability to the UF-CMA (existential unforgeability under an adaptive chosen message attack) security of the underlying Waters signature scheme, the reduction algorithm has to invoke the underlying decryption algorithm of the Linear Encryption (abbreviated as LE) (Boneh et al. 2004) up to $O(n)$ times.

1.1 Contributions

The main contribution of this paper is a revised scheme which is built on the Zhou-Lin signature scheme (Zhou and Lin 2007) (henceforth called the ZL Scheme), the compilation technique of Damgård et al. (2006), and Arita's straight-line extractable commitment scheme (Arita 2007). In comparison with the previous schemes, the new scheme enjoys the following merits:

- By using an efficient unlinkable randomizable signature (i.e., the ZL scheme) as a main building block, the resultant signature is very short and easy to verify.
- In the signing protocol, the user and the signer has only to carry out a concurrent-secure 3-round interaction.
- Thanks to the technique of straight-line extraction, the security reduction algorithm is efficient and run in strictly polynomial-time.
- Performance comparison shows that the cost (respect to communication and computation) of the new scheme is significantly lower than that of the Gjøsteen-Kråkmo scheme.
- The security of the new scheme does not depend on random oracles.

1.2 Organization

The rest of this paper is organized as follows. The next section proposes a concurrent-secure blind signature scheme without random oracles. Section 3 gives formal security proofs in the registered public key model. In Sect. 4, we show the comparison of our scheme with several recently proposed schemes in terms of the main properties as well as the complexity of signature generation/verification. Finally, we give conclusion remarks in Sect. 5.

2 The Proposed Scheme

Common Information. $ComInfo = (p, G_1, G_2, G_T, g, \tilde{g}, \hat{e})$, such that $\hat{e} : G_1 \times G_2 \to G_T, G_1 = \langle g \rangle, G_2 = \langle \tilde{g} \rangle, ord(G_1) = ord(G_2) = ord(G_T) = p$.

Common Reference String. It is assumed that the signer and the user can access to the trusted functionality F_{CRS}^D. To generate the common reference string $CRS = (CRS_1, CRS_2)$, F_{CRS}^D computes as follows:

1. Select $e_1, d_0, d_1 \in_R Z_p$, select $g_0 \in_R G_1$, compute $g_1 = g_0^{e_1}, h_0 = g_0^{d_0}, h_1 = g_1^{d_1}$, set $CRS_1 = \langle g_0, g_1, h_0, h_1 \rangle$, and discard the corresponding $Trapdoor_1 = \langle e_1, d_0, d_1 \rangle$.
2. Select $\bar{h} \in_R G_1$, select $\alpha_1, \alpha_2 \in_R Z_p^*$, compute $\bar{g}_1 = \bar{h}^{1/\alpha_1}, \bar{g}_2 = \bar{h}^{1/\alpha_2}$, set $CRS_2 = \langle \bar{g}_1, \bar{g}_2, \bar{h} \rangle$, and discard the corresponding $Trapdoor_2 = \langle \alpha_1, \alpha_2 \rangle$.

Key Setup for Signer. The signer S invokes the trusted functionality F_{reg}^{KS} defined in (Damgård et al. 2006) to register his public key (pk_p, c) and stores the corresponding private key (sk_P, e) secretly, where (pk_P, sk_P) denotes a keypair of the underlying Paillier encryption scheme (Paillier 1999), i.e., pk_P is a RSA modulus n and sk_P is the factorization of n. Also, let c be a Paillier encryption of e under pk_P, where e denotes the challenge of the underlying $\sum$ protocol (Damgård et al. 2006) selected by the signer.

Key Setup for User. Similarly, the user U invokes the trusted functionality F_{reg}^{KS}, and gets his registered public key (pk_U, c_U) and the corresponding private key (sk_U, e_U).

Key Generation. The signer S selects $(x, y) \in_R (Z_p^*)^2$, computes $X = g^x, Y = g^y$, $\tilde{X} = \tilde{g}^x, \tilde{Y} = \tilde{g}^y$, and sets $pk_{BS} = (X, Y, \tilde{X}, \tilde{Y}), sk_{BS} = (x, y)$.

Signature Generation. Given a message m to be signed, the user U and the signer S perform the following steps:

1. The signer S selects $b_0, b_1 \in_R Z_p^*$, computes $j_0 = g_0^{b_0}, k_0 = h_0^{b_0}, j_1 = g_1^{b_1}$, $k_1 = h_1^{b_1}$, then sends j_0, k_0, j_1, k_1 and Π_1 to the user, i.e.,

$$\Pi_1 = NIZK\{(b_0, b_1) : j_0 = g_0^{b_0} \wedge k_0 = h_0^{b_0} \wedge j_1 = g_1^{b_1} \wedge k_1 = h_1^{b_1}\}.$$

2. The user U checks the validity of Π_1. If so, he selects $a_1, a_2 \in_R Z_p^*$, computes $g'_1 = g_1^{a_1}, h'_1 = h_1^{a_2}, j'_1 = j_1^{a_1}, k'_1 = k_1^{a_2}$. Then, he selects $t, r_m \in_R Z_p^*$, computes $C = g^t Y^m, M_0 = g_0^m h_0^m, M_1 = g_1'^m h_1'^m, L_0 = j_0^m k_0^m, L_1 = j_1'^m k_1'^m$. In addition, he selects $r, s \in_R Z_p$, computes $u_1 = \bar{g}_1^r, u_2 = \bar{g}_2^s, u_3 = \bar{h}^{r+s} Y^m$ and Π_2, i.e.,

$$\Pi_2 = NIZK\{(t, m, r_m, r, s) : C = g^t Y^m \wedge M_0 = g_0^m h_0^{r_m} \wedge M_1 = g_1'^m h_1'^{r_m} \wedge$$
$$L_0 = j_0^m k_0^{r_m} \wedge L_1 = j_1'^m k_1'^{r_m} \wedge u_1 = \bar{g}_1^r \wedge u_2 = \bar{g}_2^s \wedge u_3 = \bar{h}^{r+s} Y^m\}.$$

　　Finally, U sends $(g'_1, h'_1, j'_1, k'_1, C, M_0, M_1, L_0, L_1, u_1, u_2, u_3), \Pi_2$ to the signer S.
3. If $j_1' = g_1'^{b_1} = k_1' = h_1'^{b_1}$ holds and Π_2 is valid, the signer S selects $r'', s'' \in_R Z_p^* \times Z_p$, computes $a' = g^{r''}, b' = (g^x C)^{r''} g^{s'' x + xy}$, and sends (s'', a', b') to the user U.

4. The user U sets $a'' = a'$, computes $b'' = b'(a')^{-t}$ and checks whether $\hat{e}(b'', \tilde{g}) = \hat{e}(a'', \tilde{X}\tilde{Y}^m)\hat{e}(X, \tilde{Y})\hat{e}(X, \tilde{g})^{s''}$. If the equation holds, he randomizes the resultant signature (s'', a'', b'') by using the method in Zhou and Lin (2007). Concretely, he selects $r_1, r_2 \in_R Z_p$, computes $s = s'' + r_1, a = a''g^{r_2}, b = b''(XY^m)^{r_2}X^{r_1}$, and sets (s, a, b) as the ZL signature on m.

It is worth noting that the NIZK proofs Π_1 and Π_2 in the above protocol are generated by applying the compilation technique by Damgård et al. (2006) to the underlying $\sum$ protocols.

Signature Verification. On input public key $(X, Y, \tilde{X}, \tilde{Y})$, message m and signature (s, a, b), check whether the following verification equation holds, i.e.,

$$\hat{e}(b, \tilde{g}) = \hat{e}(a, \tilde{X}\tilde{Y}^m)\hat{e}(X, \tilde{Y})\hat{e}(X, \tilde{g})^s.$$

3 Security of the Construction

Theorem 1. *If the signer and the user follow the signature generation protocol, the resulting signature will be accepted by a verifier with provability 1.*

Proof. First, according to the soundness of NIZK proofs Π_1, Π_2, we can conclude that: (1) the Pedersen commitment $C = g^t Y^m$ is honestly generated. (2) M_0, M_1, L_0, L_1 are commitments to message m according to the commitment scheme in Arita (2007), where j_0, k_0, j_1, k_1 are honestly generated by the signer, i.e., the equations $j_0 = g_0^{b_0}, k_0 = h_0^{b_0}, j_1 = g_1^{b_1}, k_1 = h_1^{b_1}$ hold.

Next, we prove that the resulting tuple (s'', a'', b'') is a valid ZL signature on message m. According to $C = g^t Y^m, Y = g^y, b' = (g^x C)^{r''} g^{s''x+xy}$, we get that $b' = (g^x \cdot g^{t+ym})^{r''} g^{s''x+xy} = (g^t)^{r''} g^{(x+ym)r''+s''x+xy}$. As a result, we have that $b'' = b'(a')^{-t} = g^{(x+ym)r''+s''x+xy}$, which implies that $\hat{e}(b'', \tilde{g}) = \hat{e}(g^{r''}, \tilde{g}^{x+ym})\hat{e}(g^x, \tilde{g}^{s''})\hat{e}(g^x, \tilde{g}^y) = \hat{e}(a'', \tilde{X}\tilde{Y}^m)\hat{e}(X, \tilde{Y})\hat{e}(X, \tilde{g})^{s''}$.

Finally, we prove that the randomized signature (s'', a'', b'') is valid. As the equations $s = s'' + r_1, a = a''g^{r_2}, b = b''(XY^m)^{r_2}X^{r_1}$ hold, we can get that

$$a = g^{r''+r_2}, b = g^{(x+ym)r''+s''x+xy} g^{(x+ym)r_2+r_1x} = g^{(x+ym)(r''+r_2)+sx+xy}.$$

So, we can conclude that

$$\hat{e}(b, \tilde{g}) = \hat{e}(g^{(x+ym)(r''+r_2)}, \tilde{g})\hat{e}(g^x, \tilde{g})^s \hat{e}(g^x, \tilde{g}^y) = \hat{e}(a, \tilde{X}\tilde{Y}^m)\hat{e}(X, \tilde{Y})\hat{e}(X, \tilde{g})^s.$$

Theorem 2. *In the registered public key model, the proposed scheme is unforgeable against a concurrent run of adversaries if the underlying Zhou-Lin assumption (Zhou and Lin 2006) holds.*

Proof. Assume that $\mathbf{A^{nf}}$ is an PPT forger that can successfully break our blind signature scheme with non-negligible probability, then we can construct a PPT simulator $\mathbf{S}$, which can invalidate the underlying Zhou-Lin assumption. $\mathbf{S}$ simulates the challenger and interacts with $\mathbf{A^{nf}}$ concurrently as follows:

1. Given the bilinear group parameter *ComInfo* and the public key $pk_{BS} = (X, Y, \tilde{X}, \tilde{Y})$ of the ZL signature scheme, $\mathbf{S}$ generates $CRS_1 = \langle g_0,\ g_1,\ h_0,\ h_1 \rangle$ and keeps the corresponding $trapdoor_1 = \langle e_1,\ d_0,\ d_1 \rangle$ At the same time, $\mathbf{S}$ generates $CRS_2 = \langle \bar{g}_1, \bar{g}_2, \bar{h} \rangle$ and discards the corresponding $trapdoor_2 = \langle \alpha_1,\ \alpha_2 \rangle$ $\mathbf{S}$ generates a public/secret key pair of the underlying Paillier encryption scheme, i.e., (pk_P, sk_P), selects the challenge e for the underlying $\sum$ protocol, and computes c, i.e., the ciphertext of e under pk_P. Then, $\mathbf{S}$ sends *ComInfo*, $CRS_1, CRS_2, pk_{BS}, pk_{KS} = (pk_P, c)$ to $\mathbf{A^{nf}}$ and engages the signing protocol with $\mathbf{A^{nf}}$.
2. Assume that $\mathbf{S}$ is performing the j-th ($j \in \{1, \ldots, q\}$, where q denotes the maximum number of concurrent instance of the signing protocol) instance of the signing protocol with $\mathbf{A^{nf}}$. $\mathbf{S}$ selects $b_0, b_1 \in Z_p$, computes $j_0 = g_0^{b_0}, k_0 = h_0^{b_0}, j_1 = g_1^{b_1}, k_1 = h_1^{b_1}$, and sends $j_0, k_0, j_1, k_1, \Pi_1$ to $\mathbf{A^{nf}}$. Upon receiving the request message $(C, M_0, M_1, L_0, L_1, u_1, u_2, u_3), \Pi_2$ from $\mathbf{A^{nf}}$, $\mathbf{S}$ performs straight-line extraction in the following way:
   (a) $\mathbf{S}$ invokes the KEA1-extractor (Arita 2007) on input (g_1, j_1, g'_1, j'_1), and then $\mathbf{S}$ can extract the secret value a_1, satisfying $g'_1 = g_1^{a_1}$. By running the KEA1-extractor on another input (h_1, k_1, h'_1, k'_1), $\mathbf{S}$ can get a_2, satisfying $h'_1 = h_1^{a_2}$.
   (b) $\mathbf{S}$ sets $\alpha = e_1 a_1 \bmod p, \beta = d_0^{-1} e_1 d_1 a_2 \bmod p, \alpha' = b_0^{-1} e_1 a_1 b_1 \bmod p, \beta' = d_0^{-1} b_0^{-1} e_1 d_1 a_2 b_1 \bmod p$, and computes $g_0^m = (M_0^\beta/M_1)^{1/(\beta-\alpha)}$, $j_0^m = (L_0^{\beta'}/L_1)^{1/(\beta'-\alpha')}$.
   (c) $\mathbf{S}$ runs the KEA1-extractor on input (g_0, g_0^m, j_0, j_0^m), and get the returned value m.
3. $\mathbf{S}$ forwards m to its ZL signing oracle $O_{x,y}(\cdot)$. And $O_{x,y}(\cdot)$ returns a ZL signature (s, a, b). $\mathbf{S}$ keeps $(m, (s, a, b))$ and sets $s'' = s, a' = a, b' = ba^t$, then returns (s'', a', b') to $\mathbf{A^{nf}}$.
4. After all q concurrent instances of the signing protocol are completed, $\mathbf{A^{nf}}$ outputs a fake signature (s^*, a^*, b^*) on message m^*.

Now, $\mathbf{S}$ checks if $m^* \notin Q$ and a^* is not a unit element of G_1, where Q denotes the set of queries that $\mathbf{S}$ has given to $O_{x,y}(\cdot)$. If not true, $\mathbf{S}$ outputs $\perp$. Otherwise, $\mathbf{S}$ outputs $(a^*, b^*/X^{s^*})$, thus $\mathbf{S}$ breaks the Zhou-Lin assumption.

In the following, we show why $\mathbf{S}$ can break the Zhou-Lin assumption. As (s^*, a^*, b^*) is a valid ZL signature, we conclude that the equation $\hat{e}(b^*, \tilde{g}) = \hat{e}(a^*, \tilde{X}\tilde{Y}^{m^*})\hat{e}(X, \tilde{Y})\hat{e}(X, \tilde{g})^{s^*}$ holds. Since $a^*, b^* \in G_1$, we can rewrite them as $a^* = g^\alpha, b^* = g^\beta$. It is clear that

$$\hat{e}(b^*, \tilde{g}) = \hat{e}(g^\beta, \tilde{g}) = \hat{e}(g, \tilde{g})^\beta \tag{1}$$

$$\hat{e}(a^*, \tilde{X}\tilde{Y}^{m^*})\hat{e}(X, \tilde{Y})\hat{e}(X, \tilde{g})^{s^*} = e(g, \tilde{g})^{\alpha(x+ym^*)+xy+xs^*} \tag{2}$$

The above two equations give us:

$$\beta = \alpha(x + ym^*) + xy + xs^* \bmod p.$$

Thus we can get that

$$b^*/X^{s^*} = g^{\beta-xs^*} = g^{\alpha(x+ym^*)+xy}.$$

Obviously, $(a^*, b^*/X^{s^*})$ must be of the form of $(g^\alpha, g^{\alpha(x+ym^*)+xy})$. So we get a contradiction with the Zhou-Lin assumption.

Theorem 3. *In the registered public key model, the proposed scheme is blind if the underlying LE scheme is secure with semantically secure against a chosen-plaintext attack.*

Proof. (sketch) Assume that a PPT adversary $\mathbf{A^b}$ is successful in breaking the blindness of the proposed scheme, and then we can construct a PPT algorithm $\mathbf{B}$ to break the semantic security of the LE scheme. Due to space limitations, the complete proof is not provided.

4 Performance Considerations

In Tables 1, 2 and 3, we summarize the comparison between ours and previous typical concurrent-secure blind signature schemes, with respect to important properties, computational cost and space efficiency. The notation "CRS" means the common reference string model. The notation "RPK" means the registered public key model. The notation "RO" means the random oracle model. In Tables 2 and 3, the parameter n denotes the bit-length of the message to be signed. Now, we show the methods and the concrete parameters we take in the comparison.

1. Assume that we are aiming for achieving the security level comparable to 1,024 bit RSA. To note that our scheme is implemented in the asymmetric setting, and the scheme in Kiayias and Zhou (2006), Gjøsteen and Kråkmo (2008) are

Table 1 Comparison of our scheme to recently typical schemes

Scheme	Communication moves	Model
Kiayias and Zhou (2006)	4	CRS
Au et al. (2010)	2	RO
Gjøsteen and Kråkmo (2008)	2	RPK
Our scheme	3	RPK

Table 2 Computation cost of our scheme and recently typical schemes (in number of multiplications)

	Computation cost (user)	Computation cost (signer)	Computation cost (verifier)
Kiayias and Zhou (2006)	26,720	8,432	15,600
Au et al. (2010)	126,064	1,536	403,456
Gjøsteen and Kråkmo (2008)	$28,416n + 4,564,000$	$20,352n + 3,257,600$	7,680
Our scheme	43,904	30,032	8,160

Table 3 Space efficiency of our scheme and recently typical schemes (in bits)

Scheme	Space efficiency (user)	Space efficiency (signer)	Signature length
Kiayias and Zhou (2006)	12,768	2,720	1,536
Au et al. (2010)	160	2,048	89,248
Gjøsteen and Kråkmo (2008)	$29,184n + 4,670,272$	1,024	1,024
Our scheme	23,680	9,792	480

implemented in the symmetric setting. We make the following estimates: p is about 160 bits, elements of G_1, G_2, G_T in the asymmetric setting are $160, 960, 960$ bits respectively, and elements of G_1, G_T in the symmetric setting are $512, 1,024$ bits respectively. For our scheme and the scheme in Gjøsteen and Kråkmo (2008), a possible choice of the bit-length of the modulus of the underlying Paillier encryption scheme is 2,048 bits (Gjøsteen and Kråkmo 2008; Damgård et al. 2006). For the scheme in Au et al. (2010), the authors suggested to take the elements of Z_p, Z_q, G_1 to be $1,024, 160, 1,024$ bit respectively.

2. In the comparison of Table 2, the computational costs of users and signers are all counted in the number of multiplications. By "$ExpCost(x)$", we mean the number of multiplications required to compute an exponentiation with an x-bit exponent. Similarly, "$ExpCost^n(x)$" means the number of multiplications needed to compute the product of n exponentiations with x-bit exponents. According to Peng et al. (2007), we get that:

$$(i)\,ExpCost(x) = 1.5x;\;(ii)\,ExpCost^n(x) = x + 0.5nx.$$

3. As shown in Le et al. (2009), the computation cost of a pairing is estimated to be five 512-bit exponentiations, i.e., $5ExpCost(512)$.

From the above comparisons, it is clear that our scheme significantly improves the computational cost and space efficiency of the Gjøsteen-Kråkmo scheme (Gjøsteen and Kråkmo 2008). Compared with the scheme in Kiayias and Zhou (2006), our scheme requires less communication moves and has better efficiency in verification. Moreover, in contrast with the very recent scheme by Au et al. (2010), the resultant signature of our scheme is much shorter and easier to verify.

5 Conclusions

This paper presents a concurrent-secure blind signature scheme without random oracles in the registered public key model, which achieves all the properties required by blind signature schemes, i.e., correctness, blindness and unforgeability. Performance comparison shows that the cost (respect to communication and computation) of the new scheme is significantly lower than those of the Gjøsteen-Kråkmo scheme. Furthermore, by using the up-to-date technique of straight-line extraction, the security reduction algorithm is efficient and run in strictly polynomial-time.

References

Abe M (2001) A secure three-move blind signature scheme for polynomially-many signatures. In: Pfitzmann B (ed) Eurocrypt 2001. LNCS, vol 2045. Springer, Heidelberg, pp 136–151

Arita S (2007) A straight-line extractable non-malleable commitment scheme. IEICE Trans Fundam Electron Commun Compu Sci E90-A(7):1384–1394

Au MH, Susilo W, Mu Y (2010) Proof-of-knowledge of representation of committed value and its applications. In: Steinfeld R, Hawkes P (eds) ACISP 2010. LNCS, vol 6168. Springer, Heidelberg, pp 352–369

Bellare M, Namprempre C, Pointcheval D, Semanko M (2003) The one-more-RSA-inversion problems and the security of Chaum's blind signature scheme. J Cryptol 16(3):185–215

Boneh D, Boyen X, Shacham H (2004) Short group signatures. In: Franklin MK (ed) CRYPTO 2004. LNCS, vol 3152. Springer, Heidelberg, pp 41–55

Camenisch J, Koprowski M, Warinschi B (2005) Efficient blind signatures without random oracles. In: Blundo C, Cimato S (eds) SCN 2004. LNCS, vol 3352. Springer, Heidelberg, pp 134–148

Damgård I, Fazio N, Nicolosi A (2006) Non-interactive zero-knowledge from homomorphic encryption. In: Halevi S, Rabin T (eds) TCC 2006. LNCS, vol 3876. Springer, Heidelberg, pp 41–59

Fischlin M (2006) Round-optimal composable blind signatures in the common reference string model. In: Dwork C (ed) CRYPTO 2006. LNCS, vol 4117. Springer, Heidelberg, pp 60–77

Gjøsteen K, Kråkmo L (2008) Round-optimal blind signatures from waters signatures. In: Baek J, Bao F, Chen K, Lai X (eds) ProvSec 2008. LNCS, vol 5324. Springer, Heidelberg, pp 112–126

Hazay C, Katz J, Koo C-Y, Lindell Y (2007) Concurrently-secure blind signatures without random oracles or setup assumptions. In: Vadhan SP (ed) TCC 2007. LNCS, vol 4392. Springer, Heidelberg, pp 323–341

Juels A, Luby M, Ostrovsky R (1997) Security of blind digital signatures. In: Burton S, Kaliski R Jr (eds) CRYPTO 1997. LNCS, vol 1294. Springer, Heidelberg, pp 150–164

Kiayias A, Zhou H-S (2006) Concurrent blind signatures without random oracles. In: Prisco RD, Yung M (eds) SCN 2006. LNCS, vol 4116. Springer, Heidelberg, pp 49–62

Le D-P, Bonnecaze A, Gabillon A (2009) Multisignatures as secure as the Diffie-Hellman problem in the plain public-key model. In: Shacham H, Waters B (eds) Pairing 2009. LNCS, vol 5671. Springer, Heidelberg, pp 35–51

Okamoto T (2006) Efficient blind and partially blind signatures without random oracles. In: Halevi S, Rabin T (eds) TCC 2006. LNCS, vol 3876. Springer, Heidelberg, pp 80–99

Paillier P (1999) Public-key cryptosystems based on composite degree residuosity classes. In: Stern J (ed) Eurocrypt 1999. LNCS, vol 1592. Springer, Heidelberg, pp 223–238

Peng K, Boyd C, Dawson E (2007) Batch zero knowledge proof and verification and its applications. ACM Trans Inf Syst Secur 10(2):1–28

Waters B (2005) Efficient identity-based encryption without random oracles. In: Cramer R (ed) Eurocrypt 2005. LNCS, vol 3494. Springer, Heidelberg, pp 114–127

Zhou S, Lin D (2006) Shorter verifier-local revocation group signatures from bilinear maps. In: Pointcheval D, Mu Y, Chen K (eds) CANS 2006. LNCS, vol 4301. Springer, Heidelberg, pp 126–143

Zhou S, Lin D (2007) Unlinkable randomizable signature and its application in group signature. In: Pei D, Yung M, Lin D, Wu C (eds) Inscrypt 2007. LNCS, vol 4990. Springer, Heidelberg, pp 328–342

Information Retrieval in University Student's Self-Regulated Learning

Kang Yanyu

Abstract As an emerging learning way, Online Self-regulated Learning (OSL) is accompanying the booming of China's Internet. Digital library as a powerful tool of Internet assisted learning is so important in enhancing the learning, but there are not many researches about the importance of it in China, most of them are not thorough. In this study we considered the domestic educational situation, chose university students as samples, did some researches about theirs OSL's situation and how the digital library influence the learning, after those the suggestions would be given. To study the current situation and impact factors of university student's OSL and the influence of digital library, the research took qualitative research which is good at examining details, meanwhile the researcher selected eight university students as respondents, intensively interviewed and observed the respondents regarding the current situation and impact factors.

Keywords Online self-regulated learning (OSL) • Digital library • Qualitative research

1 Background

Accompanying with the development of various fields (materials technology, information technology, communication etc.) in mid twentieth century, computer networks, especially the Internet, was developing rapidly. In 1960s, the computers were connected by wires through serial ports in low speed connection. At that time, the "network" just provided sample services such as remote printing. Network users by now, especially Internet users, can access high bandwidth Internet in different

K. Yanyu (✉)
Library of Chongqing Three Gorges University, Wanzhou District, Chongqing, China
e-mail: kangyanyu80@gmail.com

ways. Through the network, users can do many things include search information, download and share the files. The China's Internet is getting boom after 2000. According to the statistics of CNNIC (China Internet Network Information Center), by Jan. 15th, 2010, the Internet users in China reached 38,400 million, 28.8% of them are students, and most of students were university students.

Meanwhile, in the learning theory research, self-regulated (self-directed, self-monitored) learning came to one important issue (McCombs 1989). There were many researches about it have been done by the scholars around the world, and many results were achieved. Those studies were focused on the essence of self-regulated learning, internal mechanisms of its development, impact factors etc (Dong Qi and Zhou Yong 1996). In recent years, the high bandwidth Internet becomes popular, it changes the life of us (Butler and Winnie 1995). The connection between self-regulated learning and Internet is getting close (Allen and Seaman 2006): One the one hand, Internet provides learner mass learning materials and convenient data search. Through the Internet learner can search almost every material (video, paper, digital books etc.) what he/she needs in learning process, and the convenient access makes the academic communication much easier. On the other hand, Online Self-regulated Learning promotes the development of Internet in a way. Because of learners' request about learning materials, vast related databases were established, links to the search engine, learners can retrieve information expediently. It made more and more learners use Internet as a learning way. Some learning forums and FTP severs provide mass attractive files to learners, that made a lot of learners who never or seldom use Internet before change their learning ways.

The development of Internet changed the library at the same period (Zhang Xiaoyan 2007). Thanks to the Internet, the library is more "digital" than before. Compared with the traditional library, the digital library can serve the users around the clock that traditional one can do, and more users can get what they need in digital serve way. What's more, the Internet made library's service available any place that you can use the net.

Self-regulated learning is the popular learning way of university students. Self-regulated learning is a process that learners making learning goals, choosing learning methods, monitoring the process, evaluating the results by themselves (Pang Weiguo 2000). In order to enhance students' study efficiency, students have to get strong self-discipline, that is the reason why self- regulated learning is major learning way of university students. As mentioned previously, library, especially university library has changed its way to serve the users because of the development of information and communication technology. In order to avoid the disadvantage (such as working time, location etc.) of traditional library, the school has to increase its budget to purchase digital resource (journals, books etc.), storage device, network device and sever etc. Further more, through the campus network and VPN (Virtual Private Network), university library provide students and staffs local or long distance data service which is free to access in the campus (Moore 1991). Meanwhile, most of schools have information retrieval course, students can learn how to do online information retrieval. Also, it makes more and more students

Fig. 1 Case list. Note: The 1st letter is the code of school, the 2nd is department, and the 3rd is gender (Howland and Moore 2002)

Code	Sex	Major	Grade
SFF01	F	Costume Designing	2
SFM01	M	Costume Designing	3
JXM01	M	Psychology	4
CZF01	F	Chinese Education	3
CHF01	F	Material Chemistry	3
SGM01	M	Engineering	2
SGF01	F	Engineering	2
SWM01	M	English Education	2

aware that they can use more convenient and efficient way to do self-regulated learning except traditional learning way. The popularity of Internet and personal computer makes increasingly students use Internet easily. To the university students, self-regulated learning is the best learning method, that's the reason why we focus on the how does digital library influent the self-regulated learning (Barnard et al. 2008).

2 Procedures

After review the related literature, found there is few scholars did research about it, most of study is literature study. Considering the situation of domestic related studies, as well as the objects' understanding of digital library and Internet assisted learning, in this study the literature study and case study (interview and observation) were taken (Chen Xiangming 2000). In the study, we define the OSL as the self-regulated learning on the Internet surrounding. Based on the definition and eight dimensions of self-regulated learning (Dong Qi and Zhou Yong 1994), the semi-structured interview outline was developed. After that, we pick up some students with purposeful sampling (Fig. 1) (Cavanagh and Romanoski 2004).

In order to improve the reliability of the study, during the process of choosing the influence of gender, grade, major was considered. The interview based on the situation of OSL and the how did the digital library influent the process of OSL. The whole interview was recorded in audios and noted. The follow steps, all the audios were transformed into documents, then coded and categorized. After generic analysis and scenario analysis, the conclusion was developed.

3 Findings

Compared with the popularity of Internet, OSL is in poor awareness level, also the learning activity. But the digital resource of the library help the student's OSL a lot. On one hand, to the university students in China learning in the class is still

the major learning method, and this kind of situation will exist for a long period of time. On the other hand, because of the trend of library's digitization, library's digital resource is the major source.

3.1 Poor Preparation

Because have been influenced by many factors, most of students didn't care or know how to develop the learning surroundings during the OSL, especially the internal surrounding (psychological surroundings) was ignored. Comparatively speaking, in the OSL students did plenty of preparation, different individual had their own preferences in choosing their own learning resources. For most of students the library is the best place to learn, quiet and academic. The digital library makes learning can be available in various place, it means more options are available for students.

3.2 Lack of Planning

There is not proper planning in their OSL for most of university students. Most of them barely can set reasonable learning goal. Because their learning was used to be negative, most of them don't know how to manage their learning tasks.

3.3 Average Consciousness

The students knew what were their learning objects and how did the OSL influence the learning. But about how to set learning goal and task allocation, most of their awareness were not so clear.

3.4 Diversity of Methods

During the OSL, students learnt in various ways. The verbal, audio and video's searching was the major way in retrieval. In the retrieval students could get almost everything they needed through search engine and public or specific database. Otherwise, some of them used teaching video and digital book etc. as learning materials. During the OSL process, students would take some task strategies, such as extra tasks and retrieving for research interests. Obviously the best way to get learning materials was access the digital resource available in the library.

3.5   Seek Help Without Internet

In the learning process, when the students got questions, they prefer to solve it by themselves, ask for help was the last option in the retrieving on the Internet. Students were good at seeking help by E-mail, forum etc. They prefer the fast feedback, such as face to face talking or IM.

3.6   Lack of Implementation

Even university students realized that the interference from the outside or Internet itself influence the whole learning process negatively, but they got few methods to resist it. Compare with the interference (online games, soap show etc.), learning is lack of attraction.

3.7   Lack of Self-Monitoring

University students seldom monitored OSL process, knew few way to evaluate the learning effect. Some times those steps were ignored. When the learning effect turned dissatisfied, even the students noticed that they had done something, but still could found solution. Otherwise, the way they did summary was random. Even so, most of them will review the gain and loss in the learning process.

3.8   High Dependence in Library's Digital Resource

In the OSL, university students would search learning materials from different search engine, but the major resource was the digital resource of Library in the school. Follow reasons were given by them: The digital resource in the campus network is free, they can save expense; The digital resource of library can satisfy their needs in learning and research; The service of digital library make their learning more convenient, they can use the classified files through the library; It's a accustomed learning behavior to use digital resource of library, the local resource will be the first choice when they need learning resource.

4   Suggestions

After study, found there is something can be improved:

4.1   Change the Concept of Learning Way

For most of university students learning is reading printed matter or listening lessons in the classroom, network is just a tools to have fun, the only relation to learning is retrieval when we writing paper or doing homework. Of course the Internet provides us plenty of entertainment stuffs, those addicted some students, but also it is a powerful learning platform, we have to admit this. And through it, we can learn something is not in textbooks. Obviously, this kind of platform will enhance the whole learning, that is the reason why the students have to change their concepts about learning, try to learning how to learning in digital era, to make a win-win situation for entertainment and learning.

4.2   More Education Has to Be Given About How to Learning Through the Net

To the university students, they totally understand the positive influence of Internet to their learning. But in China, classroom learning is still the major learning way to the students, other learning ways were still not developed adequately, especially the Internet which got plenty of learning resources. For this reason, school should strengthen the education of basic skills about online learning. The follow questions have to be solved: how to use the software, how to do retrieval, where to find the learning resources.

4.3   Developing Good Surroundings for Study

It is the responsibility of the school to construct the campus network to develop a good learning surrounding for students' OSL. The infrastructure of campus network made Internet access available almost everywhere in the campus, it makes OSL in the quiet place available. Furthermore, the cost of online learning has to be reduced. The resources in the campus network are almost free for use, so the construction of digital library is so important, not only the hardware construction but also the database purchasing. Library should find more available free resources for the students in without break the law. The inter-library loan is a good way for the students gets more quality learning resources around the world with few cost.

4.4   More Direction and Proper Support

Educator and parents should aware that the Internet benefits the students' learning, which is the prerequisite of objective attitude. Parents should not be so anxious about Internet using, they should understand the university students are grown-ups, and they can deal the relationship between Internet entertainment and learning. School should guide students how to use Internet in learning, develop their OSL habit. More learning task can be assigned to students related online learning to support the OSL.

4.5   Encouraging Learning Exchange

Self-regulated learning is not do everything by learner itself, the exchange is a good way to find our disadvantage and get necessary help to make progress in learning. The education should focus on how to improve students' skills of learning exchange, especially the skills through the Internet. The IM, E-MAIL, forum are the good ways to do learning exchange.

4.6   Developing Better Learning Habit

The temptation in the OSL is much stronger than traditional self-regulated learning, so the good learning habits ensure learning well. The habits are routines of behavior that are repeated regularly and tend to occur subconsciously, without one being conscious about them. Therefore, the students have to keep their learning planning, eliminating interference, evaluating the learning periodically and correct the fault, do summary of whole learning process, realize the weakness and find the way to improve themselves.

References

Allen E, Seaman J (2006) Making the grade: online education in the United States. The Sloan Consortium. Retrieved 19 Oct 2008, from http://www.sloanc.org/publications/survey/pdf/making_the_grade.pdf

Barnard L, Paton VO, Lan WY (2008) Online self-regulatory learning behaviors as a mediator in the relationship between online course perceptions with achievement. Int Rev Res Open Distance Learn 9(2):1–11

Butler DL, Winne PH (1995) Feedback and self-regulated learning: a theoretical synthesis. Rev Educ Res 65:245–281

Cavanagh RF, Reynolds PS, Romanoski JT (2004) Information and communication technology learning in the classroom: the influence of students, the class-group, teachers and the home. Paper presented at the 2004 annual meeting of the Australian Association for Research in Education, Melbourne [Online]. http://www.aare.edu.au/04pap/cav04445.pdf. Accessed 8 Sept 2006

Chen Xiang-ming (2000) How to be a qualitative researcher. Educational Science Publishing House, Beijing

Dong Qi, Zhou Yong (1994) The self-monitoring of students' learning. J Beijing Norm Univ Soc Sci Edit 1:8

Dong Qi, Zhou Yong (1996) Self-monitoring and intelligence. Zhejiang People's Publishing House, Hanzhou

Howland JL, Moore JL (2002) Student perceptions as distance learners in Internet-based courses. Distance Educ 23:183–195

McCombs BL (1989) Self-regulated learning and academic achievement: a phenomenological view. In: Zimmerman BJ, Schunk DH (eds) Self-regulated learning and achievement: theory, research, and practice. Springer, New York, pp 51–82

Moore MG (1991) Distance education theory, DOESNEWS 1:25

Pang Wei-guo (2000) The advances of self-regulated learning abroad in 1990s. Adv Psychol Sci 4:12

Zhang Xiao-yan (2007) Investigation and analysis of the situation about digital resource utilization in university library. Libr Sci Res Work 1:22–23

Tracing Data Provenance Based on Inverse Mechanism in ETL

Tao Wang and Chao-fan Dai

Abstract Data integration may have no meaning without data provenance. ETL is an important step of data integration. To trace the provenance of ETL process is significant for data integration. A method based on inverse functions is described in this paper. Firstly, the ETL process is divided into three layers; then we analyzed each layer and designed the inverse mechanism, and described the algorithms of provenance tracing based on this. At last, we take an instance to show the result.

Keywords Data provenance • Inverse functions

1 Introduction

Extract, Transformation and Loading of data set is a common process of data integration. Data provenance is vital for the integration result, some one records the whole process of ETL to trace provenance, it may be perfect when the amount of data set is small, but it fails to solve the large amount of data set. Actually, It is unnecessary to record all of the data for some process which can be reversed, what it needed is just the process, not the whole data set.

The tracking methods of provenance information are mainly annotation-based (Tan et al. 2005) or reverse-based (Poulovassilis 2002) methods. Other scholars also have proposed to use bit vector to locate memory location (Siddharta 2005) or use graph theory and special query language to track (Karvounarakis 2009) it. Denotes (Siddharta 2005) is a typical system using annotation. Literature Cui and Widom

T. Wang (✉) • C. Dai
Science and Technology on Information Systems Engineering Laboratory, College of Information, System and Management, National University of Defense Technology, Changsha 410073, China
e-mail: wangtao@nudt.edu.cn; cfdai@263.net

(2003) discussed the reverse tracking mechanism of data provenance in database in detail. Literature Cui and Wiener (2000), Green et al. (2007) proposes one method of getting data provenance by analyzing the query syntax and mathematical structure separately. Mixed strategy is another choice. Literature Buneman and Tan (2002) provided a solution from one side. Literature Cong and Geerts (2006) proposes the storing problem of complex annotations which contain the operations of keywords. Literature Cui and Widom (2003) tracks the weak reverse functions, and makes some comprehensive analysis.

In this paper, we firstly decomposed the process of tracing and transformation; after that, we designed the tracing process and algorithms; at last, we show our design in a case.

2 Transformation and Tracing Decomposition

In data integration field, tracing process is the opposite side of transformation in some point. Transformation and tracing process are decomposed in this section. Two concepts about reverse should be discussed before we decompose those processes, one is "Full Reverse', and the other is "Condition Reverse". The transformations which can not be reversed are not discussed in this paper.

Definition 1(Full Reverse)

$\forall D \in \{T, M, P\}$, D is a process, $\exists D(I) = O$, if $\exists D^{-1}$ satisfied follow conditions:

(1) $D - 1(D(I)) = O$
(2) $D(D - 1(O)) = O$

So the D can be reversed, and its reverse process is signed as D^{-1}.

Definition 2(Conditional Reverse)

$\forall D \in \{T, M, P\}$, D is a process, $\exists D(i_1) = o_1, T(i_2) = o_2$, if $\exists D^{-1}$ satisfied follow conditions:

(1) $D^{-1}(D(i_1)) = o_1, D^{-1}(D(i_2)) \neq o_2$
(2) $D(D^{-1}(o_1)) = o_1, D(D^{-1}(o_2)) \neq o$

And there is no D' satisfied $D'(D(i_2)) = o_2 \& D(D'(o_2)) = o_2$. So the D can be conditional reversed, and its conditional reverse process signed as D^{-1}.

2.1 The Decomposition of Transformation Process

Transformation can be divided into three layers, transform (T) is on the top, one transform included several mappings, and one mapping (M) is composed by one or more operations (P). The relations between them can be showed as Fig. 1. That is the composition of transformation.
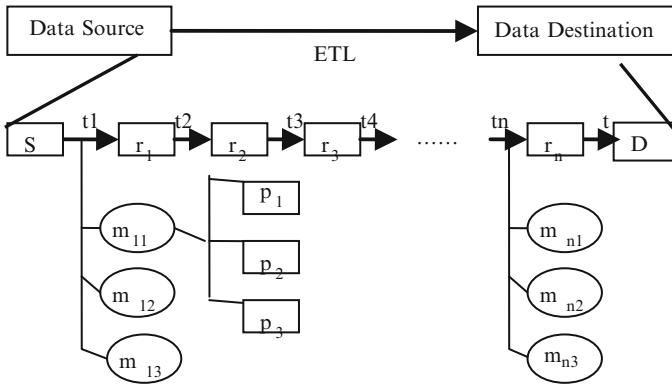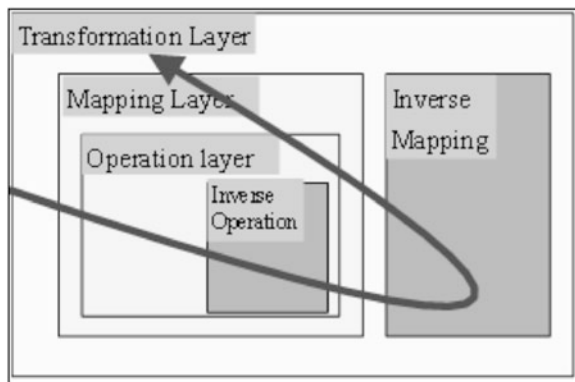
Fig. 1 The relations between ETL, T, M and P

Fig. 2 The relations between reverse and T, M, P

Definition 3(Operation)

One operation is composed by three parts, signed as P ::= <a,b,p>. ∀P satisfied follow conditions:

(1) a represents the input parameters list;
(2) b represents the output object;
(3) p represents operation rules, satisfied p (a) = b;

There are some other conceptions such as transform, mapping, inverse mapping etc. Actually, there is one concept in one layer for any type of reverse. For more details, please reference thesis (Chaofan Dai 2002). Actually, the relations between reverse and T, M, P as showed in Fig. 2.

2.2 The Decomposition of Tracing Process

Tracing tasks can be divided into four layers for ETL process, as we know, the ETL process is transforming the data from data sources to data destinations, and the

source or destination includes different types such as relation data base etc., so we treated the process of tracing as different layers, the top layer is database layer, as DB layer for short; the second is relation layer; then column or tuple layer is the third, the last is item layer.

3   Tracing Design

There are ETL and tracing layers, and then what we need are the contents of each layer, the inverse functions and tracing algorithm. These problems are discussed in this section.

3.1   The Contents of Each Layer

As we discussed above, operation layer is the foundation of the whole process. Taking SSIS (Microsoft SQL Sever 2005 Integration Service) which is an common ETL tool for example, there are six types of operation, including math operation, string operation etc., and there are several operations for each type, such as "Reverstring" belonged to string operation which reverse one string to its opposite order, it is easy to define its reverse function, obviously, it is itself. Because the operation is the bottom of the whole process, its own status totally decides whether it can be reversed or not. Of course there are some operations which can be condition reversed, such as ABS operation, which gets the expression's absolute value, so if its result is 0, and then we can infer that its input is 0. Those operations which can not be reversed are not our concern.

There are some differences for mapping layer and transformation layer, their reverse status are decided by the elements which composing them, if one of them can not be reversed, the whole transformation or mapping can not be reversed. It is the same situation for condition reverse. If the result satisfied reverse conditions, then it joins the reverse elements group.

3.2   The Inverse Functions

The inverse function is irreplaceable for tracing, no inverse function, no tracing. If one operation can be reversed, it must have an inverse function, or else, it can not be reversed as well. For operation layer, each operation which can be reversed has an inverse function, there is an example of "Reverstring" in section A, and the "ABS" case is an example of condition reverse, its inverse condition is that the result equals 0. The inverse function are defined by the designer if there are no existed functions in the assemble of operations. The description of functions can be saved as xml files as follow format:

```
<Function>
        <FunctionID>FunctionID</FunctionID >
        <FunctionName>FunctionName</FunctionName>
        <CounterFunction>CounterFunctionName</ CounterFunction >
        <FunctionType>The type return</FunctionType>
        <IDE>Develop&Excute environment</IDE>
        <FunctionDeveloper>Developer</FunctionDeveloper>
        <FunctionRegisterDate>RegisterDate</FunctionRegisterDate >
        <FunctionFunInfo>The function of the function</FunctionFunInfo>
        <ParamsCount>the number of the parameters</ParamsCount>
        <FunctionPath>location of the function</FunctionPath >
        <FunctionNameSpace>namespace of the function</FunctionNameSpace>
        <FunctionClass>the class name of the function</FunctionClass >
        <Parameter>
                <ParaNumber>1</ParaNumber>
                <ParaName>the first parameters</ParaName>
                <ParaType>type of the parameters</ParaType>
                <ParaInfo>the meaning of the parameters</ParaInfo>
        </Parameter>
        ………..
</Function>
```

The funtionLocation is the location where the function itself is. Actually, the standard functions need not this item; it just designed for user-defined functions. When we trace a target, the inverse functions are indexed by these description files.

There is no uniform format of inverse function for mapping or transformation which can be reversed; its inverse function can be described as an assembly of inverse operations or mappings.

3.3   Tracing Algorithm

Tracing process is a process of combining overall resources above all. The tracing targets are traced step by step; the result of pre-step is the foundation of next step until the tracing process reach to the end. The process can be described as follow algorithm:

Input: Tracing target
Output: Data provenance of Tracing Target
Begin

Step 1. Initialize every situation, including all of the operations which can be reversed or condition reversed, the reverse condition for condition reverse operation, the inverse function's information etc.

Step 2. Judging the layer of the tracing target, if the target belongs to DB layer, trace the provenance of the DB target; if this process can be reversed, then turn to Step 8, or else, turn to Step 7.

Step 3. Tracing the provenance of the relation layer. If the target belongs to the relation layer, and this relation process can be reversed, then turn to Step 8, or else, turn to Step 7.

Step 4. Tracing the provenance of the column layer. If the target belongs to the column layer, and this process can be reversed, then turn to Step 8; if the target belongs to the tuple layer and this is not the end of the columns, then turn to step 6; or else, turn to Step 7.

Step 5. Combine the result of each column, and then turn to Step 4.

Step 6. Trace the provenance of the item layer. If this process can be reversed, then find the inverse functions, and treats the expression's result as input of inverse function to get output, turn to Step 8, or else, turn to Step 7.

Step 7. Return −1.

Step 8. Serialize the result and save them, return the tracing result in user's preference, then visualize the result to user.

End

4 Tracing Example

We designed a provenance system based on SSIS, the frame of this is shown as Fig. 3, left part of this frame is the ETL module, and the other side is the tracing module. The ETL module can be divided into three parts, one is the data sources such as relation database, excel files and so on; the second part is the core part of ETL, they worked together by many rules. The third part is the data destination, the same as the first part. What the ETL module's output is the description file-the package, it has its own format. This package is the input of the tracing module, when it received the package, it would be serialized at first, and then parses the serialized package by parse engine. With the help of inverse functions assemble, reverse engine receive user's tracing request and return the visualization tracing result in different layers.

Based on this system, a typical case about data integration is described as follow:

Supposed that the university A combine university B, so the students' information of B should add to A's database, it is an ETL process. Their table struts and data are listed as Tables 1, 2, 3, and 4. The total record of Table 3 results from average record plus to the number of courses of Table 1.
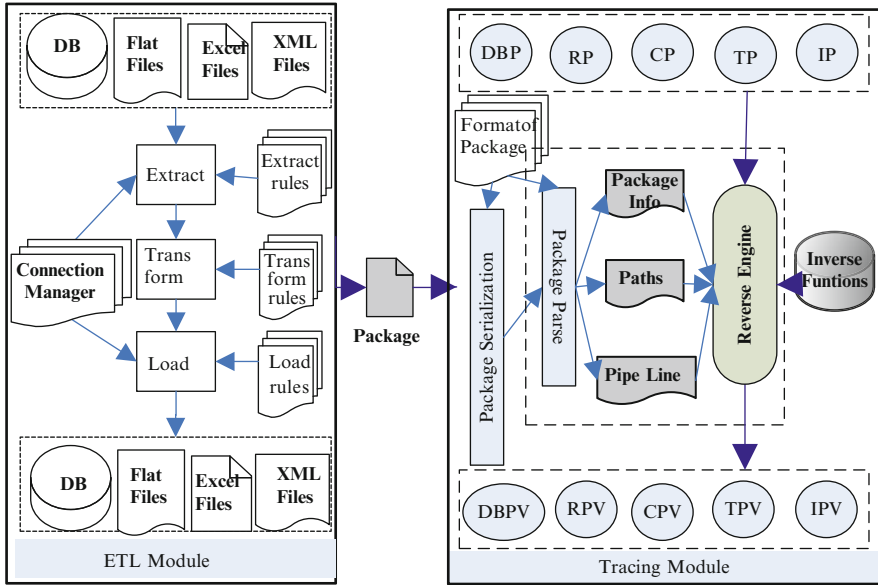
Fig. 3 Tracing system design

Table 1 The strut of students

Column	Name	Gender	Average Record	The number of courses
Data type	nvarchar(50)	nchar(10)	int	int

Table 2 The data of students

Name	Gender	Average Record	The number of courses
Tom	M	82	3
Andy	F	73	4
Peter	M	92	5

Table 3 The strut of student new

Column	Name	Gender	Total Record	Number of courses
Data type	nvarchar(50)	nchar(10)	numeric(6, 2)	int

Table 4 The data of student new

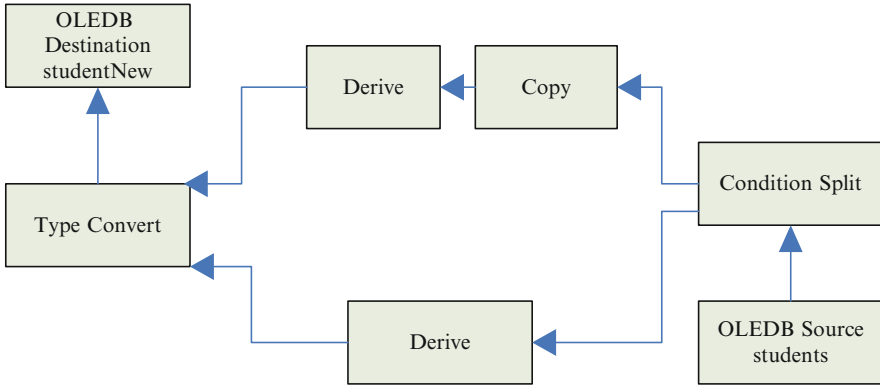| Name	Gender	Total Record	The number of courses
Tom	M	246.00	3
Andy	F	292.00	4
Peter	M	460.00	5

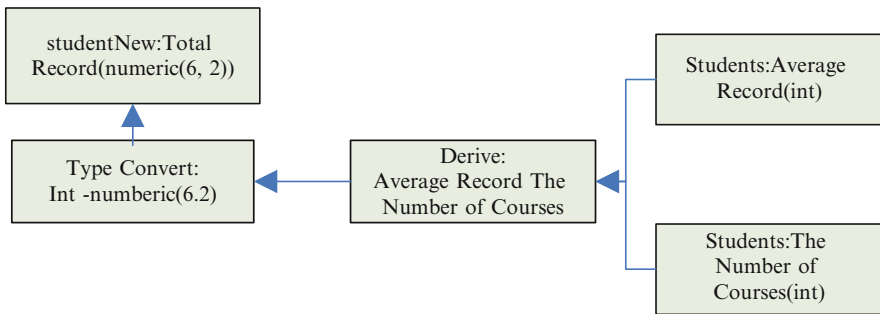**Fig. 4** The provenance of relation layer

Fig. 5 The provenance of column layer

When the user want to know where the result "292.00" is came from, they can use the system to trace the process and the data source which generated this result. The system shows the provenance step by step as follows:

Firstly, the system shows the outline of this result, namely the provenance of relation layer. It just shows the whole process of the ETL. The result is described as Fig. 4. Secondly, the provenance of column layer is showed as Fig. 5, it showed some details in control flow of ETL. In this case, it showed that "Average Record The Number of Courses = Total Record". At last, the user can get the provenance of data items, it showed the details of the data item came from. In this case, it showed that there are two sources, one is the "Average Record", namely 73, the other is "The Number of Courses", namely 4, then the value of "Total Recode" equals to the "Average Recode"*"The Number of Courses" namely 4*73 = 292, after that, the data type of column "Total Record" convert from int to numberic(6,2), namely 292 to 292.00. It is the end of provenance. It explained the whole process of the data item "292.00" (Fig. 6).
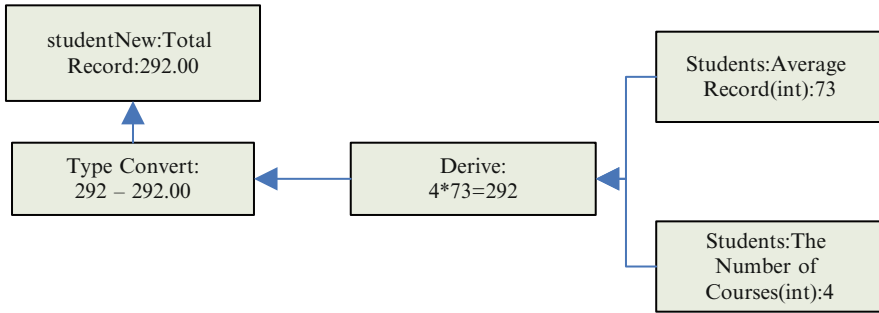
Fig. 6 The provenance of item layer

5 Conclusions

This paper described one method to trace provenance for ETL process without recording the temporary results, and showed the process of tracing by an example. It showed the promising application of data provenance in data integration field.

References

Buneman SKP, Tan WC (2002) On propagation of deletions and annotations through views. In: The ACM SIGMOD-SIGACT-SIGART symposium on principles of database systems (PODS), Madison, pp 150–158

Chaofan Dai (2002) Theories and approach of data lineage tracing in data warehouse environment. ChangSha, China

Cong WFG, Geerts F (2006) Annotation propagation for key preserving views. In: ACM international Conference on Information and Knowledge Management (CIKM), Arlington, VA, USA, pp 632–641

Cui Y, Widom J (2003) Lineage tracing for general data warehouse transformations. The International Journal on Very Large Data Bases 12(1):41–58

Cui JWY, Wiener JL (2000) Tracing the lineage of view data in a warehousing environment. ACM Trans Database Syst 25(2):179–227

Fan H (2002) Tracing data lineage using automed schema transformation pathways. Adv Database 44–55

Gadang S (2007) Bit vector based provenance tracking systems. Doctoral dissertation, University of Arkansas, Fayetteville

Green TJ, Karvounarakis G, Tannen V (2007) Provenance semirings. In: ACM SIGMOD-SIGACTSIGART symposium on principles of database systems (PODS), Arlington, VA, USA, pp 675–686

Karvounarakis G (2009) Provenance in collaborative data sharing. In: Computer and information science. PhD thesis, University of Pennsylvania

Tan WC, Chiticariu L, Vijayvargiya G (2005) DBNotes: a post-it system for relational databases based on provenance. In: Proceedings of the ACM SIGMOD international conference on management of data (SIGMOD), Santa Cruz, pp 942–944

Phase Unwrapping for Large Discontinuous Surface Based on Color Fringe Projection

Wenguo Li, Shaojun Duan, and Zhihong Yin

Abstract A phase unwrapping method based on color fringe projection is presented. Color pattern is utilized to identify absolute fringe order and calculate wrapped phase distribution, simultaneously, which is time saving. The proposed method can overcome phase discontinuous problem. Three channel values in HSV space are employed to calculate pattern color intensity, the relative position between adjacent pixel's hue channel value is used to decide absolute phase fringe order. The hue channel value is designed applying pseudorandom multi-valued sequence. The value channel intensity in HSV space is changed with the sinusoid. The detailed phase unwrapping approach is explained, the proposed algorithm can be applied to measurement for 3D object with large height steps and/or spatially isolated surfaces. The simulation and true object experiments validate the correctness of the proposed method.

Keywords Phase unwrapping • Color fringe projection • Pseudorandom multi-valued sequence • Image segmentation • HSV color space

W. Li (✉) • Z. Yin
Faculty of Mechanical & Electrical Engineering, Kunming University of Science and Technology, No. 68, Wenchang Road, 121 Street, Kunming 650093, People's Republic of China
e-mail: lwg2709@sina.com

S. Duan
Department of Foreign Languages, Kunming University, No. 2, Puxin Road, Kunming Economic and Technological Development Zone, Kunming 650214, People's Republic of China

1 Introduction

One normally carries out phase unwrapping by comparing phases at neighboring pixels and adding or subtracting 2π to bring relative phase between two pixels into the range $-\pi$ to π, this cause problems when the technique is applied to real objects in engineering or medical measurement, because such objects often contain edges or discontinuities that cause phase jumps greater than π (Burton et al. 1995; Zhao et al. 1994). It can become impossible to unwrap correctly across the image, and the wrong phase will propagate along the unwrapping path.

Many methods have been presented such as temporal phase unwrapping (Huntly and Saldner 1993; Saldner and Huntly 1997), measuring phase maps at different sensitivities etc (Zhao et al. 1994). Huntley and Saldner (1993; Saldner and Huntly 1997; Huntley and Saldner 1997) proposed temporal phase unwrapping algorithm for measuring object with surface discontinuities. Zhang H et al. proposed spatiotemporal phase unwrapping for measurement of discontinuous objects in dynamic fringe-projection phase shifting profilometry (Zhang et al. 1999). In Sansoni et al. (1997, 1999), combination of phase shifting and gray-code projection has been used to perform phase unwrapping, but this technique has disadvantage that it is time-consuming, because at least nine fames projected need be to recorded to get the unwrapped phase value.

Color information can also be utilized to assist phase unwrapping, which has many applied examples. In Zhang and Huang (2006) and Zhang and Yau (2008), an high-resolution, real-time three-dimensional shape measurement system was presented. The detail example about application for color in the coded-structure light phase unwrapping can read paper (Salvi et al. 2004). Skydan et al presented a method with up to three fringe patterns projected on the three primary color channels from three different video projectors at different viewpoints to overcome shadowing effects on the object (Skydan et al. 2005). However, the approaches described by Huang et al. (1999) and Skydan et al. (2005) only produce wrapped phase maps, so spatial phase unwrapping methods are required to generate 3D shape of an object and hence these methods cannot be applied to objects with large slope changes or discontinuities.

The key to color information is used to implement phase unwrapping is that a special structured pattern need to be designed, and the color information is employed as a tool to calculate wrapped phase and absolute fringe order. Especially for object with discontinuous surface and isolate point/edge, color information is embedded in color-structured light pattern, and then is utilized to decide continuous phase distribution, which make difficult unwrapping process become very trivial process.

In this paper, we present a phase unwrapping method employing color in fringe analysis. The method is based on HSV color space using primary hue, intensity, saturation value to create up to three separable channels for the simultaneous projection of fringe patterns. Doing this allows us to significantly reduce some of the common problems encountered in fringe analysis.

2   Color Fringe Pattern Design

### 2.1   *Pseudorandom Multi-valued Sequence Design*

A pseudorandom multi-valued sequence (PRMVS) has multi-valued entries taken from an alphabet of q symbols, where q is a prime or a power of a prime. Such a $(q^n - 1)$ -term sequence is generated by an n-position shift register with a feedback path specified by a primitive polynomial $h(x) = x^n + h_{n-1}x^{n-1} + \ldots + h_1 x + h_0$ of order n with coefficients from binary Galois field GF(q) (Trif et al.; MacWilliams and Sloane 1976). Many primitive polynomials over GF(q) are given in Table 1 (Trif et al.) for GF(3), GF(4), GF(8), and GF(9). When q is prime, the integers modulo-q form the Galois field GF(q) = (0, 1, 2, ..., p-1) in which addition, subtraction, multiplication and division are carried out modulo-q. When q is a power of a prime, $q = p^m$, the integers modulo-q do not form a field and the Galois field elements are expressed as the first q-1 powers of some primitive element, labeled here for convenience by the letter A: GF(q) = (0, 1, A, A2, ..., Aq − 2). The primitive polynomials used for different PRMVS generation depend on the nature of addition/subtraction and multiplication/division tables adopted for each particular Galois field.

In this paper, a new algorithm based upon linear feedback shift register is proposed to generate (m, n) de Bruijn sequences for m = 4 and n = 3. The prime m denotes pixel hue value in HSV color space, n denotes order. Based on Table 1 (Trif et al.), we can know its primitive polynomial $h(x) = x^3 + x^2 + x + A$. Over GF(4) = (0, 1, A, A^2), and relation $A^2 + A + 1 = 0, A^2 = A + 1, A^3 = 1$ is applied. Pseudorandom multi-valued sequence can be obtained by the following few steps: Firstly, set initial value to three registers. Because order is equal to 3, the highest power is 5, therefore, a 3*6 matrix Q[i][j] (i < 3 and j < 6) is used to store register value. Since the initial value for all register cannot be zero at the same time, therefore, at least one register's initial value is set to no-zero. Secondly, accumulative sum is calculated after each register shift, accumulative sum is regarded as next output. Thirdly, the power of coefficient for each Galois field

Table 1 Primitive Polynomial

m	q=3	q=4	q=8	q=9
2	$x^2 + x + 2$	$x^2 + x + A$	$x^2 + Ax + A$	$x^2 + x + A$
3	$x^3 + 2x + 1$	$x^3 + x^2 + x + A$	$x^3 + x + A$	$x^3 + x + A$
4	$x^4 + x + 2$	$x^4 + x^2 + Ax + A^2$	$x^4 + x + A^3$	$x^4 + x + A^5$
5	$x^5 + 2x + 1$	$x^5 + x + A$	$x^5 + x^2 + x + A^3$	$x^5 + x^2 + A$
6	$x^6 + x + 2$	$x^6 + x^2 + x + A$	$x^6 + x + A$	$x^6 + x^2 + Ax + A$
7	$x^7 + x^6 + x^4 + 1$	$x^7 + x^2 + Ax + A^2$	$x^7 + x^2 + Ax + A^3$	$x^7 + x + A$
8	$x^8 + x^5 + 2$	$x^8 + x^3 + x + A$		
9	$x^9 + x^7 + x^5 + 1$	$x^9 + x^2 + x + A$		
10	$x^{10} + x^9 + x^7 + 2$	$x^{10} + x^3 + A(x^2 + x + 1)$		

Fig. 1 The designed hue
channel pattern employing
designed PRMVS with SHC
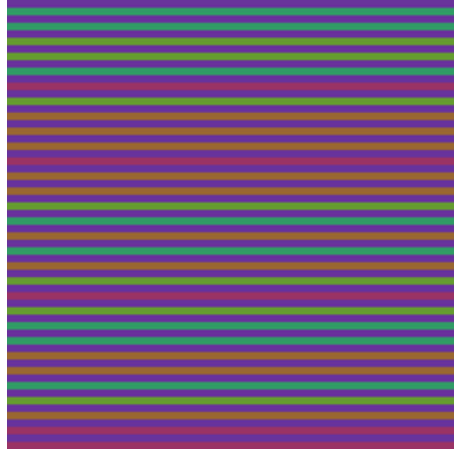
elements s[i] is played down, and the relation $A^2 + A + 1 = 0$, $A^2 = A + 1$, $A^3 = 1$
is applied. Fourthly, the value s[i] taken modulo 2 and the register output are
calculated.

2.2 Hue Channel Intensity Calculation and Hue Pattern Design

After PRMVS has been obtained applying process on Sect. 2.1, hue channel
intensity for color pattern can be calculated by PRMVS, but within the PRMVS,
there are many consecutive adjacent output sequences sharing same hue value,
which will increase difficult to distinguish output position for image process,
therefore, between two consecutive adjacent output hue code, a separate hue code
(here called by SHC) is insert and the hue pattern is shown in Fig. 1. Figure 1
shows that any adjacent output code values are different, which will take a great
convenience to later image process.

2.3 Value Channel Intensity Calculation for HSV Color Space

To increase the accuracy for pixel's intensity distribution, the intensity is designed
to change with sinusoidal. By incorporating phase-shifting detection technique into
phase measurements, a high accuracy is automatically achieved. Since position
information is embedded in a continuous intensity distribution, every pixel can
function independently in phase sampling. This results in an extremely high spatial
sampling density, which is critical for measuring local distortion effects.

2.4 Pixel Intensity Transformation from HSV Space to RGB Space

After finishing the design of hue channel intensity and value channel intensity, we can calculate color image pixel intensity in HSV space, and then change pixel intensity from HSV space to RGB space.

3 Detailed Phase Unwrapping Approach

3.1 Extraction of Hue Channel Stripes and Value Channel Fringes

The designed color pattern is calculated in HSV color space, which combines hue stripes and value fringes. After the patterns are projected on object surface, the recorded image pixel intensity is denoted in RGB color space, therefore, the recorded fringe image need to be converted from RGB space to HSV space, only after color space transformation, hue channel stripe and value channel fringe can be extracted.

3.2 Image Segmentation and Region Labeling

The extracted hue channel color stripes image denotes the PRMVS code. To decode all the PRMVS code, the hue image need to be segmented into several adjacent regions, and the region need to be labeled with different number, it is unique label corresponding to the hue channel value. All the pixels within one region share the same hue value, therefore, the image process algorithm need be utilized to implement the task.

3.2.1 Hue Channel Image Segmentation

Threshold algorithms are used to implement initial segmentation, threshold value is calculated by histogram of hue channel image. After segmentation, each pixel is distributed one hue segmentation (HS) value, and the pixels allot same HS value denote they may lie in same region, which is the candidate pixel to be within one region, but it is not always true because the pixels that have same hue value may lie in different regions and are not adjacent to each other.

3.2.2 Region Label Initialization

After hue image is segmented, within one segmentation, the pixels have same HS, and adjacent segmentations have different HS, there are some jump edges (HS edges) between adjacent segmentations, therefore, we employ the jump edge to implement region label process.

After region label initialization, there are many errors, within one region exist different HS value. therefore, the region existing error need to be divided into many small regions to assure that there is only one HS value within one region.

3.2.3 Region Splitting

Before region splitting, the region to be divided need to be marked, which will tell us which region need to be processed. The proposed process is that drawing HS value histogram for all pixels within one segmented region, if the values on histogram exist more than one wave crest, it denotes the region need to be split.

3.2.4 Region Growing

After region splitting, within each region, there is only one HS value, however, there are many pixels which have same HS value but there are not located on same region, therefore, it is necessary that the region need to be grown so that all the pixels that is adjacent to the region and have same HS value should be included into the region.

3.2.5 Region Mending

Employing above process, many pixels are still not distributed label, therefore, some process need to be done so that a region label can be slot to those pixels, here, region adjacent relation is used to calculate new region label, and region label must keep increase or decrease monotonously between adjacent region.

After above five step process, all the pixels are distributed one region label and meet increase or decrease monotonously along one region search direction.

3.3 Wrapped Phase Calculation

While image pixel intensity is transformed from RGB space to HSV space, we can utilize value channel intensity of pixel to calculate wrapped phase. By four-step method, Eq. (1) can be used to calculated wrapped phase, where $\varphi(i, j)$ is wrapped

phase at pixel (i, j), and V_1, V_2, V_3, V_4 are value channel intensities corresponding to four color fringe image after pixel intensity is changed from RGB space to HSV space.

$$\varphi(i, j) = \tan^{-1}\left[V_4 - V_2 / V_1 - V_3\right] \qquad (1)$$

3.4 Phase Absolute Fringe Order Determination

After color fringe image is segmented, each segmentation region is allot a label number by the process on Sect. 3.2, to obtain continuous phase distribution, absolute fringe order calculation is necessary, the process is as follows: **Firstly**, one or lots of straight-line vectors are decided manually and their directions are regarded as region search direction. **Secondly,** along region search direction, three continuous adjacent region label number v1, v2, v3 are recorded and combined as a group such as (v1, v2, v3), which can be regarded as a three dimension vector. **Thirdly,** the corresponding HS value vector (HS1, HS2, HS3) for (v1, v2, v3) is compared with the designed color pattern HS value, and the searched color HS value group (vs1, vs2, vs3) on designed color pattern is obtained. **Fourthly,** the searched color HS value group (vs1, vs2, vs3) on designed color pattern is utilized to calculate absolute fringe order $n(i.j)$, which is the corresponding sequence number of PRMVS code.

3.5 Phase Absolute Fringe Order Amendment

Employing above process, for most pixel (i, j), its absolute fringe order $n(i.j)$ are correct, but since the error caused by computation, image process, phase calculation etc, there are many pixels which are allot the wrong $n(i.j)$, therefore, it is necessary that an amendment method need to be implemented to correct wrong absolute fringe order $n(i.j)$, the process follows next three steps,

Step1: the wrapped phase jump pixels is marked, and these jump points are labeled as a new segmentation region edge, the process is that if the difference between the pixel's phase and its four-neighbour phase is larger than 180° or smaller than $-180°$, then the pixel is labeled as jump pixel.

Step2: the segmentation based on edge is implemented, a new segmentation region can be obtained, and the new region number is labeled as $n1(i.j)$

Step3: within one region $n1(i.j)$, if exists lots of region labels, then one of the region $m(i.j)$ corresponding to the biggest region label number need to be grown and then merging adjacent pixel until arrive the edge of region $n1(i.j)$, the grown and merging region m($i.j$) is regarded as final region label of region $n(i.j)$.

Step4: repeat step1, step2, step3, until there is only one region label within all region $n1(i.j)$, then the amendment process stop, and the corrected region label $n(i.j)$ is regarded as the final absolute fringe order.

3.6   Unwrapped Phase Calculation

After wrapped phase and absolute fringe order have been obtained, the unwrapped phase value can be calculated by the following equation:

$$\phi_b(i,j) = \phi(i,j) + 2\pi n(i,j) \tag{2}$$

Where $\phi(i,j)$ can be determined by Eq. (1), $\phi_b(i.j)$ is continuous phase distribution after phase unwrapping, and $n(i.j)$ is jump integer of phase which is the decoded absolute color fringe order calculated BY the process on Sects. 3.4 and 3.5.

3.7   Unwrapped Phase Complement

Because error caused by system setup, image process, color adjustment, there are some phase values calculated by Eq. (2) exist lots of error, and phase distributions show discontinuous, therefore, some complement process need to be done.

Phase on reference plane is theoretically continuous and phase difference between adjacent pixel is changed from $-\pi$ to $+\pi$, therefore, we can finish unwrapped phase complement process by comparing adjacent unwrapped phase value. We can firstly implement phase unwrapping process for reference plane, and then phase unwrapping process is implemented for pixels located on measured object utilizing condition that within one region having same region label and same phase jump value.

3.7.1   Unwrapped Phase Complement for Reference Plane Pixel

The phase on reference plane is theoretically continuous and phase difference between adjacent image pixel is changed from $-\pi$ to $+\pi$, we can finish phase unwrapping process by comparing adjacent unwrapped phase value.

3.7.2   Calculating Continuous Phase Distorted by Object

After the phase is unwrapped correctly for reference plane, it is an easy thing for calculation of unwrapped phase for pixel on distorted object since pixel lies on same

segmentation region share same region label, and they have same absolute fringe order. The process is as follows,

$$\phi_{object}(i, j) = \phi_b(i, j) + 2\pi n_{object}(i, j) \tag{3}$$

$$n_{object} = n_{reference} = k(i, j) \tag{4}$$

Where $\phi_{object}(i.j)$ is continuous phase distribution after phase unwrapping, $\phi_b(i, j)$ can be determined by Eq. (2). $n_{object}(i.j)$, $n_{reference}(i, j)$ are absolute fringe order on distorted object and on reference plane for same segmentation region, respectively, which is equal to region label $k(i, j)$ after unwrapped phase complementation.

4 Experiments

Few experiments are implemented to testify the correctness. **Firstly**, the simulation experiment is done, the simulated object: $F(x, y) = 20 - x*x/4 - y*y/9$, the color image distorted by simulated object while phase shift is $270°$ is showed on Fig. 2. Applying the proposed approach, Fig. 3 is the result obtained by process on Sects. 3.1, 3.2, 3.3, 3.4, 3.5 and 3.6. From Fig. 3, we can see that image pixels especially those located on the object are allot the wrong phase values. Figure 4 is the retrieved continuous phase distribution obtained by the process of Sect. 3.7. Figure 4 shows the proposed approach is correct, and can retrieve the object's continuous phase distribution. **Secondly**, a true object experiment is done, the measured object is a big container, The recorded fringe image is shown in Fig. 5, the final unwrapped phase is shown on Fig. 6. Figure 6 shows our proposed unwrapping algorithm can unwrap the object with large discontinuous surface.
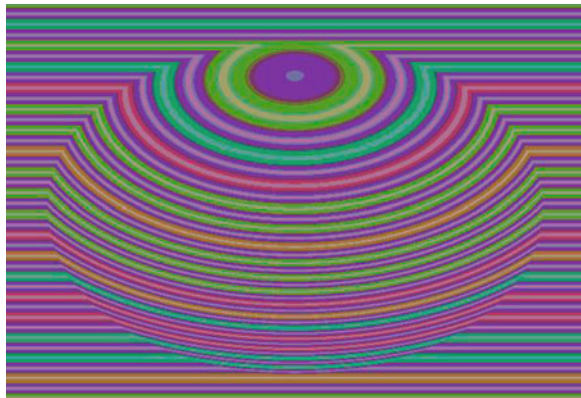

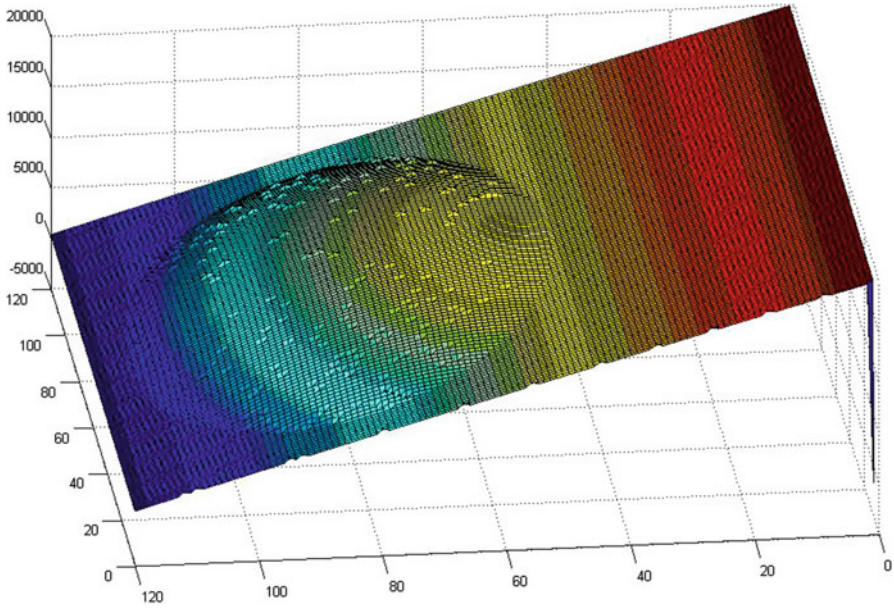**Fig. 2** The simulated color fringe image, phase shift $= 270°$

Fig. 3 Unwrapped phase without complement process on Sect. 3.7 for the simulated object
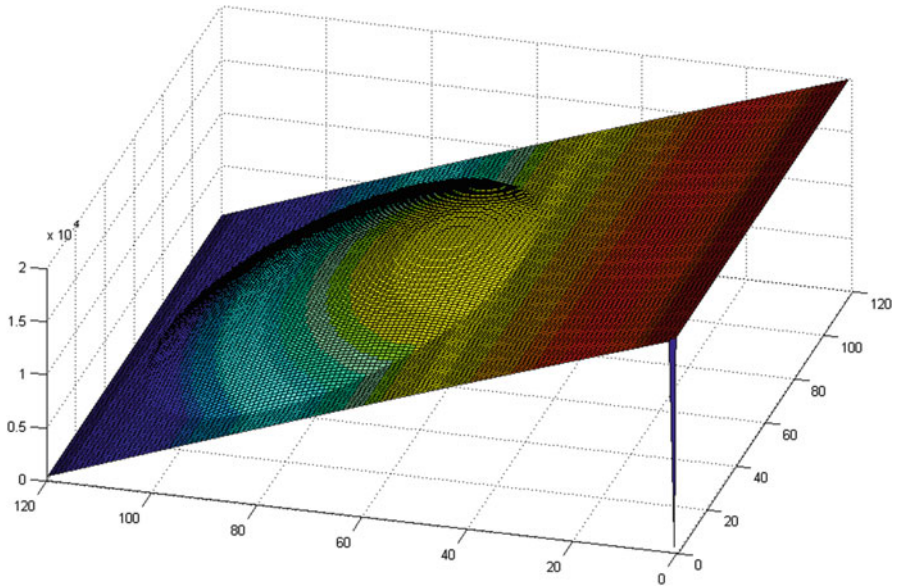

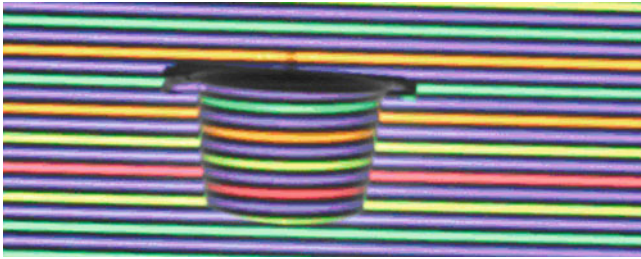**Fig. 4** Unwrapped phase after complement process on Sect. 3.7 for the simulated object

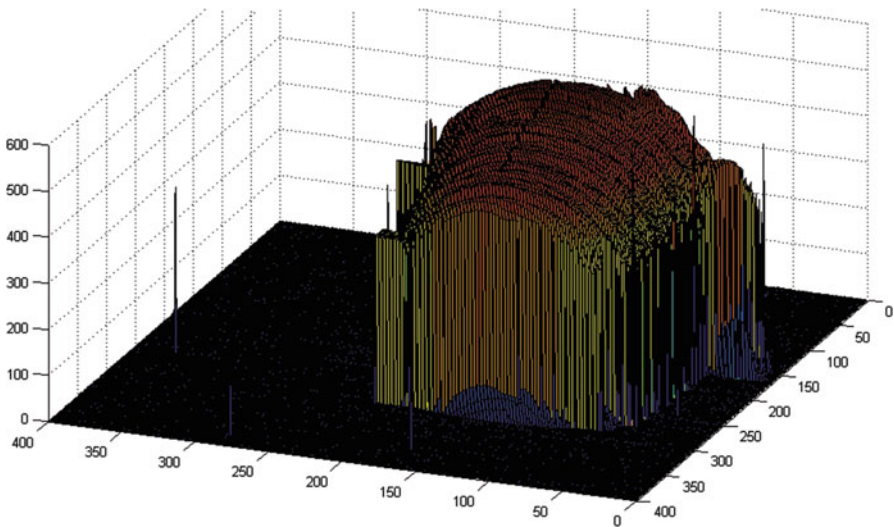**Fig. 5** The recorded color fringe image for phase shift $= 270°$

Fig. 6 The retrieved absolute phase distribution for measured object

5 Conclusion

In this paper, color fringe patterns are used to implement phase unwrapping, it is time saving. Its main principle is that adjacent region pixel's hue vale is utilized to calculate absolute fringe order of phase value, the proposed algorithm can overcome the difficult unwrapping problem for object with large height steps and/or spatially isolated surfaces. Color encoding is embedded in hue channel and sinusoidal intensity is in value channel of HSV color space, respectively. The relative phases of intensity distributed inside each fringe, moreover, can be solved by phase shift processing and then unwrapped process based on the correction of absolute phases on each stripe. Experiment results show that our algorithm can obtain accurate unwrapped phase value for object with large discontinuous surface. Compared with other phase unwrapping algorithm based on color fringe projection,

our algorithm without edge detect and without wavelet transform, therefore, the proposed algorithm can be implemented with higher efficiency. Simulation and true object experiments have validated the correctness of proposed phase unwrapping algorithm.

References

Burton DR, Goodall AJ, Atkinson JT, Lalor MJ (1995) The use of carrier frequency shifting for the elimination of phase discontinuities in Fourier transform profilometry. Opt Lasers Eng 23: 245–257

Huang PS, Hu QY, Jin F, Chiang FP (1999) Color-encoded digital fringe projection technique for high speed three-dimensional surface contouring. Opt Eng 38:1065–1071

Huntley JM, Saldner HO (1997) Error reduction methods for shape measurement by temporal phase unwrapping. J Opt Soc Am A 14(12):3188–3196

Huntly JM, Saldner HO (1993) Temporal phase-unwrapping algorithm for automated interferogram analysis. Appl Opt 32:3047–3052

MacWilliams FJ, Sloane NJA (1976) Pseudo- random sequences and arrays. Proc IEEE 64(12):1715–1729

Saldner HO, Huntly JM (1997) Temporal phase unwrapping: application to surface profiling of discontinuous objects. Appl Opt 36:2770–2775

Salvi J, Pages J, Batlle J (2004) Pattern codification strategies in structured light systems. Pattern Recognit 37:827–849

Sansoni G, Corini S, Lazzari S, Rodella R, Docchio F (1997) Three dimensional imaging based on Gray-code light projection: characterization of the measuring system for industrial applications. Appl Opt 36:4463–4472

Sansoni G, Carocci M, Rodella R (1999) Three dimensional vision based on a combination of gray-code and phase shift light projection: analysis and compensation of the systematic errors. Appl Opt 38:6565–6573

Skydan OA, Lalor MJ, Burton DR (2005) Using colored structured light in 3-D surface measurement. Opt Lasers Eng 43:801–814

Trif N, Petriu EM, McMath WS, Yeung SK (1993) 2-D pseudo- random encoding method for vision-based vehicle guidance. In: Proceedings of the intelligent vehicles '93 symposium, vol 5836, pp 297–302

Zhang S, Huang PS (2006) High-resolution, real-time three dimensional shape measurement. Opt Eng 45(12):123601

Zhang S, Yau S-T (2008) Three-dimensional shape measurement using a structured light system with dual cameras. Opt Eng 47(1):013604

Zhang H, Lalor MJ, Burton DR (1999) Spatiotemporal phase unwrapping for the measurement of discontinuous objects in dynamic fringe projection phase-shifting profilometry. Appl Opt 38(16):3534–3541

Zhao H, Chen W, Tan Y (1994) Phase-unwrapping algorithm for the measurement of three-dimensional object shapes. Appl Opt 33:4497–4500

An Adaptive Algorithm for Image Denoising Based on Wavelet Transform

Guo Peng, Yang Ping-xian, and Wang Wei

Abstract In view of the traditional wavelet de-noising edge is easily destroyed, which causes the useful detail information of image drop-out problem, this article proposed one kind of algorithm that based on the wavelet transform image auto-adapted de-noising. Firstly, this algorithm carries on the piecemeal match to the image, constructs each similar block the data set; Secondly, it carries on the wavelet transform to it, and takes the noise variance iteration as the foundation; Finally, it makes auto-adapted de-noising processing separately with the soft and hard threshold function to the high or low frequency coefficient. The experimental result shows that after the improvement method applies in the image de-noising, can retain more detail information well, enhance the image peak signal-to-noise ratio (PSNR) and the visual effect has improved.

Keywords Wavelet transform • Soft threshold function • Hard threshold function • PSNR • Detail information

1 Introduction

Noise is the main factors that affect the image quality. Easily affected by noise when using computer and other equipments carrying on image processing, cause the image quality drop, it also greatly affect people extract the useful information from image. In order to improve the image quality and prominent image expectation characteristics, de-noising processing is necessary.

In the image de-noising domain, people have been seeking various de-noising algorithms (Awate and Whitaker 2004). In wavelet domain and frequency domain,

G. Peng (✉) • Y. Ping-xian • W. Wei
Department of Automation and Electronic-Information, Sichuan University of Science & Engineering, Zigong, Sichuan 643000, People's Republic of China
e-mail: guopeng111111@126.com; ypingx@163.com; wangzhiruiwangwei@163.com

people utilize retain low frequency component method to reduce influence of noise. The basic idea of the method is based on the image information mainly distribute in low frequency part, the noise is in high frequency part. These methods in certain extent may eliminate parts of the noise, but will inevitably lose much high frequency information (image edge information) for the image. This for contain rich edge information image is very severe influence. In order to overcome this influence, a nonlinear wavelet transform method (Kaiqi Huang et al. 2004; Amolins et al. 2007; Vese and Osher 2004) was proposed by Donoho and Johnstone, selectively reduce or eliminate high resolutions part, namely wavelet contraction threshold algorithm. Along with the wavelet theory research unceasingly thorough, in image de-noising domain, the deepening of research for wavelet threshold method will obtain great development, such as soft and hard threshold methods, has obtained widespread application, and also get better effect.

In the hard threshold function de-noising processing, because it is not continual in wavelet domain, there will have discontinuous points. Simultaneously, complete remove less than the threshold of the wavelet coefficients, larger than the threshold of the wavelet coefficients, retention without any processing, which will produce bigger deviation when restructuring image, affect the de-noising effect; the soft threshold function derivative is discontinuous, so difficultly to compute its higher-order derivative. At the same time, there have constant deviation between estimated wavelet coefficient and wavelet transform coefficients, so when restructuring image, it will produce bigger standard deviation, affect the de-noising effect. In view of the shortcomings for above methods, an adaptive algorithm for image de-noising based on wavelet transform was presented.

2 The Principle for Wavelet Transform De-noising

Wavelet transform is multi-scale, multi-resolution decomposition for image. It can gather any details of the image, so it is called the mathematics microscope. Wavelet transform basic idea in image processing is to multi-level decompose image, decomposes image into different frequencies and different space sub-image. For wavelet image, wavelet transform produce data amount are equal to original image data amount, the produced wavelet image and original image has different characteristics, mainly in the image energy mainly concentrate in the low frequency part, but the energy is less in horizontal, vertical and diagonal part, and it characterization the edge information in horizontal, vertical and diagonal part for original image, it has obvious directional characteristic. The low frequency part can be called brightness image, and the horizontal, vertical and diagonal parts can be called detail image.

The process of decomposing an image (Gonzalez and Woods 2000) is shown in Fig. 1:

Moreover, those sub-signals still can be further decomposed into smaller sub-signals, and it is called n-level decomposition.
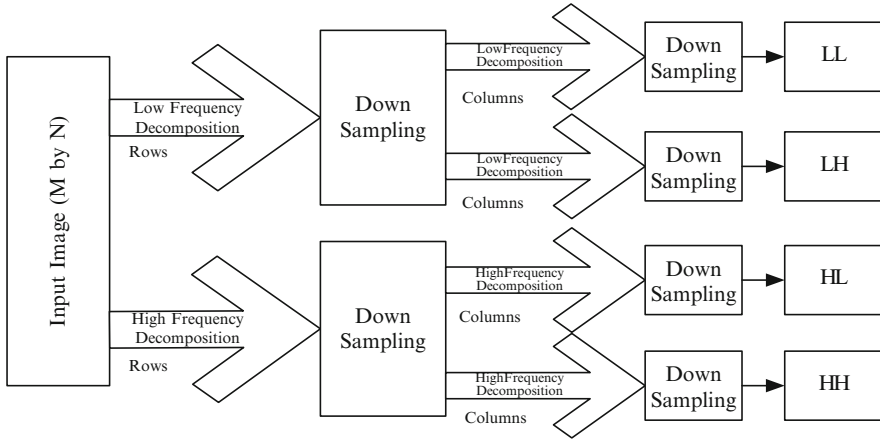
Fig. 1 Wavelet decomposition process

In practical application, image noise is generally the additive white Gaussian noise, establishes the noise image model as follows:

$$g(x, y) = f(x, y) + u(x, y) \tag{1}$$

$f(x, y)$ is the original image, $u(x, y)$ is the noise source that satisfies the Gaussian distribution, and $g(x, y)$ is the add noise image. Because the linear characteristics for wavelet transform:

$$W_g(a, b) = W_{f+u}(a, b) = W_f(a, b) + W_u(a, b) \tag{2}$$

The high-frequency part of $W_f(a, b)$ for the image which after wavelet decomposition, has the sparse distribution, the bigger wavelet coefficient correspond to the odd change position and the important information for the original image $f(x, y)$, But each sub-frequency band for $W_u(a, b)$ that gauss noise after wavelet decomposition, is still approximate Gaussian distribution. In each scale the distribution is uniform, And as the scale increase, the coefficient amplitude value reduce, Therefore for noise image wavelet decomposition high-frequency unit, may establish a threshold value, remove less than threshold of wavelet coefficients, and retain more than the threshold of wavelet coefficients, so, the wavelet coefficients which preserved are mainly produced by image decomposition, finally, after the wavelet inverse transform, we can get the de-noising image.

Wavelet transform, especially orthogonal wavelet transform remove the data relevance strongly. In wavelet domain, it can make the signal energy concentrated in some larger coefficients; but the noise energy actually distributes in the entire wavelet domain. Therefore, after wavelet decomposition, the amplitude value of signal wavelet coefficient must bigger than the amplitude value of the noise

coefficient, so use the threshold method can possible to retain the signal coefficient, and remove the majority of noise coefficient. Threshold formula has many, the unity threshold formula (Donoho and Johnstone 1994; Donoho 1995) for Johnstone and Donoho is used quite commonly, and the effect is well. The formula as follows:

$$T = \sigma \sqrt{2 \ln(N)} \tag{3}$$

σ is the noise standard deviation, and N is the signal length.

The threshold function is divided into two kinds, included soft threshold function and hard threshold function:

1. The soft threshold function is:

$$W'(j,k) = \begin{bmatrix} sign(W(j,k)) * (|W(j,k)| - T) & |W(j,k)| \geq T \\ 0 & |W(j,k)| < T \end{bmatrix} \tag{4}$$

The absolute value of wavelet coefficient compares with the threshold, the pixels that less than or equal to the threshold becomes zero; other pixels become the differential value for the pixel value and threshold, and keep the symbol.

2. The hard threshold function is:

$$W'(j,k) = \begin{bmatrix} W(j,k) & |W(j,k)| \geq T \\ 0 & |W(j,k)| < T \end{bmatrix} \tag{5}$$

The absolute value of wavelet coefficient compares with the threshold, the pixels that less than the threshold becomes zero, and the other pixels remains the same.

Although the soft and hard threshold method has obtained the widespread application, and also get better results, but these methods have the following flaws:

In the de-noising process by hard threshold function, because it is not continual in wavelet domain, there will have discontinuous points. Simultaneously, complete remove less than the threshold of the wavelet coefficients, larger than the threshold of the wavelet coefficients, retention without any processing, so when restructuring image, it will produce bigger standard deviation, affect the de-noising effect.

The soft threshold function which is continual in wavelet domain, it is not have the question of discontinuous points. But its derivative is discontinuously, Therefore when computing the higher derivative will encounter the difficulty.

3 Algorithm of This Paper

In this paper, has improved, and an auto-adapted algorithm for image de-noising based on wavelet transform was presented. This algorithm, firstly carry on block match for the add noise image, construct data set for each block, and the structure

data sets carry on wavelet transform; secondly, based on noise variance iteration, choose soft threshold function to do de-noising processing for image vertical component sub-band(LH) and horizontal component sub-band(HL), choose hard threshold function to do de-noising processing for image smooth approximation sub-band (LL) and diagonal component sub-band(HH); finally, each block's wavelet coefficient to carry on inverse transform, and reconstruct the image, get the de-noising image.

The algorithm steps as follows:

1. To do image block match and construct data set for each block;
2. For the obtained data block to carry on the wavelet transform, and each sub-block of data sets to carry on the two-dimensional wavelet decomposition, get the wavelet coefficients of each frequency band;
3. Adaptive to get the initial threshold;
4. Use hard threshold method to do de-noising processing for LL and HH , and use soft threshold method to do de-noising processing for LH and HL;
5. Wavelet coefficients of each block to carry on inverse transform, and reconstruct the image.

4 Experimental Simulation and Results' Analysis

Generally use objective measures to measure de-noising effect, such as MSE, SNR and PSNR (Barniv and Kella 1987; Bar-Shalom 1978). In this paper we choose PSNR. The PSNR value is bigger, the de-noising effect is better. The formula as follows:

$$PSNR = 10 * \log_{10} \left(\frac{L^2}{MSE} \right) \tag{6}$$

MSE defines as follows:

$$MSE = \frac{1}{m * n} \sum_{i=1}^{m} \sum_{j=1}^{n} (f(i, j) - g(i, j))^2 \tag{7}$$

In the formula, m, n are the total pixels rows and columns, $f(i, j)$, $g(i, j)$ respectively is the gray value of the original image (Fig. 2) and test image, and L is the maximum value of the image gray level.

Each algorithm de-noising effect as shows:

In the visual characteristic, from Figs. 3, 4, 5, 6, and 7 we might clear see that this algorithm can retain massive detail information during de-noising process, and this algorithm has the best visual effect. It shows that this algorithm has made better de-noising effect.
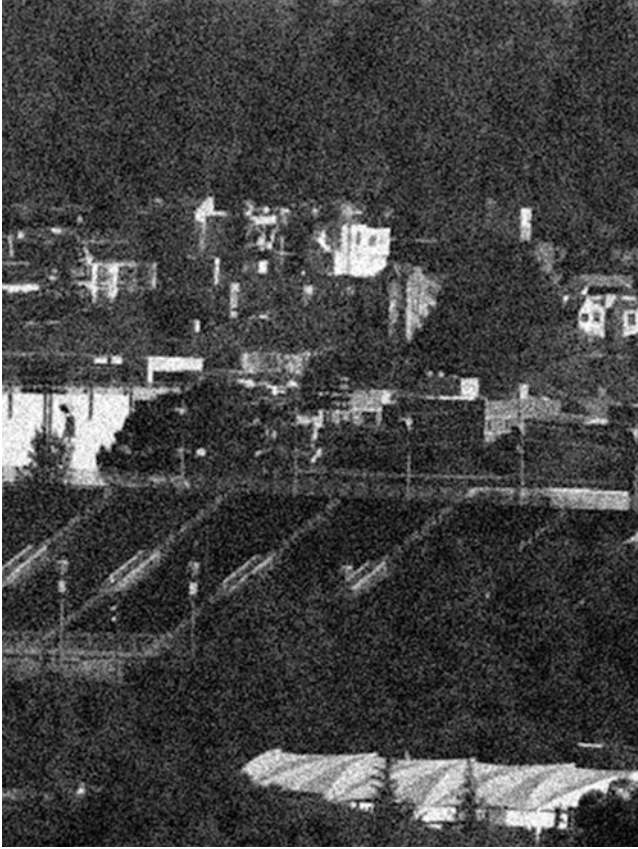
Fig. 2 Image add noise

Fig. 3 Soft threshold

Fig. 4 Median filter

Fig. 5 Hard threshold

Fig. 6 Average filter

Fig. 7 Algorithm for this paper

To the adds noise image, with the above several algorithms to carry on de-noising processing, and use formula (6) to calculate the image PSNR. Result as shown in Table 1:

Comparing the data in Table 1, we can clearly see that the PSNR for the image which was processed by this algorithm apparently higher than other four algorithms. It fully explained that this algorithm de-noising effect is obvious better than other four algorithms.

Table 1 The image's PSNR

The image	PSNR (dB)
Add noise	20.182
Soft threshold	25.862
Hard threshold	25.364
Median value filter	25.884
Average value filter	26.736
This paper algorithm	30.054

5   Conclusion

Image de-noising plays an important role in image processing, how to effectively remove the noise is the key to guarantee the quality of the reconstruction image. In this paper, an auto-adapted algorithm for image de-noising based on wavelet transform was presented according to the defects of soft and hard threshold function. The threshold is determined by the image noise in this algorithm, and choose soft threshold function to do de-noising process for image vertical component sub-band and horizontal component sub-band, choose hard threshold function to do de-noising process for image smooth approximation sub-band and diagonal component sub-band, then it effectively overcome the shortcomings which soft and hard threshold function separately de-noising. The experimental result shows that this algorithm can effectively retain the image's detail information while carry on de-noising process, and de-noising effect is better.

References

Amolins K, Zhang Yun, Dare P (2007) Wavelet-based image fusion techniques, an introduction, review and comparison. Photogram Metric Remote Sens 62(5):249–263

Awate SP, Whitaker RT (2004) Image de-noising with unsupervised, information-theoretic, adaptive filtering under review. IEEE Trans Pattern Anal Mach Intell 28(3):364–376

Barniv Y, Kella O (1987) Dynamic programming solution for detecting dim moving targets part II: analysis. IEEE Trans Aerosp Electron Syst 23(6):776–788

Bar-Shalom Y (1978) Tracking methods in a multi-target environment. IEEE Trans Autom Control 23(4):618–626

Donoho DL (1995) De-noising by soft-threshold. IEEE Trans Inf Theor 41:613–627

Donoho DL, Johnstone IM (1994) Threshold selection for wavelet shrinkage of noisy data. In: Proceedings of the 16th annual international conference of the IEEE, Piscataway, NJ, USA, vol 1, pp 24–51

Gonzalez RC, Woods RE (2000) Digital image processing. Prentice-Hall, Inc., Upper saddle River, p 07458

Kaiqi Huang, Qiao Wang, Zhenyang Wu (2004) Color image enhancement and evaluation algorithm based on human visual system. ICASSP 3:721–724

Vese LA, Osher SJ (2004) Image de-noising and decomposition with total variation minimization and oscillatory functions. J Math Imaging Vis 20:7–18

Video Watermarking in H.264/AVC Compressed Domain Using *CAVLC* Mapping

Da-wen Xu

Abstract In this paper, a fragile watermarking algorithm for H.264/AVC compressed video using *CAVLC* mapping is proposed. By extensive analysis with several video sequences, we observed that not all code words appear in the bit stream. Watermark bits can be embedded by mapping used *CAVLC* code words to unused code words. During the embedding process, the eligible code words are first identified, and then the mapping rules between these code words and the watermark bits are established. The watermark information can be extracted directly from the encoded stream without resorting to the original video, and merely requires decoding the *CAVLC* code from bit stream rather than decoding the whole video. Experimental results show that the proposed watermarking scheme can effectively embed information with little bit rate increase and almost no quality degradation.

Keywords Compressed domain watermarking • Context-adaptive variable length coding (CAVLC) • Code-space • Code words mapping • Integrity verification

1 Introduction

With the advancement of information and internet technologies, digital videos achieve a broad range of applications, such as broadcast, DVD, video conferencing, video-on-demand, video surveillance, high-definition TV, etc. However, Digital videos can be easily tampered, altered or forged by un-authorized users with video editing tools. Consequently, issues of authenticity and integrity verification of digital videos are becoming more and more attractive. Digital watermarking technology provides useful solution for such problems by embedding the watermark

D. Xu (✉)
School of Electronics and Information Engineering, Ningbo University of Technology, Ningbo, China
e-mail: dawenxu@126.com

information behind a cover, and the embedded watermark can be detected or extracted later from the cover video used for verification.

H.264/AVC is a new video coding standard jointly developed by the ITU-T VCEG and the ISO/IEC MPEG standards committees. It represents a number of advances in standard video coding technology, in terms of both coding efficiency enhancement and flexibility for effective use over a broad variety of network types and application domains (Wiegand et al. 2003). It is, therefore, highly desirable to develop watermarking algorithms that work entirely in compressed domain.

For H.264/AVC video authentication, Wang and Hsu (2010) proposed a fragile watermarking scheme for H.264 video authentication by embedding the watermark information in the last nonzero quantized coefficient in each discrete cosine transform block. In Profrock et al. (2005) some of the skipped macroblocks are used to embed the authentication data which consists of the encrypted hash value and a certificate with the public key. The advantage of the method is possibility to erase the watermark and to reconstruct the original H.264 video. In Kuo and Lo (2009), a fragile video watermarking scheme to authenticate the H.264 video content is presented. Watermark information is embedded in motion vector by exploiting its fragile nature. These algorithms describing compressed-domain watermarking process do not, in fact, operate on the compressed bitstream. These methods are often called partial decompression methods.

A new class of applications requires the watermarking to be applied directly to an entropy encoded bitstream without the partial decompression and re-encode steps. Mobasseri and Berger (2005), Mobasseri et al. (2010) exploits the fact that only a fraction of JPEG code space is actually used by available encoders. Data embedding is performed by mapping a used variable length code (VLC) to an unused VLC. More recently, Zou and Bloom (2008) described a method for applying a watermark directly to a CAVLC entropy coded H.264/AVC stream. Watermark embedding is performed by changing Intra-prediction mode in the mb_type field. In Zou and Bloom (2010), Zou described another stream replacement method for video watermarking on an H.264 CABAC encoded bitstream. However, prior to watermarking, these methods need to pre-analysis the content which is computational complex. Two types of metadata (embedding metadata and detection metadata) should be created in the pre-analysis, and the detection metadata should be sent via some external channel to the detector. Mobasseri and NaikRaikar (2007) proposed a fragile watermarking scheme by mapping a used CAVLC code word to the unused but valid portion of the code space. One of the main problems in the scheme is that the payload is too small.

This paper introduces a fragile watermarking scheme based on CAVLC code words mapping. Watermark embedding is performed by mapping used *CAVLC* code words to unused code words, which is based on the mapping rules between these codes and the watermark bits. Because our scheme performed directly in entropy encoding, watermark extraction can be achieved in real-time. The rest of the paper is organized as follows. In Sect. 2, we present an overview of *CAVLC*

in H.264/AVC encoding system. In Sect. 3, we describe the proposed watermark embedding scheme. Experimental results and analysis are given in Sect. 4. Finally in Sect. 5, conclusions are drawn and future studies are explored.

2   CAVLC in H.264/AVC Encoder

In H.264/AVC, *CAVLC* employs five syntax elements shown in Table 1.

   CAVLC code word can be expressed as the following format:

$$\{Coeff\_token, Sign\_of\_TrailingOnes, Level, Total\_zeros, Run\_before\}$$

The global code space G can be defined as all valid combinations of five syntax elements.

$$G = \{c_t | t = 1, 2, \cdots, U\} \tag{1}$$

where c_t denotes valid code word, U is the total number of valid code words. Suppose X is a video collection:

$$X = \{V_j | j = 1, 2, \cdots, M\} \tag{2}$$

where M is the total number of video samples, V_j is a video sample. *CAVLC* code word for each video can be expressed as:

$$C_j = \{cavlc_k | k = 1, 2, \cdots, \Psi_j\} \tag{3}$$

where Ψ_j is the number of used *CAVLC* code words in video V_j. The utilization rate P_j of *CAVLC* code words can be calculated as:

$$P_j = \frac{\Psi_j}{U} \tag{4}$$

Table 1 CAVLC syntax elements

| Syntax elements	Description
Coeff_Token	Both the total number of non-zero coefficients (*TotalCoeffs*) and the number of trailing ±1 values (*TrailingOnes*)
Sign_of_TrailingOnes	A sign bit for each *TrailingOnes* in reverse zig-zag order
Level	The values of non-zero coefficients except for *TrailingOnes*
Total_zeros	The total number of zeros before the last coefficient
Run_before	The number of zeros proceeding each non-zero coefficient in reverse order
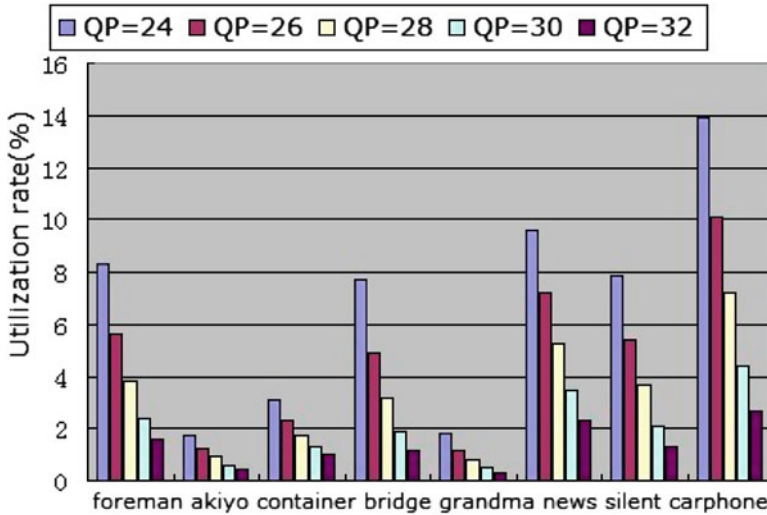
Fig. 1 Utilization rate of code words

Fig. 2 CAVLC code space
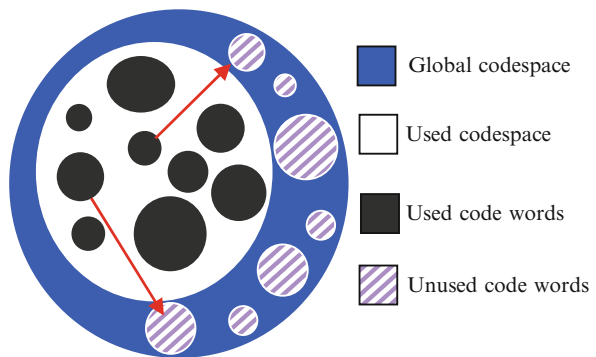
The utilization rate for different video samples is shown in Fig. 1. Apparently code space is not used fully for each video sample. Thus, watermark embedding can be implemented by mapping used *CAVLC* code words to unused code words (see Fig. 2).

3 Proposed Watermarking Algorithm

In order to describe our embedding algorithm more clearly, we have the following notations.

1. *global codespace*: The combination of $\{TrailingOnes, Sign\_of\_Trailing Ones, Level\}$ is considered. It is denoted by gcs.

2. *used CAVLC*: Denoted by u sec. All used CAVLC code words in video stream.
3. *umused CAVLC*: Denoted by *unuc*. Obviously, $gcs = u$ sec $\cup unuc$.
4. *valid CAVLC*: With one occupation and at least one *Level* with absolute value of three or less. It is denoted by *valc*.
5. *watermarkable CAVLC*: Denoted by *wmc*.

Watermark embedding is implemented by mapping *valc* to u sec. The absolute value of a level is decreased by one during mapping. When first *Level* is changed, *TrailingOnes* should be adjusted to match. Because multiple CAVLC codes may be mapped to the same unused codes, it is not possible to recover the watermark accurately. For example, $\{5, 2, 1, -1, -2, 4, 3\}$, $\{5, 1, 1, -2, -1, 4, 3\}$ and $\{5, 0, 2, -1, -1, 4, 3\}$ can be mapped to $\{5, 3, 1, -1, -1, 4, 3\}$ according to the mapping rule. In $\{5, 3, 1, -1, -1, 4, 3\}$, 5 is the total number of non-zero coefficients, 3 is the number of *TrailingOnes*. It is obvious that not all *valc* can be used for watermarking. Thus we should first detect *watermarkable CAVLC* whose mapped codes only occur once in the steam.

Watermark embedding is performed directly in *CAVLC* entropy coding through the following steps:

Step 1. The eligible code words *wmc* is identified by parsing the H.264/AVC bit stream.
Step 2. If watermark bit is 1, watermark bit can be embedded by mapping *wmc* to *wmuc*. Otherwise, *wmc* of current block is left unchanged, where *wmuc* is a valid *unused CAVLC*. In a word, watermark bit w_i is embedded by modulating the *CAVLC* codeword as

$$wm\_cw = \begin{cases} wmuc & if \ w_i = 1 \\ wmc & if \ w_i = 0 \end{cases} \qquad (5)$$

where *wm_cw* is the watermarked *CAVLC* codeword.
Step 3. *wmc* is delivered as the secret key.

Watermark extraction procedure is described as follows:

Step 1. The eligible codeword *wmc* is obtained according to the secret key.
Step 2. The watermarked *CAVLC* codeword wm_cw' is identified by parsing the H.264/AVC bit stream.
Step 3. Watermark w_i' can be extracted according to the mapping rule. If wm_cw' is the same with *wmc*, the extracted bit is 1; otherwise, the extracted bit is 0.

$$w_i' = \begin{cases} 1 & wm\_cw' \neq wmc \\ 0 & wm\_cw' = wmc \end{cases} \qquad (6)$$

Obviously, a unique codeword *wm_cw* is assigned to each *wmc*, that is, the original bit stream can be recovered with lossless.

Fig. 3 Original frames

4 Experimental Results and Analysis

The proposed watermark embedding scheme has been implemented in the
H.264/AVC JM-8.6 reference software. Four standard video sequences (i.e.,
Container, *News*, *Foreman and Carphone*) in QCIF format (176×144) at the
frame rate 30 frames/s are used for our simulation. The first 300 frames in each
video sequence are used in the experiments. The GOP (Group of Pictures) structure
is "IPPPP (QP: I 28, P 28)".

To evaluate the imperceptibility of the proposed scheme, a series of tests have
been performed. An original frame from each video is shown in Fig. 3. The
corresponding watermarked frames are shown in Fig. 4. In the experiments, no
visible artifacts can be observed in all of the test video sequences. Besides subjective
observation, PSNR (Peak Signal to Noise Ratio) is usually taken to evaluate the
perceptual quality. However, for different video contents, *PSNR* cannot be reliable

Fig. 4 Watermarked frames

methods for assessing the video quality, because they are not very matched to perceived visual quality as they are based on pixel to pixel difference calculation and ignore human perception and viewing condition.

Another objective metric is adopted to evaluate the perceptual quality: *SSIM* (structural similarity index) (Wang et al. 2004). The SSIM index lies in the range between 0 and 1, where 0 indicates zero correlation, i.e. the reference image is entirely different than the target, and 1 indicates that they are identical. From Table 2, we can see that SSIM values are above 0.95. It is generally hard to detect the degradation in video quality caused by watermark embedding.

In order to further evaluate the performance of the proposed scheme effectively, bit rate variation (R_{var}) is introduced.

$$R_{\text{var}} = \frac{R' - R}{R} \times 100 \, (\%) \tag{7}$$

Table 2 Experimental results for video quality and bit rate

Sequence	QP	SSIM	R_{var}	Sequence	QP	SSIM	R_{var}
Foreman	24	0.9955	−0.17	*CarPhone*	24	0.9927	−0.12
	26	0.9929	−0.09		26	0.9931	−0.10
	28	0.9869	−0.08		28	0.9910	−0.18
	30	0.9897	−0.08		30	0.9925	−0.09
	32	0.9722	−0.06		32	0.9835	−0.09
Container	24	0.9891	−0.17	*News*	24	0.9913	−0.18
	26	0.9852	−0.25		26	0.9898	−0.20
	28	0.9866	−0.68		28	0.9886	−0.22
	30	0.9831	−0.23		30	0.9865	−0.19
	32	0.9780	−0.20		32	0.9853	−0.16

Table 3 Comparison results between the proposed method and that of Mobasseri et al. (2007)

Sequence	QP	R_{var}		*wmc*	
		Proposed method	Method in Mobasseri et al. (2007)	Proposed method	Method in Mobasseri et al. (2007)
Foreman	24	−0.17	−0.18	1, 967	1, 260
	26	−0.09	−0.10	1, 409	905
	28	−0.08	−0.09	956	571
	30	−0.08	−0.08	570	337
	32	−0.06	−0.08	395	236
Container	24	−0.17	−0.18	1, 076	719
	26	−0.25	−0.25	816	522
	28	−0.68	−0.68	602	378
	30	−0.23	−0.23	467	279
	32	−0.20	−0.25	347	208
CarPhone	24	−0.12	−0.13	3, 163	2, 003
	26	−0.10	−0.11	2, 251	1, 400
	28	−0.18	−0.19	1, 722	1, 015
	30	−0.09	−0.10	980	579
	32	−0.09	−0.09	627	369
News	24	−0.18	−0.20	2, 671	1, 774
	26	−0.20	−0.23	1, 957	1, 264
	28	−0.22	−0.25	1, 419	898
	30	−0.19	−0.23	891	589
	32	−0.16	−0.18	594	361

where R' is the bit rate generated by the modified encoder, and R is the bit rate generated by the original encoder. The test results for three video sequences are shown in Table 2. We can see that the bit rate of our method decreased slightly.

Table 3 presents the comparison results between our proposed method and the method of Mobasseri et al. (2007). As shown in Table 3, bit rate variation (R_{var}) is very small for both methods. Apparently, the capacity of our proposed method is higher than that of Mobasseri et al. (2007).

5   Conclusions

In this paper, a low complexity watermark embedding algorithm for H.264/AVC is proposed. The eligible code words are first identified, and then the mapping rules between these code words and the watermark bits are established. Watermark bits can be embedded by CAVLC code words mapping based on the established mapping rules. Watermark extraction eliminates the need for full or even partial decompression thus making it ideal for real-time video applications. Experimental results show that the proposed watermarking scheme provides high embedding rates while preserving visual quality. However, the proposed algorithm is fragile, i.e., if the embedded video is decoded and re-encoded, the embedded data will be lost. The fragile watermarking algorithm may be appropriate for many applications such as covert communication, content authentication, and tamper proofing, etc.

References

Kuo TY, Lo YC (2009) Fragile video watermarking technique by motion field embedding with rate-distortion minimization. J Commun Comput 6(1):16–23

Mobasseri BG, Berger RJ (2005) A foundation for watermarking in compressed domain. IEEE Signal Process Lett 12(5):399–402

Mobasseri BG, NaikRaikar YJ (2007) Authentication of H.264 streams by direct watermarking of CAVLC blocks. In: Proceedings of the SPIE security, steganography, and watermarking of multimedia contents VIII, vol 6505, San Jose, CA, pp 1–5

Mobasseri BG, Berger RJ, Marcinak MP, NaikRaikar YJ (2010) Data embedding in JPEG bitstream by code mapping. IEEE Trans Image Process 19(4):958–966

Profrock D, Richter H, Schlauweg M, Muller E (2005) H.264/AVC video authentication using skipped macroblocks for erasable watermark. In: Proceedings of the conference Visual Communications and Image Processing (VCIP 2005), vol 5960, Beijing, China, pp 1480–1489

Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612

Wang CC, Hsu YC (2010) Fragile watermarking scheme for H.264 video authentication. Opt Eng 49(2) 027003:1–9

Wiegand T, Sullivan GJ, Bjntegaard G, Luthra A (2003) Overview of the H.264/AVC video coding standard. IEEE Trans Circuits Syst Video Technol 13(7):560–576

Zou DK, Bloom JA (2008) H.264/AVC stream replacement technique for video watermarking. In: 2008 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008), Las Vegas, USA, pp 1749–1752

Zou DK, Bloom JA (2010) H.264 stream replacement watermarking with CABAC encoding. In: 2010 IEEE International Conference on Multimedia and Expo (ICME 2010), Singapore, pp 117–121

A Novel Replica Placement Strategy for Data Center Network

Xin Huang and Yuxing Peng

Abstract With the rapid development of the cloud and data-intensive computing, many data center network are growing more large scale, and the number of servers is increasing at an exponential rate. As commodity-class PCs is used in the current DCN, failures caused by node failure, rack failures, link failures and routing failures become a very common phenomenon in the parallel processing of data blocks, tasks, and job scheduling. DCell is a novel network structure adapted to the DCN, which uses a recursively-defined structure to interconnect servers. Although DCell is fault tolerant, and addresses various failures, it has many questions. If the destination nodes are the failure, no matter what cannot route to them. The existing technology of DCell fail to guarantee durability and reliable for the data on the node. This paper proposes a novel idea that a replica is placed in the node of the layer by the fixed-point single-source shortest path, and there is at least one replica in each layer of the DCell. According to the proposed placement strategy of replicas, we suppose a mechanism to restore lost replica based on the neighboring nodes in the same path. Finally, we simulate and validate algorithm of placement and repair for replica. Experiments show that our algorithm has obvious effect to improve the reliability of data.

Keywords DCN • DCell • Replica • Reliability

X. Huang (✉)
Computer School of Wuhan University, Wuhan, China

Communication School of Wuhan Textile University, Wuhan, China
e-mail: bossmehx@126.com

Y. Peng
School of Computer Science, National University of Defense Technology, Changsha, China

1 Introduction

In recent years, with the rapid development of the cloud and data-intensive computing, many data center network (DCN-data center networking) are growing more large scale, and the number of servers is increasing at an exponential rate. It is such high cost of a huge number of servers that DCN have to be built with the general commodity-class PCs than the high-performance and high-end server, and which is not the pursuit of peak performance of a single server but the trade-off price between costs and performance (Barroso et al. 2003). As is the resulting problems: (1) Hardware failure is the norm rather than the exception. Node failure is inevitable in DCN. There will be a large number of concurrent node failures. The single server that is used commodity-class PCs, considered from reliability, has the 3 years (3 * 365 days) of MTBF. If each subsystem is MTBF1, MTBF2, . . . , MTBFn, the system's MTBF = 1/(1/MTBF1 + 1/MTBF2 + . . . + 1/MTBFN), the MTBF of system that is made of the 1,000 servers MTBF is 1.095 days. The MTBF of 10,000 servers is the only 2.628 h. (2) Caused the loss of data due to a server failure. (3) Performance problems for node failure, link failure and route rack failure.

To solve aboved questions in DCN, many novel technique and strategy are proposed: (1) DCN network topology with the novel, more robust and fault-tolerant routing protocol. For example, Fat-Tree (Al-Fares et al. 2008), BCube (Guo et al. 2009), FiConn (Li et al. 2009) and DCell (Markus Kliegl 2009; Chuanxiong Guo et al. 2008; Kliegl et al. 2010). (2) The distributed file systems and programming models adapting to cloud computing and data-intensive computing. For example, GFS (Sanjay Ghemawat et al. 2003), MapReduce (Dean and Ghemawat 2005), BigTable (Chang 2006), hadoop (Pallis et al. 2005) and so on. (3) The durability and high availability of data object in the storage system. In a distributed system, replication is the key technology to provide high availability, improve performance and enhance the reliability, and is widely used in storage systems to reduce data loss and increase data availability (Pascal Felber and Andre Schiper 2001; ByungGon et al. 2006). However, many data in cloud computing and data-intensive computing are cut into the middle data or a lot of chunks that be distributed storage in a huge amount of the serves in the DCN, which can parallel processing them with the all-to-all communication pattern. As commodity-class PCs is used in the current DCN, failures caused by node failure, rack failures, link failures and routing failures become a very common phenomenon in the parallel processing of data blocks, tasks, and job scheduling.

DCell (Markus Kliegl et al. 2009; Chuanxiong Guo et al. 2008; Kliegl et al. 2010) is a novel network structure adapted to the DCN, which uses a recursively-defined structure to interconnect servers. Although DCell is fault tolerant, and addresses various failures at link, server and server-rack levels, it has many questions. If the destination nodes are the failure, no matter what cannot route to them. The existing technology of DCell fail to guarantee durability and reliable for the data on the node.

This paper is mainly studied the replica placement strategy for the DCell. We propose a novel idea that a replica is placed in the node of the layer by the FPSSSP (fixed-point single-source shortest path), and there is at least one replica in each layer of the DCell. We suppose a strategy of the route adjacent node to repair a replica which is lost for serve failure. When a copy is lost, a new replica is created by the re-selected node using the recovery mechanisms. Under normal circumstances, the $k + 1$ number of replica can be created and placed if DCell have k layers. Of course, if there are t_k nodes in DCell, the maximum number of replica is t_k in theory. The numbers of replica are too small to ensure durability and availability of data. But, the larger number of replica is, the higher overhead is. The correct number of extra replicas can be determined by a compromise between the number and availability.

2 Placement Strategy of Replica DCell

As the size of data centers network continue to expand, the failures of severs and switch became the norm. Fault-tolerant in the DCN should be required to have redundancy in the physical topology and a more robust mechanism in the protocol. But the only physical redundancy and fault-tolerant protocol does not guarantee durability and available of the data, for the failure of the nodes result in loss of data and unavailable.

This paper presents a novel placement strategy of the replicas by the fixed point single-source shortest path in DCell. When data is written into the node s, to prevent data loss from the failure, it must be replicate multiple replicas, and placed them in different nodes. First, the node s calculates its own layer, assuming that it belongs to the level-k where have t_{k-1} numbers of DCell$_{k-1}$s. The node s can select a nearest node f_1 from itself on the same sub-layer DCell$_{k-1}$s, which is used to place the first replica r_1. Then, node f_1 replicate it to form the second replica r_2, and place r_2 into the node f_2 which is a nearest node from different sub-level DCell$_{k-1}$s in the same level-k. The node f_2 single out a nearest node in the sub-layer DCell$_{k-1}$s (or DCell$_{k+1}$s) along the level $(k - 1)$ (or layer $k + 1$) path to place the third replica r_3. There is only a replica in each level of DCell$_k$ by so recursively selecting a node as placing the replicas. Obviously, all replicas except r_1 are placed in the node on the shortest path which is formed from top to bottom(or from bottom to top) by using f_2 as a starting point, and which passes through each level of DCell to place a replica.

Suppose there are k levels throughout DCell, we can at most put up to $k + 1$ replicas in accordance with the above approach. When the degree m of the replica is greater than $k + 1$, the remaining $m-(k + 1)$ replicas should be how to place it? There are two strategies to place the remaining replicas of:

- The node f1 place the replica $k + 2$ in the node f_2 which come from the second nearest DCell$_{k-1}$s in the same level, then f_2 ' as a starting point from the shortest
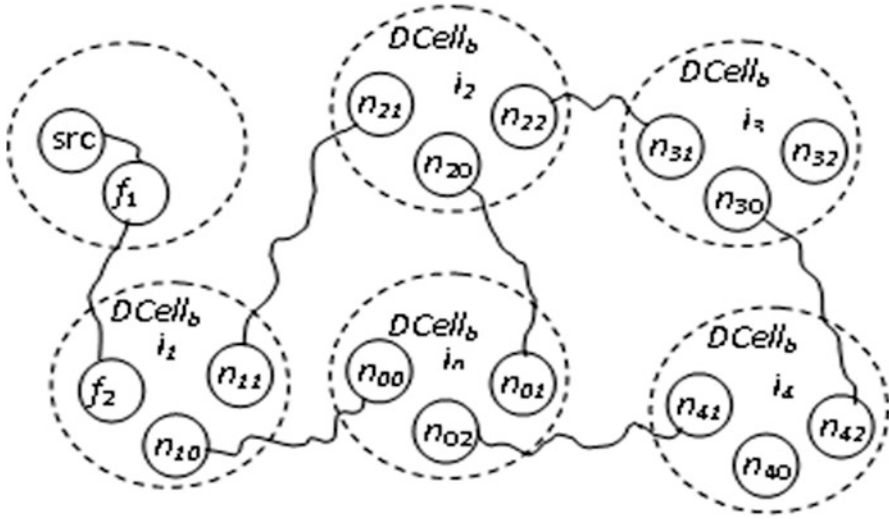
Fig. 1 Optimal replica placement of the DCell

path to the node in the DCell$_0$, which recursively place a replica in the node passing through each level until all the remaining $m-(k+1)$ replicas.

- Randomly selected $m-(k+1)$ node to place the remaining replicas.

The first method is easy to query and recovery the replicas, for they is on the same shorter path. However, it may load to low load-balance of nodes. The second method is simple and can randomly place the remaining replicas in the node, but it is more difficult to lookup and recovery them. Taking into account $k+1$ replica that has been placed in the node of each level along the shortest path, and the remaining $m-(k+1)$ replicas can be placed in accordance with the second strategy to obtain high load-balance.

Figure 1 shows that replicas can be placed in the DCell by the fixed-point single-source shortest path strategy. When the data are stored on the Ssrc, it must create replica for reliability and durability. First, node Ssrc finds a nearest node f_1 to place the first replica r_1 in the same sub-level DCell$s$. Then f_1 is fixed source point to find the shortest path to the other node in its sub-units of the adjacent level with the greedy algorithm, and select one of the smaller load node as a point placed replica. In Fig. 1, DCell$_b$ is a sub-DCell$s$ in same level, assuming that f_2 is selected to place to the second replica r_2. Nodes f_2 create and place the third replica r_3 in the same way. It is recursive until all the layers have a replica so far.

```
ReplicaPlace()
/* s is datanode denoted using the (k+1)-tuples
n is the number of nodes in a DCell_0
t_k is the number of servers in a DCell_k
m is the number of replica
```

```
R is node set of placement replica
*/
k←len(s);
R ←Φ;
do
  f₁←Single-Rources(DCell_{k-1}b,s);
  R←{f₁}∪R;
  s←f₁;
k←k−1, m←m−1;
While (k≠0)
/*Select m=m−(k+1) number node to place other replica */
If (m≠0)
  f₁←RandSelect({t_k});
  R←R∪{f₁};
  m←m−1;

return
```

The *ReplicaPlace*() shows the algorithm of the replica placement. Single-Rources can be directly achieved with Dijstra or Bellman-Ford algorithm. Rand-Select use to randomly select a node stored replica from tk numbers of node.

If there is no any limit, how many replicas can be placed at most in the DCell? According to the given boundary of the number of servers DCell in the Markus Kliegl et al. (2009), we can easily come to the upper bound of the degree of replica.

Theorem 1 *m is denoted the upper bound of the degree of replica in Dcell.*

$$m = \left\lfloor c^{2^k} \right\rfloor \tag{1}$$

Where

$$c = \left(n + \frac{1}{2}\right) \prod_{i=0}^{\infty} \left(1 + \frac{1}{4(t_i + \frac{1}{2})}\right)^{1/2^{i+1}}$$

The theorem 1 has proved by the Markus Kliegl et al. (2009), Chuanxiong Guo et al. (2008). For example, when $n = 2$, $k = 2$, we can draw the maximum degree of the replica $m = 42$. If the degree of replica is too small, the data is easily lost, and is not reliable. But, the degree of replica is too large that its cost is also high. In practice, the size degree of the replica is affected by many factors, for example, the size of the file, storage space and the number of available nodes, placed strategy, etc. Ming Zhong et al. (2008) achieves a calculation formula of degree of the replica based on the popularity of the replica, the file size, and failure rate of nodes. $k_i = C + \log 1/p(r_i/s_i)$, where C is a constant. r_i is the popularity of the object $_i$. the

file size is denoted by s_i, and p is the failure rate of node. Therefore, the size of the degree of replica must be a compromise between reliability and costs to obtain the optimal performance.

3 Repair Strategy of Failure Replica

Failures are quite common in current data center network (Barroso et al. 2003; Sanjay Ghemawat et al. 2003). Although the replication is able to ensure that the data is available in the case of the failure, to keep the system reliable and durable available, the failure issue of replicas is need to consider caused by any failure of the node, rack and link. When replica is lost, how many replicas are available? If a file has m replicas, of which n is available in considering the circumstances of the failure, $m-n$ replicas is not available. The availability of each replica is α, then the total probability of availability:

$$P_r \; [n < m] = \sum_{i=0}^{m-1} \binom{n}{i} a^i (1-a)^{n-i} \tag{2}$$

When only one replica exists in the system, replica must be restored immediately. Otherwise, the data will be lost permanently. According to the previous placement strategy of replicas, we propose a mechanism to restore lost replica based on the neighboring nodes in the same path. When the replica is lost, the neighboring node selects the node in the same path and creates a new copy using the greedy algorithm. If a replica of $DCell_b$ in the same level is lost due to various failures, the successor node (or predecessor node) on the path is responsible to re-create a new replica. When successor (or predecessor) node detects that its predecessor (or successor) fails, it firstly find a node with a greedy algorithm which is the shortest distance with itself in same DCellb. Then, it transmits their data to the selected node to create a new replica. Finally, the information of new replica in selected node is broadcast to the other node in same $DCell_b$. If the successor node (or predecessor node) on the path does not find a node with the greedy algorithm, it is clear that all the nodes in same $DCell_b$ fails, and select a node to create a new replica in another $DCell_b$.

The recovery mechanisms of replica can be illustrated in Fig. 1. The node f_2 is considered down if no DCellBroadcast messages is received by n_{21} (successor node) and f_1 (predecessor node) before timeout (5 s by default). If the node f_2 fails, the loss of its data is inevitable. Then, n_{21} (or f_1) fixed source point began to select the node n_{11} based on the shortest path arithmetic in the sub-level i_1 of the same $DCell_b$, and create a new replica for lost replica of data in the node n_{11}. All node of the sub-level i_1 are considered down if n_{21} can't find a node. Then, n_{21} will look for the right node in its adjacent sub-level i_0 and i_2.

4   Experiment and Simulation

In this section, we run simulations of a DCN based on DCell to see if replica placement indeed provides better reliability under server node, rack and link failure. The simulation is done in an event-driven model. There are two kinds of events pushing the virtual time forward. One is node failure events, which are triggered by exponentially distributed and independent failures for nodes. The other is a repair event for the replica. In all the simulations, the DCell network structure is constructed using servers $t_k = 176,820$ with $n = 4$, $k = 3$, and DCell1 is a rack by the connection rule in Markus Kliegl et al. (2009), Chuanxiong Guo et al. (2008), Kliegl et al. (2010). Each node manages averagely 1,024 files whose average size is 10 KB. Data reliability is variable from 0.995 to 0.99995 for different replica placement with node failure. The results are shown in Fig. 2. We observe the data availability into a downward trend with the increase of node failure. When the node failure rate is greater than 17%, the maximum is only 0.995 for data availability in the case without replication. Using random placement strategy, data availability is 0.9995, while the placement strategy used FPSSSRS data availability to 0.99995.

Figure 3 plots MTTF versus the failure ratio of replica under repair ratio $\mu = 0.1$, 0.5. We see that MTTF decreases with the failure ratio. MTTF is declining, when the failure ratio λ is greater than the repair ratio μ. There is a serious data loss and has been unable to repair the data under the failure ratio $\lambda = 0.35$. This is because the DCell topology has been destroyed, most of the queries cannot be completed. As shown in Fig. 4, MTTF versus repair ratio with the failure $\lambda = 0.01, 0.02, 0$. We observe that MTTF increases as repair ratio.
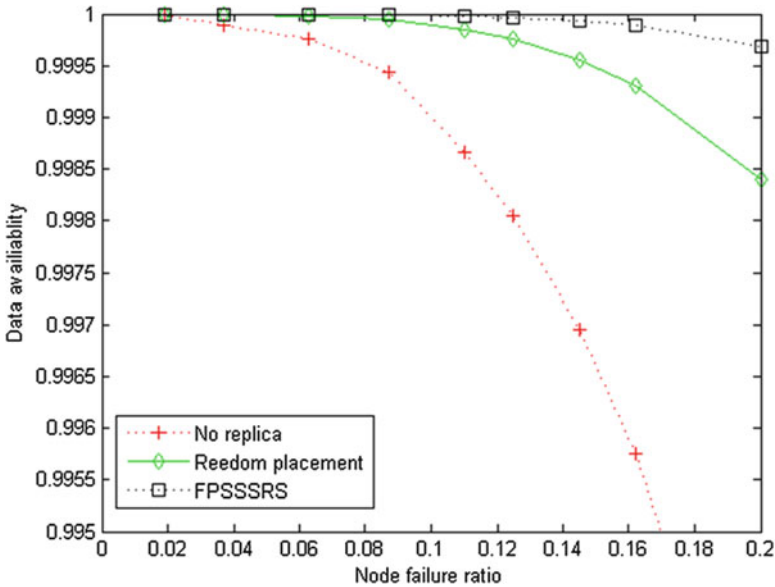

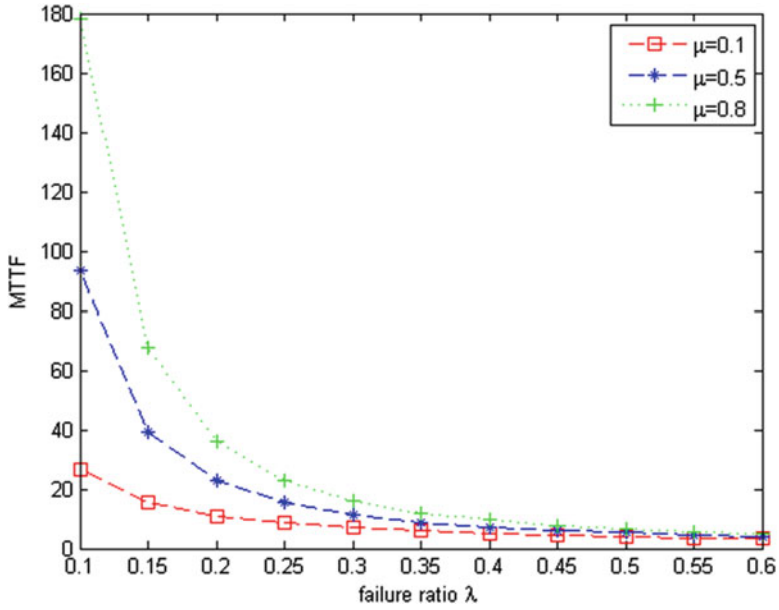**Fig. 2** Data availability for different replica placement with node failure

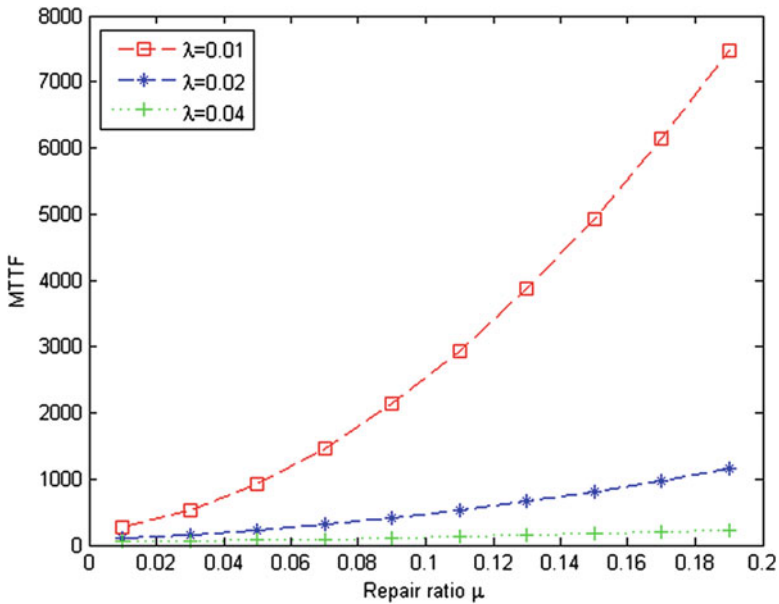**Fig. 3** MTTP vs. failure ratio

Fig. 4 MTTR vs. replica ratio

5 Related Work

Replication is a key technology in distributed systems for improving both availability and performance, and has been studied and deployed in different distributed architectures: DBMS, P2P, Data Grids, and Web service. In this paper, we study reliability of the replica placement on DCN. Previous works has focused on achieving high throughput and low latency. Jianliang Xu et al. (2002) examines the replica placement strategies for applications with both read and write operations, and proposed two efficient optimal placement algorithms for replicating data on the proxies with and without replica number constraint respectively. Replication proxies are located alongside some selected routers and form a tree structure. Zhao Zhang and Weili Wu (2008) study the optimal placement problem in a ring network with majority voting protocol and even numbers of replicas. Veronika Rehn-Sonigo ((2007)) propose a efficient algorithm called optimal replica placement to determine locations for placement replicas in the replica placement problem in cluding QoS and bandwidth. John A. Chandy (2008) present a strategy for placement of replica fragments particularly with respect to the minimizations of data communication latencies.

In the distributed storage systems, HotSpot uses client load to decide placement policy (Qiu et al. 2001) while HotZone improve on the computational complexity of HotSpot in optimizing access latency (Szymaniak et al. 2005). Pallis et al. (2005) develops a heuristic method to optimize placement but does require popularity statistics

In P2P file-storing system like TotalRecall (Ranjita Bhagwan et al. 2004) and Chord (Ion Stoica et al. 2001), replicas have been recognized in two ways: global randomly chosen set of nodes and ID neighborhood node. Pastry uses the active replication, randomly placed in the network to improve data availability, and expects to make the user get closer replica by replication. CFS system relies on a static replication factor coupled with an active repair policy. Replicas are propagated around the ring structure on Chord. Usually, structured overlays are used to place replicas to ID neighboring node. The data object is copied and store to the k neighboring nodes, which has two reasons: First, the adjacent node ID can be quickly found, but the adjacent node ID Often evenly distributed in the network, they also less likely to fail, thus increasing data availability.

6 Conclusion

DCell is a novel network structure adapted to the DCN, which uses a recursively-defined structure to interconnect servers. When data is written into the node, it must be replicate multiple replicas, and placed them in different nodes to prevent data loss caused by the failure. This paper is mainly studied the replica placement strategy for the DCell. The main content is involved: (1) replica placement strategies and

algorithms; (2) the re-recovery strategy, how to re-create after the failure. (3) How to dynamically select the number of replica, and to guarantee reliability of data, but also not to waste too much storage space.

We propose a novel idea that a replica is placed in the node of the layer by the fixed-point single-source shortest path, and there is at least one replica in each layer of the DCell. According to the proposed placement strategy of replicas, we suppose a mechanism to restore lost replica based on the neighboring nodes in the same path. When the replica is lost, the neighboring node selects the node in the same path and creates a new copy using the greedy algorithm. Finally, we simulate and validate algorithm of placement and repair for replica. Experiments show that our algorithm has obvious effect to improve the reliability of data.

References

Al-Fares M, Loukissas A, Vahdat A (2008) A scalable, commodity data center network architecture. In: Proceedings of the ACM SIGGOMM, Seattle, Washington, USA

Barroso L, Dean J, Holzle U (2003) Web search for a PLANET: THE Google cluster architecture. IEEE Computer Society, Los Alamitos

ByungGon C, Dabek F, Haeberlen A, Sit E, Weatherspoon H, Frans Kaashoek M, Kubiatowicz J. Morries R (2006) Efficient replica maintenance for distributed storage systems. In: Proceedings of the NSDI'06, San Jose, CA

Chandy JA (2008) A generalized replica placement strategy to optimize latency in a wide area distributed storage system. In: Proceedings of the ACM, DADC'08, Boston, MA, USA, 24 June 2008

Chang F, Dean J, Ghemawant S, Hsieh WC, Wallach DA, Burrows M, Chandra T, Fikes A, Gruber RE (2006) Bigtable: a distributed storage system for structured data. In: Proceedings of the OSDI'06. hadoop. http://hadoop.apache.org

Chuanxiong Guo, Haitao Wu, Kun Tan, Lei Shi, Yongguang zhang, Songwu Lu (2008) DCell: a scalable and fault-tolerant network structure for data centers. In: Proceedings of the ACM SIGCOMM'08, New York, NY, USA, 17–22 Aug 2008

Dean J, Ghemawat S (2005) MapReduce: simplified data processing on large clusters. Google, Inc. Retrieved 6 Apr 2005

Guo et al. (2009) ACM SIGCOMM. Association for Computing Machinery, Inc., Barcelona, Spain, October 2009

Ion Stoica, Robert Morris, David Karger, Frans Kaashoek M, Hari Balakrishnan (2001) Chord: a scalable peer-to-peer lookup service for internet. In: SIGCOMM'01, San Diego, CA, USA

Jianliang Xu, Bo Lin, Dik L. Lee (2002) Optimal replica placement on transparent replication proxies for read/write data. pcc. In: 21st IEEE International Performance, Computing, and Communications Conference (IPCCC 2002), Phoenix, Arizona, pp 103–110

Kliegl M, Lee J, Li J, Zhang X, Guo C, Rincon D (2010) Generalized DCell structure for load balanced data center networks. In: Infocom 2010, Institute of Electrical and Electronics Engineers, Inc.

Li et al. (2009) IEEE Infocom. IEEE Communications Society, April 2009

Markus Kliegl, Lee J, Li J, Zhang X, Guo C, Rincon D (2009) The generalized DCell network structures and their graph properties, Microsoft Research Technical Report, MSR-TR-2009-140, Oct 2009

Ming Zhong, Kai Shen, Joel Seiferas (2008) Replication degree customization for high availability. In: Proceedings of the ACM EurroSys'08, Glasgow, Scotland, UK, 1–4 Apr 2008

Pallis et al. (2005) Proceedings of the 3rd Latin American Web Congress (La-Web 2005). IEEE Press, Buenos Aires

Pascal Felber, Andre Schiper (2001) Optimistic active replication. In: ICDCS'01, Phoenix, AZ, USA

Qiu L, Padmanabhan VN, Voelker GM (2001) On the placement of web server replicas. In: Proceedings of the IEEE INFOCOM, Anchorage, AK, USA

Ranjita Bhagwan, Kiran Tati, Yu-Chung Cheng, Stefan Savage, Geoffrey M. Voelker (2004) Total recall: system support for automated availability management. In: Proceedings of the NSDI'04, University of California, San Diego

Sanjay Ghemawat, Howard Gobioff, Shun-Tak Leung (2003) The Google file system. In: Proceedings of the ACM SOSP'03, Bolton Landing, New York, USA, 19–22 Oct 2003

Szymaniak M, Pierre G, van Steen M (2005) Latency-driven replica placement. In: Proceedings of the international Symposium on Applications and the Internet (SAINT), Trento, Italy

Veronika Rehn-Sonigo (2007) Optimal replica placement in tree networks with QoS and bandwith constraints and the closest allocation policy. http://www.ens-lyon.fr/LIP

Zhao Zhang, Weili Wu (2008) Proceeding IEEE Computer Society, ICDCS'08, the 28th international conference on distributed computing systems, 2008, pp 387–394

Hybrid Communication Method for Data Gathering in Wireless Sensor Networks

Quan Huang, Huiyong Yuan, and Yeqing Yi

Abstract When cluster heads transmit their data to the sink via multi-hop communication, the cluster heads closer to the sink are burdened with heavy relay traffic and tend to die early. On the contrary, if all cluster heads transmit data's to the sink via single-hop communication, the cluster heads further from the sink will die much more quickly than those closer to the sink. In this paper, we first develop an analytical model to derive the optimal cluster radius. Then we propose a hybrid communication method where the nodes can transmit data to the sink in either single-hop or multi-hop. Finally, we conduct extensive experiments and show that our method outperforms LEACH and HEED in terms of network lifetime by balancing energy consumption.

1 Introduction

Wireless sensor networks have attracted much research attention in recent years and can be used in many different applications, including battlefield surveillance, machine failure diagnosis, biological detection, inventory tracking, home security, smart spaces, environmental monitoring, and so on (Akyildiz et al. 2002). A wireless sensor network consists of a large number of tiny, low-power, cheap sensor nodes.

Q. Huang (✉) • H. Yuan • Y. Yi
Hunan Institute of Humanities, Science and Technology, Loudi Hunan, China
e-mail: hnldhq@126.com; hnldyhy@126.com; yeqingyi@126.com

The sensor nodes in a wireless sensor network are usually deployed randomly inside the region of interest or close to it. The sink give commands to all the sensor nodes and gather data from the sensor nodes.

Being different from the traditional wireless networks, the sensor nodes are limited in energy, so the method for data gathering and routing must be energy efficient in order to prolong the lifetime of the entire network (Dasgupta et al. 2003). In order to prolong network lifetime, the sensor nodes can be organized hierarchically by grouping them into clusters. A better clustering algorithm may average the workload on each sensor node. The important factor that impacts the energy consumption is the cluster radius (Chen and Megerian 2006). If the difference in the cluster radius is big among all the generated clusters, the cluster-head having large cluster radius will consume a larger amount of energy than the others.

In clustered wireless sensor networks, the sensor nodes do not transmit their collected data to the sink, but to designated cluster heads which aggregate the data packets and send them to the sink via single-hop or multi-hop communication mode. For single-hop mode, the cluster heads furthest away from the sink are the most critical nodes, while in multi-hop mode, the cluster heads closest to the sink are burdened with a heavy relay traffic load and die first (Olariu and Stojmenovic 2006). In case of sensor nodes failure or malfunctioning around the sink, the network connectivity and coverage may not be guaranteed. No matter how many remaining sensor nodes are still active, none of them can communicate with the sink. As a result, the system lifetime becomes short.

In this paper, we develop an analytical model to estimate optimal cluster size and propose an optimal communication method for periodical data gathering applications in sensor networks. In our model, the cluster heads are allowed to communicate with the sink with hybrid communication method instead of only multi-hop or only single-hop.

2   Related Work

Many protocols have been proposed for data gathering or communication in wireless sensor networks. The LEACH (Heinzelman and Chandrakasan 2002) is a routing protocol for forming clusters in a selforganized homogeneous sensor network when the sink is located far from the sensor nodes. In LEACH, some nodes are elected as cluster heads while the other nodes communicate with the sink through cluster heads. This protocol randomly rotates the job of cluster heads based on the node's remaining energy in order to uniformly balance the energy consumption throughout the network.

The LEACH allows only single-hop clusters to be constructed. On the other hand, the authors Su and Zhang (2006) propose the similar clustering protocol where sensors communicate with their cluster-heads in multi-hop communication

mode. The HEED (Younis and Fahmy 2004) extends LEACH by incorporating communication range limits and cost information. In HEED, the initial probability for each sensor node to become a cluster head is dependent on its residual energy. Later on, sensor nodes that are not covered by any cluster heads double their probability of becoming a cluster head. This procedure iterates until all sensor nodes are covered by at least one cluster head.

The authors Yu et al. (2008) study the problem of constructing a data gathering tree over a wireless sensor network in order to minimize the total energy for compressing and transporting information from a set of source nodes to the sink. They investigate a tunable data compression technique that enables effective tradeoffs between computation and communication costs. They derive the optimal compression strategy for a given data gathering tree and investigate the performance of different tree structures for networks deployed on a grid topology.

In Yang et al. (2010), the authors present a cluster-based data gathering and transmission protocol (CDAT) for wireless sensor networks. The CDAT achieves a good performance in terms of lifetime by a clustering method of balancing energy consumption and data prediction transmission strategy. The initial probability of sensor node for cluster head election is derived from mathematical relation between application's seamless coverage ratio and numbers of required cluster heads.

The authors Tharini and Vanaja Ranjan (2010) propose a scheme for data gathering in wireless sensor networks. This scheme exploits temporal correlation in sensor data and uses this characteristic to efficiently gather at the sink using joint compression and prediction algorithm.

Our work differs from the above works since we focus on how to prolong the lifetime of sensor networks using hybrid communication method. We first derive the optimal cluster radius and then propose a hybrid communication method for data gathering in wireless sensor network.

3 System Model and Problem Formulation

3.1 Energy Consumption Model

In this paper, we assume there is an energy-efficient MAC protocol in the underlying MAC layer, energy will be consumed only when performing sensing task, processing raw data, and transmitting and receiving data for itself and other sensor nodes. We also assume all sensor nodes have power control and can use the minimum required energy to send information to the recipients.

We use a simplified radio model shown in Heinzelman and Chandrakasan (2002) for the radio hardware energy dissipation. Both the free space and the multi-path fading channel models are used in the model, depending on the distance between the transmitter and receiver. The energy spent for transmission of a l-bit packet over distance d is

Fig. 1 Sensor network model
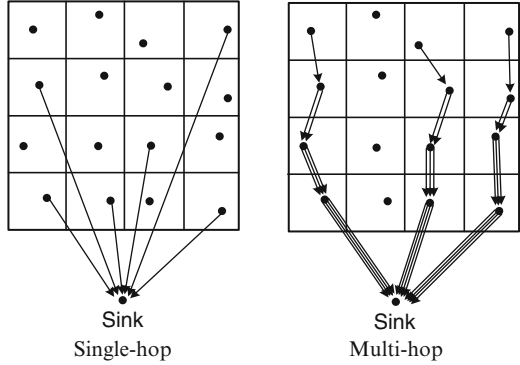
$$E_{Tx}(l, d) = E_{elec}(l) + E_{amp}(l, d)$$

$$\begin{cases} l \times E_{elec} + l \times efs \times d^2 \ , & d \leq d_0 \\ l \times E_{elec} + l \times emp \times d^4, & d > d_0 \end{cases} \tag{1}$$

and to receive this message, the radio expends energy is

$$E_{Rx} = l \times E_{elec} \tag{2}$$

We assume that the radio channel is symmetric, which means the cost of transmitting a message from A to B is the same as the cost of transmitting a message from B to A.

3.2 System Assumptions

In this paper, we assume that N sensors are randomly located in the square $W \times W$ observation region, and are stationary that they do not change their locations once deployed, the sink is located outside the region at position $(W/2, W + D)$. The sensor nodes are grouped into clusters. With clustering, sensor nodes transmit their sensed data to their cluster heads. Cluster heads aggregate the received data and forwards it to the sink. Figure 1 depicts an application where sensor nodes periodically transmit their data to the sink. The figure illustrates that cluster heads transmit the aggregated data in single-hop and multi-hop communication mode.

We assume that data aggregation technique is available, and data collected from sensor nodes in a cluster is packed into one packet. Furthermore, the sensor nodes are location-unaware and have the same characteristics. It's also assumed that the sensed data is highly correlated, thus the cluster heads can always aggregate the data gathered from its members.

3.3   Hybrid Communication Method

Notice that in the multi-hop mode, the smaller distance between cluster heads and the sink, the more energy consumes. On the other hand, in single-hop mode, cluster heads which are farther from the sink dissipate more energy than those closer to the sink.

To balance the energy consumption, our network model employs the hybrid communication method which consist of the single-hop and multi-hop mode. The cluster heads use single-hop mode in some rounds but multi-hop mode in the other rounds. We use parameter λ to measure how often the single-hop mode is used. Suppose T is the total rounds that the network can perform, Ts is the number of rounds that single-hop mode is used and Tm is the number of rounds that multi-hop mode is employed. Let $\lambda = Ts/T$ be the frequency with which the single-hop mode is used.

3.4   Problem Statement

Our objective is to find the optimal cluster radius r and the parameter λ for the hybrid communication method to maximize the network lifetime.

In order to extend the lifetime of the sensor network, we need to maximize the lifetime of the node who dies first. The problem of maximizing the lifetime can be written as

$$Objective : \max\{ T \}$$
$$Subject\ to\ \ 0 \leq \lambda \leq 1 \tag{3}$$
$$T_s \times E_{single\_hop} + T_m \times E_{multi\_hop} \leq E_{init}$$

where we denote E_{single_hop} and E_{multi_hop} as the average energy consumption of single-hop and multi-hop mode in a round, E_{init} as the sensor initial energy. The constraints indicate that we can use hybrid communication method.

4   Solutions for the Optimization Problem

4.1   The Optimal Cluster Radius

Mhatre and Rosenberg in (2004) define characteristic distance as that distance which when used as the inter-sensor distance, minimizes the energy consumption for sending a data packet from a source sensor to a destination sensor. This characteristic distance, d_{char}, is

$$d_{char} = \sqrt{2E_{elec}/efs} \tag{4}$$

Hence, we set the width of each ring is d_{char}. Only width d_{char} can the energy consumption rate be minimized. A good clustering algorithm may average the workload on each sensor and the generated clusters should have the same cluster sizes. Thus, we set the optimal cluster radius $r = d_{char}$.

4.2 Solutions for the Problem

In this section, we first study the energy consumption. Let us consider a single cluster with radius r. The average number of sensor nodes in the cluster is

$$N_{clu} = \frac{N}{W \times W} \times \left(\sqrt{2}r\right)^2 = \frac{2Nr^2}{W \times W} \tag{5}$$

To keep the total energy dissipation within the cluster as small as possible, the cluster head should be positioned at the centroid of the cluster area S. In this case, the square of distance between cluster members and the cluster head is given as:

$$d_{clu}^2 = \iint_S (x^2 + y^2)dxdy = r^2/2 \tag{6}$$

The transmitter energy consumption of cluster members is given by

$$E_1 = (N_{clu} - 1) \times len \times (2E_{elec} + efs \times d_{clu}^2) \tag{7}$$

where we denote len as the length of the data packet.

Each cluster head receives the data from all member nodes in the cluster, aggregates and transmits to the sink. The sink is located at the coordinate (x_0,y_0), the distance between the sink and a cluster head with the coordinate (x_i,y_i) can be expressed as

$$d_{to\_sink} = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \tag{8}$$

The energy required for the transmission of aggregation data is

$$E_2 = len \times (E_{elec} + emp \times d_{to\_sink}^4) \tag{9}$$

The number of clusters can be given by

$$N_{cluster} = \lceil W \times W/(2r^2) \rceil \tag{10}$$

Thus, the average energy consumption of single-hop mode in a round is

$$E_{single\_hop} = (E_1 + E_2) \times N/N_{cluster} \tag{11}$$

Let us consider the average energy consumption of multi-hop transmission. The total transmitter energy consumption of cluster members is

$$E_3 = (N - N_{cluster}) \times len \times (2E_{elec} + efs \times d_{clu}^2) \tag{12}$$

The number of transmission data among cluster heads is given by

$$N_c = \sum_{i=1}^{\lceil W/(\sqrt{2}\,r)\rceil - 1} i \times \left\lceil W/(\sqrt{2}\,r) \right\rceil \tag{13}$$

The energy consumption of transmission data among cluster heads can be expressed as

$$E_4 = N_c \times len \times (2E_{elec} + efs \times (2r)^2) \tag{14}$$

The energy required for the transmission of the data from cluster heads to the sink is

$$E_5 = len \times (E_{elec} + emp \times D^4) \tag{15}$$

where we denote D as the distance between the sink and observation region border. Thus, the average energy consumption of multi-hop in a round is

$$E_{multi\_hop} = (E_3 + E_4 + E_5)/N \tag{16}$$

Hence, the original problem (3) is reformulated as:

$$
\begin{aligned}
&Objective : \max\{\,T\,\} \\
&Subject\ to\ \ 0 \le \lambda \le 1, \quad T_s = \lambda \times T, \quad T_m = T - T_s \\
&\qquad\quad T_s \times E_{single\_hop} + T_m \times E_{multi\_hop} \le E_{init} \\
&\qquad\quad constraints\ (11),\ (16)
\end{aligned}
\tag{17}
$$

Note that problem (17) is a linear programming problem, which is easier to solve than the problem (3).

4.3 Clustering Algorithm

Clustering a wireless sensor network means partitioning its nodes into clusters, each one with a cluster head and some sensor nodes as its members.

```
state←candidate;
broadcast Residual_Energy_Msg
receive Residual_Msg;
update neighborhood table NT[];
t←the broadcast delay time for competing a cluster head;
while (the timer1 (T) for cluster head election is not expired)
{     if (CurrentTime<t)
      {        if (a Head_Msg is overheard from a neighbor NT[i])
               {     state←plain;     NT[i].state=head;     }
               else   continue;
      }
      else    if (state=candidate)
      {           state←head;        broadcast Head_Msg;   }
}
while (the timer2 for cluster join is not expired)
{       if (state=plain && have not sent Join_Msg)
               Send (Join_Msg to the nearest cluster head);
        else   receive (Join_Msg from its neighbor plain nodes);
}
```

Fig. 2 Cluster head selection pseudocode

Our algorithm is a distributed cluster heads competitive algorithm, here cluster head selection is primarily based on the residual energy and position of each node. Every sensor become a tentative cluster head with the same probability p which is a predefined threshold. The algorithm pseudocode for an arbitrary sensor node i is given in Fig. 2.

5 Simulation and Results

We compared the performance of hybrid communication method with LEACH and HEED in this section. We used the same energy consumption model as Heinzelman and Chandrakasan (2002). Every result shown below is the average of 50 independent experiments.

We first use the number of alive sensor nodes to compare the performance of the three methods. Figure 3 shows the comparison results. It is shown that the first sensor node halts for the starvation of battery in LEACH and HEED is earlier than in our method. As shown in Fig. 3, our method clearly improves the lifetime of sensor nodes.

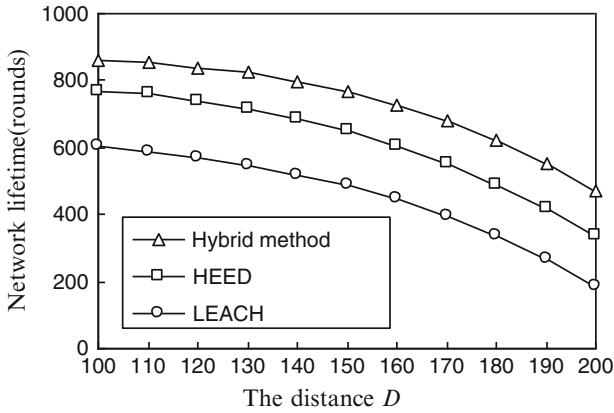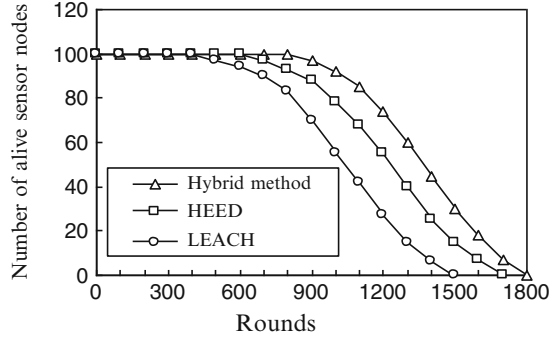**Fig. 3** Number of alive sensor nodes comparison, we assume $D = 100$ m

Fig. 4 Network lifetime comparison

We then compare the network lifetime to evaluate the performance of the three methods. We define the network lifetime to be the number of rounds until the first sensor node is drained of its energy. Figure 4 shows the network lifetimes of the three methods when the distance D changes from 100 to 200 m. However, our method outperforms LEACH and HEED in terms of network lifetime.

6 Conclusion

In this paper, we present a hybrid communication method where the cluster heads can transmit data in either single-hop or multi-hop communication. Simulation experiment results show that the hybrid communication method can efficiently balance the energy consumption and prolong the network lifetime.

References

Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E (2002) Wireless sensor networks: a survey. Comput Netw 38:393–422

Chen H, Megerian S (2006) Cluster sizing and head selection for efficient data aggregation and routing in sensor networks. In: Proceedings of the WCNC 2006, Las Vegas, USA, pp 2318–2323

Dasgupta K, Kalpakis K, Namjoshi P (2003) An efficient clustering-based heuristic for data gathering and aggregation in sensor networks. In: Proceedings of the WCNC 2003, New Orleans, USA, pp 1948–1953

Heinzelman WR, Chandrakasan A (2002) An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wirel Commun 1(4):660–669

Mhatre V, Rosenberg C (2004) Design guidelines for wireless sensor networks: communication, clustering and aggregation. Ad Hoc Netw J 2(1):45–63

Olariu S, Stojmenovic I (2006) Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting. In: Proceedings of the INFOCOM 2006, Barcelona, Spain, pp 1–12

Su H, Zhang X (2006) Energy-efficient clustering system model and reconfiguration schemes for wireless sensor networks. In: Proceedings of the CISS 2006, Princeton University, Princeton, USA, pp 99–104

Tharini C, Vanaja Ranjan P (2010) An efficient data gathering scheme for wireless sensor networks. Eur J Sci Res 43(1):148–155

Yang J, Zhang DY, Zhang YY (2010) Cluster-based data aggregation and transmission protocol for wireless sensor networks. J Softw 21(5):1127–1137

Younis O, Fahmy S (2004) HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans Mobile Comput 3(4):336–379

Yu Y, Krishnamachari B, Prasanna VK (2008) Data gathering with tunable compression in sensor networks. IEEE Trans. Parallel Distributed Systems, 19(2):276–287

The Study and Application of "Trinity System" Based on Multi-living Agent Method

Dong Zhong and Chun Chen

Abstract "Trinity System", which is consisted of "110,122,119" three police command stations merging into a police command center, the system is to achieve a unified alarm, unified command, joint action and to provide emergency services and urban public security. Multi-living Agent Method is a new method to analysis the complex system. This paper proposes to use a new method of multi-living agent to design "Trinity System" based on analyzing the traditional "Trinity System", according to the guiding theory of the system. The paper designs the whole framework and system software architecture of "Trinity System", it provides a good research direction to enhance the capabilities of emergency response for "Trinity System".

Keywords "Trinity System" • Multi-living Agent • System theory

1 Introduction

With the rapid development of the economy and society, the city public security has become an increasing important issue. Because urban disasters and emergencies can not be avoided completely and a single sector of the government subordinate

D. Zhong (✉)
Department of Electronic and Information, Hubei University of Science and Technology, Xianning, China
e-mail: Zhongdong0129@163.com

C. Chen
The Commission of Xianning City Housing and Urban and Rural Construction, Xianning, China
e-mail: chenchunzd@163.com

agencies is difficult to deal with the disasters and emergencies, meanwhile solving the memory and distinguish of special service code, we need to build emergency response mechanism. The key is to achieve "110,122,119" three police command units merging into a police command center, this is the "Trinity System", it can achieve a unified alarm, unified command, joint action to ensure emergency services and city public security. "Trinity System" (Jotshi et al. 2009) is an integrated system, which is consist of several subsystems, meanwhile "Trinity System" is a complex information system. This paper proposes to use a new method of multi-living agent to design "Trinity System" based on analyzing the traditional "Trinity System", according to the guiding theory of the system. The paper designs the whole framework and system software architecture of "Trinity System", it provides a good research direction to enhance the capabilities of emergency response for "Trinity System".

2 The Basic Discussion of "Trinity System", the Interposing of "Multi-living Agent" Method

2.1 "Trinity System" Is Related Closely to and System Theory

In addition to the general system characteristics, "Trinity System" has the characteristic of modular hierarchical system architecture design, the system software core is based on a unified messaging service platform, it merges independent species (110,119,122) departments, multiple data, various business into a system, it can select the number of subsystems according to the actual needs (Yuanming 2006).

The core principles of system theories, such as the movement of the formation about open dissipative system layer self-organizing mechanism, evolution of the system and environment, and so on. These system theories summarized highly the trend of the dynamic development. They have practical significance for the study of "Trinity System" (Yue Wang 2006). For example, the improving of the police efficiency is related with the response alarm systems, the environment of command and control system, a unified monitoring and management platform. It requires the system and environment co-evolution to achieve our expected requirements. In this case, In this case, this paper proposes the new method of multi-living agent to analyze and study the whole framework and system software architecture, it has important significance to improve the emergency response capacity of "Trinity System".

2.2 The Basic Idea of "Multi-living Agent"

In the first, the "living" in "Multi-living agent" only comes from the word of the function (action), then it became a specific technical word by re-define the specific

examples after the application (ZhongDong and Chen chun 2011; Atkey 2008). It can be divided into three cases:

1. The word can be used only in a qualitative sense, which is corresponding with resting state or non-live (moving) of (inactivity);
2. The word has been defined as a measure of activity intensity. If it has been given the meaning of the activity, we should use the activation in the general; if the activity meaning has been removed, then we should use the non-activation or inactivation and other words;
3. To the physiological processes occurring in the organism or in active state or property; "Living" means a nature of activity for the vitality which is going on a tenacious.

"Living" gives the vigor to the system, it has showed the character of the survival and development for "Trinity System", it plays a important role in various functions. The "more" in "Multi-active agent" reflects the multiple functions in "Trinity System" (Stüber 2001); (Lai Hong Xiao et al. 2007), it has the essential attributes of multiple location, multiple levels and a number of related subsystems, the "proxy " in "multi-living agent" integrated the tendency of the human, the system function, the system's basic structure and operation of the system based on "living" sense, mapping the future vision into reality system or a representative system, it also can be mapped to a specific system, namely "Trinity System", this system has the man-machine integration, scheduling on-site emergency response center and the combination of characteristics(Atkey 2008). This paper proposes the whole framework and system software architecture based on the above basic characteristics of multi-active agent.

3   The Design of the Whole Framework and System Software Architecture for "Trinity System" Based on Multi-living Agent Method

3.1   The Design of the Whole Framework for "Trinity System" Based on Multi-living Agent

This paper designs the whole framework for "Trinity System" based on multi-living agent, the system is constituted of a unified monitoring and management platform, the shared platform of the plans for management, GIS/GPS sharing platform, the exchange platform for information sharing, communication integration services platform, database sharing platform. This multi-agent system is mainly to provide daily communication support for the activity of the Public Order (110) Patrol (122), Fire (119) emergency services. In addition, it is to avoid duplication investment of the government, it also can provide daily scheduling tool for urban management, forest fire, flood and other government departments. Figure 1 has shown the whole framework:
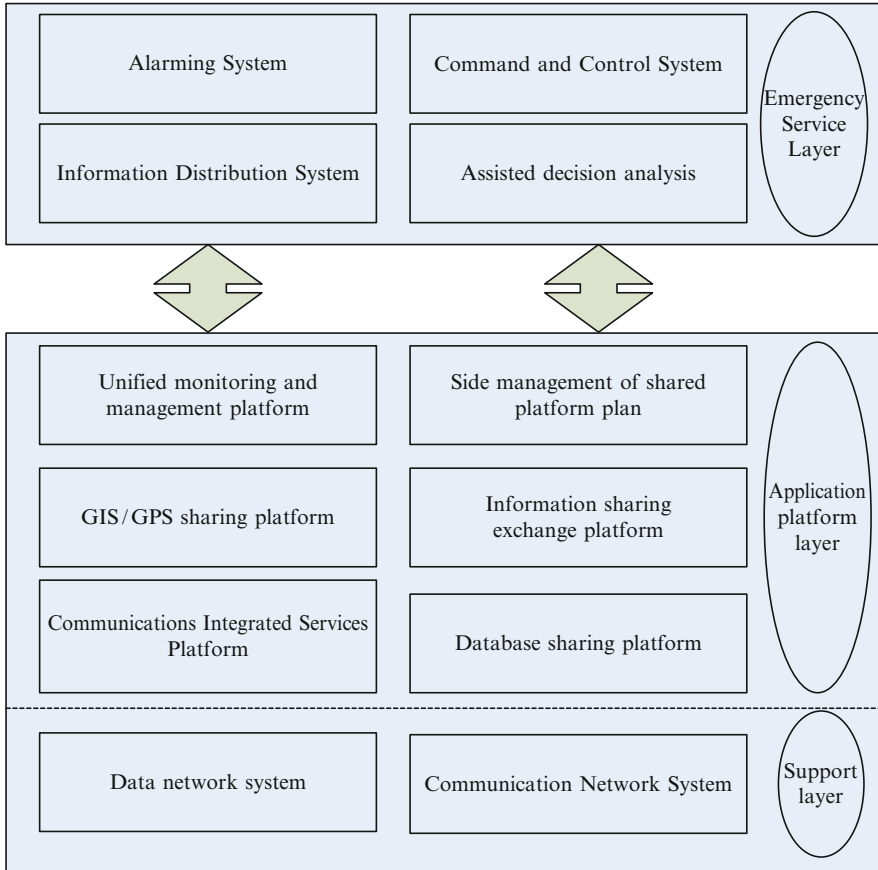
Fig. 1 Three more than one active agent system block diagram of the overall framework

For the whole framework of "Trinity System", this paper has a preliminary study based on multi-living agent approach, it is equivalent to using the multi-living agent approach to study a complex information system. For the "Trinity System", from the point of system theory, this is a complex information system and it is associated with assistant scheduling decision function, meanwhile it is a further high-level decision problem of remote dispatching. In this complex information system, "living" means in the top can be understood to make great effort to protect the network management and remote dispatching center to take the most efficient way to provide daily services for emergency communications support. At the same time, "living" also represents the timeliness and effectiveness measures of the communication when a problem comes from the system or communication handicaps occurs. "Multi-living" represents the functioning coordination among the various business systems platform agent layer (for example, response alarm system, information distribution system, a unified monitoring and management platform, GIS/GPS sharing platform,
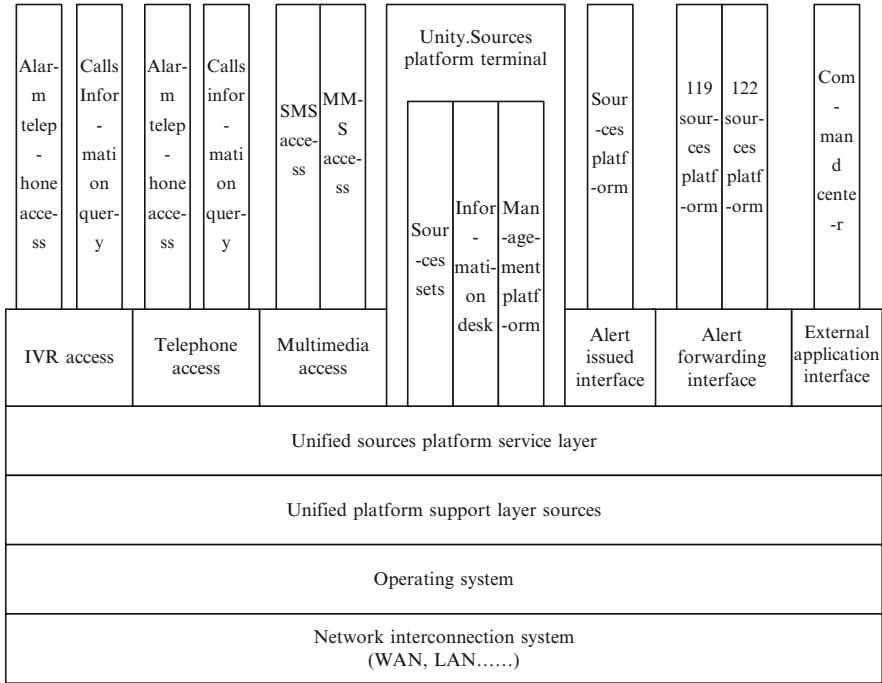
Fig. 2 Three more than one active agent software system block diagram

the exchange platform for information sharing, communication integration services platform, database, and various sub-sharing platform system). Meanwhile, under the construction of each "living agent", the agents can be divided into various sub-agents, such multi-level decomposition until a layer of specific hardware and software agents.

3.2 The Design of System Software Architecture for "Trinity System" Based on Multi-living Agent

The work model of "Trinity System" is a combined of unified alarm, unity and reunification alarm response alarm, the United Source police combination. The alarm command center dispatches directly to the fire command center, traffic police, patrol, police station and other police. Alarming unified platform is one of the main objectives of system implementation, Alarming platform software components are divided according to application level, we design the multi-living agent software system block diagram of "Trinity System" shown in Fig. 2:

The software system has been design in modular idea, the alarm system is consists of a unified platform services layer, support layer, the operating system layer, networking layer. The system should guarantee the access of the underlying, business processes and business systems are independent interfaces each other. From the viewpoint of system theories, this multi-living agent software architecture is also a complex information system, the system is mainly to achieve a unified response alarm, "110,122,119" merging into a unified response alarm platform. The paper divides the basic modules of the software system in accordance with the application of the software, it can ensure the independence among each module. The meaning of "living" in the multi-living agent software system can be understood to achieve the unity of response alarm, it can ensure the scheduling decision. At the same time, "living" also represents the business in case of alarm, the police intelligence information access interface, and police issued the interface, alarm forwarding interfaces and the external application interface are to ensure a unified platform software alarm information exchanging with the outside timeliness and effectiveness of the entire system to ensure data security. "Multi-living" represents the coordination and support of each business layer proxy of multi-agent software architecture (such as the access layer, interface layer, unified alarm platform services layer, unified platform support layer, the operating system level, network interconnection system layer, and so on.). The multi-living agent software architecture has the basic functions and features of following: Network connectivity layer proxy and the operating system layer proxy are the underlying software architecture of multi-living agent. It can ensure the system's security, stability and versatility. Support system layer proxy is running for the underlying application.

4   Conclusions

This paper proposes to use a new method of multi-living agent to design "Trinity System" based on analyzing the traditional "Trinity System", according to the guiding theory of the system. The paper designs the whole framework and system software architecture of "Trinity System", it provides a good research direction to enhance the capabilities of emergency response for "Trinity System".

References

Atkey DA (2008) Practical methodology for the measurement of screening effectiveness of trunking. In: IEEE Conference Publication, USA, pp 77–83
Jotshi A, Qiang Gong, Batta R (2009) Dispatching and routing of emergency vehicles in disaster mitigation using data fusion. Socioecon Plann Sci 43:1–24
Lai Hong Xiao, Cui Yingan, Cui Duwu (2007) Call center architecture based on generic sound card. Comput Eng 22(33):73–75

Stüber GL (2001) Principles of mobile communication, 2nd edn. Kluwer Academic Publishers, New York, USA

Yuanming Sun (2006) Solutions to the problems of urban emergency response system in China. Chongqing Adm 32(5):75–77 (in Chinese)

Yue Wang (2006) The new method for the construction of complex information system – multiple active agents. China Eng Sci 8(5):30–34

ZhongDong, Chen chun (2011) The study of scheduling function optimization for trunking communication system in CERS. Sci Technol Eng 18(3):211–213

Improved Video Object Segmentation Method Based on Background Registration

Wenxiu Fu, Xingmin Wang, Yuhang Wang, and Bo Fan

Abstract In allusion to the common occlusion problems in video segmentation methods, a novel improved video object segmentation method based on background registration is proposed. At first, the high frequency message and most noise of the images were eliminated by 1-D wavelet transform. Then, calculate the front background modeling and the back background modeling by comparing the static index and the static threshold, and employ the secondary frame difference to optimize the front background modeling and the back background modeling, then the initializing background modeling can be obtained. Accumulated frame difference was used to optimize the initializing background modeling to reduce the probability of background that is misjudged in foreground, which makes further improvement on the segmentation results. Last pick-up the video object is obtained by mathematic morphologic method. Experimental results show that the background registration method was very non-sensitive to static threshold and has strong robustness. The algorithm can overcome effectively the occlusion and quickly giving accurate segmentation results from the sequence with static background for video object segmentation, it's also can preferable the method in spatial accuracy.

Keywords Video segmentation • Background registration • Accumulated frame difference

W. Fu (✉) • X. Wang • B. Fan
School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China
e-mail: wxfu@bjtu.edu.cn; 09120308@bjtu.edu.cn; 09120282@bjtu.edu.cn

Y. Wang
College of Petroleum Engineering, China University of Petroleum, Beijing, China
e-mail: wangyuhang.11@163.com

1   Introduction

Nowadays, the moving video object segmentation is becoming the focus of various researches, it's applied widely in various fields such as indexing and retrieval of video information (Lin and Kung 2000), the video intelligent surveillance system (Nascimento and Marques 2006), video coding (Porikli 2004), video analysis (Foresti 1999) and so on. Typical video segmentation algorithms are optical flow algorithm (Zhang Ze-xu and Cui Ping-yuan 2008), frame difference method (Meier and Ngun 1999), background subtraction method (Fan Ya-chun et al. 2009) and so on. Though the frame difference method is simple, there will be several holes when a large area of consistent region exists within moving video object or the occlusion problems will arise when the moving object is moving fast. Though background subtraction method can detect dynamic change well, it's only fit to the video sequence which could extract the whole background. In consideration of the limitation of background subtraction method, T. Meier and K.N. Nagan propose a mixed video segmentation algorithm whose key is filter stationary background (Meier and Nagan 1998). On these bases, the algorithm has been improved by Yang who proposes a background registration method based on inverse process of filter stationary background (Yang Gao-bo and Zhang Zhao-yang 2003). Because Yang's method need not motion estimation and has less calculation, Yang's method is selected to produce background modeling, and is improved too.

In this paper, we propose a novel video object segmentation method based on improved background registration. We deal with $2N + 1$ images each time. According to the top ten images, the front background modeling is generated by methods of frame difference, threshold, the secondary frame difference and background registration. Similarly, the back background modeling is created by the back ten images. The background modeling which can be obtained by the front and the back background modeling has high precision and can overcome effectively the occlusion.

2   Preprocessing

The images always suffer from various kinds of noises and external disturbances, which has adverse effect on the succeeding treatment. Therefore, denoise is indispensable to the stage of image preprocessing. We use wavelet transform (Yang Gao-bo and Zhang Zhao-yang 2005) to reduce noises.

We can assume a signal: $f(x) \in L^2(R)$, its continuous wavelet transform is defined as follows.

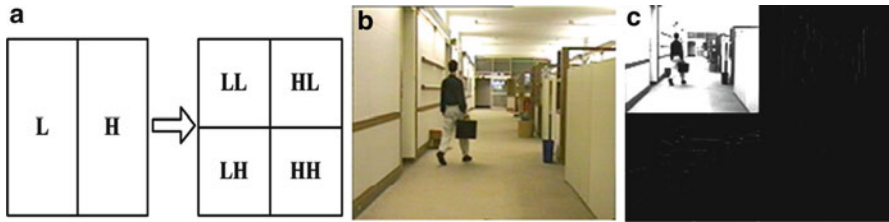$$W_f(a, b) = \frac{1}{\sqrt{|a|}} \int f(x) \psi^* \left(\frac{t - b}{a} \right) dx \tag{1}$$

Fig. 1 The wavelet one-resolution of image. (**a**) The principle diagram, (**b**) the original image, (**c**) the result of wavelet

$\psi^*(t)$ is the conjugation of $\psi(t)$. In this paper, harr wavelet radix is used to filter. Figure 1 illustrates the wavelet one-resolution for image. LL is the wavelet coefficients, which keep the low spectra information of the image and eliminates most of the image noise and the high spectra information.

3 The Background Registration Algorithm

Background registration is the process that extract background model from the video sequence. The more accurate background model, the final results is more accurate. Yang's background registration method extracts the background model by the method that compares the static index with the static threshold. So, choose a proper static threshold is very crucial. The background registration algorithm we proposed uses the front background modeling and the back background modeling to create the background modeling, which can overcome effectively the occlusion and reduce the influence of the static threshold.

3.1 Generate the Initial Background Modeling

In this paper we produce a novel approach of front-back background registration. We deal with $2N+1$ images each time. The selection of N has to do with the static threshold (TH_{SI}) and satisfying the conditions: $N = TH_{SI} + 2$. First, the front background modeling can be generated by the front ten images, and the back background modeling can be generated by the back ten images. Then we can obtain the initial background modeling by the front and the back background modeling. This background modeling which can overcome effectively the occlusion is good for the post-treatment.

3.1.1 Generate the Front Background Modeling

First, the front frame difference should be computed. The basic idea of the frame difference method is that if the object in the image is moving, then the gray value of the corresponding position also changed, but changed a little in the background area. So, we can get the motion information by compute the frame difference of the consecutive frame. Suppose the image sequences after wavelet transform are: $f_n(x, y)(1 \leq n \leq 2N - 1)$, which should be processed by threshold:

$$Z_{qn}(x, y) = \begin{cases} 1, if \ |f_n(x, y) - f_{n-1}(x, y)| > T \\ 0, if \ |f_n(x, y) - f_{n-1}(x, y)| \leq T \end{cases}, n = 2, 3, \ldots, N \qquad (2)$$

T is the threshold. This paper uses the secondary frame difference to reduce the error. The secondary frame difference (Zhao Yan-ling et al. 2004) is that get the intersection of the two adjacent frame difference images. So we can obtain the frame difference modeling of the n-frame:

$$D_{qn} = Z_{qn} \bigcap Z_{q(n+1)}, \ n = 3, 4, \cdots, N \qquad (3)$$

Then the background registration begins. The static index record number of times that one pixel of a position belong to the background in the successive frames. So the static index of the pixel (x, y) in the n-frame is:

$$SI_{qN}(x, y) = \begin{cases} SI_{q(N-1)}(x, y) + 1, & if \quad D_{qN}(x, y) = 0 \\ 0, & else \end{cases} \qquad (4)$$

Then we can obtain the front static modeling of the n-frame:

$$BG_{qN}(x, y) = f_N(x, y), \ if \quad SI_{qN}(x, y) \geq TH_{SI} \qquad (5)$$

TH_{SI} is the static threshold.

3.1.2 Generate the Back Background Modeling

The back background modeling is created by the back N images, and this method is similar to the method of the front background modeling. First, the frame difference modeling can be obtained as follows:

$$Z_{hn}(x, y) = \begin{cases} 1, if \ |f_n(x, y) - f_{n+1}(x, y)| > T \\ 0, if \ |f_n(x, y) - f_{n+1}(x, y)| \leq T \end{cases}, n = N, N + 1, \cdots 2N - 2$$

$$(6)$$

Z_{hn} is the back frame difference of the n-frame. The back frame difference modeling can be obtained after the secondary frame difference.

$$D_{hn} = Z_{hn} \bigcap Z_{h(n+1)}, \ n = N, N+1, \cdots, 2N-3 \qquad (7)$$

Then the back static index of the n-frame is:

$$SI_{hN}(x, y) = \begin{cases} SI_{h(N+1)}(x, y) + 1, & if \ D_{hN}(x, y) = 0 \\ 0, & else \end{cases} \qquad (8)$$

So the back background modeling of the n-frame is:

$$BG_{hN}(x, y) = f_N(x, y), \ if \ SI_{hN}(x, y) = TH_{SI} \qquad (9)$$

3.1.3 Generate the Initial Background Modeling

The initial background modeling is the union of set of the front and the back background modeling. The union of the set is as follows:

$$BG_N = BG_{qN} \bigcup BG_{hN} \qquad (10)$$

BG_N is the initial background modeling. Figure 2 illustrates the moving object which is extracted by each background modeling. (a) illustrates the moving object which is extracted by the front background modeling. (b) illustrates the moving object which is extracted by the back background modeling. (c) illustrates the moving object which is extracted by the initial background modeling.

3.1.4 Optimization of the Initial Background Modeling

From the Fig. 2 we can see there are also some errors in the initial background modeling. The main source of the errors is that the background is judged as foreground by mistaken. So we improve the initial background modeling as follows: Calculate the front and the back accumulated frame difference of the N-frame.

The front accumulated frame difference is:

$$D_q = \sum_{n=3}^{N} D_{qn} \qquad (11)$$

The back accumulated frame difference is:
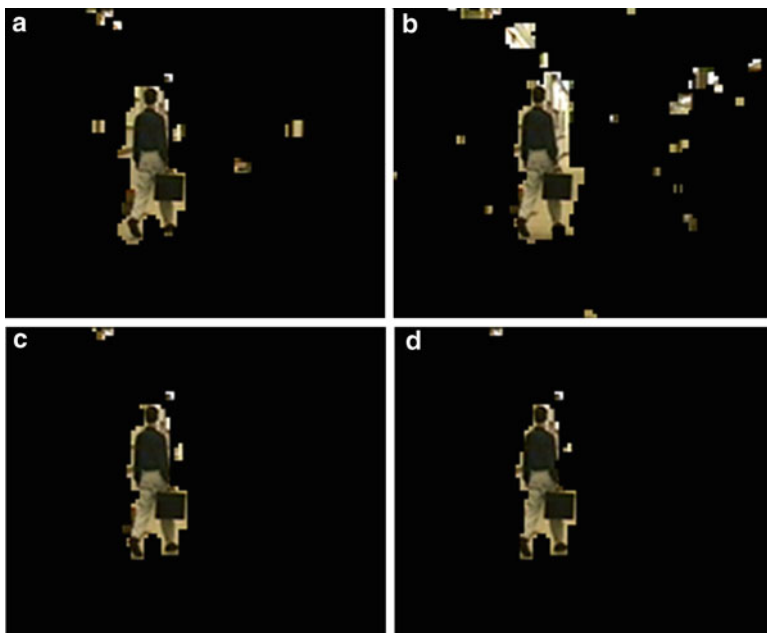
$$D_h = \sum_{n=N}^{2N-3} D_{hn} \qquad (12)$$

Fig. 2 Background modeling. (**a**) Front background modeling, (**b**) back background modeling, (**c**) initial background modeling, (**d**) final background modeling

So the accumulated frame difference can be obtained as follow:

$$D = D_q + D_h \tag{13}$$

According to the initial background modeling, we can find the pixels which satisfy the following conditions in the area which have been judged as the moving object.

$$SI_{qN}(x, y) \neq TH_{SI}, SI_{hN}(x, y) \neq TH_{SI}, \ D(x, y) \leq TH \tag{14}$$

So, $BG(x, y) = f_N(x, y)$. $BG(x, y)$ is the final background modeling, TH is the accumulated threshold and satisfy $TH < TH_{SI}$.

Figure 2 show that we can correct some errors after the optimization process. It's because part of the noise is reduced after the accumulated threshold process, so the background modeling is more exact.

3.2   Comparison of the Algorithms

The background modeling which is obtained by the improved algorithm is little affected by the static threshold. In order to prove the advantage of the improved

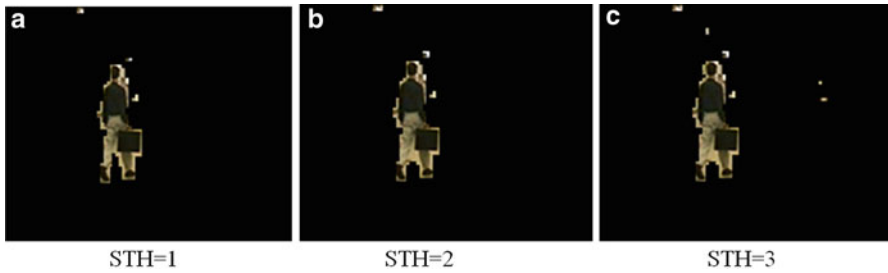**Fig. 3** Yang's algorithm. (**a**) STH $= 1$, (**b**) STH $= 2$, (**c**) STH $= 3$

STH=1 STH=2 STH=3

Fig. 4 The improved algorithm. (**a**) STH $= 1$, (**b**) STH $= 2$, (**c**) STH $= 3$

background registration algorithm, we use the two algorithms to obtain the background modeling of the 47th frame of the Hall Monitor video sequence. The results are shown in Figs. 3 and 4.

The three pictures in Fig. 3 are the background models which are obtained by the initial method and Fig. 4 shows the results which are obtained by the improved method. There are two advantages in the improved background registration. First, because of the improved algorithm adopts front-back background model, so the occlusion phenomenon of the two directions cancel each other out and improves the precision of the background model. Second, the improved background registration is not sensitive to the static threshold. The background model almost no change with the increase of the static threshold. In a word, the background model generated by the improved algorithm can effectively solve the occlusion problem and the static threshold has little effect on the occlusion problem.

3.3 Create the Initial Foreground Modeling

After we get the background model, the initial object model can be extracted directly from the background model as follow:

$$CDM(x, y) = \begin{cases} 0, & \text{if } BG(x, y) = f_N(x, y) \\ 1, & \text{else} \end{cases} \qquad (15)$$

3.4   Post-treatment

After these series of processing, an initial object model has been generated. But there are still some residual noises as shown in Fig. 4. So we can filter out the noises by morphological filter and morphological opening and closing operation.

4   The Algorithm Evaluation and Experiment Results

4.1   Evaluation Criterion

In order to correctly evaluate the performance of an algorithm, we adopt PCM (pixel classification based measure) evaluation criterion (Yang and Zhang 2004). PCM reflects the probability of that the pixels of the foreground are judged as background by mistaken and the pixels of the background are judged as foreground by mistaken. It's defined as follow:

$$PCM = 1 - \frac{Cardi(B_{ref} \bigcap F_{seg}) + Cardi(F_{ref} \bigcap B_{seg})}{Cardi(B_{ref}) + Cardi(F_{ref})} \qquad (16)$$

$Cardi()$ is the number of pixels we have taken; B_{ref} is the referenced background and B_{seg} is the reality background; F_{ref} is the referenced foreground and F_{seg} is the reality background.

The other evaluation of video segmentation algorithm is the time consistency, which reflects the stability of the space accuracy between adjacent frames. It's defined as follow:

$$TC(t) = 1 - |SA(t) - SA(t-1)|, \quad t = 2, \cdots, n \qquad (17)$$

$SA(t)$ is the space accuracy of the t-frame and $SA(t-1)$ is the space accuracy of the t-1 frame.

4.2   Experiment Results

Two typical MPEG-4 video sequences are simulated and analyzed by using the improved algorithm. The hardware environment of the simulation is Genuine Intel(R) T2080 @1.73 GHz CPU, 1 GB memory, and the software platform is Windows XP system, Matlab 7.0.

The hall monitor is a typical indoor surveillance video test sequence. Its background is complex and the speed of video object is quick. Akiyo is a typical indoor sequence and there exist large coherence area in video object and a little movement in the head. The simulation results are shown in Figs. 5 and 6.
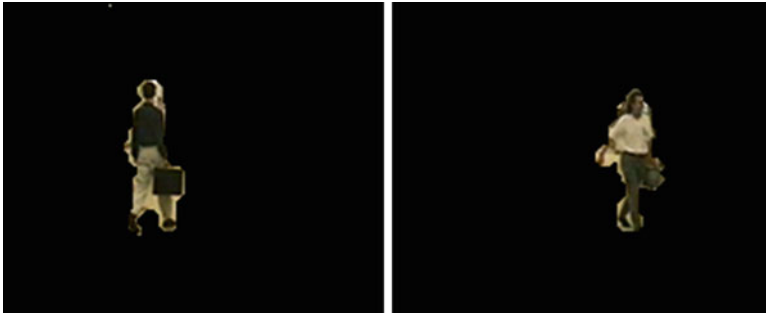
Fig. 5 The segmentation results of Hall Monitor sequence

Fig. 6 The segmentation results of Akiyo sequence

In this paper, we use the Yang's segmentation results as the reference model, and the static threshold is 3. We deal with nine images each time. The evaluation results of the space accuracy and the time consistence of Hall Monitor and Akiyo sequences are shown in Fig. 7. From the evaluation map we can see the space accuracy of the improved algorithm has been enhanced and the time consistence of the two algorithms show little difference. Therefore, the improved method can effectively solve the occlusion problem for the sequences with large speed of the video object and the performance of its anti-noise becomes better. So the improved algorithm can be used in both the sequence with complex background noise and in the sequence with high speed.

5   Conclusion

In this paper we propose a new algorithm of background registration. We can obtain the initial background modeling by the front and the back background modeling. This background modeling can overcome effectively the occlusion. Accumulated
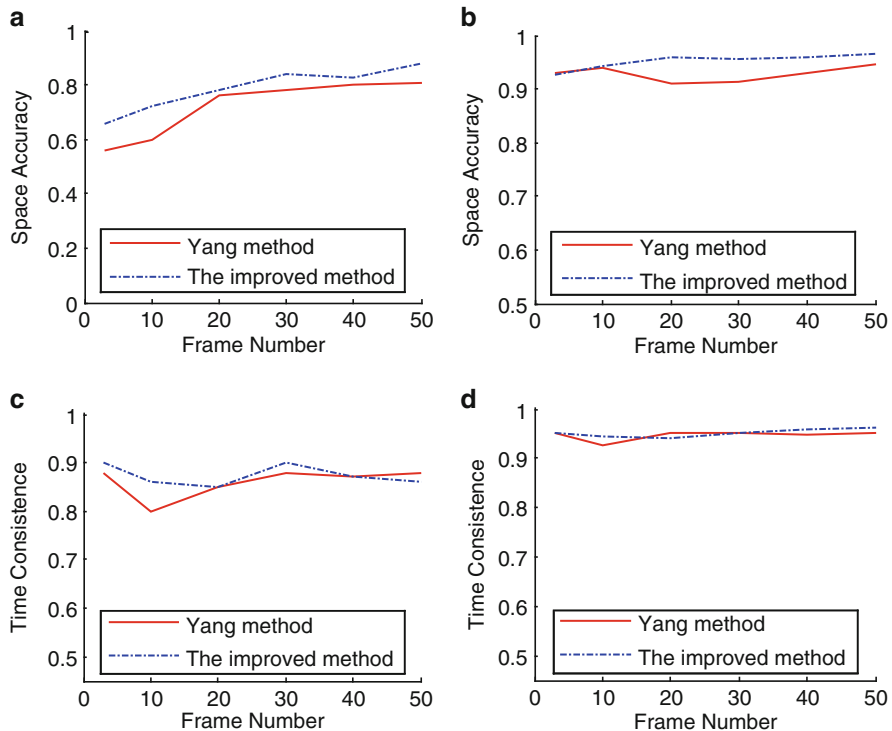
Fig. 7 Evaluation of video sequence segmentation results. (**a**) Hall Monitor space accuracy, (**b**) Akiyo space accuracy, (**c**) Hall Monitor time consistence, (**d**) Akiyo time consistence

frame difference was used to optimize the initializing background modeling to reduce the probability of background that is misjudged in foreground, which makes further improvement on the segmentation results. Finally, we can filter out the noises by morphological filter and morphological opening and closing operation. The experimental results show that this method can effectively solve the occlusion problem which is lead by the large speed of the video object and the performance of its anti-noise is better. This method has accuracy segmentation results and is also simple and easy.

References

Fan Ya-chun, Zhou Ming-quan, Geng Guo-hua (2009) Background cut algorithm leaving illumination infection. J Image Gr 14(7):1413–1417

Foresti GL (1999) Object recognition and tracking for remote video surveillance. IEEE Trans Circuit Syst 9(7):1045–1062

Lin J, Kung SY (2000) Video object extraction and representation: theory and application. Springer, Berlin, p 155

Meier T, Nagan KN (1998) Automatic segmentation of moving object for video object plane generation. IEEE Trans Circuit Syst Video Technol 8(5):525–538

Meier T, Ngun KN (1999) Video segmentation for content-based coding. IEEE Trans Circuit Syst Video Technol 9(8):1190–1203

Nascimento JC, Marques JS (2006) Performance evaluation of object detection algorithms for video surveillance. IEEE Trans Multimed 8(4):761–774

Porikli F (2004) Real-time video segmentation for MPEG encoded video sequences. SPIE conference on real-time image VIII, San Jose: [s. n.], vol 5297 pp 195–203

Yang Gao-bo, Zhang Zhao-yang (2003) An efficient video object segmentation algorithm based on background registration and change detection. J Shanghai Univ Nat Sci Edit 9(2):95–99

Yang G, Zhang Z (2004) Objective performance evaluation of video object segmentation algorithms with ground truth. J Shanghai Univ 8(1):71–74

Yang Gao-bo, Zhang Zhao-yang (2005) Automatic video object extraction algorithm based on wavelet and watershed. J Commun 26(3):7–13

Zhang Ze-xu, Cui Ping-yuan (2008) A reliable method of image registration based on optical flow field and feature extraction. Chin J Electron 17(1):12–15

Zhao Yan-ling, Zhang Zhi-chao, Gao Zhen-ming (2004) Simple and workable moving objects segmentation algorithm. Infrared Laser Eng 33(6):611–614

Construction and Study of Finite Element Model of Thoracolumbar Spine Based on CT Images

Shaobin Li, Xingmin Wang, Wenxiu Fu, and Zhongtian Man

Abstract   Directly reading 228-layer 0.75-mm-thick consecutive CT Dicom format images of human body thoracolumbar into the Mimics, and then the three-dimensional geometric surface mesh model was acquired after the redundancy data was cleaned. Then the surface mesh model could be introduced into Ansys, and volumetric mesh in Ansys. Then the volumetric mesh was introduced into mimics to assign material, and the three-dimension finite element model was generated in Ansys. The triangles with low quality could be repaired in the mimics. This method is rapid and high accuracy. The three-dimension finite element model is very helpful to study biomechanics.

Keywords   CT image • Mimics • Spine • Ansys • Finite element model

1   Introduction

In recent years, with the maturing of rapid development of computer and image processing technology, three-dimension reconstruction of medical image has become a hotspot. Through the software of three-dimensional finite element we can generate the finite element model with realistic appearance, structural integrity, good geometric similarity and high simulation, which play an important role in surgical simulation (He Ya-qi et al. 2008; Li Xiao-lin 2009), minimally invasive treatment, diagnosis and medical education and also have a high research value in clinical medicine (Xiao-lin Li 2009) and biomechanics (Bao Chun-yu and Liu Jin-tao 2009; Bao Chun-yu and Meng Qing-hua 2009). In this paper we use Mimics and Ansys to reconstruct the three-dimension finite element model of thoracolumbar vertebral spine more quickly based on the CT image, and solve difficult problems common

S. Li (✉) • X. Wang • W. Fu • Z. Man
School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China
e-mail: shbli@bjtu.edu.cn; 09120308@bjtu.edu.cn; wxfu@bjtu.edu.cn; 09120376@bjtu.edu.cn

in the modeling process, which laid the foundation for the subsequent rapid and accurate establishing of three-dimensional finite element model and the study of biomechanics.

2 Data

Data source: An adult female volunteer. We use spiral CT to continuous scan subjects lumbar spine, slice thickness is 0.75 mm, the total number of layers is 228, the scanning data were stored in dicom format.

Software: Mimics10.01 (Belgium, Materialise Company) and Ansys11.0 (U.S., Ansys company).

3 Generate Surface Mesh Model

3.1 Data Input

The 228 CT dicom format images were directly read into Mimics, then we adjust the appropriate window width and position, and position the image.

3.2 Generation and Optimization of the Surface Mesh Model

Defined threshold range which is $226 \sim 1{,}371$ HounsField units, and use regional growth to generate the original model of lumbar spine. We must fill holes with each CT image, and then three-dimensional finite element model of lumbar spine can be generated after three-dimension calculation. Figure 1 shows the original model of lumbar spine.

Then we can use migics9.9 to optimize the surface mesh model. The quality and density of mesh determine the accuracy of the final model. Mesh optimization is very important in the entire generation of the three-dimensional modeling, including smoothing, reducing the number of grids, improve the quality of grid, reduce the number of grid under the premise of guaranteeing quality. If there is still a small amount of low-quality triangles, we can manually remove low-quality triangles, and add high-quality triangles. Finally, all mesh have reached high quality. Figure 2 shows the optimized surface mesh model and the quality histogram.

From the above chart we can see that the optimized surface mesh model quality generally high, we can achieve the desired results using the above methods to optimize.
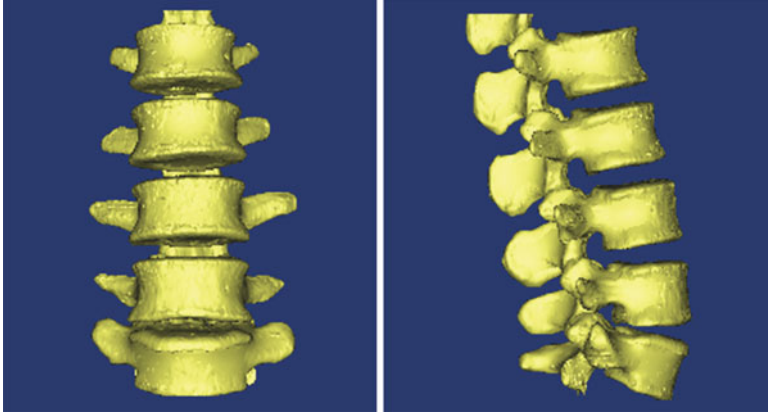
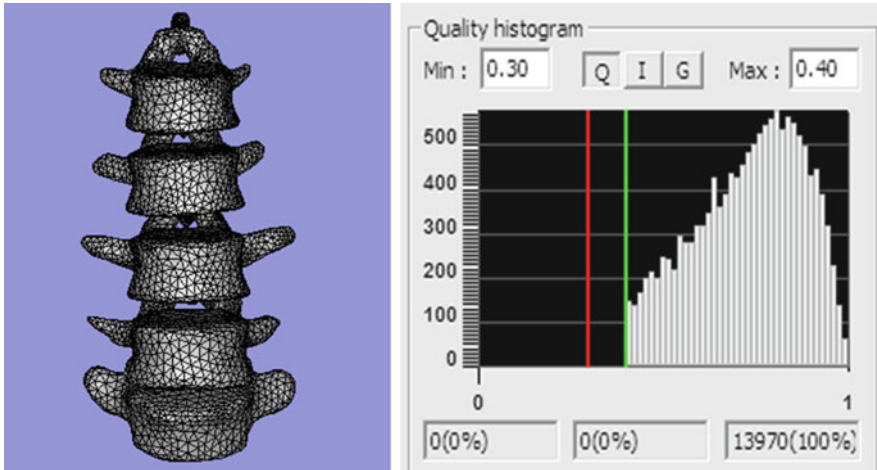**Fig. 1** The original model of lumbar spine

Fig. 2 The optimized surface mesh model and the quality histogram

Finally, after the triangle intersection test we can see the intersecting triangles have disappeared. If there is still a small amount of cross-cutting of the triangles, they can be handled manually. Then we get the surface mesh model which meets the requirements.

The optimization surface mesh model will be output and save as the suffix .lis file.

4 Generate the Volume Mesh Model

4.1 Divide the Volume Mesh Model

The .lis file which was generated by the previous step is read into Ansys, after distribute the property of the unit we can divide the volume mesh model.

The type of the unit is defined as Solid92. The unit has a secondary displacement characteristic and is three-dimension 10-node tetrahedron elements and is adequate for the division of irregular geometric solid.

Then we can get the nodes of the volume mesh model, the unit file, and according to the nodes and the unit file, we can write the Prep7 file. This model comprises a total of 110,464 nodes, 70,201 tetrahedron elements.

4.2 Material Assignment

The Prep7 file is lead into Mimics, then we use the empirical formula (Rho et al. 1995; Kopperdahl et al. 2002) and according to the CT value to define the elastic modulus and Poisson's ratio of the thoracolumbar spine, defined as 10.

4.3 Generate the Volumetric Mesh Model

After assign material we lead the model into Ansys, and then the three-dimension finite element model of thoracolumbar spine will be generated. As shown in Fig. 3, respectively, thoracolumbar spine model for the front and side shots.

5 Problems and Solutions During Generate the Model

After generate the surface mesh in Mimics, there always have error message when we lead the Lis file into Ansys, as shown in Fig. 4.

View from the literal meaning, the meaning of the error was "The Jacobian of the 2347th triangular element is zero or negative. If it contains the middle node, it is likely the location of middle node is too poor."

Although mesh optimization has been carried out in Mimics, and mesh quality requirements are met, but that is only in terms of triangles. The actual meaning is there have convex and concave surfaces with a mesh block. As shown in Fig. 5.

This error is hard to find in Mimics, because it is not easy to observe the error from the surface of the surface mesh, we must remove some surface grid and then we can see the error. The solution is to find the corresponding error position according
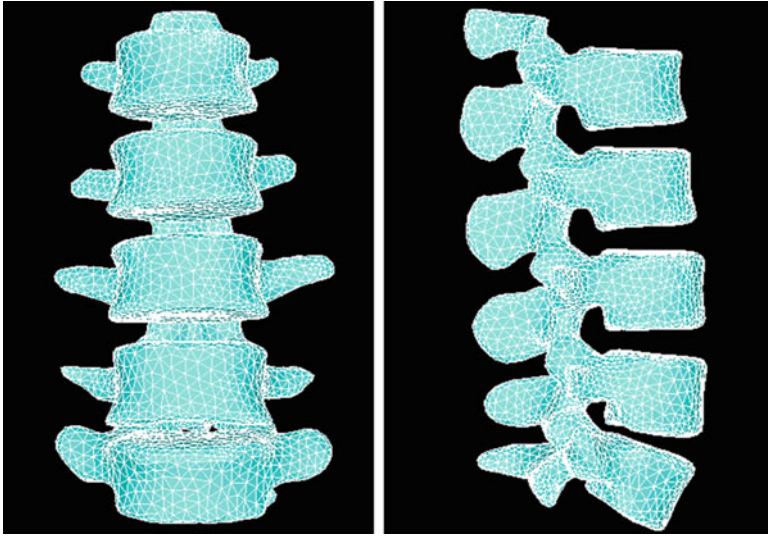
Fig. 3 Three-dimensional mesh model

Fig. 4 Recurrent error message
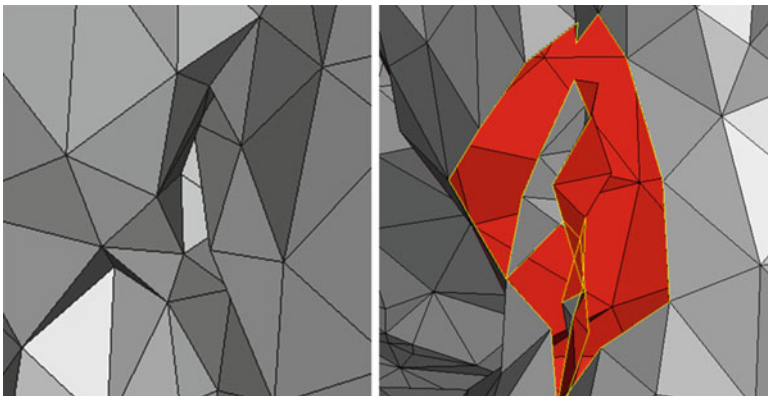

Fig. 5 Part of the uncorrected surface mesh

Fig. 6 Part of the corrected surface mesh
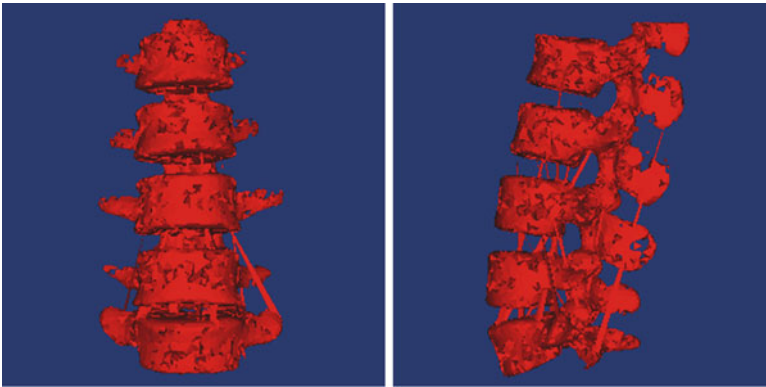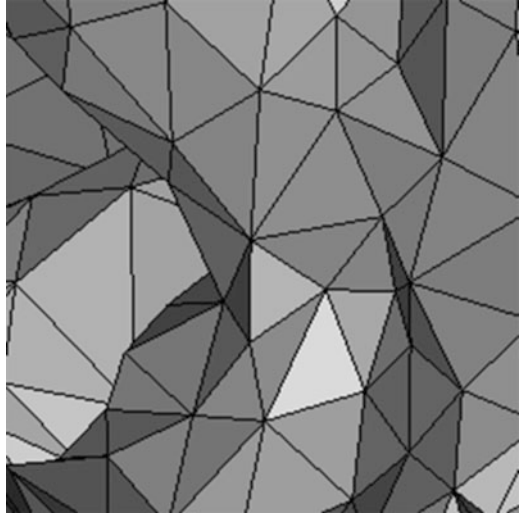
Fig. 7 The error message in Mimics

to the unit number which is get from the error message after leading a good surface mesh model into Ansys, and then manually corrected in Mimics. The correction mesh is shown in Fig. 6.

After generate the volumetric mesh model in Ansys, there perhaps some errors when we lead the Prep7 file into Mimics as shown in Fig. 7.

This is because the type of the unit file which was written in Ansys is wrong. When the number of the node is more than 99,999, the unit should be prepared for long file type. This is an easily overlooked error.

6 Conclusion

In this paper, we use CT image data to generate the three-dimension finite element model of the human thoracolumbar spine, the speed is high and the model is accuracy. From Fig. 2 we can see the quality of most of the surface mesh model is close to one which almost close to perfect, using the above methods to generate the surface mesh model lay a good foundation for the three-dimension finite element model. From Fig. 3 we can see, the model looks realistic and have structural integrity, good geometric similarity and high simulation. This model can realistic simulate the various movements and their structure and material properties of the spine, and it is stress analysis also, it establish a good experimental platform for biological mechanics and clinical medicine.

References

Bao Chunyu, Liu Jintao (2009) Lumbar spine subject stress characteristic for operators using chain saw in forest harvesting and finite element analysis. Sci Silvae Sinicae 45(3):96–100

Bao Chunyu, Meng Qinghua (2009) The mechanics characteristic of lumbar spine for weightlifter with snatch: based on finite element theory. J Tianjin Univ Sport 24(1):42–44

He Ya-qi, Zhang Xue-lin, Tang Bing-hang (2008) Construction of three-dimensional finite element model of hipbone based on DICOM data. J Huazhong Univ Sci Technol Health Sci 37(2): 251–254

Kopperdahl DL, Morgan EF, Keaveny TM (2002) Quantitative computed tomography estimates of the mechanical propeties of human vertebral trabecular bone. J Orthop Res 20(4):801–805

Li Xiao-lin (2009) Method and clinical significance of reconstructing a three-dimensional finite element model of human thoracolumbar spine based on sophisticated CT scanning. Shandong Med J 49(14):8–10

Rho JY, Hobatho MC, Ashman RB (1995) Relations of mechanical properties to density and CT numbers in human bone. Med Eng Phys 17(5):347–355

Xiao-lin L (2009) Reconstructing a bone finite element model of human thoracolumbar spine based on CT images by using Mimics software. J Clin Rehabil Tissue Eng Res 13(29):7619–7622

MLB-Wrapper: Distributed High Scalable BP-Wrapper

Yongrui Xu and Yongguo Han

Abstract In distributed system, concurrent accesses lead to lock contention which greatly increases response time. Cache can reduce response time, but cannot gain scalability. Currently, the common solution is to modify the caching algorithm itself, although the scalability can be improved, the hit ratio hurt. In this paper, we propose a system framework, called MLB-Wrapper, which resolves the scalability issue in distributed systems. The framework does not modify the cache algorithm itself. Therefore, it preserves high cache hit ratio about the replacement algorithms. Multi-levels distributed cache is used in MLB-Wrapper which adopts batching technique in every level to eliminate the lock contention. Furthermore, MLB-Wrapper is flexible because different replacement algorithms can be used in each level. We describe the main idea and the detailed implementation of MLB-Wrapper in this paper. Experiment results show that, in the intensive concurrent environment, MLB-Wrapper can not only reduce lock contention but also have higher scalability.

Keywords Distributed systems • Lock contention • Multi-level cache • High scalability

1 Introduction

Nowadays, with the continuous rapid development of various applications, the data which need to be managed are sharply increasing. Meanwhile, the system faces continuous growth in the number of users and a large number of concurrent accesses. Many persistent storage systems of applications are becoming overburdened, response deteriorated, and other major performance issues arisen, especially the

---

Y. Xu (✉) • Y. Han
Department of Computer Science, Southwest University of Science and Technology,
Mian Yang, China
e-mail: wowwowxyr@126.com; hansir@swust.edu.cn

SNS (Social Network System) applications such as Facebook and Twitter. They need to store hundreds of PB's data and face millions of concurrent accesses per second. In order to solve the scalability issue, Google designed Percolator (Peng and Dabek 2010), MapReduce (Dean and Ghemawat 2004), Microsoft and Yahoo! also designed their own platforms such as Dryad (Isard et al. 2007), (Yu et al. 2008), S4 (Neumeyer et al. 2010). On the other hand, for the mass data storage management, there are lots of researches which have been made (e.g (Baker et al. 2011); (Giuseppe et al. 2007); (Ghemawat et al. 2003); (Weil et al. 2007); (Chang et al. 2008); (Cooper et al. 2008)). Recently, Stanford University proposed a new idea named RAMClouds (John et al. 2010) which is based on the Cloud technology to solve the mass data storage management issue.

Although these methods can solve the scalability and mass data storage management issue, most of which rely on the cheap disk devices which have expensive cost in operation. In order to reduce data access delay, they need a high hit ratio replacement algorithm to construct the caching system. So far, there are lots of researches which have been made about replacement algorithms, from LRU to 2Q (Johnson and Shasha 1994), ARC (Megiddo and Modha 2003) and LIRS (Jiang and Zhang 2002). But, none of them can solve the lock contention issue. It is just like the "Single point of failure" in a distributed system and an unscalable caching system limits the scalability of the whole system.

In order to solve this issue, many researches focus on improving the replacement algorithm. For example, modifying the LRU replacement algorithm to the Clock (Corbato 1969), ARC to the CAR (Bansal and Modha 2004) and LIRS to the Clock-Pro (Jiang et al. 2005) algorithm. The lock contention issue can be resolved by these proposals, but all of them hurt the hit ratio of original algorithms.

For this reason, some researches did not focus on modifying the algorithm, but on adding additional layers to handle the scalability issue. One of them is BP-Wrapper (Ding et al. 2009). BP-Wrapper not only preserves the high hit ratio of original replacement algorithm but also separates the strategy of concurrent control with implement of the algorithm itself.

Every coin has two sides. BP-Wrapper is no exception. It was designed for the traditional DBMS buffer management and there is only one centralized cache in traditional DBMS. Many distributed systems in real world have lots of decentralized cache, so BP-Wrapper is not suitable for them.

In this paper, we propose a system framework named MLB-Wrapper which can solve the scalability issue of caching system in the intensive concurrent environment. As same as BP-Wrapper, MLB-Wrapper also does not modify the replacement algorithm. So, it not only preserves the high hit ratio of original replacement algorithm but also is suitable for almost any replacement algorithms like BP-Wrapper. Compared with BP-Wrapper, MLB-Wrapper divides cache into multi-level and in every level we use batching technology to reduce lock contention. That's why we named it MLB-Wrapper(Multilevel Batching-Wrapper). The multi-level cache further reduces the lock contention and it also greatly reduces the communication frequency between client and server. Less communication frequency means much less lock contention and it is an important feature to improve

system scalability in the intensive concurrent environment. On the other hand, the design of multi-level cache makes the caching system more flexibility. The experimental results indicate that the proposed framework is able to reduce 30% access latency compared with BP-Wrapper in the intensive concurrent environment.

The rest of this paper is organized as follows. Section 2 describes the background knowledge of the lock contention. Section 3 describes the main idea and the implementation details of the MLB-Wrapper. Section 4 evaluates the proposed system framework. Section 5 discusses related work and Sect. 6 is the conclusion and future work.

2   Background

In DBMS memory address space, there are lots of fixed size pages which need to be accessed by many concurrent transactions. Due to the physical implementation of DBMS, these pages are stored in the I/O devices. In order to reduce the access latency, DBMS must cache these pages. At the same time, the consistency of cached pages must be ensured. In traditional way, when a transaction accesses the cached pages it must lock them to prevent other concurrent transactions from accessing the same pages. In this way, every request must execute their operations in serial ways and all of them cannot be paralleled. It greatly limits the scalability of the system.

So, in order to improve the scalability of the DBMS buffer manger, Xiaoning Ding etc. proposed the system framework BP-Wrapper in 2009. In this framework, they provided two key scalability supports, one is batch execution and the other is prefetching. That's why the framework named BP-Wrapper. In fact, the benefit of prefetching technology is limited. When a page is prefetched before actually used, during this time, other request updates this page, the page has been stale for the original request. So, it must do the extra invalidating operations. In many systems, especially in distributed systems, this operation is much more expensive than lock operation. In Fig. 1, we only show the batch execution technology of BP-Wrapper.

In BP-Wrapper, every concurrent transaction has an extra corresponding FIFO which records the request information. When cache hits, append current request information into the FIFO which belongs to the current transaction and return the requested page. Otherwise, request must wait until acquiring the lock, then, uses the corresponding FIFO's history informations to update the cache in a batching fashion. In this manner, only cache misses we need to lock the replacement algorithm and lock only once.

As we said before, BP-Wrapper is well suitable for solving the scalability issue about DBMS buffer management. But, the centralized design such as the hash table is not suitable for distributed systems. Although we can have huge buckets in the hash table to avoid unnecessary lock contention, that's not enough in intensive concurrent environment. Besides this, there is one important fact that has been left out of consideration. In most transactions internal, there are some pages that are "hot spots", but between different transactions, the requested pages are often
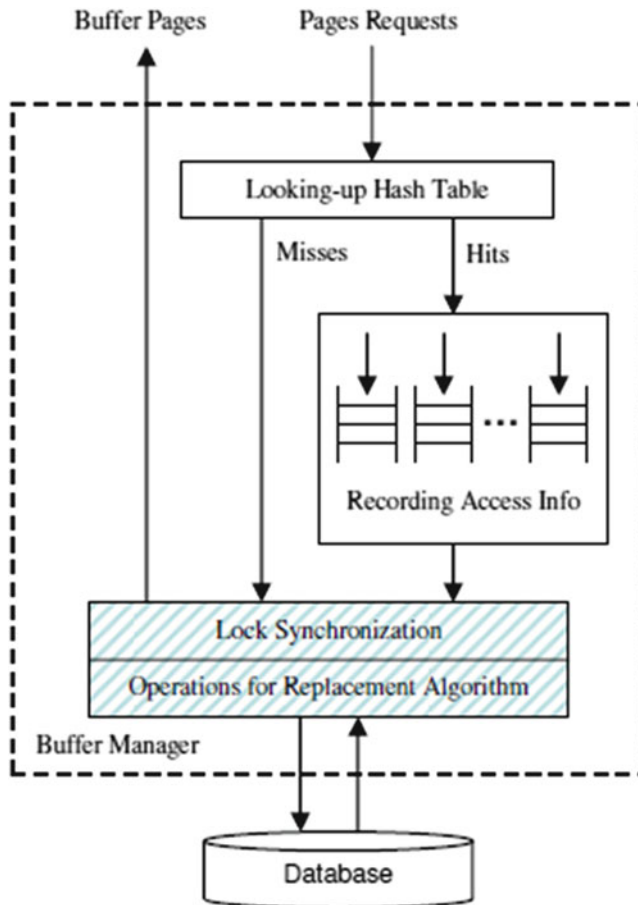
Fig. 1 Batching technique in BP-Wrapper (Ding et al. 2009)

irrelevant. So, if we insist on centralized design, the caching system may have higher probability of occurrence about the phenomenon of "Thrashing (Denning 1968b)". In a word, single centralized cache with BP-Wrapper is not enough. In order to further improve the scalability, we must turn to distributed solution.

3 Description of MLB-Wrapper

As above said, In order to make caching system get higher scalability with lower latency in the intensive concurrent environment, we propose the MLB-Wrapper framework. Figure 2 shows the architecture of it.

Fig. 2 The architecture of MLB-Wrapper

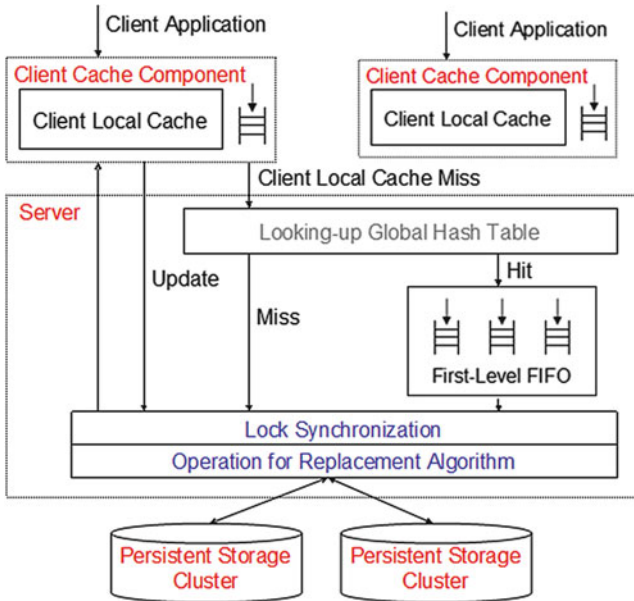Contrast Fig. 1 with Fig. 2, the biggest difference between BP-Wrapper and MLB-Wrapper is that, in MLB-Wrapper, we distribute the cache in the client and the server, which are called client-side wrapper and server-side wrapper. In the client-side wrapper, including a client own small capacity cache which is independent of server cache and a client FIFO which records request information that corresponds to the client. We name the client FIFO in the client-side wrapper as Second Level FIFO, meanwhile, we name the FIFO in the server-side wrapper corresponding to the client as First Level FIFO. When the client application accesses the client own cache, there is no lock contention. More importantly, because this cache is relevant to the client application, it can avoid "Thrashing" to the fullest extent. Only when the Second Level FIFO is full, the number of access information reaches a pre-determined threshold or the client own cache misses, it needs to communicate with server and update the server cache. Otherwise, there is neither any lock contention nor communication between client and server. The First Level FIFO in the server-side wrapper is used when the client-side cache misses. The details of the MLB-Wrapper can be seen in the section of "A Detailed Implementation Description".

In real world, we not only use MLB-Wrapper in the client/server scene, but also use MLB-Wrapper to construct our cache server cluster. In cache server cluster scene, some cache servers play the client role and other ones play the server role. We use an example to explain this idea. In this example, we assume that the total data range is from 0 to 299 and in every server, the capacity of the server can only cache

50 items. In our design, one cache server caches data range from 0 to 149 which is named server 0, the other one caches data range from 150 to 299 which is named server 1. They both play the client role in MLB-Wrapper. We have the third server which is named server 2 and it caches data range from 0 to 299. Actually, the most common design way is every cache server caches 100 data range, we call this design as "flat design". In flat design, if the server corresponding to the requested data cache misses, it has to face the expensive cost about cache misses. In our design, though the data range that every server must face is increased, due to the FIFO in every level which records the request information, we can make maximum use of every cache server. In other words, if we request data between 0 and 99 frequently, in flat design, server 1 and server 2 are idle, but only using server 0 we cannot get high cache hit ratio. On the other hand, compared with the centralized design of original BP-Wrapper, we can get much better scalability because we can easily scale-out the cache server.

3.1 The Benefits of the MLB-Wrapper Framework

Compared with the centralized design of original BP-Wrapper, there are many benefits in the MLB-Wrapper, These benefits include, but are not limited to:

- Significantly reducing lock contention about the hash table.
  Because of the client-side wrapper, the client application will access the server-side hash table only when the client own cache misses. The access frequency of the hash table dropped significantly that leads to much less lock contention about the hash table.
- Greatly reducing the client-side and the server-side communications.
  In the centralized design of original BP-Wrapper, the cache is only at the server. So, every client request must communicate with server. Even if the communication way is shared memory, such as the buffer management in DBMS, the communication cost is not cheap. Because of the client-side cache in MLB-Wrapper, the numbers of communication are greatly reduced.
- The flexibility of the client-side cache.
  Thanks to the independence of the server-side cache, the client-side cache can use different replacement algorithms. It also can be different between different clients. But, due to the Second Level FIFO in the client-side wrapper, the history information is not lost and does not affect server-side cache hit ratio. HTML5 (http://www.w3.org/TR/html5/) is the future of the web client, the one of the most important feature in HTML5 is "Web Storage" (http://dev.w3.org/html5/webstorage/). In MLB-Wrapper, we can use SQLite (http://www.sqlite.org/) to implement the "Web Storage" and make it as the client-side cache, even accelerate the cache with CUDA (Bakkum and Skadron 2010a), (b). In MLB-Wrapper, nothing is impossible, because of the flexibility of the client-side cache.

3.2 A Detailed Implementation Description

In this section, we will detail the MLB-Wrapper work flow.

First of all, the client application accesses the client-side cache, if hits, returns the resource to the application, updates the client-side cache and appends the request information to the Second Level FIFO. If the number of access information in the Second Level FIFO reaches a pre-determined threshold, try to acquire the lock of the server-side cache, if succeeds, execute the operations defined by the replacement algorithm once for all the accesses in the First Level FIFO in a batching fashion and then in the Second Level FIFO; if the number of access information in the Second Level FIFO reaches the maximum capacity, the client must wait until it acquires the lock and updates the cache with the First Level FIFO and then Second Level FIFO in a batching fashion. When we update the cache, we must use the First Level FIFO first and then Second Level FIFO, that's because the accesses in the First Level FIFO are always before the ones in the Second Level FIFO. As said above, not difficult to know that the capacity of client-side cache, the maximum capacity of the FIFOs and the pre-determined threshold of the two FIFOs can decide the performance of the MLB-Wrapper.

Otherwise, the client-side cache misses. At this time, if the available capacity of the First Level FIFO in the server-side corresponding to the client is enough, we move the records from Second Level FIFO to the First Level FIFO, if it is not enough, we first do the update operation with First Level FIFO and then move the records. The update operation must acquire the lock, but in real world, the capacity of the First Level FIFO in the server-side is much larger than the Second Level FIFO in the client-side, so this situation is rare. After moving the records from Second Level FIFO to the First Level FIFO, it then looks up the hash table of the server-side to see whether the resource is in the server cache. There are two cases: (1) server cache hits. This time, it returns the resource to the client and then updates the client own cache. Meanwhile, the server asynchronous checks the First Level FIFO corresponding to the client to see whether the FIFO reaches the threshold or the maximum capacity and decides to update server cache or not. The client doesn't need to wait for the update operation finished, so the latency is very low. (2) server cache misses. This time, the operation is the same as BP-Wrapper when the cache misses. The only difference is that we must update the client-side cache with the resource at last.

4 Performance Evaluation

4.1 Experimental Method and Settings

In order to accurately verify the MLB-Wrapper that has the less lock contention than the BP-Wrapper, we use one process to simulate the server-side wrapper and
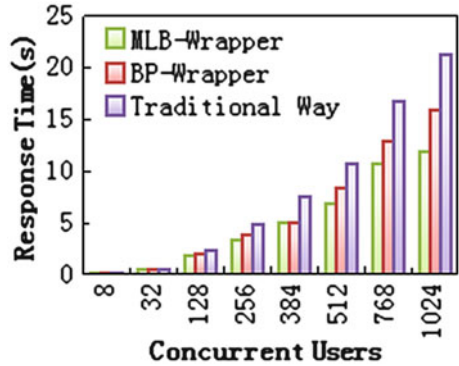
use lots of concurrent threads in this process to simulate the client-side wrapper.
We use every thread's TLS(Thread Local Storage) to construct the client-side cache
and the Second Level FIFO. In our experiment, one thread stands for one user. The
reason why we do this is that, if the client-side wrapper and the server-side wrapper
live in the different node or live in the same node but not in the same process, the
expensive communication costs between them will dominate the results. If we use
one process and multithreads, the communication cost is very cheap because all
the things are in the same memory address space. This time, different results are
primarily caused by the lock contention, the less time means less lock contention.

So, we wrote C++ code to verify our MLB-Wrapper framework. The server
cache replacement algorithm was LIRS which has high hit ratio. In our code, we
can adjust the capacity of the cache and the ratio of LIR. The First Level FIFO and
the Second Level FIFO we can also adjust the capacity and the threshold. In every
client own cache, we use LRU replacement algorithm. We set the number of the
total resource is 1,000 and every client will access 2,000 times. If there is no special
description, the capacity of server cache is 200 which is 20% of the total resource,
the client cache is 40, the capacity and the threshold of the First Level FIFO is 40
and 20, 10 and 5 in the Second Level FIFO.

4.2   Experiments

- The average response time of clients finishing their requests.
  According to Fig. 3, it is not difficult to find that, the more concurrent users
  the more lock contention. In each case, the average response time of MLB-
  Wrapper is the lowest. With the more concurrent users, the slowest growth rate
  about response time means MLB-Wrapper has the best scalability in the large
  concurrent environment.
- Cache misses with different concurrent users.
  According to Fig. 4, we think batching executing can avoid the cache "Thrash-
  ing". So, MLB-Wrapper can also get better cache hit ratio than the traditional

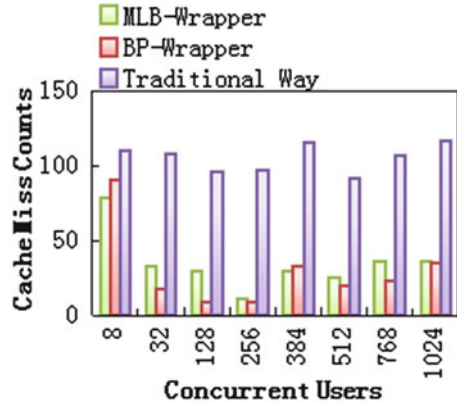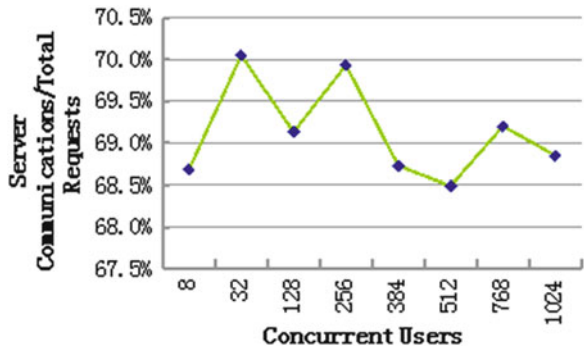**Fig. 4** Average cache misses with different users

Fig. 5 The ratio is the number of requests which need to communicate with the server divided by the total number of requests. At this time, we use the same capacity of the client-side cache with different concurrent users

way which doesn't use batching technology. The number of current users has little influence on the cache hit ratio except the number is too small. We believe that, the number of users within a certain range, the more users means the hot resource is accessed more frequently and it can promote the cache hit ratio (Fig. 4).

- Communications between client-side wrapper and server-side wrapper in the MLB-Wrapper.

Based on intuition, the more capacity of client-side cache, the less communications the client need to do with server. Figure 6 confirms it. But that is not always true. If the capacity exceeds out of a range, little advantage comes from larger client-side cache. On the other hand, we use the default capacity of the client-side cache, and observe the results with different concurrent users. According to Fig. 5, it proves that there is nearly no influence between the number of concurrent users and the ratio which is the number of requests that need to communicate with the server divided by the total number of requests.

Fig. 6 The ratio is the same as Fig. 5. We can see that with the more capacity of client-side cache, the ratio is becoming lower
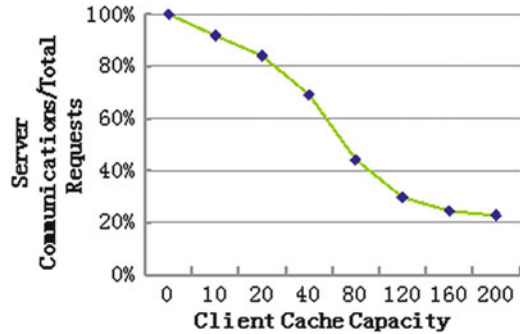
5 Related Work

In the software version management systems such as SVN (http://svnbook.red-bean.com/en/1.5/svn-book.pdf), in order to reduce the lock contention, they adopt optimistic concurrency control (Kung and Robinson 1981) lock strategy. In some NoSQL storage system such as BigTable, they use granular lock to reduce lock contention. For example, BigTable only supports the single-row transactions, not general transactions across row keys. Another way is reducing the lock holding time. In the kernel 2.6 on Linux, the interrupt subsystem divides the hardware interrupts into the "Top Halves" and "Bottom Halves" parts. In the "Top Halves", the kernel is locked and cannot response other hardware interrupts. The kernel defers the unnecessary work in the hardware interrupt to the "Bottom Halves". So it reduces the lock holding time and make the kernel response other hardware interrupts in time.

6 Conclusion and Future Work

In this paper, we proposed the MLB-Wrapper to solve the scalable issue in the intensive concurrent environment. In order to further reduce the lock contention, we add another client-side wrapper between client application and the server-side wrapper. On the other hand, Like BP-Wrapper, the framework does not modify the cache algorithm itself. Therefore, it preserves the advantage of the high cache hit ratio about the algorithm itself and also can be applied to almost all the cache replacement algorithms. What's more, our framework can greatly reduce the communications between client applications and the server, which is one of the most important features in the intensive concurrent environment.

As for future work, we plan to evaluate MLB-Wrapper on larger distributed systems. We will continue to research on how to divide the cache levels that we can have less lock contentions and get better scalability.

References

Baker J, Bond C, Corbett JC, Furman JJ, Khorlin A, Larson J et al (2011) Megastore: providing scalable, highly available storage for interactive services. In: 5th Biennial Conference on Innovative Data Systems Research (CIDR'11), Monterey, CA, pp 223–234

Bakkum P, Skadron K (2010a) Accelerating SQL database operations on a GPU with CUDA. In: GPGPU'10: proceedings of the third workshop on general-purpose computation on graphics processing units, New York, NY, USA, pp 94–103

Bakkum P, Skadron K (2010b) Accelerating SQL database operations on a GPU with CUDA: extended results, University of Virginia Department of Computer Science technical report CS-2010-08, May 2010

Bansal S, Modha DS (2004) CAR: clock with adaptive replacement. In: FAST'04 proceedings of the 3rd USENIX conference on file and storage technologies, San Francisco, CA

Chang F, Dean J, Ghemawat S, Hsieh WC, Wallach DA, Burrows M (2008) Bigtable: a distributed storage system for structured data. J ACM Trans Comput Sys (TOCS) 26(2):1–26

Cooper BF, Ramakrishnan R, Srivastava U, Silberstein A, Bohannon P, Jacobsen H et al (2008) PNUTS: Yahoo!'s hosted data serving platform, VLDB'08, Auckland, New Zealand, 24–30 Aug 2008

Corbato FJ (1969) A paging experiment with the multics system. In: In Honor of Philip M M, Feshbach, Ingard (eds) MIT Press, Cambridge, MA, pp 217

Dean J, Ghemawat S (2004) MapReduce: simplified data processing on large clusters. In: OSDI'04: sixth symposium on operating system design and implementation, San Francisco, CA, December 2004

Denning PJ (1968b) Thrashing: its causes and preventnon. In: Proceedings of the AFIPS national computer conference AFIPS, San Francisco, California, pp 915–922

Ding X, Jiang S, Zhang X (2009) BP-Wrapper: a system framework making any replacement algorithms (almost) lock contention free. In: ICDE'09, pp 369–380

Ghemawat S, Gobioff H, Leung ST (2003) The google file system. In: SOSP'03 proceedings of the nineteenth ACM symposium on operating systems principles, Bolton Landing, NY, 19–22 Oct 2003

Giuseppe DeCandia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati, Avinash Lakshman, Alex Pilchin et al (2007) Dynamo: Amazon's highly available key-value store. In: SOSP'07, Stevenson, WA, pp 205–220, 14–17 Oct 2007

HTML5 http://www.w3.org/TR/html5/. Accessed 13 Jan 2011

Isard M, Budiu M, Yu Y, Birrell A, Fetterly D (2007) Dryad: distributed data-parallel programs from sequential building blocks. In: Proceedings of the 2nd European conference on computer aystems (EuroSys), Lisbon, Portugal, pp 59−72

Jiang S, Zhang X (2002) LIRS: an efficient low interreference recency set replacement policy to improve buffer cache performance. In: Proceedings of the 2002 ACM SIGMETRICS conference on measurement and modeling of computer systems, (SIIMETRICS'02), Marina Del Rey, CA, 15–19 June 2002

Jiang S, Chen F, Zhang X (2005) CLOCK-Pro: an effective improvement of the CLOCK replacement. In: Proceedings of the annual USENIX technical conference'05, Anaheim, CA

John O, Parag A, David E, Christos K, Jacob L, David M et al (2010) The case for RAMClouds: scalable high-performance storage entirely in DRAM. ACM SIGOPS Op Sys Rev 43(4): 92–105

Johnson T, Shasha D (1994) 2Q: a low overhead high performance buffer management replacement algorithm. In: Proceedings of the 20th VLDB conference, Santiago, Chile, pp 439–450

Kung HT, Robinson JT (1981) On optimistic methods for concurrency control. ACM Trans Database Syst 6(2):213–226

Megiddo N, Modha DS (2003) ARC: a SELF-TUNING, LOWOVERHEAD REPLACEMENT CACHE. In: Proceedings of 2nd USENIX conference on file and storage technologies, San Francisco, CA, pp 115–130

Neumeyer L, Robbins B, Nair A, Kesari A (2010) S4: distributed stream computing platform. In: ICDMW IEEE international conference on data mining workshops, Sydney, Australia, pp 170–177

Peng D, Dabek F (2010) Large-scale incremental processing using distributed transactions and notifications. In: Proceedings of the 9th USENIX symposium on operating systems design and implementation, Vancouver, BC, Canada

SQLite http://www.sqlite.org/

SVN http://svnbook.red-bean.com/en/1.5/svn-book.pdf

Web Storage http://dev.w3.org/html5/webstorage/

Weil SA, Brandt SA, Miller EL, Long DDE, Maltzahn C (2007) Ceph: a scalable, high-performance distributed file system. In: OSDI'06 Proceedings of the 7th symposium on operating systems design and implementation, Berkeley, CA

Yu Y, Isard M, Fetterly D, Budiu M, Erlingsson Ú, Gunda PK et al (2008) DryadLINQ: a system for general-purpose distributed data-parallel computing using a high-level language. In: OSDI'08 Proceedings of the 8th USENIX conference on operating systems design and implementation, Reykjavík, San Diego, CA

The Hardening Characteristic Spectrum Analysis of Vascular Wall Elasticity Noninvasive Dynamic Information

Geng Haofei, Gan Ping, Han Sheng, Huai Xian, Chen Lili, Deng Shixiong, and Xiong Xingliang

Abstract In the analysis of vascular wall dynamic information, typical oscillogram of pressure displacement wave of vascular wall (PDWVW) and speed wave information were gained from normal and patient groups, which were based on analysis theory of power spectrum. The analyzed results showed that the parameter G and H of signal A and D not only had statistical significance, but also were proportional to the elasticity of the vessel wall. So the two signals could be considered as the main characteristic parameters in quantitative diagnosis of the vascular walls elasticity diseases. It may have some clinical significance in quantitative diagnosis of early vascular diseases.

Keywords Vascular wall dynamic information • Power spectrum analysis • Degree of elasticity

1 Introduction

The occurrence of angiocardiopathy often closely contacts with human vascular wall elasticity and geometric configuration. The geometric configuration of the vascular and mechanical elasticity will have some change, when the pathological process takes place. When angiocardiopathy is diagnosed such as hypertension, atherosclerosis etc., it's a clinical sign that the moderate or serious pathological process of vascular wall has taken place. It is one of the challenges at present in the clinical diagnosis how to early detect and diagnose the pathological process of vascular wall elasticity and geometric configuration as in Chen et al. (2008).

G. Haofei • G. Ping • H. Sheng • H. Xian • C. Lili • D. Shixiong • X. Xingliang (✉)
Laboratory of Forensic Medicine and Biomedical Information, Chongqing Medical University, Chongqing, China
e-mail: ghf029@foxmail.com; xxlsober@sina.com

The human vascular wall dynamic information (VWDI) is the important parameter of medical information which relates to kinestate and elasticity of human vascular wall. It mainly contains pressure displacement wave of vascular wall (PDWVW), pressure speed wave (PSV), pressure displacement acceleration wave, blood volume wave and so on, which are all noninvasive dynamic medical information as in Gan and Hu (2006).

In order to study biomedical signal deeply, the time domain signal is often transformed into frequency domain, which is used to reveal the implicated clinical significance. The spectrum analysis of the signal PDWVW and PSV have been done by the dynamics information measurement system of vascular wall based on remote virtual instrument developed in the preliminary study as in Chen et al. (2009).

2   Vascular Wall Dynamic Information Spectrum Analysis

Much medical information is implicated in the spectrum of the signal in the biomedical engineering. Consulting a large number of documents, the author has not found that spectrum analysis is used to research elasticity degree of vascular wall. So, a preliminary analysis has been done to dynamic information of the vascular wall based on spectrum analysis theory, in order to find the method which can quantitatively analyze hardening degree of vascular wall.

As can be seen from Fig. 1, multi-channel waveform of dynamic information of vascular wall can be got by the measurement system. The first wave is the signal PDWVW, which would be focused on in the result analysis. The second wave is the blood volume wave, the third the signal ECG and the fourth the signal PCG.
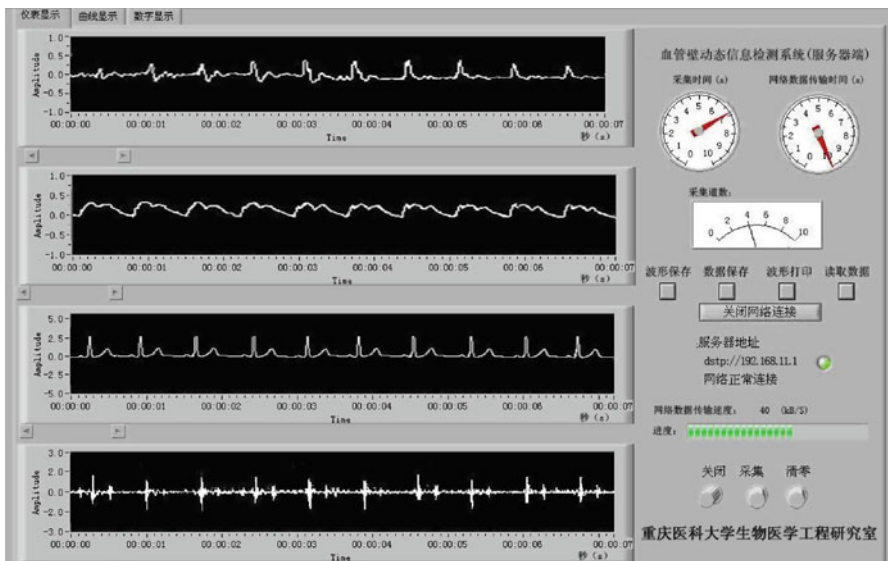

Fig. 1   The interface figure of VWDI measurement system

2.1 Medical Signals Power Spectrum Analysis

The spectrum analysis of medical signal can reveal implicit periodicity of medical signal as well as some useful information. Especially, the acrometron of power spectrum and peak often indicate important medical information of some special diseases in the biomedical engineering.

The power spectrum analysis is used to observe periodic variation in heart rhythm activity, which closely associates with elasticity degree of vascular wall for cardiovascular system. So, the characteristic parameter of medical signal spectrum analysis is the important criterion which estimates cardiovascular system and degree of vascular wall elasticity.

2.2 PDWVW Spectrum Analysis

From the perspective of signal analysis, almost all the actual signal can be expressed by linear non-periodic signal, and if it is decomposed, we will find that the measured PDWVW is composed of the different amplitude and phase harmonic as in Nie (2005), whose frequency is continuous. This non-periodic motion process can be expressed in a non-periodic function. Non-periodic waveform isn't directly changed into Fourier series. However, it can be seen as a periodic signal which has an infinite period. Using amplitude or phase for the ordinate and frequency for the abscissa would receive the spectrum curve of the signal PDWVW. By doing this, the time domain signal is changed into a frequency domain signal.

2.3 Research Processes and Objects

Firstly, using sampling and analog-digital conversion obtains the PDWVW signal. And then intercept the analysis signal by window treatment and do the FFT. Eventually, operate the power spectrum analysis.

The study object is divided into normal group and patient group. The normal group is that vascular disease can be excluded by strictly clinical examination and the patient group is which has been clearly confirmed in patients with vascular disease including primary hypertension, secondary hypertension, atherosclerosis, cardiac insufficiency, premature beat, blood hyperviscosity, abnormal heart rhythm and so on. In accordance with the principles of medical statistics analysis, collecting the required sample size is five times or more than observed variable, the sample is sixty-two which come from volunteers detected as in Chen (2009).

2.4   Intercept Method of Time Domain Signal

The signal PDWVW and PSV are treated with by the six methods interception as follow.

- Three seconds signal before and after the maximum point of the signal PDWVW is referred to as signal A.
- One second signal before and after the maximum point of the signal PDWVW is referred to as signal B.
- The first two seconds in the three seconds before the maximum point of the signal PDWVW is referred to as signal C.
- Three seconds signal before and after the maximum point of the signal PSV is referred to as signal D.
- One second signal before and after the maximum point of the signal PSV is referred to as signal E.
- The first two seconds in the three seconds before the maximum point of the signal PSV is referred to as signal F as in Feng et al. (2006).

3   Results

3.1   Typical Power Spectrum

Normal group interception signals power spectrum typical figure: The interception signal power spectrum whose sample number is 20 are showed as Fig. 2. Patient group interception signals power spectrum typical figure: The interception signal power spectrum whose sample number is 18 are showed as Fig. 3.

3.2   Power Spectrum Results

Take the frequency corresponding to the largest energy and the second largest energy as the characteristic parameters of the power spectrum analysis. The frequency of the largest energy corresponding is referred to as a parameter G and the frequency of the second largest energy corresponding as a parameter H.

To make sure that whether the parameter G and H of six kinds of signals in the normal group and patient group have statistically significant, use the t test and $\alpha = 0.05$ by the SPSS statistical analysis software and the results are showed as Tables 1, 2, 3, 4, 5 and 6. The P_M means the equalizing value the parameter G, S_M the standard deviation of the parameter G, P_N the equalizing value the parameter H and S_N the standard deviation of the parameter H.
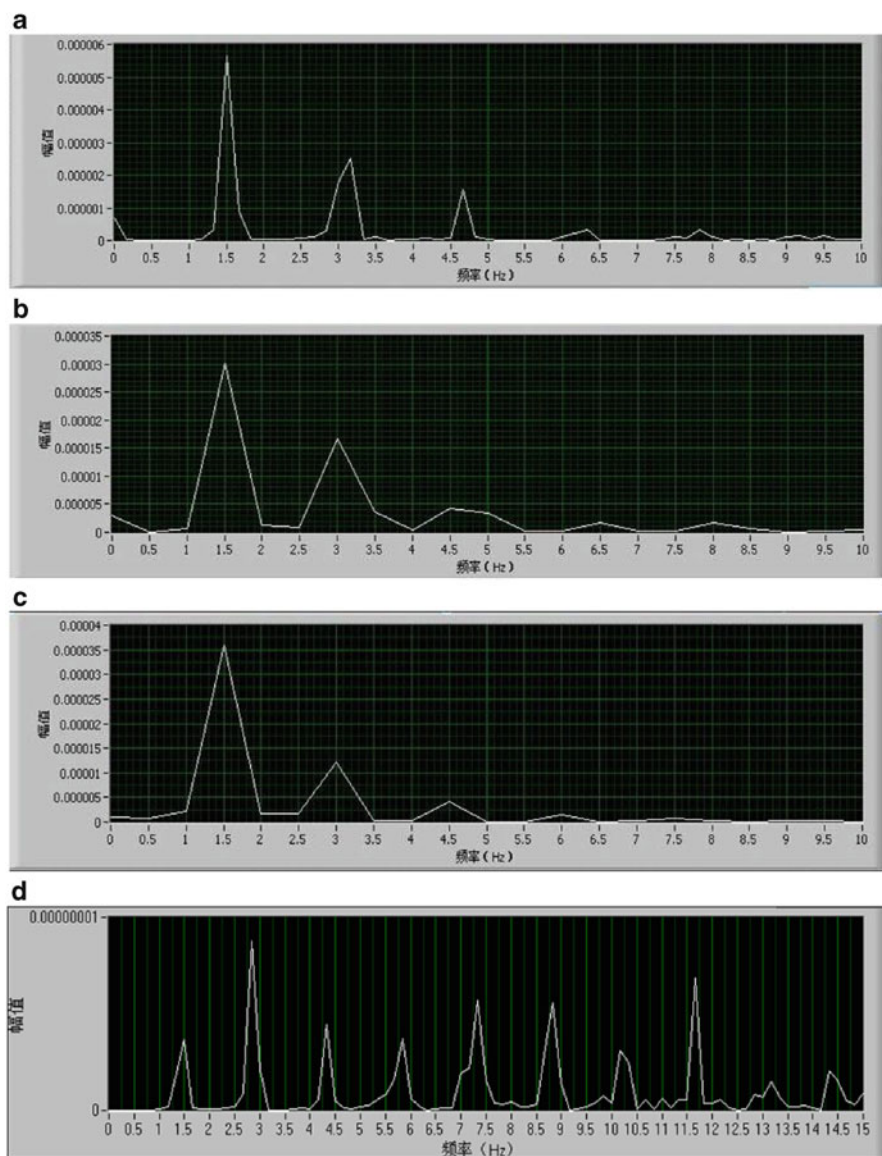
Fig. 2 Normal group power spectrum figure: (**a**) to (**f**) is the power spectrum figure of signal A to F

e

f

Fig. 2 (continued)

a

b
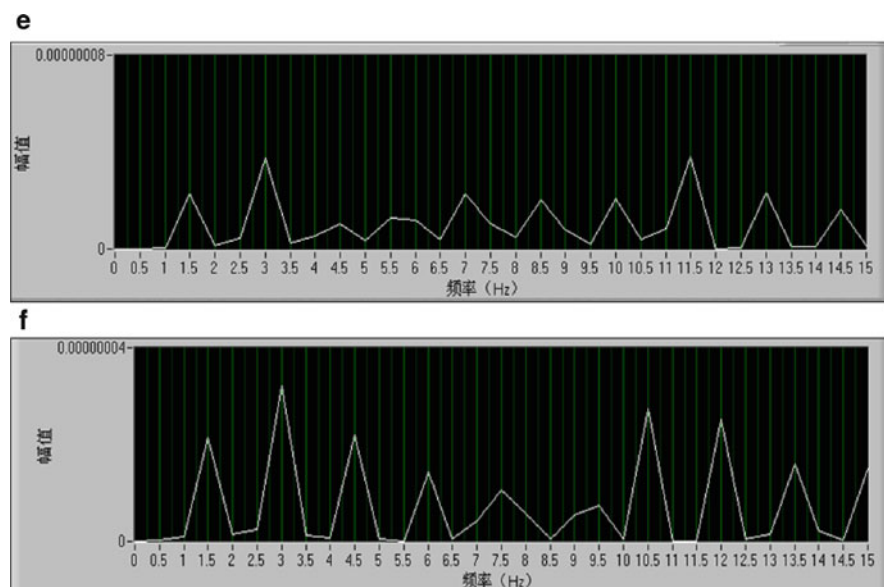
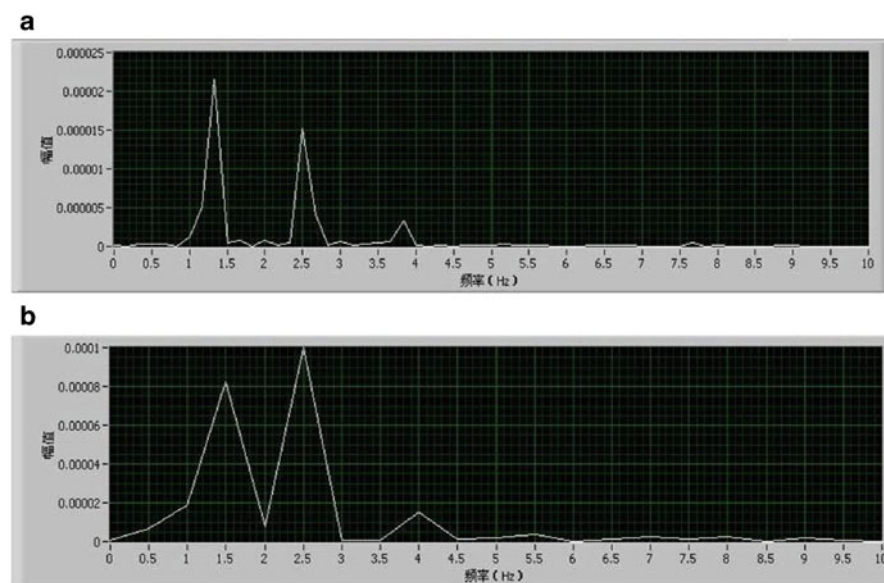**Fig. 3** Patient group power spectrum figure: (**a**) to (**f**) is the power spectrum figure of signal A to F
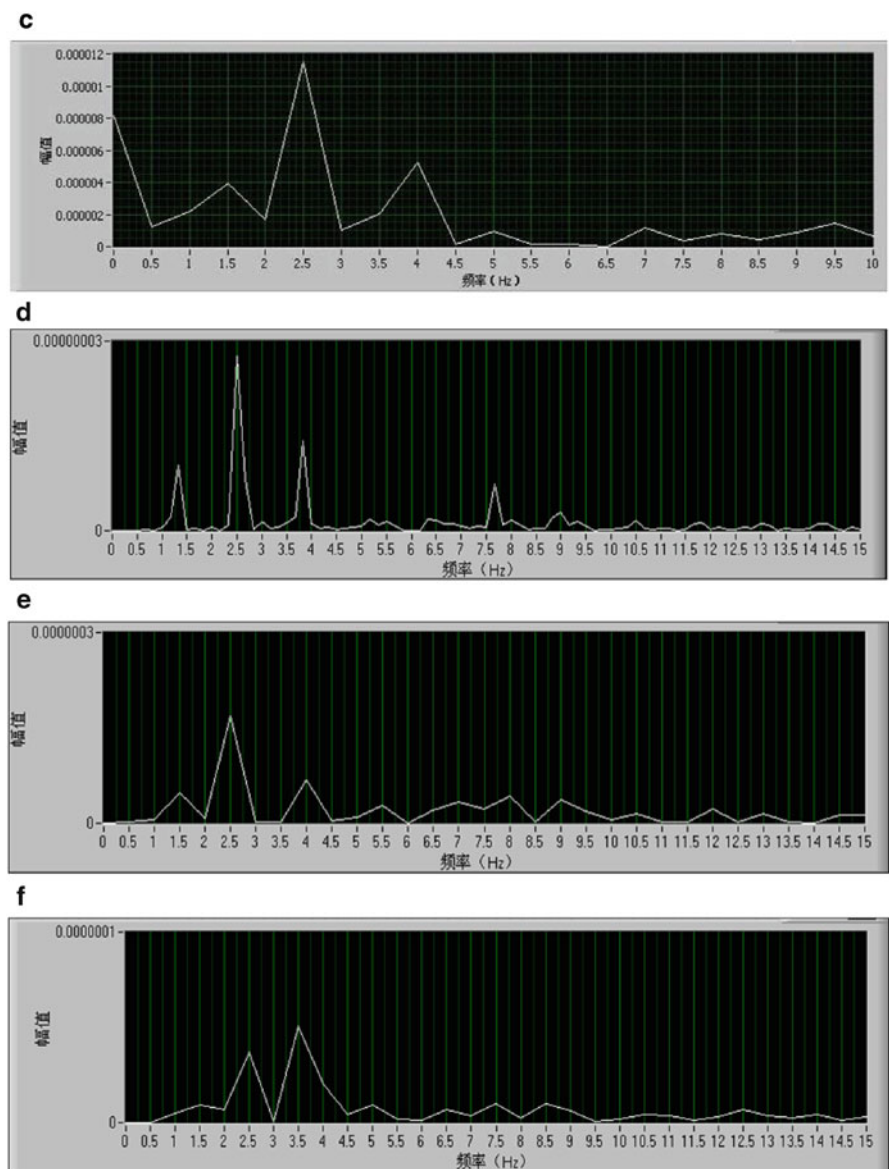
Fig. 3 (continued)

Table 1 The power spectrum results of signal A

Power spectrum parameters	Normal group(Hz)	Patient group(Hz)	Difference
$P_M \pm S_M$	1.39 ± 0.20	1.14 ± 0.14	$P < 0.05$
$P_N \pm S_N$	2.98 ± 0.25	2.37 ± 0.21	$P < 0.05$

Table 2 The power spectrum results of signal B

Power spectrum parameters	Normal group(Hz)	Patient group(Hz)	Difference
$P_M \pm S_M$	1.29 ± 0.18	2.20 ± 0.16	$P < 0.05$
$P_N \pm S_N$	2.72 ± 0.16	1.24 ± 0.19	$P < 0.05$

Table 3 The power spectrum results of signal C

Power spectrum parameters	Normal group(Hz)	Patient group(Hz)	Difference
$P_M \pm S_M$	1.41 ± 0.25	2.45 ± 0.22	$P > 0.05$
$P_N \pm S_N$	2.92 ± 0.30	4.03 ± 0.34	$P < 0.05$

Table 4 The power spectrum results of signal D

Power spectrum parameters	Normal group(Hz)	Patient group(Hz)	Difference
$P_M \pm S_M$	2.91 ± 0.30	2.48 ± 0.20	$P < 0.05$
$P_N \pm S_N$	11.80 ± 0.21	3.84 ± 0.34	$P < 0.05$

Table 5 The power spectrum results of signal E

Power spectrum parameters	Normal group(Hz)	Patient group(Hz)	Difference
$P_M \pm S_M$	11.60 ± 0.20	2.49 ± 0.23	$P < 0.05$
$P_N \pm S_N$	2.93 ± 0.20	3.99 ± 0.28	$P < 0.05$

Table 6 The power spectrum results of signal F

Power spectrum parameters	Normal group(Hz)	Patient group(Hz)	Difference
$P_M \pm S_M$	2.93 ± 0.27	2.59 ± 0.26	$P > 0.05$
$P_N \pm S_N$	10.61 ± 0.22	3.43 ± 0.25	$P < 0.05$

3.3 Power Spectrum Analysis Results

- The equalizing value and standard deviation of the normal group were higher than the patient group about the signal A. That the parameter G and H of the signal A were proportional to the degree of vascular wall elasticity showed that the degree of vascular wall elasticity normal group was higher than that of patient group. And that P were less than 0.05 in comparison of the parameter G and H showed that the parameter G and H of the signal A had statistically significant. So the parameter G and H of the signal A can be used as one of the characteristic quantities in the quantitative diagnosis of vascular wall elasticity degree.

- That P were less than 0.05 in comparison of the parameter G and H showed that the parameter G and H of the signal B had statistically significant. The normal group parameter G (P_M) was lower than the patient group but the parameter H (P_N) higher, so only the parameter H of the signal B was proportional to the degree of vascular wall elasticity.
- That P was less than 0.05 in comparison of the parameter H showed that only the parameter H of the signal C had statistically significant. That the equalizing value and standard deviation of the normal group were lower than the patient group about the signal C shows that the parameter G and H of the signal C were both inversely proportional to the degree of vascular wall elasticity.
- The results of the signal D were exactly the same as the signal A. So the parameter G and H of the signal D can be used as one of the characteristic quantities in the quantitative diagnosis of vascular wall elasticity degree.
- That P were less than 0.05 in comparison of the parameter G and H showed that the parameter G and H of the signal E had statistically significant. The normal group parameter G (P_M) was higher than the patient group but the parameter H (P_N) lower, so only the parameter G of the signal E was proportional to the degree of vascular wall elasticity.
- That P was less than 0.05 in comparison of the parameter H showed that the parameter H of the signal F had statistically significant. That the equalizing value and standard deviation of the normal group were higher than the patient group about the signal F showed the parameter G and H of the signal F were both proportional to the degree of vascular wall elasticity.

4 Conclusion

The difference of all signal parameters were significant except the parameter G of the signal C and F in the results of six kinds of signals power spectrum analysis of normal group and patient group. Therefore, choosing the parameter G and H as the characteristic parameters of vascular wall elasticity disease diagnosis is feasible.

In many cases, the parameter G or H was proportional to the degree of human vascular wall elasticity in six kinds of signals power spectrum analysis. And the degree of elasticity of normal group was higher than that of patient group. Especially, the parameter G and H of signal A and D not only had statistically significant, but also were proportional to the degree of human vascular wall elasticity. So, the two signals can be used as one of the main characteristic quantities in the quantitative diagnosis of vascular wall elasticity degree.

The medical clinical interpretation of the power spectrum analysis needs further study.

References

Chen LL (2009) Vascular wall dynamic information detection and diagnosis system based on remote virtual instrument technique. Chongqing Medical University, Chongqing

Chen LL, Gan P, Zeng ZF (2008) Vascular wall dynamic information detection and diagnosis system based on virtual instrument technique. J China J Bioinform 7:55–59

Chen LL, Gan P, Tan HM (2009) Study on the measurement of noninvasive dynamic information of human vascular wall. J Biomed Eng 26:157–169

Feng J, Gan P, Hu GH (2006) Collecting method for dynamic information of vascular wall. J Chin J Sens Actuator 19:629–632

Gan P, Hu GH (2006) Research on collecting method for dynamic information of vascular wall. J Transducer Microsyst Technol 25:18–21

Nie N (2005) Technology and applications of biomedical digital signal processing. Science Press, Beijing

Robust Localization Based on Measurement Reliability

Xiaoping Wu and Shili Tan

Abstract As the key technology, localization is a fundamental research that has been extensively studied in the literature of wireless sensor networks. Improving the accuracy and decreasing the costs of localization is the mainly goal for node localization. To improve the precision of located nodes, we design the wTLS localization approach, which relies on the distance reliability and makes good use of the distance information. Based on the variance of different hop counts, we apply the approach to DV-Hop localization. The experiment results show that wTLS outperforms the state-of-arts approaches with high accuracy, efficiency, and consistency of performance and gains more excellent results than TLS and LS.

Keywords Wireless sensor network • Robust localization • Measurement reliability

1 Introduction

Wireless Sensor Networks (WSNs) are widely proposed to apply in environment observation and forecasting, health caring, battle-field surveillance. For most applications, sensor data without spatial and temporal coordinates is of very limited use. Sensor nodes have to be aware of their location to be able to specify "where" a certain event takes place. Therefore, the problem of localizing the sensors is of paramount importance for many classes of sensor network applications.

The location information of sensor nodes is an indispensable element in various applications, such as fire risk evaluation, canopy closure estimates (Mo et al. 2009), microclimate observation, and search and rescue in the wild. The most

X. Wu (✉) • S. Tan
School of Mechatronic Engineering and Automation, ShangHai University,
Shanghai 200444, People's Republic of China
e-mail: wuxiaoping05@gmail.com; tanshili@shu.edu.cn

existing works on sensor network localization can be classified as based on range-based or range-free approaches. In the range-free approaches, the algorithms do not need ranging hardware support and are immune to range measurement errors while providing less accurate localization results. In the range-based approaches, the algorithms require more sophisticated range hardware support while providing more accurate localization results than the range-free algorithms when the range measurement is accurate. The range-based algorithms include the robust least-square localization, DV-based approach (Shang et al. 2003), and collaborative multilateration. Among them, DV-based approach, SDP and MDS-MAP also deal with the range-free case (Ji and Zha 2004). Much less work has been done on utilizing other range related information for sensor network localization. Received signal strength indication (RSSI) is a measure of the RF energy received and is closely related to range. For the localization purpose, the information provided by RSSI or similar types of measurements is less than range but more than a connectivity-only hop count, and it can be used to improve the accuracy of connectivity only localization algorithms. Summing up the above methods, each approach needs the distance information, so the measurement distance reliability will be helpful for improving localization precision.

Ranging quality is the key that determines the overall localization accuracy. Bearing this point in mind, recent proposals focus more on error control and management. Some of those proposals enhance the localization accuracy by deliberately reducing the contribution of error-prone nodes to the localization process. The rest schemes improve localization by identifying large ranging errors and outliers, relying on the topological or geometric properties of a network (Hazas and Ward 2002; Moore et al. 2004). The localization error comes from the vertex error and the error in distance measurement (Liu et al. 2003). When GPS or other device assisted location technology is not available due to shadowing or cost, it is attractive to develop inexpensive localization approaches. In node localization, most existing methods start with a small number of anchor nodes whose locations are known. Other nodes localize themselves with respect to these anchor nodes. For anchor nodes, the vertex error is mostly zero. Bur for regular measurements, a Gaussian white noise is always happened, which is hard to removed for localization. For many occasion, the ranging reliability could be conscious through the efficient measurement analysis. RSSI measurement always leads to the outlier distance caused by a node in a pit, with a faulty antenna, or at the edge of radio range to other nodes (Jian et al. 2009). Different hop counts have different variance stated by measurement results when locating a node with DV-Hop. For improving the localization accuracy, we amend the conventional least square localization approach and design wTLS approach. The contributions of our work are summarized as follows.

1. We propose a weight Taylor least squares scheme (wTLS), which use reasonable weight coefficient based on the measurement reliability. With wTLS, the estimated node locations are less error than those output by other schemes.

2. For better understanding wTLS, we apply the approach for DV-Hop. We put forward the means of gaining the weight and give the experiment results of the approach.
3. Except for DV-Hop, wTLS can also solve the outlier problem of measurement results, which makes use of more information than purely removing the outlier data. We validate the approach by experiments.

This paper presents a robust distributed node localization based on the measurement reliability, which uses ratio value based on different measurements data reliability. The rest of this paper is structured as follows. Section 2 represents the motivations from the original localization technology. Section 3 describes the design of robust localization algorithm. Section 4 analyzes the simulation about actual problem and evaluates the results.

2 Motivation

Radio Signal Strength Indication (RSSI) is based on the fact that the strength of radio signal diminishes during propagation. As a result, the understanding of radio attenuation helps to map the signal strength to the physical distance. In theory, radio signal strengths diminish with distance according the power law. A generally employed model for wireless radio propagation is as follows

$$P(d) = P(d_0) + 10\eta \log(\frac{d}{d_0}) + X_\delta \tag{1}$$

Where $P(d)$ is the received power at distance d, $P(d_0)$ is the received power at some reference distance d_0, η is the path-loss exponent, and X_δ is a log-normal random environment noise followed $N(0, \delta_x^2)$ that accounts for fading effects. Hence, if the path-loss exponent for a given environment is known, the received signal strength can be translated to the signal propagation distance. However, in practice, RSSI-based ranging measurements contain noises on the order of several meters. The ranging noise occurs because radio propagation tends to be highly dynamic in complicated environments. On the whole, RSSI based ranging is a relatively cheap solution without any extra devices, as all network nodes are supposed to have radios. It is believed that more careful physical analysis of radio propagation may allow better use of RSSI (Kung et al. 2008). Nevertheless, the breakthrough technology is not there today. We randomly select a time to collect the RSSI readings from all the 100 nodes. RSSI are then used for range measurements based on Eq. (1). By comparing the converted distances with the ground-truth ones, Fig. 1 plots the mean error of range measurements between every node and its neighbors. Generally, if a node has line of sight connections with all its neighbors, the mean error of its range measurements is small. If a node lies in a pit, has a faulty antenna, or is relatively far from all its neighbors, the mean error is likely to be large.
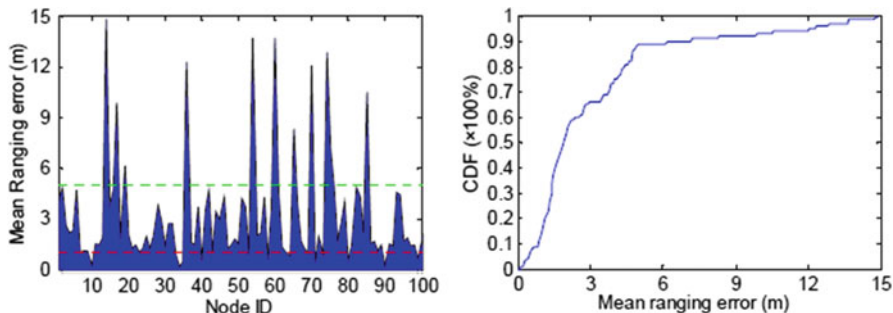
Fig. 1 Errors of range measurements on the nodes

Table 1 Mean and variance under different hop counts

| Hop counts	1	2	3	4	5
Mean	19.24	21.25	22.08	22.10	21.8
Var	51.5	13.58	7.22	3.44	2.30

The anisotropic problem exists large randomness in the deployed fields. When using a rang-free approach (e.g. DV-Hop) without considering this problem, nodes are located at the same position if their hop counts to the landmarks are equal. Nevertheless, they might be at a distance from each other in reality. The anisotropic problem also brings severe challenges to range-based localization (Shang et al. 2004), and we test the mean and variance in different hop counts of DV-Hop. The experiment results are shown in Table 1, and we can find the distance between nodes has the different reliability due to different hop counts. The far nodes distance variance are smaller than the near, so their reliability is higher (Li et al. 2005).

3 Design

The distances from an unknown node to several references constrain the presence of this node, which is the basic idea of the so called multilateration. Figure 2 shows an example of trilateration, a special form of multilateration which utilizes exact three references. A to-be-located node (node 0) measures the distances from itself to three references (nodes 1, 2, 3). Obviously, node 0 should locate at the intersection of three circles centered at each reference position. The result of trilateration is unique as long as three references are non-linear. Supposing the location of the unknown node is (x, y) and it is able to obtain the distance estimates $\hat{d}_i$ to the i-th reference node locating at (x_i, y_i), $i = 1, 2, \cdots, n$, let d_i be the actual Euclidean distance to the i-th reference node, i.e.,

$$(x_i - x)^2 + (y_i - y)^2 = d_i^2 \quad i = 1, 2, \cdots, n \tag{2}$$
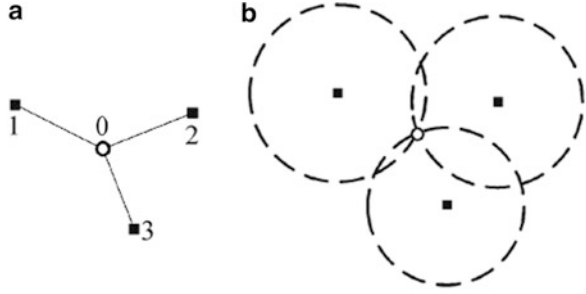
The difference between the measured and the actual distances can be represented by $\hat{d}_i - d_i$ ($\hat{d}_i$ is the distance calculated from the estimated sensor position, d_i is the measured distance). In the localization of wireless sensor network, it is the mainly means to measure the distance or angle between the unknown node and the anchors. If Eq. (2) minus each other, we can obtain the linear equation $AX = b$, the least squares (LS) solution to the linear system is $X = (A^T A)^{-1} A^T b$ X (Liu et al. 2003). The matrix of b is $x_i^2 - x_j^2 + y_i^2 - y_j^2 + d_i^2 - d_j^2$. If the measured d_i and d_j exist error, the subtraction by Eq. (2) brings to bigger error. The pure subtraction of Eq. (2) will lose more measurement information, so we introduce Taylor Least Squares (TLS) here.

$$f(x, y) = \sqrt{(x_i - x)^2 + (y_i - y)^2} \ i = 1, 2, \cdots, n \tag{3}$$

Expanding Eq. (3) using Taylor formula at the point (x_0, y_0) and obtain

$$f(x, y) = \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2} + \frac{(x_0 - x_i)}{\sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2}} h$$

$$+ \frac{(y_0 - y_i)}{\sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2}} k \tag{4}$$

Eq. (4) is also expressed as matrix

$$\mathbf{J} \cdot \Delta\mathbf{x} = \Delta\mathbf{d} \tag{5}$$

Where

$$\mathbf{J} = \left(\frac{(x_0 - x_i)}{\sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2}}, \frac{(y_0 - y_i)}{\sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2}} \right)$$

$$\Delta\mathbf{x} = (h, k), \ \Delta\mathbf{d} = d_i - \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2} \ i = 1, 2, \ldots, n \tag{6}$$

Considering the parameters (x_i, y_i) and d_i, the vertex error of anchor nodes is mostly zero, so the main error of located results comes from d_i because of the measurement noises. Sometimes the measurement reliability can be estimated or predicted. Tab. 1 lists the mean and variance value with different hop counts. From the table, we can find the reliability in different hop count is differential, so it is efficient for improving precision to give the each node different weight to each equation. We substitute for the equation with

$$\mathbf{w} \cdot \mathbf{J} \cdot \Delta \mathbf{x} = \mathbf{w} \cdot \Delta \mathbf{d} \tag{7}$$

Where

$$\mathbf{w} = diag(w_1, w_2, \cdots, w_n) \tag{8}$$

we call $\mathbf{w}$ as weight matrix, w_i value is decided by d_i reliability. The approach is called as weight Taylor Least Squares (wTLS). RSSI range measurements are easy-to-implement and thus get popular in practice. Empirical models of signal propagation are constructed to convert absolute RSSI to distance. However, noise and outlier is inevitable in distance ranging for environmental factor, adversary attacks, hardware malfunction or failure. A feasible means is removing the outlier data for improving precision, but in fact the outlier data also include some useful information, purely getting rid of the outlier data will partly lose information. Weakening the outlier data with small weight coefficient could improve the location accuracy effectively. Many papers introduce the identification of the outlier nodes, the distance difference $\hat{d}_i - d_i$ can judge the outlier. If the connectivity is high and the solution is unique, $\hat{d}_i - d_i$ are closely related to the actual distance difference.

4   Evaluation

DV-Hop utilizes the connectivity information to estimate node locations. Every node counts its hop counts (i.e. the number of hops on the shortest path between two nodes) to at least three landmarks. The distance between a node and a landmark is calculated as the product of the hop count between them and the per-hop distance, which is a pre-determined constant for all the nodes. The location of a node is then calculated by using Least Square Estimation in WSNs. We divide the DV-Hop into two steps, computing the distance between beacon nodes and unknown nodes, and solving for the unknown coordinate with the distance data. Here we are concerned with the second step, which is how to improve the accuracy using the distance data computed by the first step. In this group of simulations, 100 nodes are randomly deployed in the area of 100×100 m, the beacon node number is 20, the experiment shows that communication radius is less than 15, the beacon node density is too sparse to compute the percent hop distance, at radius 15, a 95% of node is
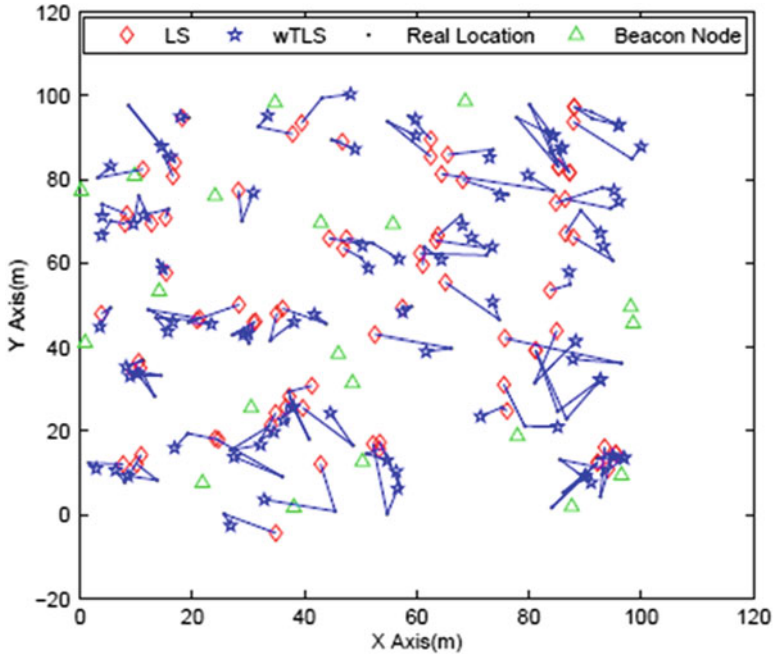
Fig. 3 Localization results of LS and wTLS

connected. Specifically, DV-Hop performs very poor when the node density is low. This is because the anisotropic problem appears to be more serious with sparse deployments. When the node density increases, the anisotropic problem is mitigated and the localization accuracy is thus improved.

Figure 3 plots the localization results of the three approaches. Least-Square have bigger error for being unaware of the uncertainty. Taylor Least Squares decrease the error by considering least $\hat{d}_i - d_i$. wTLS can get the least mean localization error of three approaches using the weight according the reliability of ranging. The weights selecting is critical for the accuracy, an appropriate weight could decrease the error effectively. In our experiment, we use the subsection weight to distinguish the ranging reliability. It is easy to see that wTLS performs better than LS and TLS. Figure 3 shows that for almost all the nodes, wTLS achieves higher localization accuracy than other two approaches.

From Fig. 3, we can see the performance of LS is the worse one. Actually, we observe in the experimental results that many different nodes are estimated to the same locations by LS, because they have same hop counts to the landmarks, but their real locations are far from each other. With a higher node density, it is also easier for wTLS to obtain enough good range measurements for localization. Nevertheless, as the node density goes up, the mean localization error quickly drops and then gradually converges.

Fig. 4 Impact of the number of landmarks

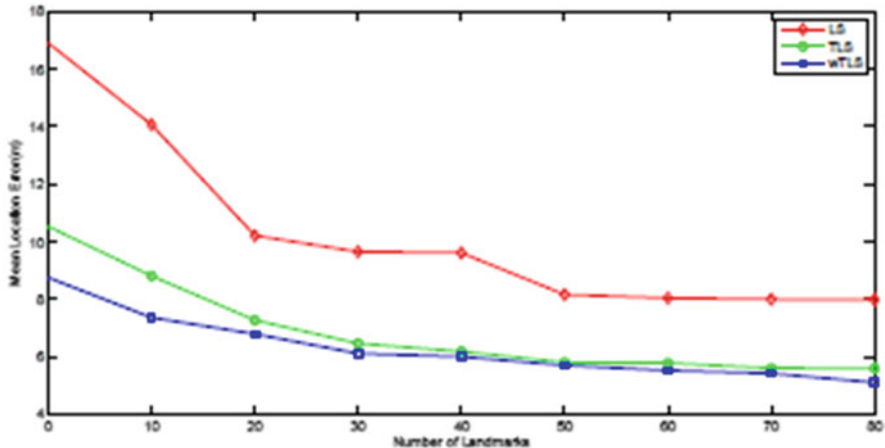Power-based localization, which relies on range estimation via received signal strength of RF signals, is convenient in ad hoc networks of wireless nodes. We consider the problem of using 802.11 radio ranging to localize ad-hoc nodes in an outdoor, open-space environment. The solution is straightforward when the terrain is flat and nodes can all hear each other well. However, this may not be the case in real-word applications. For example, a node in a pit, with a faulty antenna, or at the edge of radio range to other nodes may cause large errors in ranging measurements due to a weak signal from the node or received by the node. In addition, heterogeneous radio equipment may also introduce inconsistency in measurement errors.

In the general approach, removing the bigger error distance data is the universal method, but in fact the bigger error also includes some useful information. Purely getting rid of the data will lose the information. wTLS can make good use of the distance information including the outlier data with the appropriate weights. We conduct the experiment, Fig. 4 shows the result, wTLS also achieves less error than the conventional least square approach. The LS approach achieve the max error is 18 m, likewise, the max error of wTLS is within 11 m. From Fig. 4, we can see the performance of LS is worse than wTLS.

Considering the ubiquitous ranging errors and poor overall ranging quality, the robustness of a localization approach against such interfering factors is the last but not least metric we want to evaluate. For this purpose, we conduct a group of simulations with 1,000 nodes. Node density $N_d = 12$ and the number of landmark is set to 12. We use another two parameters to control the degree of ranging errors. The first one is the percentage of bad links which is respectively set at 0, 10, 20, 30, 40, and 50%. Another parameter is the relative ranging error. We assume in the simulations that the links on a node are either all good or all bad. The relative ranging error of a link conforms to a Gaussian distribution $N(\mu_{bad}, 0.2\mu_{bad},)$ where μ_{bad} denotes the average of relative ranging error and
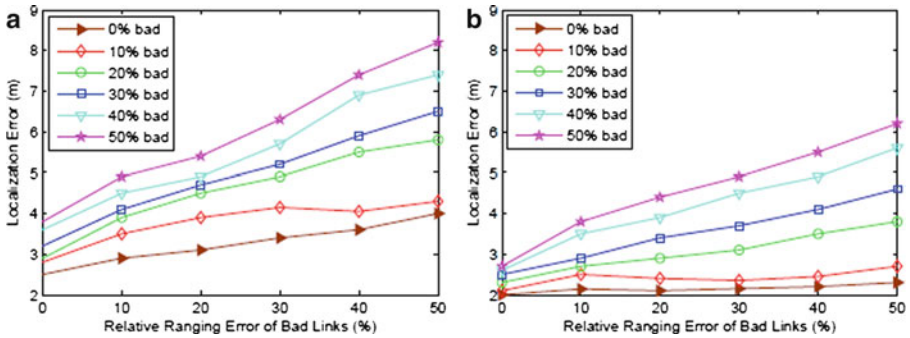
Fig. 5 Comparison of localization errors: (**a**) purely removing the big error measurement; (**b**) wTLS

set at 0, 10, 20, 30, 40, and 50%, respectively. Meanwhile, we assume the links are asymmetric. Figure 5 plots the mean localization errors of purely removing the big error measurement and using wTLS approach, under different settings. We find the former is relatively insensitive to the changes of ranging errors, because it mainly relies on the measurement information of rest node instead of all. The approach results have localization error of more than 2.5 m, even when all links are good, the max error can achieve 4 m. Its performance seriously degrades when the percentage of bad links gets above 20%, compared with the approach, wTLS has even better performance. When all the links are good, its localization errors are almost near 2 m. Even when there are 50% bad links, wTLS still perform robustly enough, the localization error is no more than 6 m. This group of simulation shows the remarkable advantages of wTLS in extremely complex environments.

References

Hazas M, Ward A (2002) Robust distributed network localization with noisy range measurements. In: Proceedings of the 4th international conference on ubiquitous computing. IEEE Press, Philadelphia

Ji X, Zha H (2004) Sensor positioning in wireless Ad-hoc sensor network using multidimensional scaling. In: Conference of 23rd IEEE INFOCOM. IEEE Press, Hong Kong

Jian L, Yang Z, Liu Y (2009) Beyond triangle inequality: sifting noisy and outlier distance measurements for localization. In: Proceedings of 28th IEEE Infocom. IEEE Press, Brazil

Kung HT, Lin C-K and Vlah D (2008) Localization with snap-inducing shaped residuals (SISR): coping with errors in measurement. In: Proceedings of 14th ACM MobiCom. ACM Press, San Francisco

Li Z, Trappe W, Zhang Y, Nath B (2005) Robust statistical methods for securing wireless localization in sensor networks. In: Proceedings of 4th ACM/IEEE IPSN. ACM/IEEE Press, Los Alamitos, CA

Liu J, Zhang Y, Zhao F (2003) Robust distributed node localization with error management. In: Proceedings of 4th ACM MobiHoc. ACM Press, Maryland

Mo L, He Y, Liu Y, Zhao J, Tang S, Li X-Y, Dai G (2009) Canopy closure estimates with GreenOrbs: sustainable sensing in the forest. In: Procedding of the 7th ACM SenSys. ACM Press, Berkeley

Moore D, Leonard J, Rus D, Teller S (2004) Robust distributed network localization with noisy range measurements. In: Proceedings of 2nd ACM SenSys. ACM Press, Baltimore

Shang Y, Ruml W, Zhang Y (2003) Localization from Mere Connectivity. In: Proceedings of 4th ACM MobiHoc. ACM Press, Maryland

Shang Y, Ruml W, Zhang Y (2004) Localization from connectivity in sensor networks. In: IEEE transactions on parallel and distributed systems, vol 15, No 11. IEEE Press, New York, pp 961–974

Log Mining of Virtual Host Website Based on Association Rule

Xi Guangwen and Long Gonglun

Abstract Website log recorded the users' behavior of accessing and browse. Users' accessing and browsing behaviors are recorded in web logs, by mining which those behaviors can be analyzed and hence provide valuable data for optimizing web structures. But most virtual hosting websites have no log-recording functions, it brings great inconvenience to log mining. This paper introduces a method of creating logs on virtual hosting website and mining them with associate rules would help in discovering users' visiting patterns, which improves site structures and optimizes user experiences.

Keywords Log mining • Association rules • Virtual hosting • Apriori

1 Introduction

Along with the swift development of the Internet, more and more enterprises and people have been build up websites on the Internet and gaining strategic advantages by internet applying on crucial missions. Each day large amount of logs are created reflecting users' behaviors on webs. A proper analysis of such data can generate visiting patterns of users and hence enable better decisions on market strategies, improve efficiency of market activities. And it can also provide useful information to web owners and customized services to appointed users (Liu Li-Zhen et al. 2003; Cheung et al. 1996a, b). Many websites are hosting on virtual hosting and not every

X. Guangwen (✉)
Network Center, Chongqing Three Gorges University, Wanzhou, Chongqing, China
e-mail: irains@vip.qq.com

L. Gonglun
Center of Information and Technology, Chongqing Normal University, Chongqing, China
e-mail: 65003723@qq.com

hosting provider offers log function, which brings great difficulty on log mining. So for creating logs we develop a separate module to provide data source for web log mining.

2 Virtual Hosting

Virtual hosting is a method with special hardware and software technologies for cutting a real computer host into multiple logic storing units (logic host). The unit has no real entity but like a real host can act functions like individual domains, IP addresses, storing spaces and complete internet servers. This virtualized logic hosting is named virtual hosting. Different virtual hosting are completely independent from one another. It seems virtual hosting performs exactly like an independent hosting.

Since multiple virtual hosting share resources of one real hosting the utilization of servers and cables are largely increased, thanks to which no extra servers or cables cost is necessary since in one server IP addresses can be set up in without any conflicts or sites with independent domains can be built up. Meanwhile we can save the cost on special server supervisors and don't need to worry about server maintaining problems.

The low cost and high efficiency of virtual hosting are greatly attractive for small medium organizations and web fans. Now virtual hosting has already developed itself as an important platform for enterprises information publicizing, products promoting and personal presenting.

3 Log Mining on Virtual Hosting Webs

A basic web log mining procedure include data preparing, mining rules applying,[1] pattern analyzing and visualization (Lv Jia 2006).

Normally virtual hosting webs have no logs so we design a log generating module for creating users accessing logs. A data made for mining rule applying is produced by reading logs (Savasere et al. 1995; Toivonen 1996). Then strong association rules are generated from Apriori (show in Fig. 1).

3.1 Association Rules

Association rules, invented by Agrawal and others in 1993 (Agrawal et al. 1993), are the most active research method used for data mining. Based on lots of business

---

[1] http://www.99to.com/help/hosting.asp
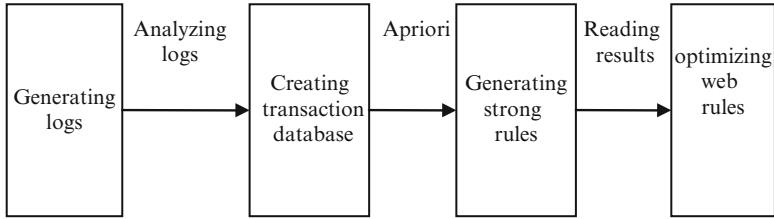
Fig. 1 Log mining procedure for virtual hosting webs

records useful association rules are found to be helpful on various business decisions and have been widely applied in analyzing shopping cart and transactions and building customized webs.

Association rules is a implication expression just as $A \Rightarrow B$, $A \subset I$, $B \subset I$, and $A \cap B = \emptyset$. Rule $A \Rightarrow B$ is established in the transaction set D, with the support S, where S is the percentage of D which contains $A \cup B$, it is the probability of $P(A \cup B)$. Rule $A \Rightarrow B$ has the confidence C in the transaction D, if D contains transaction A also includes transaction B which the percentage is C, it is the conditional probability of $P(B|A)$. The rules which are satisfied the minimum support threshold (min_sup) and the minimum confidence threshold (min_conf) referred as a strong rule (Agrawal and Shafer 1996; Agrwal and Srikant 1994a, b). We used the classical Apriori algorithm to mine transaction database.

3.2 Creating Logs

A module is added in each web to record visitors' IP addresses, visiting time, visited webs, their internet browsers, etc. and those information will be saved in txt documents. P2 shows part of the logs of Chinese Engine Web which uses virtual hosting. From row1 to row5 is respectively date, time, visited webs, requested parameters and their IP addresses. (show in Fig. 2)

3.3 Creating Database

To apply association rules we need to create a database including three sheets of IP sheet of the structure Ip(IpID, Ip), web sheet of the structure Page(PID, Page, Hits-total visits of one page) and transaction sheet of the structure Tran(TranID-transaction ID, IpID-reflecting IpID in the IP sheet, PageList-page sequence showing ID visited by IP of IpID).

```
#Fields: date time cs-uri-stem cs-uri-query c-ip
2010-10-01 00:00:41 /view.asp page=13&subid=21&subtype=%D0%C5 59.74.45.49
2010-10-01 00:01:03 /index.asp - 61.128.167.2
2010-10-01 00:01:30 /Article/ShowClass.asp ClassID=1&page=5 203.208.60.73
2010-10-01 00:01:51 /Article/ShowArticle.asp ArticleID=9856 203.208.60.73
2010-10-01 00:02:25 /index.asp - 202.202.218.243
2010-10-01 00:02:54 /office/readwj.asp id=4231 202.160.179.92
2010-10-01 00:03:46 /Article/ShowClass.asp ClassID=1 202.202.218.246
2010-10-01 00:03:49 /index.asp - 202.202.218.243
2010-10-01 00:03:51 /index.asp - 113.251.196.110
2010-10-01 00:03:51 /index.asp - 113.251.193.54
2010-10-01 00:04:02 /Article/ShowArticle.asp ArticleID=10525 202.202.218.243
```

Fig. 2 Part of logs

To create the transaction database,

1. Scan logs, remove the data and time, put row3 and row4 together to form strings: StrIP, StrPage.
2. See if StrIP exists in the IP sheet. If not add new IP record to the IP sheet and generate NumIpID for the IP.
3. See if StrPage exists in the Page sheet. If not add new Page record or otherwise revise the original record Hits + 1 then NumPageID for this Page is generated.
4. Find records those IpID equal to NumIpID in the table of Tran, if present, we need to analyze NumPageID is whether exist in PageList, if NumPageID does not exist, add NumpageID into Pagelist; If not, add a new transaction record in to Tran.

3.4   Researches and Analysis

We added a log creating module for each page of the website of Chongqing Three Gorges University and mined the logs of 1–31, October.

After several experiments, we took 10% as the support, so we have been frequent itemsets $l = \{1,31,113,126,172,266\}$, the collection element1,31,113,126,172,266 means pageID respectively, index.asp (Home), /yzlyb/imgchk/validatecode.asp (Code), Photo/ShowPhoto.asp? PhotoID $= 2$ (the campus scenery), Article/ShowArticle.asp? ArticleID $= 10516$ (School Profile), Article/ShowClass.asp? ClassID $= 1$ & page $= 2$ (Announcement), officee/yzap/yzap.asp (week of the session.) Set confidence level is 60%, by the confidence formula:

$$confidence(A \Rightarrow B) = P(A|B) = \frac{support\_count(A \cup B)}{support\_count(A)}$$

Strong association rules are as follows:

$$31 \wedge 113 \wedge 126 \wedge 172 \wedge 266 \Rightarrow 1, confidence = 100\%$$

$$31 \wedge 113 \wedge 12 \wedge 266 \Rightarrow 1 \wedge 172, confidence = 77.78\%$$

$$31 \wedge 113 \wedge 172 \wedge 266 \Rightarrow 1 \wedge 126, confidence = 100\%$$

$$1 \wedge 126 \wedge 172 \wedge 266 \Rightarrow 31 \wedge 113, confidence = 63.64\%$$

$$31 \wedge 126 \wedge 172 \wedge 266 \Rightarrow 1 \wedge 113, confidence = 100\%$$

$$31 \wedge 126 \wedge 266 \Rightarrow 1 \wedge 113 \wedge 172, confidence = 63.64\%$$

$$1 \wedge 113 \wedge 172 \wedge 266 \Rightarrow 31 \wedge 126, confidence = 70\%$$

$$31 \wedge 172 \wedge 266 \Rightarrow 1 \wedge 113 \wedge 126, confidence = 63.64\%$$

$$1 \wedge 113 \wedge 126 \wedge 172 \wedge 266 \Rightarrow 31, confidence = 77.78\%$$

$$113 \wedge 126 \wedge 17 \wedge 266 \Rightarrow 1 \wedge 31, confidence = 77.78\%$$

$$1 \wedge 31 \wedge 172 \wedge 266 \Rightarrow 113 \wedge 126, confidence = 63.64\%$$

$$113 \wedge 172 \wedge 266 \Rightarrow 1 \wedge 31 \wedge 126, confidence = 70\%$$

$$1 \wedge 31 \wedge 126 \wedge 266 \Rightarrow 113 \wedge 172, confidence = 63.64\%$$

$$1 \wedge 31 \wedge 126 \wedge 172 \wedge 266 \Rightarrow 113, confidence = 100\%$$

$$126 \wedge 172 \wedge 266 \Rightarrow 1 \wedge 31 \wedge 113, confidence = 63.64\%$$

$$1 \wedge 31 \wedge 113 \wedge 172 \wedge 266 \Rightarrow 126, confidence = 100\%$$

$$1 \wedge 31 \wedge 113 \wedge 126 \wedge 266 \Rightarrow 172, confidence = 77.78\%$$

By reading the above we learnt the six pages were highly associated with one another and the confidence level of four association rules were 100%, which indicated those pages were good options of presenting links of other pages. We visited the website and found in the notice page there was no link to the main page. It would be a good choice to add the main page link to the notice page to optimize website structures and improve visitors' experience.

4 Conclusion

But the normal Apriori algorithm generate large candidacy itemsets and needs to scan databases repeatedly, which costs too much for both time and spaces. We can use "itemsets count base on Hash" and other methods to improve the efficiency (Jiawei Han and Micheline Kamber 2006; Weihong 2004; Park et al. 1997). In the future, we will research how to improve the efficiency of Apriori algorithm.

References

Agrawal R, Shafer JC (1996) Parallel mining of association rules. IEEE Trans Knowl Eng 8(6):962–968

Agrawal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large databases. In: Proceedings of 1993 ACM SIGMOD international conference management of data, Washington, D.C, May 1993, pp 207–216

Agrwal R, Srikant R (1994a) Fast algorithms for mining association rules in large databases. In: Research Report RJ 9839, IBM Almaden Research Center, San Jose, Canada, June 1994

Agrwal R, Srikant R (1994b) Fast algorithms for mining association rules in large databases. In: Proceedings of the 20th international conference very large databases, Santiago, Chile, September 1994, pp 487–499

Cheung DW, Wong CY, Han J (1996a) Maintenance of discovered association rules in large databases: an incremental updating technique. In: Proceedings of the 12th international conference on data engineering, New Orleans Louisiana, pp 106–114

Cheung DW, Ng VT, Fu AW, Fu YJ (1996b) Efficient mining of association rules in distributed databases. IEEE Trans Knowl Data Eng 8(6):910–921

Jiawei Han, Micheline Kamber (2006) Data mining concepts and techniques, vol 4. Mechanical industry publishing house, Beijing

Liu Li-Zhen, Song Han-Tao, Lu Yu-Chang (2003) The applied research of web usage mining. Comput Sci 30(9):47–48

Lv Jia (2006) Research on application of web log mining technology. J Chongqing Norm Univ (Natural Science Edition) 23(4):39–45

Park JS, Chen MS, Yu PS (1997) Using a hash-based method with transaction trimming for mining association rules. IEEE Trans Knowl Data Eng 9(5):814–825

Savasere A, Omiecinski E, Navathe S (1995) An efficient algorithm for mining association rules in large databases. In: Proceedings of the 21st international conference on very large databases, Zurich, Switzerland, September 1995, pp 432–444

Toivonen H (1996) Sampling large databases for association rules. In: Proceedings of the 22nd international conference on very large databases, Bombay, India, pp 134–145

Weihong YU (2004) Implementation of data mining based on apriori algorithm using VB. Comput Eng 30(2):196–197

A Dynamics Model on Support System for Spare Parts

Weipeng Wang

Abstract The logistics of spare parts shows the characteristics of dynamics. We establish a support system based on demand-pull model. This paper analyses the feedback control mechanism of the system and the mathematical principles of the dynamics of the supply–demand model of spare parts based on system dynamics. Then we build a system state model, analyzing the dynamic actions of the support system. Our study shows that: (1) Excessive logistics delay may affect the stability of the system; (2) The stability and performance of the support system of spare parts based on the "demand-pull" model are closely related to the supply chain management system and technical conditions, also depend on the interactions of multiple factors. Then we get the value threshold of the key parameters to make the system stable, and test the impacts of logistics delay on the stability of the system. Finally, through system dynamics simulation, we optimize the system.

Keywords Spare parts • System dynamics • Organizational level • Demand—pull model

1 Introduction

According to the study of Clerk Maxwell (1892), Eason et al. (1955), Bernhard and Marion (2000), Huiskonenj (2001), Xiong Jingqi (2005), the logistics of spare parts shows the characteristics of dynamics in the implementation process. This is because the flow of spare parts in the supply system will cause changes in the quantities of spare parts inventory. To draw up a supply plan of spare parts based on demand more accurately, keeping the supply and demand balance, we must make a profound analysis of the structure of the logistics system of spare parts based

W. Wang (✉)
School of Economics and Management, Weifang University, Weifang, China
e-mail: wangweipeng1974@126.com

on demand-pull model, finding the law of the dynamic changes of the difference between supply and demand and the relationship between system performance and system parameters, so as to support basis for policy analysis and system design. However, how to reveal the objective law of the flow of spare parts from the angle of dynamic mechanism, so as to support the basis for the supply decision-making and risk control, which is one of the frontier theories in the current research area of the support of spare parts. Most existing studies on the problem of the security supply of spare parts focus on the analysis and processing of the external data of the supply security system, but with little consideration of the factors leading to the flow of spare parts within the system and the changing characteristics of the factors. They cannot reveal completely and deeply the essential law of the flow of spare parts, so it is difficult to adjust flexibly the order rate of spare parts, and cannot meet the needs of modern production. Compared to the maintenance of other level, the organization—level maintenance has particularly significant impacts on the conditions of equipment. So this article studies the organization—level inventory control of spare parts (including order, stock—in and stock—out) and usage activities, analyzing the inventory control mechanism of spare parts based on demand—pull model. For the inventory control, product reliability and maintainability are the important factors affecting demand rate of spare parts. This article will take the demand rate of spare parts as input to the model, studying the logistics law of spare parts in the supply support activities, finding the optimum strategy for spare parts inventory control.enlargethispage*12pt

2   System Dynamics Principles and Simulation

2.1   Causality Analysis

We think that the supply—and—demand feedback control mechanism of the support systems of spare parts is that: to observe and understand the actual activities of the support system of spare parts, obtaining the match information of supply and demand, and then to adjust the actual supply activities of spare parts according to the feedback information of actual supply and the deviation between supply and demand. The causal relationship among the factors of the support system of spare parts is shown in Fig. 1.

2.2   A System Dynamics Model of Supply and Demand of Spare Parts on the Organizational Level

Assume that without considering the impacts of the demand information delay on the supply activities, then the structure of the supply—demand system of spare parts can be described by the system dynamics flow diagram shown in Fig. 2.
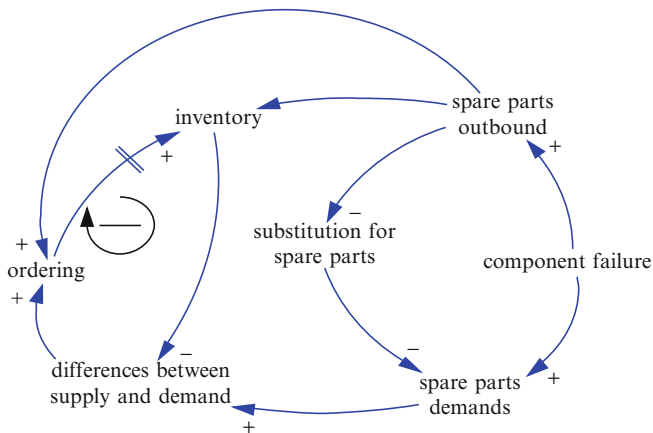
Fig. 1 Causal relationship of feedback control of supply and demand of spare parts
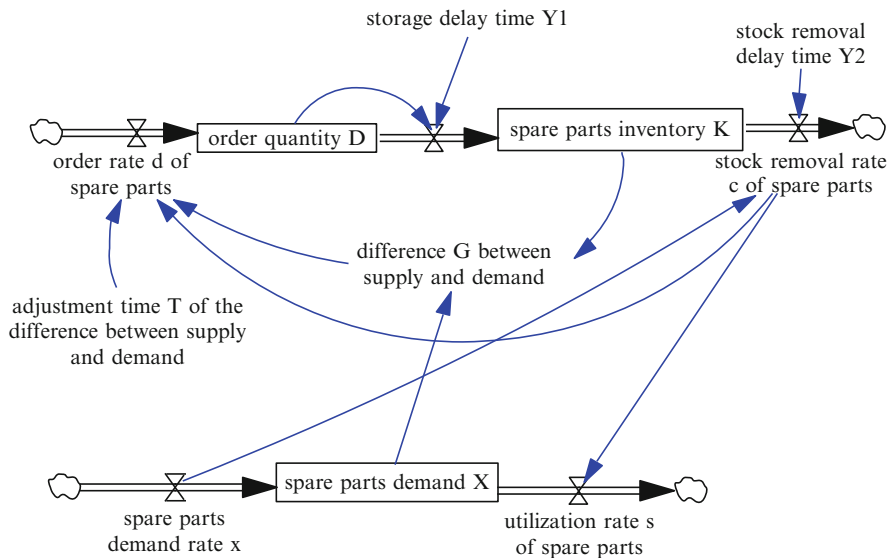

Fig. 2 The spare parts support SD model considering logistic delay

It not only expresses the logistics structure and feedback control relationship more clearly, but also defines the quantitative relationship between the variables through the mathematical equations.

$$K\,(t+1) = K(t) + [r(t) - c(t)] \tag{1}$$

$$X(t+1) = X(t) + [x(t) - s(t)] \tag{2}$$

Here, t and t + 1 are the two adjacent nodes with the interval of one time unit. The bellows are the same.

The quantity of order of spare parts D means the quantity of spare parts which are in the supply process but not in warehouse after making the order decision, the order rate d is the input velocity of D, the storage rate r is the output velocity.

$$\text{Abstract. } D\,(t+1) = D\,(t) + [d(t) - r(t)] \tag{3}$$

The variable D reflects that there exists a logistics delay time Y1 between order decision-making and stock—in. But the supply of spare parts cannot be stock—out immediately, so the delivery before the time $t - Y1$ must be in warehouse after the time delay of Y1. That is,

$$r(t) = d(t - Y_1) \tag{4}$$

At first, there is no demand for spare parts, demand rate of spare parts $x = 0$, the spare parts demand $X = 0$, stock—out rate of spare parts $c = 0$, spare parts inventory K is equal to the initial inventory of spare parts. If there exists the demand for spare parts, then the demand rate $x > 0$. If the inventory is sufficient, the stock—out of spare parts is on-demand; if the inventory is not sufficient, the stock—out rate determines by the inventory K. For simplicity, this article assumes that the inventory was sufficient, considering the factor of time delay, the stock—out of per unit of time is equal to the demand generated before the Y2 units of time. So

$$c(t) = x(t - Y_2) \tag{5}$$

Excluding the replacement time for spare parts, the spare parts usage rate is approximately equal to the stock—out rate. That is,

$$s(t) = c(t) \tag{6}$$

Order rate d depends on the support system and strategy. In the support system of spare parts, in order to ensure a certain degree of satisfaction rate of spare parts, it is a reasonable approach: to take the feedback information of the stock—out rate c of spare parts as one of the basis for the order decision-making. Considering the current demand X and inventory K and the difference G between them, then we can decide whether to adjust the orders and the adjustment range on the basis of the stock—out rate c, When $X > K$, increases the order, and the order rate d will be higher than the stock—out rate c of the previous cycle; When $X < K$, do the opposite. The adjustment range of the order rate d is proportional to the difference G between the supply and demand, and the time spent on complementing the difference between the supply and demand is set to the adjustment time T. The mathematical formula is:

$$G(t) = X(t) - K(t) \tag{7}$$

$$d(t) = c(t - 1) + \frac{G(t)}{T} \tag{8}$$

From the analysis of the dynamic process above, we can see that in the demand-pull model, inventory managers adjust the order through the difference between actual inventory and demand, this is a negative feedback system regulated by a difference. In addition, because the equipment failure rate and replacement rates lead to the randomness of the spare parts demand, the inventory K is not constant, but fluctuates according to demand.

3 Analysis of the Dynamic Actions

3.1 Description of the System State Space

Add the formula 4, formula 5, formula 7 and formula 8 into the formula 1, then we will have

$$
\begin{aligned}
K(t) = {}& K\,(t-1) \\
& + \left(\left[X\,(t-1-Y_1) - K\,(t-1-Y_1)\,T\right]/T\,\right) \\
& - x\,(t-1-Y_2) + x\,(t-2-Y_1-Y_2)
\end{aligned}
\tag{9}
$$

Add the formula 5 and formula 6 into the formula 2, then we have

$$
X(t) = X(t-1) + x(t-1) - x(t-1-Y_2)
\tag{10}
$$

Add the formula 4, formula 5, formula 7 and formula 8 into the formula 3, then we will have

$$
D(t) = D(t-1) + \left(
\begin{array}{c}
[X(t-1) - K(t-1)] - \\
\begin{bmatrix} X(t-1-Y_1)- \\ K(t-1-Y_1) \end{bmatrix}
\end{array}
\right) \Big/ T
\\
+ x(t-2-Y_2) - x(t-2-Y_1-Y_2)
\tag{11}
$$

Make $x1(t) = K(t)$, $x2(t) = X(t)$, $x3(t) = D(t)$, $r(t) = x(t)$, then the state space model of the supply–demand system of spare parts can be defined as

$$
\begin{aligned}
x(t) = {}& A_1 X(t-1) + A_2 x(t-1-Y_1) \\
& + B_1 r(t-1) + B_2 r(t-1-Y_2) \\
& + B_3 r(t-2-Y_2) + B_4 r(t-2-Y_1-Y_2)
\end{aligned}
\tag{12}
$$

In this equation,

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ -\dfrac{1}{T} & \dfrac{1}{T} & 1 \end{bmatrix}; \quad A_2 = \begin{bmatrix} -\dfrac{1}{T} & \dfrac{1}{T} & 0 \\ 0 & 0 & 0 \\ \dfrac{1}{T} & -\dfrac{1}{T} & 0 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad B_2 = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}; \quad B_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}; \quad b_4 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

3.2 Derivation of the System Dynamics

- Without Considering the Logistics Delay (Y1 = Y2 = 0)
  When Y1 = Y2 = 0, that is, no logistics delay.
  In this case, add the Eq. 5 into the Eq. 1, so we can have

$$K(t) = K(t-1) + [r(t-1) - x(t-1)] \tag{13}$$

Add the Eq. 5 and 6 into the Eq. 2, then we can get

$$X(t) = X(t-1) \tag{14}$$

Add the Eq. 4, 5, 6, and 7 into the Eq. 8, then we have

$$\begin{aligned} r(t) &= [-K(t-1) + X(t-1) - r(t-1)]/T \\ &+ \left[1 + \dfrac{1}{T}\right] x(t-1) \end{aligned} \tag{15}$$

Finally, the state space model of the supply–demand system of spare parts can be expressed by an equation as below:

$$y(t) = Fy(t-1) + Hr(t-1) \tag{16}$$

In this equation,

$$y(t) = \begin{bmatrix} K(t) & X(t) & r(t) \end{bmatrix}^T$$

$$F = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -\dfrac{1}{T} & \dfrac{1}{T} & -\dfrac{1}{T} \end{bmatrix}; \quad h = \begin{bmatrix} -1 \\ 0 \\ 1 + \dfrac{1}{T} \end{bmatrix}$$
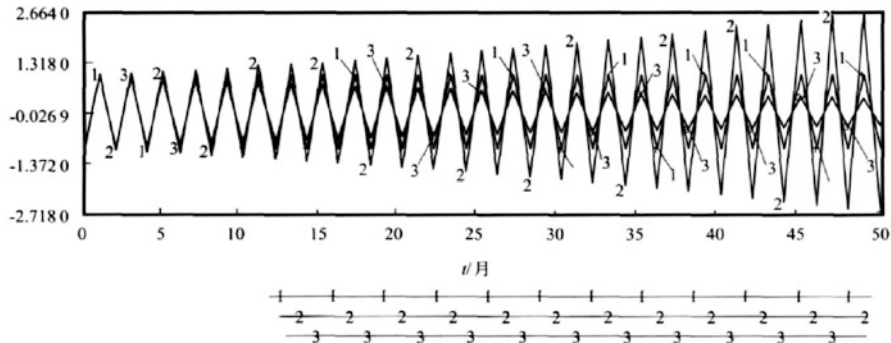
Fig. 3 The dynamic current of G without logistic delay

Its' characteristic equation is

$$D(\lambda) = |\lambda I - F|$$

$$= \begin{vmatrix} \lambda - 1 & 0 & -1 \\ 0 & \lambda - 1 & 0 \\ \frac{1}{T} & -\frac{1}{T} & \lambda + \frac{1}{T} \end{vmatrix} = 0$$

that is,

$$\lambda(\lambda - 1)\left[\lambda + \frac{1}{T} - 1\right] = 0 \tag{17}$$

Solve the Eq. 17, we get

$$\lambda_1 = 0, \quad \lambda_2 = 1, \quad \lambda_3 = 1 - \frac{1}{T}$$

According to the stability criterion of the system, the numeric area of the parameters making the system stable is $T > 0.5$.

- Considering the logistics delay (Y1 > 0, Y2 > 0)

When Y1 > 0, Y2 > 0, the system becomes a delay system. Theoretically, we can establish a new model to analyze the stability of the system, but the model will be too complicated. So in the case of the delay time being greater than zero, we will determine the conditions that make the system stable by computer simulation.

3.3 Simulation of the System Dynamics

Assume the simulation step is 1, that is, the order cycle is 1 month. When $Y1 = Y2 = 0$, the results of the simulation are shown in Fig. 3. In this figure, curve
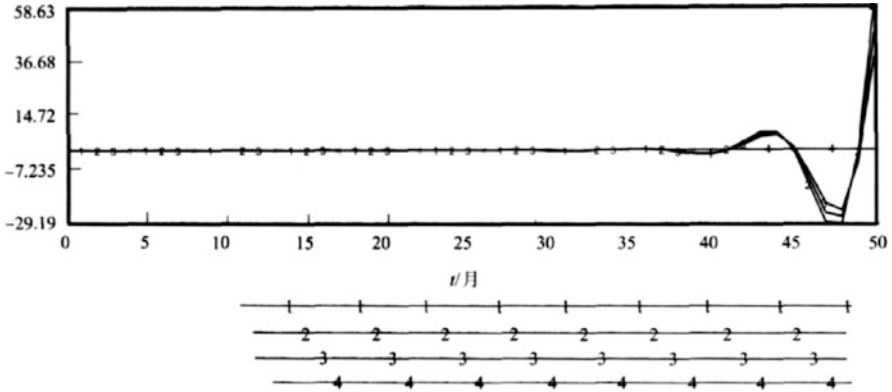
Fig. 4 The dynamic current of *G* with logistic delay

Table 1 The simulating schemes for different ordering strategy

	Order cycle/month		
T/month	1.0	0.5	2.0
1.0 × order cycle	SIMU 11	SIMU 12	SIMU 13
0.5 × order cycle	SIMU 21	SIMU 22	SIMU 23
2.0 × order cycle	SIMU 31	SIMU 32	SIMU 33

1 means T = 0.500, curve 2 means T = 0.495, curve 3 means T = 0.505. We can see that curve 1 shows the characteristics of continuous oscillation, curve 2 shows the characteristics of divergent oscillation, and curve 3 shows the characteristics of damped oscillation.

Considering the logistics delay, assume Y1 = 2, Y2 = 1, then we will get the results of the simulation shown in Fig. 4.

The results of the simulation above show that:

1. In the case of all the T are equal, the logistics delay will weaken the dynamic stability of the system, and even make the system unstable.
2. Lower the order adjustment range, extend the time of supply–demand balance, which will make the decision-makers have sufficient time to observe and respond to the changes of the differences of supply and demand, so as to effectively control the operation of the support system.

4 Optimization

Assume Y1 = 2, Y2 = 1, we design nine simulation programs (shown in Table 1).

Assume the simulation operation time is 200 months, then we get the variable G, the results are shown in Table 2.

Table 2 Simulation results of different order strategies (G)

T/month	Program								
	SIMU11	SIMU12	SIMU13	SIMU21	SIMU22	SIMU23	SIMU31	SIMU32	SIMU33
0	-1	-1	-1	-1	-1	-1	-1	-1	-1
20	0.782 072	-1305.11	9.() R9 37	1624.34	112.094	-62.9618	2.16133	24.6295	6.56251
40	691.083	$1.42225e^6$	0.223606	$-1.66287e^6$	$-1.33447e^{10}$	-1.12423	1.60067	-182.781	3.9074
60	1494.97	$1.19355e^9$	-3.12177	$1.70278e^9$	$-1.53998e^{15}$	980.841	0.870789	968.115	2.14066
80	-188258	$-1.33312e^{12}$	6.29596	$1.74365e^{12}$	$-1.73559e^{20}$	$1.22643e^6$	0.439056	1287.82	1.17452
100	-789833	$-1.09063e^{15}$	1.30944	$1.7855e^{15}$	$-1.91948e^{25}$	$6.89028e^7$	0.225845	-127737	0.644596
120	$5.05694e^7$	$1.24924e^{18}$	-3.96312	$-1.82835e^{18}$	$-2.08993e^{30}$	$2.67159e^9$	0.121157	$2.2789e^6$	0.353764
140	$9.96205e^{20}$	5.83422	$1.87223e^{21}$	$-2.24527e^{35}$	$8.17243e^{10}$	0.0666618	$-2.94689e^7$	0.194153	$3.19734e^8$
160	$-1.17032e^{24}$	-1.56286	$-1.91716e^{24}$	too big	$1.92257e^{12}$	0.0368958	$3.14047e^8$	0.106548	$-1.33691e^{10}$
180	$-9.09607e^{26}$	-4.10318	$1.96318e^{27}$	too big	$2.5901e^{13}$	0.0203648	$-2.79322e^9$	0.0584755	$-1.15254e^{11}$
200	$1.0961e^{30}$	5.75788	$-2.01029e^{30}$	too big	$-4.92356e^{14}$	0.0111942	$1.9295e^{10}$	0.0320892	$3.47304e^{12}$

By comparison, among the nine schemes above, SIMU 13, SIMU 31 and SIMU 33 can make the system stable. In which, SIMU 31 can make the system come to and maintain the supply–demand balance in short time, so SIMU 31 is the best.

5 Conclusions

- Excessive logistics delay may affect the stability of the system. If the logistics delay time is fixed, the stability of the system depends on the order strategy, nothing to do with demand. An improper order strategy may bring a larger supply–demand deviation which makes the system unstable.
- The stability and performance of the supply system of spare parts based on the "demand-pull" model are closely related to the supply management system and technical conditions, which is the results of the interactions between multiple factors. By comparing the simulation results of different programs, we can choose the optimal spare ordering strategy.

References

Bernhard JA, Marion CA (2000) System dynamics model in supply chain management:research review. In: Proceedings of the 2000 Winter Simulation Conferences, Orlando, pp 342–351

Clerk Maxwell J (1892) A treatise on electricity and magnetism, vol 2, 3rd edn. Clarendon, Oxford, pp 68–73

Eason G, Noble B, Sneddon IN (1955) On certain integrals of Lipschitz-Hankel type involving products of Bessel functions. Philos Trans R Soc Lond A247:529–551, references

Huiskonenj H (2001) Maintenance spare parts logistics: special characteristics and strategic choices. Prod Econ 71(1–3):125–133

Xiong Jingqi (2005) Control technology of computer. Publishing House of Electronica Industry, Beijing, pp 96–97, in Chinese

Study on High Sensitivity GPS Signal Acquisition Techniques Indoor Positioning

Liu Yu, Liu Ding-xing, and Tan Ze-fu

Abstract Global Position System is more and more close with people's life from start run till now. It has already been applied to various fields. Compared with the common circumstance, as outdoors, GPS signals power will be weaken much. So researches on HS-GPS indoor position technology are necessary in the extreme to implement high accuracy position in indoor environment. Weak GPS signals acquisition algorithm in low signal to noise ratio environment is studied, signals process flow and acquisition performance of algorithm, such as coherent integration, non-coherent integration and differentially-coherent integration. It can be seen from theoretical analysis and simulation that differentially-coherent integration is capable of detecting GPS signals in low signal to noise ratio environment.

Keywords GPS signals • Low signal to noise • High sensitivity • Detection

1 Introduction

GPS receiver has been widely used. Global Positioning System receivers must acquire and track the pseudorandom codes and carrier signals from several GPS satellites. In order to track and decode the information in GPS signal, an acquisition method must first be used to detect the presence of the signal (Lin 2002; Van Diggelen 2002). But the indoor circumstances include indoor places, forests and city environment and so on. The quality of GPS signals in the indoor circumstances is deteriorated seriously (Schon and Bielenberg 2008), the signal strength is less than −160dBW and the general receiver only is a sensitivity of −140dBW, and its reliability and position accuracy will be debased much (Jardak and Samama 2009;

L. Yu (✉) • L. Ding-xing • T. Ze-fu
Department of Electronic and Information Engineering, Chongqing Three
Gorges University, Chongqing 404000, China
e-mail: liuyusx@126.com; liudingxing@126.com; tanzefu@126.com

Singh et al. 2005). So researches on HS-GPS indoor position technology are necessary in the extreme to implement high accuracy position in indoor environment and extend application area of GPS, the researches on HS-GPS positioning technology have become hot points currently (Bickerstaff et al. 2006). To meet the requirement of acquisition weak GPS signal indoor, high sensitivity detection algorithms have been proposed. Weak GPS signals acquisition algorithm in low signal to noise ratio environment is studied, such as coherent integration (Dafesh and Fan 2001), non-coherent integration and differentially-coherent integration (Andreas and Andre 2005; Park and Kim 2002; Parviainen et al. 2008). It can be seen from theoretical analysis and simulation that differentially-coherent integration is capable of detecting GPS signals in low signal to noise ratio environment.

2 Model of Signal Acquisition DF Processing

The following is a model of GPS signal that has been used in this study (Chansarkar and Garin 2000):

$$y_k = Ad(t_k)c[(1 + \eta)(t_k - t_s)]\cos[w_{IF}t_k - (w_D t_k + \varphi_0)] + v_k \qquad (1)$$

In this model, the constant A is the signal amplitude. The function $d(t)$ is the GPS data stream, which is an unknown sequence of $+1$ and -1 with a period of 20 ms. The function $c(t)$ is the C/A PRN code of the received signal. The quality η is the fractional perturbation of the chipping rate due to Doppler shift, and t_s is the start time of the PRN code, which is a measure of its phase. The frequency w_{IF} is the intermediate frequency carrier to which the L1 carrier signal gets mixed in the RF front end, and w_D is the carrier signal's Doppler shift. Suppose the signals of from both channel I and channel Q are respectively correlated with the local C/A code, and then we can get the following general forms,

$$I_i = \frac{\sin(\pi \Delta f_i T)}{(\pi \Delta f_i T)} \sqrt{\frac{T_{COH}}{N_0}} AR(\tau_i)D_i \cos(\Delta \varphi_i) + I_{ni} \qquad (2)$$

$$Q_i = \frac{\sin(\pi \Delta f_i T)}{(\pi \Delta f_i T)} \sqrt{\frac{T_{COH}}{N_0}} AR(\tau_i)D_i \sin(\Delta \varphi_i) + Q_{ni} \qquad (3)$$

Where $R(\tau_i)$ is the C/A code autocorrelation function; Δf is the average frequency misalignment between the local estimation and the Doppler of lth satellite over the kthT_{COH}; $\Delta \varphi_i$ is the error of estimated carrier phase, and I_{ni} and Q_{ni} are random variables.

3 Acquisition Algorithm of Weak GPS Signal

3.1 Additional Coherent Integration Algorithm

The basic principle of coherent integration is to divide the received signals according to the length of C/A code and correspondingly overlie them into an overlying signals which have a period of length of one C/A code, and then do some relatively coherent convolutions and make a comparison between cumulative effect and the determined threshold value to decide whether the acquisition is successful or not. The process of acquisition is a two-dimensional search that is about Doppler Frequency Points and PN Code phase retardation. The functions of coherent integrate acquisition can express that:

$$P_{COH}(n, \hat{\omega}_D) = [\sum_{I=\text{Im}}^{\text{Im}+L-1} I_I(n, \hat{\omega}_D)]^2 + [\sum_{I=\text{Im}}^{\text{Im}+L-1} Q_I(n, \hat{\omega}_D)]^2 \qquad (4)$$

$I_I(n, \hat{\omega}_D)$ and $Q_I(n, \hat{\omega}_D)$ are respectively in-phase component and quadrate component of receiver of 1 ms integrate accumulation, $\hat{\omega}_D$ is the estimate of Doppler shift. When the maximum of P_{COH} is higher than the determined threshold, we can say that the GPS signals are succeeded in acquisition, now the f_d that the maximum is corresponding is the needed carrier Doppler shift, t_s is the time of code retardation.

The advantage of coherent acquisition algorithm is that it can obviously improve the signal to noise ratio, that's because the received signals have been coherently integrated r times, the signal to noise ratio is as r times as before. From this we can see: the more times and the longer time of coherent integration have, the more obvious signal to noise ratio. However, the time T_{COH} of coherent integration will affected by two factors, the first one: because the frequency of navigation is 50 Hz, and the data bit overturns every 20 ms, hence the time of coherent accumulation can't surpass 20 ms; the second: in the same condition of permitted acquisition loss, extend the time of coherent accumulation, which means reduce the step Δf of frequency dominos search and increase the times of frequency dominos search, it will far extend the search-time. On the choice of time, it's very difficult to make a compromise between development of process gain and search-time.

3.2 Non-coherent Integration Algorithm

Owing to the fact that the time of coherent integration can't be very long, otherwise it will limit process gain And then square the two branches signals of in-phase component and quadrate component which have been coherently integrated,

and then add them up, which can further develop the ratio of signal to noise. Non-coherent integration can be expressed as follows

$$P_{\text{NCH}}(n, \hat{\omega}_D) = \sum_{i=0}^{M-1} \left(I_i^2 + Q_i^2 \right) \tag{5}$$

The M is the section number of non-coherent integration. Non-coherent integration algorithm develop, the ration of signal to noise by means of increasing the data segment of coherent integration. Its process gain can rise by $10 \log \sqrt{M}$ dB, the time of coherent integration can't be limited and it can make the frequency domain step longer, and the search-time will shorter than coherent integration does.

However, the square and accumulation of non-coherent integration algorithm not only promote to increase the signal energy, but also increase the noisy energy and decrease the ratio of signal to noise, this is called the square loss. In the non-coherent integration algorithm, the square loss can disappear only in the condition of decent ration of signal to noise, usually it will exist. The worse the ration, the more serious square loss. Non-coherent integration is not very obvious to meliorate the ration. If we want to get a satisfied one, the times of non-coherent integration must be set very high and the time of search also can be longer.

3.3 Differentially-Coherent Integration Algorithm

Where we proposed use differentially-coherent combining instead of non-coherent combining of the coherent integration results, the methods of the differentially-coherent combining multiplies the coherent integration results from the two adjacent coherent integration data. It can be observed that the differentially-coherent scheme lessens the bit-transition problem with decreased squaring loss in comparison with the non-coherent combing of the COH methods. The unit decision variable for the adapted acquisition algorithm base on DF is denoted by P_{NCH}:

$$P_{NCH} = \left| \sum_{i=1}^{N} [(I_{i-1} + j\, Q_{i-1})^* (I_i + j\, Q_i)] \right|^2 \tag{6}$$

Where the operator (*) represents a complex conjugate operation. It was shown that this results in a lower combing loss within the verification stage of a double-dwell acquisition CDMA system than with non-coherent integration

The use of DF decision variables can bring many advantages. First of all, as to the GPS C/A code, if the time of coherent integration is a PRN code period, that is 1 ms, so 1 bit time of the GPS navigation data (20 ms) contains 20 results of coherent integration. Secondly DF is to avoid the square loss that caused by NCH. Compared with coherent integration, the acquisition algorithm based on DF is not

sensitive to navigation data bit transitions owing to the use of "differential". A longer coherent integration time is not required for the acquisition process owing to the "sum" after differenced. So, the acquisition algorithm based on DF has an advantage over coherent integration and non-coherent integration on the improvement of SNR.

4   Algorithm Experimental Simulation

The adapted acquisition algorithm based on non-correlation integration and differential correlation has been tested using actual GPS data. In the process of simulation, the sampling frequency is 16.368 MHz, the central frequency with which we does not take Doppler effect into account, is 4.092 MHz, Doppler frequency deviation is ± 10 K. The SNR of GPS signal that simulated is -20 dB$\smallfrown$$-40$ dB, which including three satellites, which are 6, 8 and 24 separately. The power of satellite which numbered 6, 8 is about -28 dB and 24 is merely -40 dB. The detection probability is $p_d = 0.9$, false probability is $p_f = 10^{-5}$, the time of coherent integration is 5 ms. The simulation is emphasized on numbered 8 satellites. Figure 1 shows the acquisition result of the numbered 8 satellite by traditional non-correlation integration. From Fig. 1 we can get: when the signal to noise ratio of GPS signals is -28 dB, the carrier's Doppler frequency and C/A code phase of the number 8 satellite can be succeeded in acquisition through 5 ms's coherent integration. Figure 2 shows the every frequency of the numbered 8 satellite by traditional non-correlation integration, when the SNR of GPS signal is -40 dB, from Fig. 2,we can not judge whether the C/A code phase and carrier frequency is right or not. Figure 3 shows the C/A code phase and every frequency of the numbered 8 satellite by the adapted acquisition algorithm based on differential correlation, when the SNR of GPS signal is -40 dB. From Fig. 3, we can find the C/A code phase is clearly 4,156.
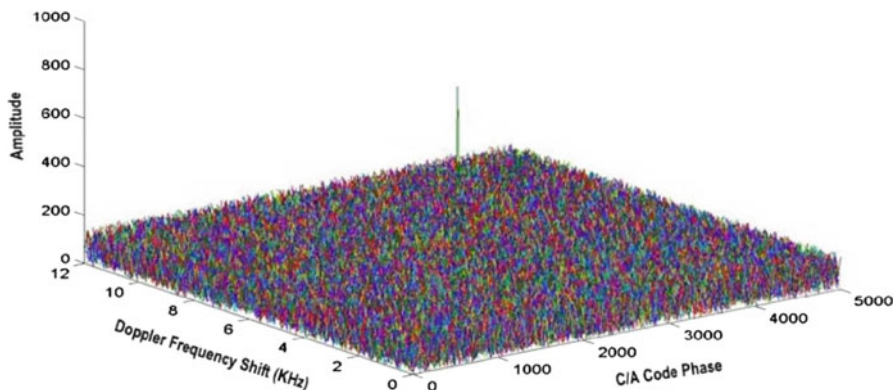
Fig. 1 Correlation result using non-correlation integration (SNR $= -28$ dB)
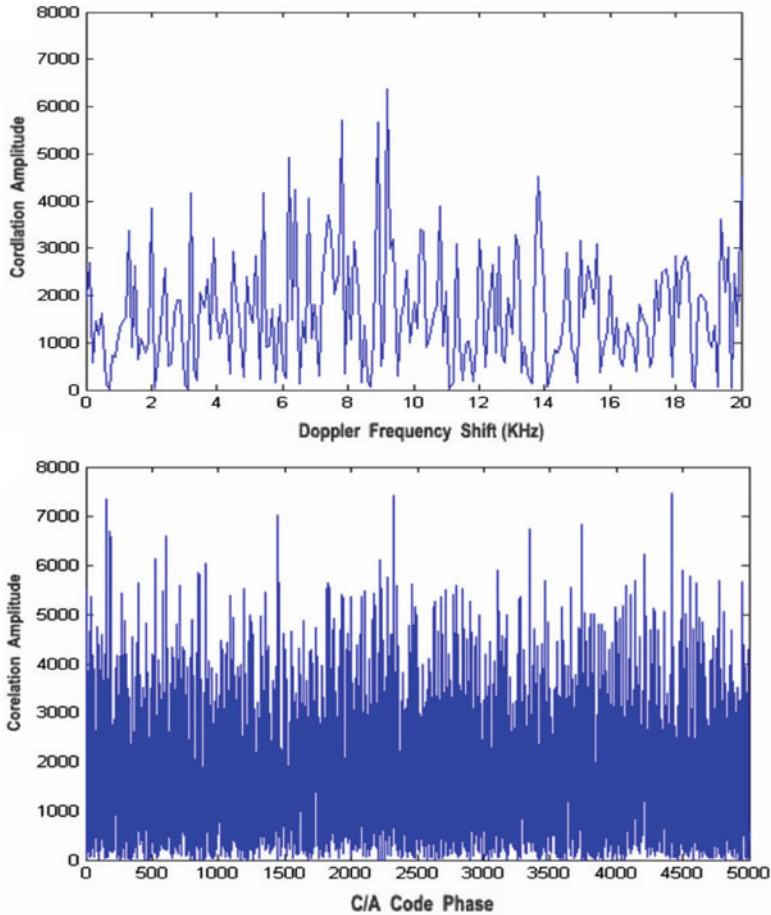
Fig. 2 Correlation result using non-correlation integration (SNR = −40 dB)

5   Conclusion

This paper present the adapted acquisition algorithm based on correlation integration, non-correlation integration and differential correlation integration. In terms of GPS signals acquisition technology of high sensitivity. On the aspect of improving the correlation peak value, the effects of coherent integration and non-coherent integration algorithm, they are basically similar, but the coherent integration is better. Differential correlation integration can acquire GPS signal with SNR −40 dB in a short time, and which avoid the square loss caused by non-correlation integration. So it can improve the detection sensitivity. In short,
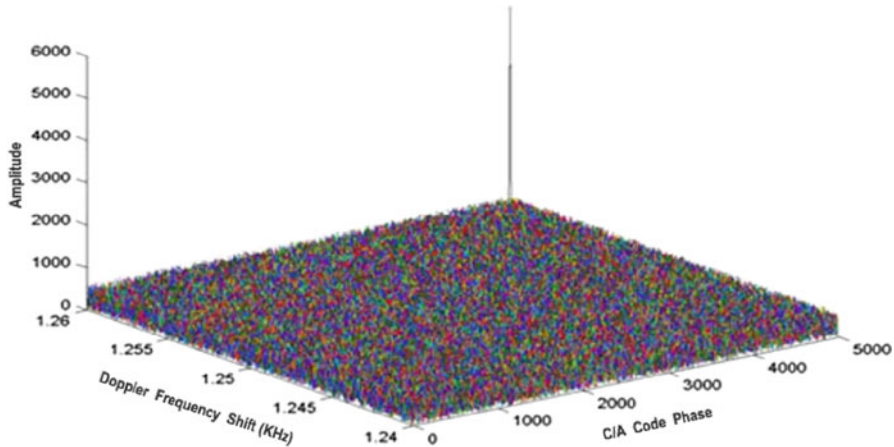
Fig. 3 Correlation result using differential correlation integration (SNR = −40 dB)

differential correlation integration has an advantage again the traditional coherent integration and non-coherent integration algorithm, this method is satisfied with indoor position in general.

References

Andreas S, Andre N (2005) Differential correlation for Galileo/GPS receivers. IEEE international conference on acoustics, speech and signal processing, vol 3. Philadelphia, pp 953–956

Bickerstaff J, Frayling R, Haddrell T (2006) Capture analysis and mitigation of multi-path in a high sensitivity GPS receiver. ION GNSS 2006, Fort Worth, TX, pp 1696–1705

Chansarkar M, Garin L J (2000) Acquisition of GPS signals at very low signal to noise ratio. In: Proceedings of the Institute of Navigation ION National Technical Meeting-2000, Anaheim, CA, pp 731–737

Dafesh PA, Fan T (2001) A novel approach to enhancing the sensitivity of C/A code GPS receivers. In: Proceeding of ION 57th annual meeting, Albuquerque, NM, pp 482–485

Jardak N, Samama N (2009) Indoor positioning based on GPS-repeaters: performance enhancement using an open code loop architecture. IEEE Trans Aerosp Electron Sys 45:347–359

Lin JB (2002) Sensitivity limit of a stand-alone GPS receiver and an acquisition method. In: Proceedings of ION GPS 2002, Portland, OR, The Institute of Navigation, pp 1663–1667

Park SH, Kim YB (2002) A novel weak signal acquisition scheme for assisted GPS. In: Proceedings of ION GPS 2002, Portland, OR, The Institute of Navigation, pp 1457–1467

Parviainen J, Kantola J, Collin J (2008) Differential barometry in personal navigation. Position location and navigation symposium, Monterey, CA, pp 148–152

Schon S, Bielenberg O (2008) On the capability of high sensitivity GPS for precise indoor positioning. Hannover, pp 121–127

Singh S, Cannon ME et al (2005) Field test assessment of assisted GPS and high sensitivity GPS receivers under weak/degraded signal conditions. ION GNSS 2005, Long Beach, CA, pp 2930–2943

Van Diggelen F (2002) Indoor GPS theory & implementation. IEEE position location and navigation symposium, Palm Springs, CA, pp 240–247

Research on Ontologies for Archives Knowledge Management

Meng Huang and Jun Luo

Abstract Within computer science, the term ontology is coined in the knowledge sharing and reuse effort, for efficient engineering of Knowledge-based systems. The archives knowledge is broadly distributed in organizations. They are heterogeneous and polymorphism. In this paper, we raise an overview of using ontology in archives knowledge management. Firstly, we discuss the ontology definition, ontology language and ontology engineering methodology. Secondly, we give the key process to implement ontological modeling of university archives knowledge. Finally, we introduce an architecture of ontology-based archives knowledge management.

Keywords Ontology • Archives knowledge • Ontology engineering methodology

1 Introduction

Knowledge Management (KM[1]) comprises a range of strategies and practices used in an organization to identify, create, represent, distribute, and enable adoption of insights and experiences. Such insights and experiences comprise knowledge, either embodied in individuals or embedded in organizational processes or practice.

An archive is a collection of historical records. Archives contain primary source documents that have accumulated over the course of an individual or organization's

---

[1]KM, http://en.wikipedia.org/wiki/Knowledge_Management.

M. Huang (✉)
Archives, Chongqing Three Gorges University, Wanzhou, Chongqing, China

College of Computer Science, Chongqing University, Chongqing, China
e-mail: jsjhcs@hotmail.com

J. Luo
College of Computer Science, Chongqing University, Chongqing, China
e-mail: lr@cqu.edu.cn

lifetime. University archives reflect their enduring culture and history. Now, the archives data resources increase rapidly. Because many resources are heterogeneous, communications among archives resources become more difficult. The semantic technology can supply more measures to solve the embarrassment.

2 Related Working

2.1 Ontology

The word "ontology" is used with different meanings in different communities. Firstly, we refer to a philosophical discipline, namely the branch of philosophy which deals with the nature and structure of "reality". Secondly, it reflects the most prevalent use in computer science. We refer to an ontology as a special kind of information object or computational artifact.

In 1993, Gruber originally defined the notion of an ontology as an "explicit specification of a conceptualization" (Gruber 1993). In 1997, Borst defined an ontology as "a formal specification of a shared conceptualization" (Borst 1997). This definition additionally required that the conceptualization should express a shared view. Also, such conceptualization should be expressed in a formal machine readable format. In 1998, Studer et al. merged these two definitions stating that "An ontology is a formal, explicit specification of a shared conceptualization."

2.2 Ontology Language

Ontology languages allow users to write explicit and formal conceptualizations of domain models. The main requirements are a well-defined syntax, efficient reasoning support, a formal semantics, sufficient expressive power and convenience of expression. We state the fashionable ontology languages, such as RDF, RDF(S), and OWL. The languages are recommendation by W3C.

Although often called a "language", RDF[2] (Resource Description Framework) is essentially a data model. Its basic element is an object-attribute-value triple, called a statement. The fundamental concepts of RDF are resources, properties, and statements. RDF schema (RDFS[3]) make assumptions about any particular application domain, which define the semantics of any domain.

---

[2]RDF. http://www.w3.org/TR/rdf-primer/.

[3]RDFS. http://www.w3.org/TR/rdf-schema/.

The Web Ontology Language (OWL[4]) is characterized by formal semantics and RDF/XML-based serializations for the Semantic Web. OWL, is endorsed by the World Wide Web Consortium (W3C), inherits most features from DAML + OIL. OWL was defined as three different sublanguages, each of which is geared towards fulfilling different aspects of these incompatible full set of requirements: OWL-Lite is designed for a classification hierarchy and straightforward constraint features; OWL-DL supports users who want the maximum expressiveness while retaining computational completeness and decidability, and OWL-Full has useful computational properties for reasoning systems with maximum expressiveness, but without computational guarantees.

2.3   Ontology Engineering Methodology

During the development of ontology engineering methodology, engineers put forward many mature methods, but these methods almost aim at special fields and haven't uniform standard. When we design ontology, familiar methods are as follows: TOVE (Toronto Virtual Enterprise), special for constructing TOVE ontology (about enterprise modeling ontology) (Gruninger and Fox 1995); METHONTOL-OGY mainly use for constructing Chemistry ontology (Fernandez et al. 1997); ENTERPRISE is for constructing ontology of enterprises modeling (Uschold and Gruninger 1996); IDEF5 is used to describe and get enterprise ontology, IDEF is acronym of "Integration Definition for Function Modeling"(IDEF5 Method Development Team for the IICE Methods Engineering Thrust); seven-step method has been designed by School of Medicine, Stanford University (Noy and Mcguinness 2002).

3   Archives Ontology Model

The common modeling tools are WebOnto, Protégé, OntoEdit and so on. Refer to seven-step method when we have constructed the ontology of university archives knowledge. The process is as follow:

1. Determine areas scope and aim of domain ontology;
2. List important terms in the ontology and define classes and class hierarchy;
3. Define class attributes, their hierarchy and characters;
4. Create instances and fill in attributes values of the instances;
5. Ontology. If the design is not up to the mustard, and we must analyze and design over again.

---

[4]OWL. http://www.w3.org/TR/2004/REC-owl-guide-20040210/.

Fig. 1 The development
process
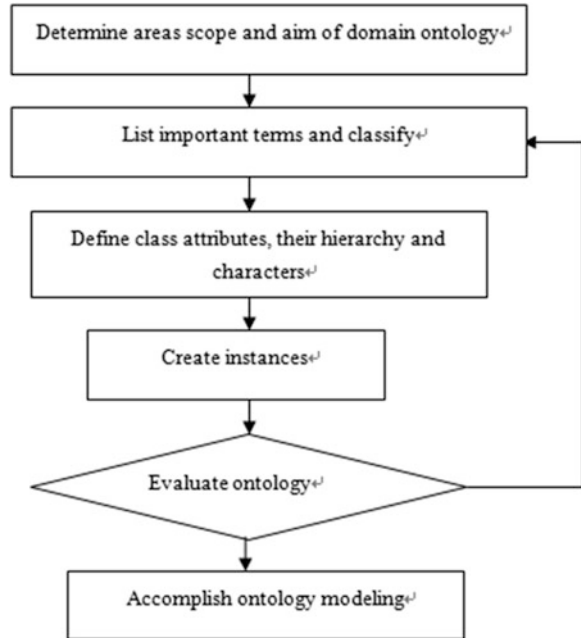
As mentioned above, the ontology development is necessarily an iterative
process. Figure 1 denotes the development process.

3.1 Listing the Concepts

We build the ontology in term of university archives entity taxonomy and archival
experts. Our aim is to assist archivist to pigeonhole exactly. According to this
purpose, we list the concepts as follow: record, record office, archives, retention
period, fonds, fonds constituting unit, file, file note, archival code, folder list, inner
file item list, file-item list and so on. And then, we define classes and class hierarchy.

3.2 Relationship Among Concepts

Within OWL (Web ontology language), relations are described by properties of
concepts. Two types of properties are object properties and datatype properties. The
object properties are relation between instances of two classes. The datatype prop-
erties are relations between instances of classes and RDF literals and XML Schema
datatype. Object properties are basically denoted by the verbs. Within university
archive ontology, object properties are "Include" and "included", "express", "make
up of" and so on. Datatype properties are archival code value, date, time, fonds code
value and so on.
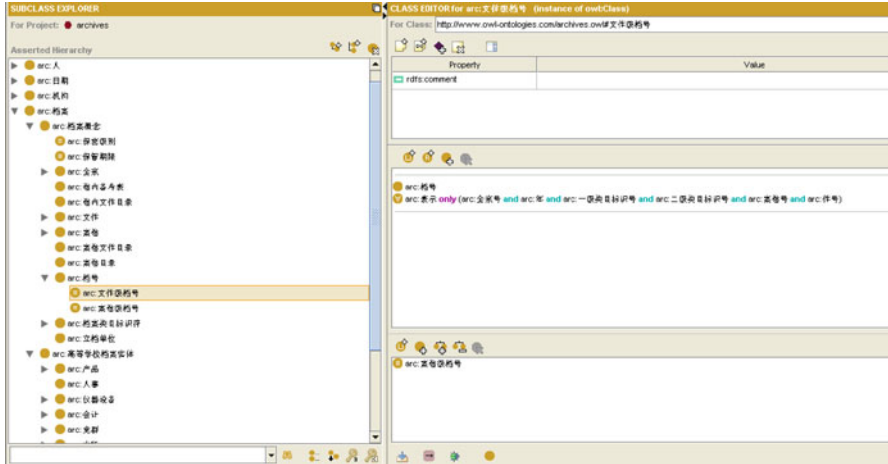
Fig. 2 University archives ontology

3.3   Instantiation

An instance (individual) is minimally introduced by declaring it to be a member of class. Within university archive ontology, department subclass of organization class has many individuals. For example, college of computer science, school of foreign language, school of economics and management and so on. The instances of archival retention period are "Permanent", "long term", "short term", "30 years" and "10 years". Individuals of archival security classification are "top secret", "general secret" and "important secret".

   To construct the ontology of university archive ontology, we use ontology editor software Protégé.[5] Modeling language is OWL, which is recommended by W3C. We store the ontology by OWL-based file format. Figure 2 shows the modeling.

4   Architecture of Ontology-Based Applications

Figure 3 shows the architecture of ontology-based applications. It composed of the ontology store layer, the ontology middleware layer, ontology modeling layer and information access layer. The architecture addresses all the key stages of the knowledge management lifecycle.

---

[5]Protégé. http://protege.stanford.edu/.
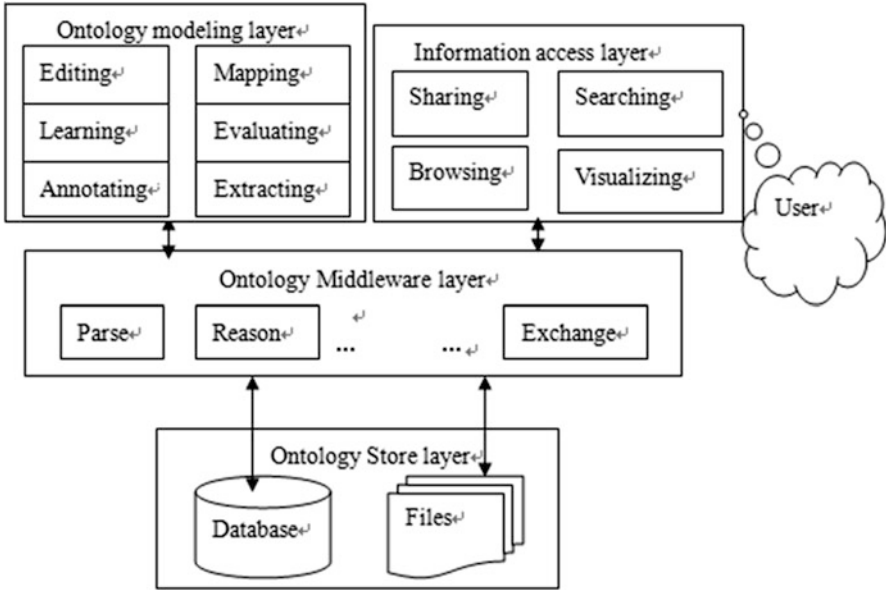
Fig. 3 Architecture of ontology-based applications

4.1 Ontology Store Layer

Ontologies are often used to improve data access. For this purpose, existing data has to be linked to an ontology and appropriate access mechanisms have to be provided. Now we adopt the relational databases and files to store ontologies.

4.2 Ontology Middleware Layer

Ontology middleware is required with support for development, management, maintenance, and use of knowledge bases. We can parse, reason and exchange ontologies by special middleware software. The layer communicates the information between the store layer and modeling layer or information access layer.

4.3 Ontology Modeling Layer

In the layer, knowledge acquisition is main function. The knowledge engineer needs to be supported by ontology editing tools which create and maintain ontologies. The large amounts of unstructured and semi-structured information exist in organization,

so automatic knowledge extraction from unstructured and semi-structured data in external data repositories is required. The ontology mapping, evaluation, and learning are significant in lifecycle of ontology engineering.

4.4 Information Access Layer

Finally, and perhaps most importantly, information access tools are required to allow end users to exploit the knowledge represented in the system. The facilities mostly apply to finding, sharing, visualizing, and browsing knowledge.

5 Conclusion and Future Works

As above mention, we give an overview of using ontology in archives knowledge management and realize a case study: ontology of university archives. We put forward an architecture of ontology-based applications. Future works may include constructing more sub-domain ontologies of university archives knowledge. More efficient and persistent storage method of ontology model can be another research point in the future.

References

Borst W (1997) Construction of engineering ontologies. PhD thesis, Institute for Telematica and Information Technology, University of Twente, Enschede, the Netherlands

Fernandez M, Gomez-Perez A, Juristo N (1997) Methontology: form ontological art towards ontological engineering, AAAI-97 spring symposium on ontological engineering, Stanford University, Stanford, 24–26 March 1997

Gruber TR (1993) A translation approach to portable ontologies. Knowl Acquis 5:199–220

Gruninger M, Fox MS (1995) Methodology for the design and evaluation of ontologies. Workshop on basic ontological issues in knowledge sharing, IJCAI-95, Montreal

IDEF5 Method Development Team for the IICE Methods Engineering Thrust. IDEF5 method report. http://www.idef.com/pdf/Idef5.pdf

Noy NF, Mcguinness DL (2002) Ontology development 101: a guide to creating your first ontology. http://protege.stanford.edu/publications/ontology-development/ontology101.pdf

Studer R, Benjamins R, Fensel D (1998) Knowledge engineering: principles and methods. Data Knowl Eng 25:161–198

Uschold M, Gruninger M (1996) Ontologies principles, methods and applications. Knowl Eng Rev 11:93–136

Dynamic Analysis of Terrorism Using Artificial Intelligence Techniques: A Case Study

Daijun Zhong and Guiqiang Hu

Abstract Apply intelligent information technologies to anti-crime and anti-terrorism is a hot topic in security bureaus of different countries. Social Network Analysis is a promising method to analyze terrorist network structures. Yet, the method is inadequate to process the dynamic property of terrorist networks. This paper proposes an approach combining Agent simulation and Social Network Analysis so that the dynamic property of terrorist networks can be handled. We collect criminal data to construct numbers of virtual terrorist networks. A series of experiments have been conducted using real terrorist network data to verify the effectiveness of the proposed approach. The experimental results show that our approach can use to evaluate a number of strategies for destabilization of terrorist organizations.

Keywords Terrorist • Organization • Artificial Intelligence • Simulation

1 Introduction

By the increasing threat of terrorism and spread of terrorist organizations, it is so important to understand the properties of such organization and to develop successful strategies to destabilize them.

A terrorist network is primarily a social network in which individuals connect with one another through various relations such as kinship and coworkers.

D. Zhong (✉) • G. Hu
Modern Education Technology Center, Chongqing University of Arts and Science,
Chongqing 402160, China
e-mail: daijunzhong@163.com

It suggests that social network analysis (SNA) (Wasserman and Faust 1994) will be useful in characterizing the structure and property of the terrorist organization. The problem of network destabilization can be partially solved by adopting SNA methods.

A terrorist organization is dynamic system and keeps changing overtime. The ideal way to capturing dynamics of terrorist networks is replicating the criminal domain in the real world many times and testing the situation in the replicated environments. However, these techniques are often expensive, unethical or impossible to complete. Multi-agent systems can serve as effective tools for reasoning about human and group behavior. Its high computing performance of MAS (Multi-agent system) (Moss and Davidsson 2001) simulation allows multiple experiments to be ran for many times with less cost.

In this paper, we integrated the methods from two different domains: social network analysis and multi-agent simulation. We collect criminal data to construct numbers of virtual terrorist networks. SNA measures are used to identify key persons in a terrorist network. Impact of the strategies is tested using a multi-agent simulation model. A large number of simulations run with different scenarios are done based on the terrorist networks to verify the effectiveness of the proposed approach. The experimental results show that our approach can find the useful strategies to destabilize the terrorist organization.

2 Background and Previous Research

2.1 SNA in Terrorist Network Analysis

SNA is used in sociology research to analyze patterns of relationships and interactions between social actors in order to discover the underlying social structure (Wasserman and Faust 1994; Scott 1991). Recently, SNA has been recognized as a promising technology for studying criminal organizations (Sparrow 1991; Klerks 2001; McAndrew 1999). In SNA studies, a network is usually represented as a graph which constrains a number of nodes (network members) connected by several links (relationships).

Using SNA methods to analyze criminal or terrorist networks, researchers have many findings. For example, Faulkner analyzed an illegal network and found that individual centrality in the network, as measured by degree, betweenness and closeness (Freeman 1979), is an important predictor of an individual's position in the organization (Carl Baker and Faulkner 1993). Krebs studied the terrorist network about the 19 hijackers in 9/11 events by SNA method and he found that Mohamed Atta, had the highest degree and acted as the ring leader of the network (Krebs 2001).

Traditional SNA are static analysis in which social networks are constructed at one point in time. With the static view of the criminal network, the analysis is limited.

2.2   *Simulation Models of Human Networks*

Terrorist network can be viewed as one kind of human networks. There are serial models introducing dynamic qualities to network analysis. The model System Dynamics study the overall system behavior by different equations modeling processes. It can achieve very high degree of proximity to reality on the overall system level. But the individual entity's behavior is not very understandable (Karnopp and Margolis 1990). Model Blanche combines modeling of social structure and individual attributes. Its drawback lies in the simplified model when applying in the real world problems (Contractor and Monge 2003). Another model developed by Snijders et al. simulates behavior of members and evolution of an organization. They used the model to get a deeper understanding of the relation between individual behavior and actions in social structures. To simulate an organization, they represented the organization as a social network and used measures known in the social network to represent the behavior of members. This work shows an example of how agent-based simulation approach is used in social studies (Snijders et al. 2007).

3   Research Methods

We examined many different networks ranging from social activist groups to violent terrorist groups and found that the Human network's structure are often adaptive and capable of rapid response. Organizations with this structure, particularly more distributed and decentralized ones, evolve continuously. That is to say, organization can be destabilized by altering their structure. In our situation, to capture the dynamic property of terrorist organization, we provide an approach based on a multi agent model CrimAS in the social network analysis perspective.

By integrating the Multi-agent system and SNA methodology, we can see the terrorist organization evolves and how its structure influences its performance by simulate the process. Therefore, we introduce (1) social networks in the model, (2) the agent-based model of simulated network, (3) agent architecture in the model, (4) task expanding and decision algorithm, (5) construction of terrorist network, (6) measures in the model.

3.1   Social Networks in the Model

In CrimAS terrorist networks are represented as a directed graph consisted of Nodes and Edges. We give the definition:

$$SN = (V, E)$$
$$V = a_i : \text{ set of agents nodes}$$
$$E = \left\{ (i, j) \left| LD [i] [j] \neq 0, \forall a_i, a_j \in V \right. \right\}$$

In the definition, *SN* is a two-tuples. *V* are sets of agents nodes, *E* is set of links between nodes. Where *LD*[*i*][*j*] represent the weight of the link. If *LD*[*i*][*j*] = 0, the node *i* and node *j* have no connection. The node and links have many attributes associated with terrorist domain.

3.2   The Agent-Based Model of Simulated Network

To simulate the activity of a terrorist organization, we use the artificial intelligence techniques enable the agents in the network to reason and plan the strategies to maximize the utilities. The methodology of our multi-agent model is based on the following principles:

- The model consists of a population of agents that can reason and plan.
- There is no rule that dictates global behaviors.
- Agents form briefs about the world and other agents from interactions.
- Agents planning the strategies of communication base on the mechanism of maximizing their utilities.
- When the network is damaged, agents have the capability to rebuild the network by creating new links.

3.3   Agent Architecture in the Model

Agent in CrimAS is primarily BDI (Brief-Desire-Intention) architecture (Bratman and Israel 1987). In the initial statement, an agent form beliefs about itself and other agents from the predefined input matrices. The input matrices which come from the empirical domain date including the human agent, knowledge, resources, task assignments and their attributes information. It can be either a binary matrix or non-binary one. The desire of agent is to complete tasks. Task in CrimAS may have subtasks and requirement such as related knowledge and resource. So we have to expand the task and made decision to solve the conflicts. Therefore, the intension and planning are formed. Then agent will communicate the agents with the related knowledge and resource by query them. And query the agents with
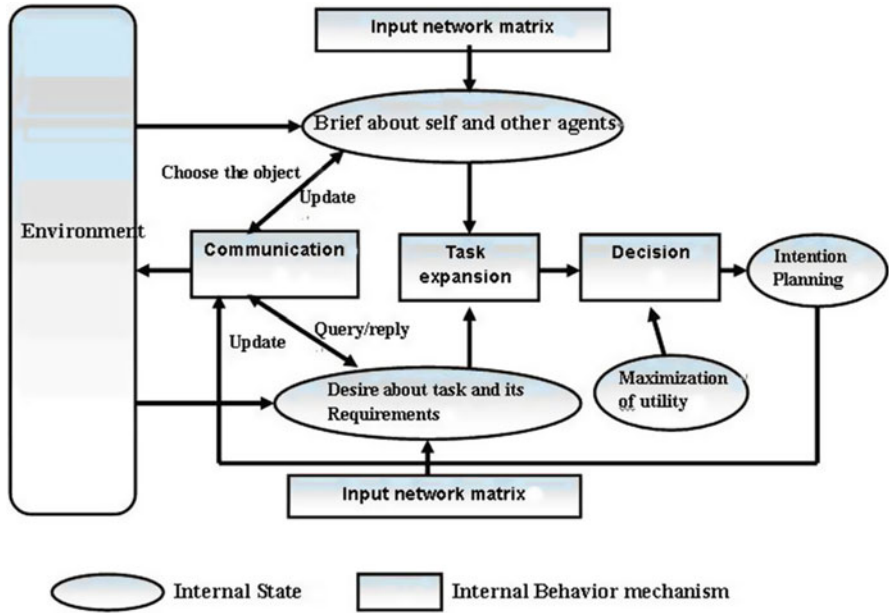
Fig. 1 Agent architecture in CrimAS

different social position will have different costs. When tasks are completed through the communication and knowledge transfer the brief and desire will be updated (Fig. 1).

3.4 Task Expanding and Decision Algorithm

In the previous researches, an organization is claimed as an information processing system structured to achieve a series of tasks. In such organizations, tasks are described as sequences of communications and actions. In terrorist domain the task often have subtask. Because of this structure, we use the Hierarchical Task Network (HTN) (March 1988) which is developed by the AI planning community to lead the tasks into an executable precedence network. In HTN tasks which can be immediately carried out if the knowledge and resource requirement are met is called primitive tasks. The task expanding works by choosing an expansion for the non-primitive tasks and resolving conflicts iteratively until a conflict free plan consisting of primitive tasks can be found. We used the algorithm generalized by Daijun Zhong and Guiqiang Hu (Erol and Hendler 1994).

In CrimAS model, the task arrangement decision is based on the mechanism of maximizing their utilities. Because the decision problem in our model is continuously, the function of utility cannot be evaluated by one state, but a series of states-Environment history.

The problem in CrimAS can be defined as a Markov decision process (MDP) (Astrom 1965). It is consist of S0, T(s,a,s') and R(s). Where S0 is the initial status, T(s,a,s') represent the probability of get to status s' when complete the action a in status s, R(s) is the reward function. We use the iteration Richard Bellman equation (Bellman 1957) to calculate the utilities, the iteration step of update is:

$$U_{i+1}(s) \leftarrow R(s) + \gamma \max_a \sum_s T\left(s, a, s'\right) U_i\left(s'\right)$$

The algorithm of Richard Bellman update is described as following:

```
function Utility-iteration(MDP,ε)
input: MDP, an Markov decision process with states S,
transition model T, reward function R, discount γ, the
maximum error allowed in the utility of any state ε
local variables: U,U', vectors of utilities for states
in S, the maximum change in the utilities of any state δ
repeat
U = U', δ = 0
for each state s in S do
```

$$U'(s) \leftarrow R(s) + \gamma \max_a \sum_s T\left(s, a, s'\right) U\left(s'\right)$$

```
if | U' (s) - U (s)|> δ then δ ← | U' (s) - U (s)|
until δ < ε(1-γ)/γ
return U
```

4 Results

4.1 *The Analysis of Different Key Persons' Isolations*

The input of the test is the terrorist network we constructed. This paper has the purpose of analyzing the difference between isolating the key persons with high centrality and the persons with high knowledge exclusivity. Therefore, we generate the scenarios based on the key persons identified by the two SNA measures. After the scenarios are generated, we run the simulation by isolate a set of agents at a designated time in the agent-based environment to test how the network will evolve. Because it is a discrete event simulation, we can measure agents' degree of knowledge diffusion at a certain simulated time-point to analyze how the performance changed according to the scenarios. So we can evaluate which strategy is better.
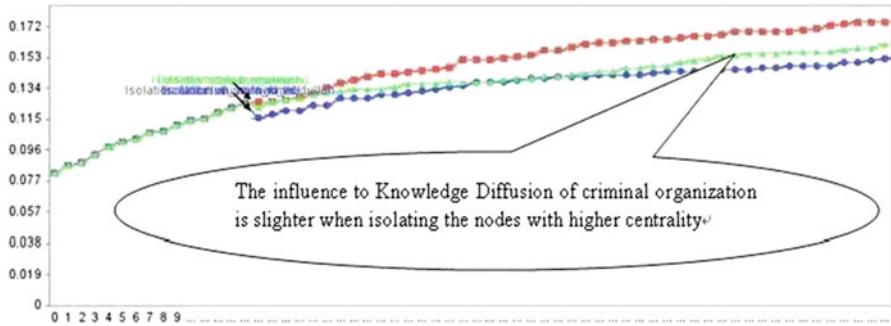
Fig. 2 The evolving knowledge diffusion rate over time under two strategies
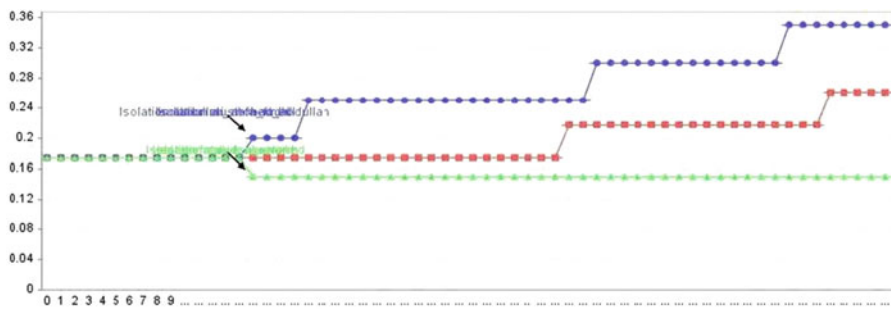
Fig. 3 The weapon experience knowledge diffusion rate over time under two strategies

By calculating the 21 agents' centrality and relative knowledge exclusivity in the terrorist network we constructed, we devise two isolation strategies with 65 time points and isolate the two groups of agents at point 15.

After running the simulation by isolate a set of agents at the designated time in the agent-based environment, we measure agents' degree of knowledge diffusion each certain simulated time-points to analyze how the performance changed according to the strategies.

As Fig. 2 showing, the impact to knowledge diffusion of the terrorist organization is slighter when isolating the agents with higher centrality.

However, the situation change when we dill down on the weapon experience diffusion. When isolating the agents with higher centrality the weapon experience knowledge diffusion decreasing greatly and cannot recover (Fig. 3).

Otherwise, when evaluating the task accuracy (Schreiber 2006) we find the impact to terrorist organization is slighter when isolating the agents with higher centrality at beginning. After some adjustment of the terrorist organization, the
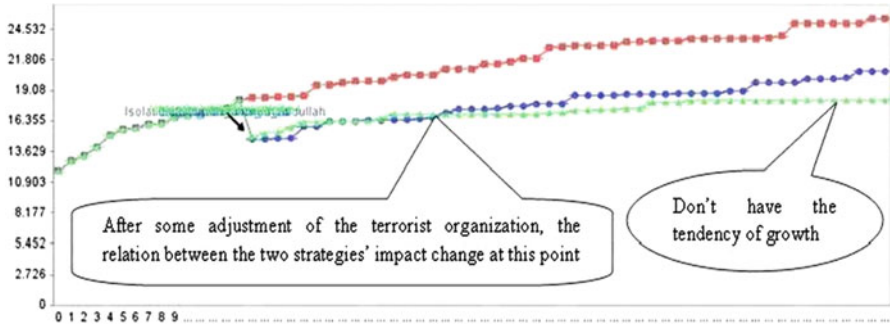
Fig. 4 The evolving task accuracy over time under two strategies

relation between the two strategies' impacts change at time point 25. The impact to strategy isolating the agents with higher centrality become obvious and the task accuracy has the tendency of un-growth (Fig. 4).

From the analysis above all, we can come to the conclusion: As to the terrorist network we constructed, we can get the useful strategies to destabilize the terrorist organization. (1) When disrupting the communication of the terrorist network, we should pay more attention on the persons with higher knowledge exclusivity. They are sometimes the expert or technologist of the organization. (2) When disrupting the weapon flow of the terrorist network, we should focus on the persons with higher centrality. They are sometimes the key actor of the organization. (3) When disrupt the ability of task completion of the terrorist network, select the persons with higher centrality to isolate is a better choice and the suppression last longer.

4.2 The Analysis of Network Structure

In order to identify the importance of overall structure in terrorist network analysis, we used another dataset which is gathered by John Gagnon in 2003.[1] It is a network including 67 prisoners in one prison. John Gagnon investigated them for years and formed a network using the information about them. It is a typical star network according to the visualization (Fig. 5).

We device two strategies about isolate the persons with highest centrality and random ones to the John Gagnon's criminal network and U.S. embassy bombing network. We set the same time points and replicate number and do procedure like Sect. 4.1. The result of knowledge diffusion rate changing over the time is showing in Figs. 6 and 7. As we can see, the knowledge diffusion rate becoming higher when isolated the random person in both the two networks. Obviously, the impact of

---

[1]http://vlado.fmf.uni-lj.si/pub/networks/data/Ucinet/UciData.htm 2008.
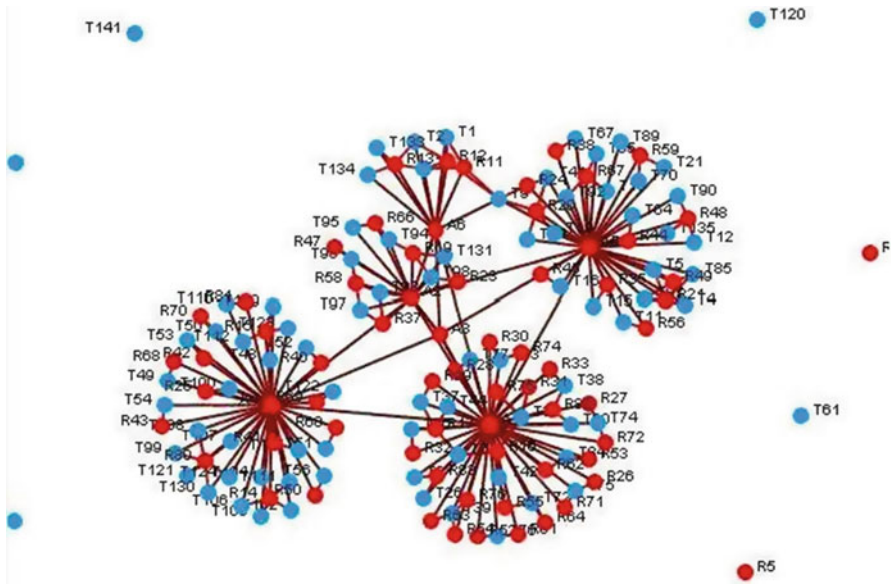
Fig. 5 The visualization of the John Gagnon's criminal network
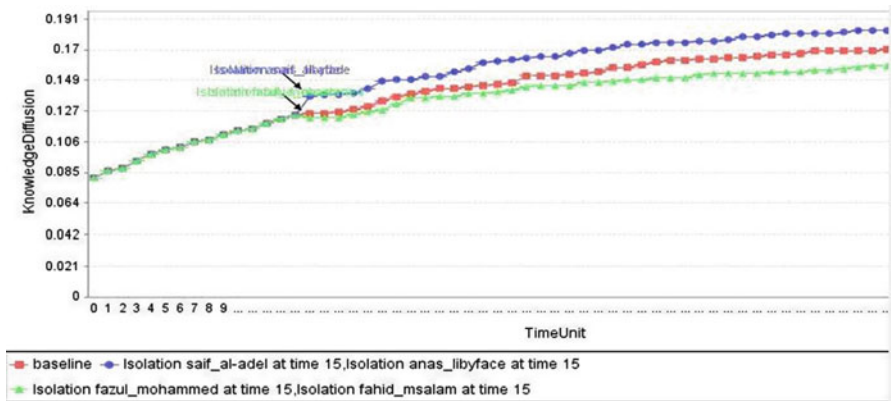
Fig. 6 The evolving knowledge diffusion rate over time in U.S. embassy

strategy using on John Gagnon's criminal network are more serious when isolating the persons with high centrality. Thus, we can draw the conclusion (1) Dispersal networks such as the U.S. embassy bombing one are very flexible because their complexity. When destabilize those networks, we should void isolating the persons without high centrality firstly. In actually, removal those nodes in the network will make it become more efficient. (2) As to the star networks, persons with high
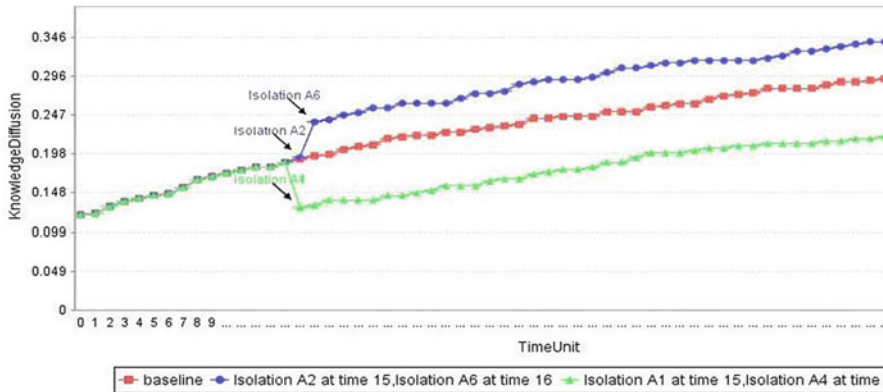
Fig. 7 The evolving knowledge diffusion rate over time John Gagnon's criminal network

centrality pay an extremely important to information flow. They are very vulnerable. The law enforcement could disrupt the criminal network to star structure step by step and destabilize it with low cost.

5 Conclusions

In this paper, we integrated the methods of social network analysis and multi-agent simulation to analyze the virtual terrorist networks we constructed. We use SNA to select the nodes to isolate. Then, CrimAS, a multi-agent simulation model, is used to evolve the network with and without the selected nodes. Experiments verify the effectiveness of the proposed approach. The experimental results show that the approach can find the useful strategies to destabilize the criminal organization.

We get the isolation strategies based on the key actors of a terrorist network. Future work is to validate whether the person captured by SNA is really the key role in the organization. Furthermore, more effective methods for anti-crime and anti-terrorism can be extended based on this framework.

References

Astrom KJ (1965) Optimal control of Markov decision processes with incomplete state estimation. J Math Anal Appl 10:174–205

Bellman R (1957) Dynamic programming. Princeton University Press, Princeton

Bratman JM, Israel DJ (1987) Toward an architecture for resource-bounded agents. Technique report CSLI-87-104, Center of Study of Language and Information, SRI and Stanford University, Canada

Carl Baker WE, Faulkner R (1993) The social organization of conspiracy: illegal networks in the heavy electrical equipment industry. Am Sociol Rev 58(12):837–860

Contractor NS, Monge PR (2003) Using multi-theoretical multi-level (MTML) models to study adversarial networks. In: Breiger R, Carley K, Pattison P (eds) Dynamic social network modeling and analysis, workshop summary and papers. The National Academic Press, Washington DC

Contractor NS, Monge PR (2009) Using multi-theoretical multi-level (mtml) models to study adversarial networks. In: Breiger R, Carley KM (eds) Summary of the NRC workshop on social network modeling and analysis. National Research Council

Erol K, Hendler J (1994) Semantics for hierarchical task network planning. Technical report CS-TR-3239, UMIACS-TR-94-31, Computer Science Dept., University of Maryland, March

Freeman LC (1979) Centrality in social networks: I. Conceptual clarification. Soc Netw 1: 215–239

Karnopp D, Margolis D (1990) System dynamics: a unied approach. Wiley, New York

Klerks P (2001) The network paradigm applied to criminal organizations: theoretical nitpicking or a relevant doctrine for investigators. Connection 24(3):53–56

Krebs VE (2001) Mapping networks of terrorist cells. Connection 24(3):43–52

March J (1988) Decisions and organizations. Basil Blackwell, Oxford

McAndrew D (1999) The structural analysis of criminal networks. In: Canter D, Alison L (eds) The social psychology of crime: groups, teams, and networks, Aldershot, Dartmouth. Offender Profiling Series, III, pp 53–94

Moss S, Davidsson P (2001) MABS 2000. LNCS (LNAI), vol. 1979. Springer, Heidelberg

Schreiber C (2006) Human and organizational risk modeling: critical personnel and leadership in network organizations. PhD thesis, Carnegie Mellon University, CMU-ISRI-06-120, USA

Scott J (1991) Social network analysis. Sage Publication, London

Snijders T Steglich, Christian EG, Schweinberger M (2007) Modeling the co-evolution of networks and behavior. In: van Montfort K, Oud H, Satorra A (eds) Longitudinal models in the behavioral and related sciences. Lawrence Erlbaum, Mahwah, pp 41–71

Sparrow M (1991) The application of network analysis to criminal intelligence: an assessment of the prospects. Soc Netw 13:251–274

Wasserman S, Faust K (1994) Social network analysis: methods and applications. Cambridge University Press, Cambridge

Optimization Used in the Collision Problems and Their Application

Tan Ying

Abstract In this paper, optimization methods are provided in order to solve the collision problems. Exactly collision time and reaction pulse could be solved using optimization methods which greatly improve the animation quality and computation time. And these methods also could simulate Newtonian physics and Coulomb friction. Moreover, unlike other papers, this one also tries to elaborate how to practically use these methods in some concrete examples.

Keywords Optimization • Collision time • Reaction pulse application

1 Introduction

Collision problems, including collision detection and collision reaction, have attracted people's attention for quite a long time. And they are still ones of most popular topics in the area of computer graphics. See (Aliyu and Al-Sultan 1994, 1995) for a survey of some solutions of these various algorithms. Many researchers have designed excellent simulation techniques for various applications. Generally speaking, some important should be satisfied in these methods, such as physical realism and numerical stability, and of course computationally efficient. But like many other problems and their solutions, sometimes one aspect should be paid less attention in order to improve others. Here, optimization method used in solving collision problems will be presented because it can greatly accelerate the speed of computation. Also a practical application, cloth simulation (Jingde et al. 2009), which will use the collision routine, will be provided in order to prove this method's efficiency.

T. Ying (✉)
Network Information & Education Technology Center,
Chongqing Three Gorges University, Chongqing, China
e-mail: ctanying@vip.qq.com

The following section will first give a brief introduction about collision problems. Section 3 will describe the optimization method and section 4 will talk about how to and the result of using these methods in a cloth simulation instance (Jingwei et al. 2010).

2 Collision Problems

Basically speaking, given two objects, how I should check whether they intersect with each other, if they do not interpenetrate each other, how far are they apart. And if overlap, how much is the amount of penetration. From these questions it is easy to find out that collision problems are everywhere. In the animation, when a policeman shoot a bullet; when a player kick a ball and so on. Therefore, collisions construct our world. If researchers want to simulate a realistic scene of our daily life, or even a virtue scene like a cartoon (Jin-ming et al. 2010), collision constraint has to be considered.

It is straightforward to understand what collision is: no penetration, however to be more exactly, more things need to be considered. The following aspects are important:

1. Find the exact collision point; this could be dependent on the resolution of the screen.
2. Two dimension or three dimension collision.
3. Collision objects are convex objects or non-convex objects.
4. Polygonal or non-polygonal objects.
5. Objects are open surfaces or closed volumes;
6. Objects are rigid bodies or deformable ones.
7. Objects are made in pair wise or multiple ones.

After listing these considerations, we may find that collision problems are not toy problems.

Generally speaking, collision problems can be divided into two categories, collision detection and collision reaction; collision detection is a routine used to detect the collision point (Jiu-feng et al. 2010): both the collision geometric information and collision temporal information. Because any collision could be boiled down to edge and edge or vertex and edge collisions, this routine could be simplified to check these collisions in the near future; collision reaction is used to give a response after detecting a collision—a soccer ball cannot immediately stop after it dropped on the floor, a proper repulse should be applied to the object so that it looks natural. It turns out that both of these two problems could be solved by some optimization methods.

3 Optimization Methods

Optimization has played a significant role in the computer graphics, here I will first give a optimization form description of collision problems and then explain it.

$$
\begin{aligned}
\min \quad & t \\
\text{constr} \quad & f(x, y, z, t) <= 0, \\
& g1(x, y, z, t) <= 0, \\
& g2(x, y, z, t) <= 0, \\
& \ldots \\
& gN(x, y, z, t) <= 0, \\
& ts <= t <= tf.
\end{aligned}
\tag{1}
$$

Here collision involves two objects, lets say A and B. We need to find the collision time t, if we can find this collision time, then the collision position could be derived by one of A, B's velocity (Tian-hu et al. 2010).

f(x, y, z, t) and gi(x, y, z, t) are description of two involved objects A, B, I use gi(x, y, z, t) to represent B because B may be divided many convex parts. Constraints like f(x, y, z, t) <= 0 means that collision points should be in this object. Based on these description we could find that this system will find the collision point because the only collision point locates on both objects meanwhile it has the smallest time t.

About the description of moving objects, it will depend on the shape and motion of object. For example, a sphere of radius 2 and center (0, 0, 0) at time 0 and moving with velocity (1, 2, 3) units/s is represented as:

$$
M(x, y, z, t) = \{(x, y, z, t) \,|\, (x - 0 - t)\hat{}2 + (y - 0 - 2t)\hat{}2 + (z - 0 - 3t)\hat{}2 <= 4\}
\tag{2}
$$

After presenting this description, we may apply this method onto any parametric objects to solve the collision time exactly when collision happens.

If collision happens, we need to give a repulse to objects so that they bounce with each other. This could be solved similarly like the collision detection routine, first simply inverse the velocity according to the collision geometric information, and then check whether another collision will happen less in a time step, if it happens, we need add some constraints such that collision only happen after a time step (Wei et al. 2011).

4 Cloth Simulations

The reason why collision is hard to deal with in cloth simulation is that cloth is deformable object, no matter self-collision or collide with other objects, many collision could happened simultaneously. We could not only solve the collision one by one, since the second collision's result could bring another collision to the first one. Conversely, we need to find a "zone" (Ya-fang et al. 2010), we find a potential area where collision will happen in the near future, then assign a similar repulsion to objects in this zone to avoid collisions.

Therefore we could use the collision detection routine quickly find the "zone" and involved objects, then give some response to objects in this zone, this way, multi-collision problem could be solved efficiently (Yan et al. 2011).

5 Conclusion

This paper focuses the collision problems, and it tries to apply optimization method into solutions of them. Also it provides a cloth simulation example to iterate the use of this method. It is easy to find that by using of this kind of method; both the efficiency and speed could be greatly improved. However, it also has limitation, only objects could be represented parametrically should apply this method. Also I try to find a direct solution to find the final position and velocity in a systematical collision, but it turns out this could not be done now; so people have to solve the collision result one step by one step, and I believe there will be some improvement in this part in the near future.

References

Aliyu MDS, Al-Sultan KS (1994) Collision detection algorithms: a survey. Technical report CCSE-029, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia
Aliyu MDS, Al-Sultan KS (1995) LP-based algorithms for detecting the collision of moving objects. J Oper Res Soc 46:854–886
Jingde X, Ye L, Huijuan C, Kun T (2009) Joint optimization algorithm of codebook size in multi-parameter, multi-mode quantization. J Tsinghua Univ (Sci Technol) (10):65–67
Jingwei T, Xiaokai W, Yanan P (2010) Virtual scene optimization algorithm based on visibility sensor node. Softw Guide (08):60–61
Jin-ming S, Chen H, Ling-ge J (2010) Channel-routing joint optimization algorithm for wireless mesh networks. J Shanghai Jiaotong Univ (8):1051–1056
Jiu-feng Z, Chun-ming J, Zheng W, Xing-lan J, Ji-wu Y (2010) PSO algorithm for back-calculation of source intensity. China Saf Sci J (CSSJ) (10):123–128
Tian-hu Y, Xing-peng M, Guo-qian W, Xiao-yang M (2010) Optimal algorithm for gradient based on edge detection. App Res Comput (1):367–370

Wei F, Jun S, Yanrui D, Wenbo X (2011) Medical image registration based on hybrid quantum-behaved particle swarm optimization. Comput Eng Appl (3):170–173

Ya-fang S, Xiao-dan W, Jun-yan X, Qing-huai L (2010) Global optimization algorithm based on visual cognition. J Chang Univ Technol (Nat Sci Ed) (6):15–18

Yan W, Bing Z (2011) An improved particle swarm optimization algorithm. Value Eng (7):171–172

A Kind of Research on Gray Image Watermarking Algorithm Based on Chaos

Hao Luo, Shigong Long, and Chaohui Jiang

Abstract This paper proposed a kind of gray image digital watermarking algorithm of combining discrete wavelet transform and discrete cosine transform with singular value decomposition based on chaos and scrambling for the purpose of raising security, invisibility and robustness of watermark. It hided the digital watermarking in the discrete cosine transform domain. The image that contained the watermark was restored and the watermark was extracted by the singular value decomposition. The experimental results demonstrate that the watermark not only has good invisibility by use of this algorithm but also has good robustness against geometric distortion, noise attack and so on.

Keywords Chaos • Digital watermarking • Discrete wavelet transform • Discrete cosine transform • Singular value decomposition

1 Introduction

The digitization of information media has provided the great convenience for our life with the development of digital technology and Internet. Meanwhile, the copy that is not authorized, hostile attack and interpolation become more and more serious. It is not enough to encrypt the multimedia by traditional encryption and decryption technology. However, digital watermarking technology can make up for the shortage that is the encryption and decryption technology can't enough to protect the information that has been decrypted.

H. Luo (✉) • S. Long • C. Jiang
School of Computer Science and Information, Guizhou University,
Guiyang 550025, China
e-mail: hluo79@163.com

Digital watermarking technology can embed the information that can prove the copyright affiliation and trace the tort behavior into the digital information. This paper puts forward a kind of digital watermarking algorithm that bases on the chaos and Hilbert scramble and combines the discrete wavelet transform (DWT) and discrete cosine transform (DCT) with singular value decomposition (SVD) (Bian and Li 2008). At first, the binary image should be scrambled and done the chaos processing. The original image should be done the DWT transformation and DCT transformation. Then the watermark that has been scrambled and encrypted should be hidden in the DCT domain. Meanwhile, the DCT domain that does not contain the watermark should be done the SVD decomposition and retain the related orthogonal matrix preparing for restoring the image that has embedded the watermark.

There are several reasons of using DCT transformation. First, the transformation coefficients are almost uncorrelated after the image is transformed. Second, the channel error and the quantization error will be disseminated into each pixel of each image block in common with the random noise and the accumulate error will not be made. Last, the energy in the data block can be compressed in the part low frequency coefficient by the DCT transformation (Liu and Jiang 2009).

2   Hilbert Curve Transformation

The Peano curve can traverse all the point that involve in the square. And the curve can traverse all the space. Hilbert who was a Germany mathematician had constructed a kind of simple Peano curve that is Hilbert curve (Wang and Xiaoshuang 2006). According to the label sequence of dividing square step by step, the definition of the Hilbert curve is given out (Wang and Xiaoshuang 2006). As shown in the Fig. 1a, b, Fig. 1a is the first level division and curve diagram. Figure 1b is the second level division and curve diagram.

The square is divided equally into four parts and the four parts is connected. Then each small square is divided equally into four small squares and each square is connected. Then the Hilbert curve that has different levels is formed.
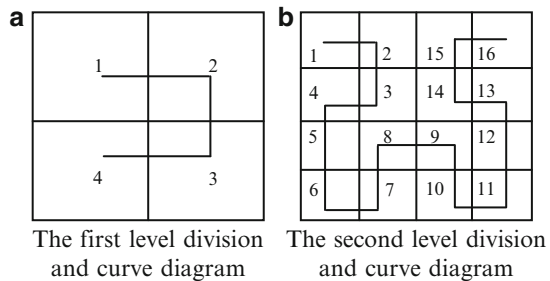
Fig. 1 Square division and curve diagram. (**a**) The first level division and curve diagram. (**b**) The second level division and curve diagram

The first level division and curve diagram

The second level division and curve diagram

According to the trend of Hilbert curve, the all points of image should be traversed. Then we can get the image of Hilbert transform.

3 The Define of Chaos Sequence

The chaos is a kind of irregular sport seemingly in the certain system (Ge et al. 2008). The chaos sequence can be generated by use of initial value. Therefore, the Logistic chaos sequence is adopted in this paper. The equation of this kind of sequence is defined as shown below.

$$x_{n+1} = \mu x_n (1 - x_n) \tag{1}$$

$x_n \in (0, 1), 0 < \mu \leq 4, \mu$ is called branch parameter.

When the μ limited in the narrow range of $3.75 \leq \mu < 4$, the system shows a kind of chaos state. Another form of Eq. (1) is defined as shown below.

$$x_{n+1} = 1 - \lambda x_n^2 \tag{2}$$

$\lambda \in (0, 2), x_n \in [-1, 1]$. Iteration have repeatedly mutation along with the enlargement of λ. The system will appears chaos state when λ equals to 1.40115. This paper has generated the chaos sequence by use of Eq. (2). The value of λ is 1.9.

4 The Singular Value Decomposition of Image

We regard an image as a matrix because the character of gray level distribution can describe the internal feature. Any matrix B of $M \times N$ can be written as show below.

$$B = U \Sigma V^T = \sum_{i=1}^{r} \lambda_i U_i V_i^T \tag{3}$$

U is the orthogonal matrix of $M \times M$. V is the orthogonal matrix of $N \times N$. Σ is a nonnegative diagonal matrix and Σ equals to $diag(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0)$. The diagonal elements include the singular value of B. And $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$. r is the rank of matrix Σ. Each column of Matrix U is the eigenvector of matrix BB^T. Each column of V is the eigenvector of matrix $B^T B$. Σ is the singular value matrix. This kind of transformation is called SVD transformation. The SVD of matrix is used in the image because the singular value has good stability. When the image is attacked, the singular value has little change. Moreover, singular value has the invariance to transposition transformation, zoom transformation and mirror image transformation.

5 The Process of Chaos Processing and Scrambling Processing to Watermarking Image

The watermarking image has been done the preprocessing of chaos and scrambling before the watermark is embedded in the original image. Then the attacker can only get the integrated processing result of chaos and Hilbert scrambling instead of the original watermarking image if the watermark is extracted. Then the security of watermark is increased (Mansouri et al. 2009).

There are several steps as follows:

1. First, the original watermarking image of $M \times N$ must be done the Hilbert transformation. The transformation times is indicated by the variable K. The transformation times equals to 20.
2. The sequence of real number is generated by use of Logistic sequence. The length of sequence is the size of the binary watermarking image. The initial condition is that x_0 equals to 0.123 and λ equals to 1.9.
3. Changing the chaos sequence into the sequence that is made of number 0 and number 1 by use of symbolic function. The sequence length is $M \times N$.
4. Changing the chaos sequence into the two dimensional matrix. Then the two-dimensional matrix and binary watermarking image that has been done the chaos and scrambling must do the exclusive OR (XOR) operation each other. The watermarking image that has been done the chaos and scrambling processing is made at last.

The whole process is shown in the diagrammatic sketch as follows (Fig. 2).

The Fig. 3a is an original watermarking image. It's size is 64×64. The result of the watermarking image that has been processed is shown in the Fig. 3b. The diagram result is shown below.
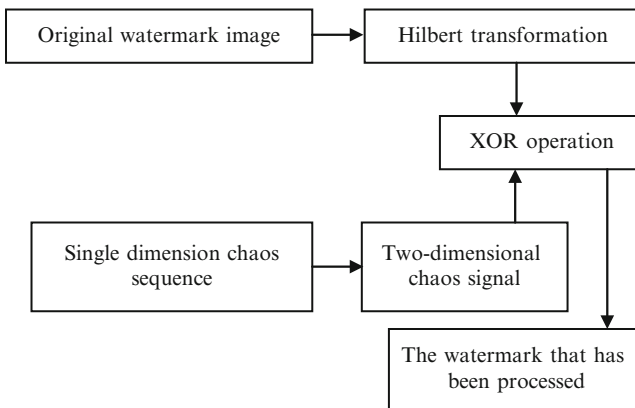
Fig. 2 The diagrammatic sketch of chaos and scrambling processing
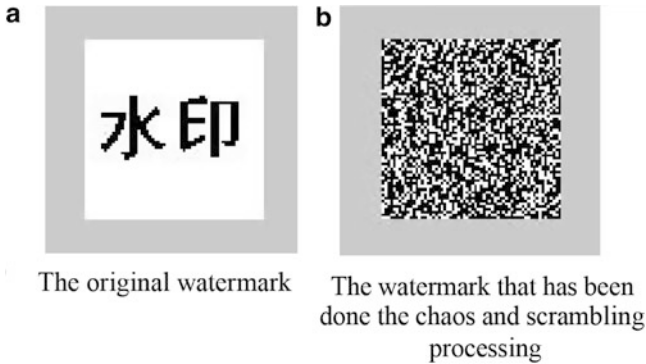
Fig. 3 The original watermark and pre process result of original watermark. (**a**) The original watermark. (**b**) The watermark that has been done the chaos and scrambling processing

If the matrix of Hilbert transformation, transformation times and the initial condition that is needed by the Logistic mapping is known, the original image can be restored.

6 Algorithm Description

6.1 The Embedded Algorithm of Watermark

The following contents are the embedded process of watermark. First, the first level discrete wavelet decomposition is done to the original image and the DCT transformation is done to the low frequency part. Second, the SVD is done to the low frequency part. Then the watermark is hided in the low frequency part. The specific step is shown as follows.

1. The discrete wavelet transformation is done to the original image. Then the original image is divided into four parts. They are the low frequency part named LL1, detailed part of horizontal direction named HL1, detailed part of vertical direction named LH1 and detailed part of diagonal direction named HH1.
2. Partition of 8×8 is done to the LL1 whose size is 256×256. DCT transformation is done to each sub block. Then each sub block is transformed into a DCT domain.
3. The part of 256×256 is done the SVD transformation based on the step (2) in order to get the orthogonal matrix U and V. The two matrixes can be used to restore image. The following formula can denote the process.

$$cLL1 = UDV^{T} \tag{4}$$

4. The watermark of 32×32 that has been done the chaos and scrambling processing is hided in the DCT domain of 256×256 based on the step (2). The watermarking information of one bit is hided in each sub block of 8×8. The coefficient of position (2,3) and position (4,1) in each sub block can be marked as B(2,3) and B(4,1). If the sub block need to hide the number zero, the coefficient of position (4,1) must be smaller than the coefficient of position (2,3). If the sub block need to hide the number one, the coefficient of position (4,1) must be greater than the coefficient of position (2,3) and the absolute value of (B(4,1)−B(2,3)) must be more than or equal to α, α is evaluated as 30 in this paper.

5. The information is hided in the part $cLL1'$ that must do the SVD transformation. The following formula can denote the process.

$$cLL1' = U_1 D_1 V_1^T \tag{5}$$

6. The singular value that has changed multiply by U and V^T is rcLL1.The following formula can denote the process.

$$rcLL1 = UD_1 V^T \tag{6}$$

7. The rcLL1 is got according to step (6). Then rLL1 is got according to the inverse transformation of partitioned DCT transformation.

8. The HL1, LH1 and HH1 are got according to the step (1). The watermarking information has been hided in the image that is got according to the rLL1, HL1, LH1, HH1 and discrete wavelet inverse transformation.

6.2   The Extracting Algorithm of Watermark

If the image I_w that contains the watermark, matrix U_1 and V_1^T is known in the process of extracting watermark, the specific step of extracting watermark is shown as follows.

1. The discrete wavelet transformation is done to the image that has been hided the watermark. Then four parts are got. They are separately the low frequency part named wLL1, the detailed part of horizontal direction named wHL1, the detailed part of vertical direction named wLH1, the detailed part of diagonal direction named wHH1.

2. The division of 8×8 is done to the low frequency part named wLL1 that has been hided the watermark. Each sub block is done the DCT transformation. Then the wLL1 is transformed to cwLL1.

3. SVD transformation is done to the cwLL1, the formula is shown as follows.

$$cwLL1 = U_0 D_0 V_0^T \tag{7}$$

4. Matrix D_0 multiply by U_1 and V_1^T is $cwLL1'$. The formula is shown as follows.

$$cwLL1' = U_1 D_0 V_1^T \tag{8}$$

5. Partition of 8×8 is done to the $cwLL1'$. According to the hiding regular, the hiding information that has one bit is extracted in each sub block. If the coefficient of position (4,1) is greater than the coefficient of (2,3), the hiding information one is got from this sub block. If the coefficient of position (4,1) is smaller than the coefficient of position (2,3), the hiding information zero is got from this sub block.
6. The watermarking sequence that has been extracted is transformed into the matrix of 32×32. The watermarking image is got after inverse transformation of chaos and scramble processing have been done.

7 Experiment Result and Analysis

The image lena of 512×512 is the original carrier image in the experimentation of this paper. The watermark is the binary image of 32×32. The peak signal to noise ratio (PSNR) can be used to test the invisibility level of watermark. The formula is shown as follows (Wang et al. 2004).

$$PSNR = 10\log_{10} \frac{D^2 MN}{\sum\limits_{x=1}^{M} \sum\limits_{y=1}^{N} (I(x, y) - I_w(x, y))^2} \tag{9}$$

D is peak value of peak signal. If the value of PSNR is bigger, the imperceptibility of watermark is better after the watermark is embedded.

The similarity of the watermark that has been extracted and the original watermark can be measured by the normalized cross-correlation (NCC). The formula is shown as follows.

$$NCC = \frac{\sum\limits_{i=1}^{M} \sum\limits_{j=1}^{N} w(i, j)w'(i, j)}{\sqrt{\sum\limits_{i=1}^{M} \sum\limits_{j=1}^{N} w^2(i, j)} \sqrt{\sum\limits_{i=1}^{M} \sum\limits_{j=1}^{N} w'^2(i, j)}} \tag{10}$$

If the value of NCC is more close to the number one, the two watermarking image is more similar.

The result of experiment is shown as follows. Figure 4a is the original carrier image. Figure 4b is the original watermarking image. Figure 4c is the carrier image that contains the watermark. Figure 4d is the watermark that is extracted by the algorithm of this paper.

The original carrier image

The original watermark

The carrier image that contains the watermark

The watermark that has been extracted

Fig. 4 The original image, the original watermark, the result of hiding watermark and extracting watermark. (**a**) The original carrier image. (**b**)The original watermark. (**c**) The carrier image that contains the watermark. (**d**) The watermark that has been extracted
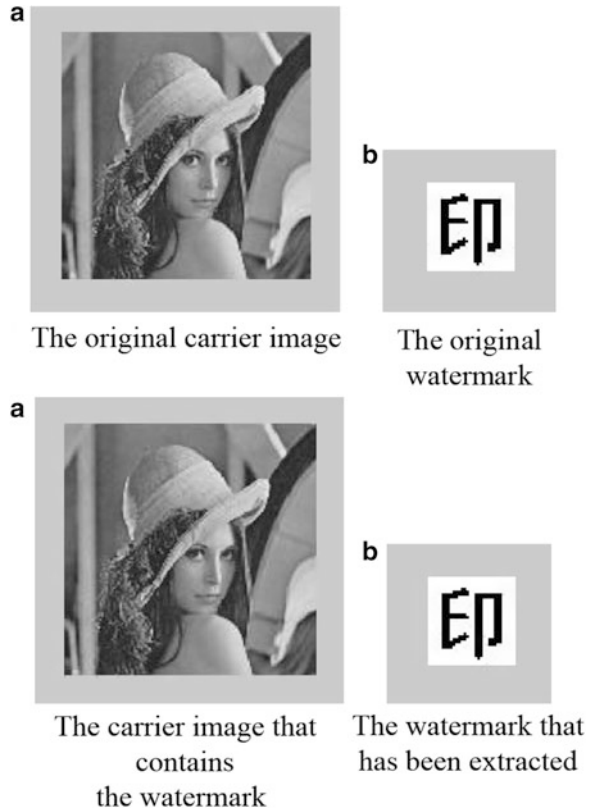
Table 1 The experiment result of noise attack

Attack method	0.02 Gaussian noise	0.05 salt and pepper noise	0.05 multiplicative noise
NCC	0.9661	0.9974	0.9987
PSNR	17.2035	18.4189	18.8725

The value of PSNR equals to 58.7508 between the Fig. 4a, c after verification. The value of NCC equals to 1 between Fig. 4c, d. That is to say, the algorithm of this paper has the good invisibility of watermark (Zhang et al. 2010).

The watermark is attacked by the noise, geometric distortion and so on. The purpose is to verify the algorithm of this paper that has the strong robustness.

According to the different assault type, the different test result can be gotten. The noise assault experiment is done at first in this paper. The NCC value and PSNR value are shown as Table 1.

The quarter cut, 30° rotation and 60° rotation are done at the center position of image in the geometric distortion experiment. The mirror vertically is done to the image that has been embedded the watermark (Zhang 2006). The process of zoom

Table 2 The experiment result of geometric distortion

Attack method	Quarter cut	30° rotation	60° rotation	Mirror vertically	Zoom operation
NCC	0.9961	1	1	1	0.9832
PSNR	11.0421	10.6574	10.1518	12.4158	28.8643

Table 3 The experiment result of JPEG compression

Attack method	Compressibility factors is 5	Compressibility factors is 10	Compressibility factors is 75
NCC	1	1	1
PSNR	27.2884	30.3330	37.4148

Table 4 The experiment result of JPEG compression

Attack method	Gauss low frequency filter of 5×5	Mean filter of 5×5	Median filter of 5×5
NCC	1	0.9884	1
PSNR	41.0947	28.3135	31.2599

operation is the image that has been embedded the watermark is compressed to 128×128. Then the image is enlarged to 256×256. The result of NCC and PSNR value are shown as Table 2.

The compressibility factors are 5, 10 and 75 in the joint photographic experts group (JPEG) compression experiment. The NCC value and PSNR value are shown as Table 3.

The gauss low frequency filter of 5×5 is adopted in the wave filter experiment. It's standard deviations is 0.5. And mean filter of 5×5, median filter of 5×5 are adopted. The experiment result are shown as Table 4.

8 Conclusion

This paper proposed a method of hiding watermark effectively. DWT algorithm, DCT algorithm and SVD algorithm are combined cleverly in this paper to complete the process of hiding watermark and extracting watermark. The watermarking image is done the Hilbert scramble before hiding the watermark. And the Logistic chaos sequence is generated. The result of Hilbert scramble and the two-dimensional signal that is transformed by the chaos sequence is done the XOR operation each other. Then the safety of watermarking system can be increased. In addition, the character of the DWT domain, DCT domain and SVD decomposition can be used well. Then the robustness and invisibility can be raised. These experiments result show that the algorithm of this paper has the strong robustness and the algorithm can resist all kinds of noise attack, JPEG compression, and so on. And the prerequisite is keeping the good invisibility of watermark.

References

Bian D, Li J (2008) A digital watermarking algorithm for images based on DWT and SVD. Comput Simul 25:87–90

Ge X, Tian H, Guo L, Han T (2008) The principle and application of information hiding. Tsinghua University Press, Beijing

Liu J, Jiang H (2009) An improved watermarking algorithm for digital image based on DCT and SVD. Comput Eng Sci 31(1):38–40

Mansouri A, Mahmoudi Azna Veh A, Torkamani Azar F (2009) SVD-based digital image watermarking using complex wavelet transform. Sadhana 34:393–406

Wang S, Xiaoshuang X (2006) A new algorithm of Hilbert scanning matrix and its matlab program. J Image Graph 11:119–122

Wang L, Guo C, Li P (2004) The experimental textbook of information hiding technology. Wuhan University Press, Wuhan

Zhang Q, Li K, Yuan Z (2010) Robust digital image watermarking algorithm based on chaos and SVD-DWT. Appl Res Comput 27:718–720

Zhang R (2006) The watermark algorithm of transform domain digital image based on chaos. Master's Thesis, Central China Normal University

Mobile Communications LTE FDD System of Performance Index and Key Technology

Chen Gang and Chen Qiang

Abstract In order to put itself in an advantageous position in the future mobile communication technologies competitive environment, to meet the user's multiplication needs and to deal with the market challenge, 3GPP has adopted the long term evolution (LTE) about the project work at the end of 2004. As one of the most important projects that 3GPP started in recent years, many new techniques have been used by LTE to provide higher user data rates, to effectively improve the cell edge performance, to improve system capacity and coverage, and to reduce latency this paper studies eNodeB MAC protocol functions at the basis of analyzing the performance targets, the key technologies and the network architecture.

Keywords Mobile communication • Long term evolution (LTE) • Performance index • Key technology

1 Mobile Communication System Development Outline

Mobile communication technology has mobility, freedom, with time and location limit etc., are deeply changing people's life and behavior. In recent years, the Internet exerts an enormous influence on communications, also influences and drive the mobile communications fields of development, mobile Internet, mobile

C. Gang (✉)
Continue Education Institute Chongqing Radio and TV University,
Jiulongpo District, Chongqing, P.R. China 400039
e-mail: chongyoucg@163.com

C. Qiang
Chongqing Municipal Electric Power Company Maintenance Branch,
Jiulongpo District, Chongqing, P.R. China 400039
e-mail: cqcq_116@163.com

multimedia will become mobile business development direction, rich data business needs a high data bandwidth and Service Quality, Quality of Service (QoS) guarantee platform.

The first generation of Mobile communication System is based on Frequency Division Multiple Access (Division Multiple Access, been FDMA) technology, the simulating communication systems, a typical representative is the American AMPS (please System, advanced Mobile Advance Mobile telephony System) and later improved System TACS (Total Communications System, full Access communication System), and Mobile Telephone, NMT (Nordic Nordic Mobile Phone) and Telegraph and Telephone NTT (Nippon, Japan Telephone and Telegraph) etc. AMPS use analogue cellular transmission 800 MHz frequency band, is our 1980s using simulation mobile communications format, use 900 MHz band. However, there exist such as limited capacity, shape.whether too much, mutually exclusive, secrecy bad, and the average quality is not high, cannot provide data services, between system cannot provide roaming etc. deficiencies.

The second generation of mobile communication system mainly adopts TDMA (Division Multiple Access, a TDMA) and Code Division Multiple Access (Code Division Multiple Access, CDMA) of digital communication technology, global basically has GSM and CDMA IS - 95 two system. GSM Standards are ETSI (European Telecommunications Institute, present the European Telecommunications Standards Institute) proposed, in the world at present most countries use the standard. The second generation mobile communication system of core business is business, core network based on the speech circuit switching technology, composed of TDM lines carrying voice and data services at low speed. Two network formats between hard realization roaming, cannot satisfy the demand of global communication. In the meantime, data communication technology progress, especially the spread of the Internet, make multimedia mobile communication become the mass is expected to enjoy service, and become a modern communication development one of the major objectives of.

With GPRS as a representative of 2.5 g technology, its realization is based on the existing circuit switching networks increase a parallel packet network. 2.5 g just in original wireless technology and communication platform was improved, the rate of increase and business provide business flexibility in such aspects as is very limited, cannot fundamentally change the second generation of mobile communication system to voice traffic and low speed data key business situation.

The third Generation mobile communication system referred to 3 G (are), is by the law, International Telecommunication Union (ITU International Telecommunication Union) in 1985, puts forward the work in 2,000 MHz frequency band, expected in around the year 2000 commercial system. Then called future land mobile communication system, 1996 officially renamed the 2000 international mobile communication plans "(abbreviation IMT - 2000).

And the second generation digital mobile communication system, 3 G system than the most important characteristic is can provide mobile multimedia business, the design goal is to provide greater than the second generation system of system capacity and better communication quality, and can globally better realize seamless
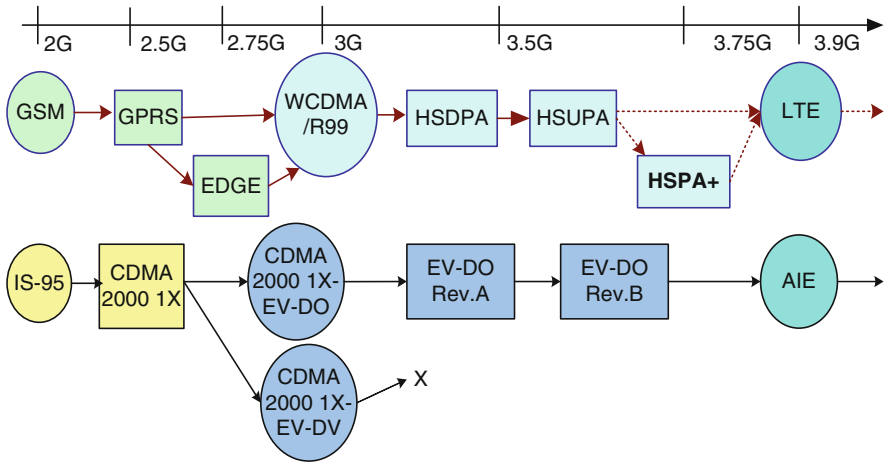
Fig. 1 Wireless communications technology development and evolution figure

roaming and provide users with including voice, data and multimedia, a variety of business, also want to consider with existing second generation system good compatibility. ITU certain IMT - 2000 technical standard mainly includes Wideband Code Division Multiple WCDMA (cdma), broadband Access, CDMA2000 (Code Division Multiple Access 2000, cdma 2000) and td-scdma (a Synchronous Code Division Multiple Access, Division cream-and-green – Synchronous cdma).

The fourth generation mobile communication system (4 G) and even more advanced mobile communication system will appear and development, considering the current existing multi-drug communication system mutual confluence, forming a multidimensional network solutions. Future fourth generation mobile communication system is the multi-functional integration of broadband mobile communication system, the fourth generation mobile communication will with broadband, IP melt, the comprehensive function of system has a variety of shape.

Mobile communication system are continuously in development, shown in Fig. 1, from today still use IS - 95, GSM, etc. > 2 G system, gradually development to support certain Data business Data Rate for EDGE (to Global optimisation techniques, Global Evolution similarities enhancement Data Rate of technology), CDMA1X, and 3 G WCDMA, td-scdma, etc. In order to meet business requirements, higher 3 G (High Speed development HSDPA Downlink Packet Access, high-speed Downlink link grouping technology), HSUPA (High Speed, High Speed Access Uplink Packet Uplink grouping technology), when the future (LTE graphs supplied for similarities, long-term Evolution), AIE (Air with similarities, Air Interface Evolution) etc. At the same time, take the WLAN (Wireless Local Area, the Wireless Local Area Network), a number WiMAX (Interoperability for Microwave Access, global Microwave Access Interoperability) technology as a representative of the Wireless Local Area Network, man technologies are constantly developing and perfecting.

2   Performance Indicators

The mobile communications market in the high speed development, accelerate the existing network improvement, and constantly meet the users' demand, therefore, to offer new business will be in the fierce competition in an invincible position in the only alternative. Meanwhile wireless communication technology develops quickly, with Orthogonal Division Multiplexing (OFDM been, Orthogonal Frequency Division Multiplexing) technology as a representative of the new technology in recent years has become a focus, and gradually industrialization. To cope with that kind of pressure, maintain the mobile communication field's leadership, the 3GPP (law are well, bought the third Generation cooperation organization) made further efforts, in November 2004 through the 3GPP LTE long-term evolution of stetting up, to develop 3 G evolution model system technical specifications as a target.

LTE is about UTRAN and UTRA improvement project, according to the 3GPP LTE research Work of the working process is divided into two stages: SI (Study on this Item, the technical feasibility Study phase) and WI (schools, and the specific technical specifications on this Item written stage). Based on the OFDM, Multiple Input Multiple Output (MIMO, Multiple Input Multiple Output) new technology such as LTE is committed to providing higher data rate, higher village capacity, lower delay time and reduce the cost of users and mobile operators.

LTE is coping with the start of the project of "other wireless communication standard" competition. Aiming at WiMAX "low mobility broadband IP access" positioning, LTE puts forward corresponding requirements, such as similar bandwidth, data rate and spectrum efficiency index, to optimize low mobility, only support PS (Packet Switching) domain, emphasize radio multicast business etc. Meanwhile, for VoIP (Voice Over IP, VoIP) and online games to the user's attention, LTE face delay is very strict.

LTE formulation main performance indexes include:

significantly increased peak rate: 100Mbps, ascending descending 50Mbps
in maintaining the current site configuration under the condition of invariable, increase "residential edge rate"
significantly improved spectrum utilization, efficiency and reach the 3GPP spectrum Release 6 2–4 times
reduce user surface and control the delay, the user face delays (one-way) less than 5 ms, control surfaces delay less than 100 ms
flexible configuration 20 MHz bandwidth. Variety of 125 ∼
support and existing 3GPP and the 3GPP system interoperability
further enhance MBMS (in Broadcast Multicast Service) business
lower cost and realize from the arrangement Release 6 of the low cost evolution
realize reasonable system and terminal complexity, cost and power consumption
support enhanced IMS (IP Subsystem, in IP Multimedia Subsystem) and core network
pursue backwards compatible, but should be carefully considered the performance improvement and backward compatibility between the balance

provide various business types of powerful support, especially in PS domain, such as VoIP

of low-speed moving optimization system, at the same time support high-speed moving

as possible the similar technique supports both pairs (paired) and the pairs (unpaired) band

as the support a simple frequency coexist

UTRA system which E - performance requirements mainly includes user throughput rate and spectral efficiency two aspects.

2.1 The User Throughput Rate

For downlink link, and a cumulative probability distribution corresponding MHz of 5% per user throughput rate should achieve R6 version HSDPA 2–3 times of. Every MHz bandwidth to the average user throughput rate should achieve R6 version HSDPA of 3–4 times. The above R6 version of the performance objectives are in the Node B use single antenna, UE Equipment, users (User Equipment) the use of performance-enhancing class 1 receiver, and when the value of the UTRA base station - E distal can use double transmitting antenna, in UE end can use double receiving antenna. Support user throughput rate should be proportional to system bandwidth.

For the uplink cumulative probability distribution, 5% per user corresponding MHz throughput rate should achieve R6 version to enhance uplink access 2–3 times. Every MHz bandwidth to the average user throughput rate should achieve R6 version to enhance uplink access 2–3 times. R6 edition and E - UTRA use UE single send antenna, the base station end double receiving antenna configuration. Support user throughput rate should be proportional to system bandwidth.

2.2 Spectrum Efficiency

E - UTRA shall not changing the base station cloth address at the same time greatly improve spectrum efficiency and improve residential edge bit rate.

Downlink spectrum efficiency should achieve below target: in load network, spectrum efficiency (bits/SEC/hz/site) should achieve R6 version HSDPA of 3–4 times. The above R6 version of performance objectives are in the Node B use single antenna, UE end use performance-enhancing class 1 receiver values, and E - UTRA base station end can use double transmitting antenna, in UE end can use double receiving antenna.

Uplink frequency spectrum efficiency should achieve below target: in load network, spectrum efficiency (bits/SEC/hz/site) should achieve R6 version enhanced uplink 2–3 times. R6 version and LTE use single antenna, send the UE base station end double receiving antenna configuration.

3   LTE Key Technology

The performance of the system for realizing LTE index, need improvement and increase 3 G system air interface technology and network structure, the 3GPP after fierce debate eventually will determine the OFDM and MIMO scheme as the wireless network of evolution of standards.

As wireless communication system foundation and the sign of the air physical interface layer technology, the 3GPP LTE system for physical downlink transmission scheme decided to adopt the spectrum utilization of high technology, but for uplink OFDM transmission scheme is from reduce costs and volume of terminal consider decided to adopt the SC - FDMA (performance - Division Multiple Access, Carrier been Single Carrier Frequency Division Multiple Access) technology to reduce the emission of peak all power than a terminal. OFDM technology is LTE system technology basis and main characteristics, and its parameters Settings of whole system performance will be decisive impact. As the most basic parameters OFDM systems carrier interval by theoretical analysis and comparative simulation for 15 KHz finally determined, data to the resource block mapping methods can adopt concentration (localized) or discrete (distributed) way. Cyclic Prefix circular Prefix (CP) in length, determines the OFDM systems anti-multipath ability and coverage ability, LTE system will use the length of two sets of circulation Prefix scheme.

Due to OFDM own characteristics of make it very suitable for wireless broadband channel under high-speed transmission can well against the frequency wireless transmission environment selective decline, get a higher frequency spectrum utilization, Because the user between channel decline of independence, can use subcarrier allocation brings multiuser diversity gain improve performance, achieves the service quality requirements. By adopting OFDMA detachable traditional design constraints, be helpful for system design parameters in a flexible and free choice, For user receiver speaking, be aimed at OFDMA air interface processing relatively simple, in greater bandwidth and high order MIMO configuration can reduce terminal complexity.

In LTE in MIMO considered achieve user average throughput rate and spectral efficiency requirements of the best technology. The basic idea is multiple-input multiple-output (MIMO) system in transceiver double end use several antenna, through space-time processing technology, make full use of space resource, improve the spectrum utilization. Usually multipath effect can cause decline, be regarded as harmful elements. Multi-input multi-output (MIMO) is one of the important characteristics can be regarded as a multipath spread the favorable factors use, improve system performance. Multi-input multi-output (MIMO) can use stochastic decline and multipath delay spread (if any) to increase transmission rate. Wireless communication system using MIMO technology, do not need additional spectrum resources, just hardware and system complexity of increase, will get good performance of the system.

As the main transmission rate improve system means, LTE system will design applicable "community, tiny area, etc.". Various kinds of environmental MIMO technology. Multi-input multi-output (MIMO) antenna number basic configuration has determined for downlink 2×2 (namely base station side two root antenna, UE lateral two root receiving antenna), 1×2 uplink (i.e. UE side a root antenna, base station side two root receiving antenna). A lot of methods for realizing the multiple-input multiple-output (MIMO) technology, including air, Division Multiplexing receiver by Multiplexing (SDM), air multi-access Division Multiple Access (receiver by SDMA), preforming, encoding (pre-alpha - for), Rank adaptive (Rank adaptation), smart antenna and ring-opening launch diversity, etc.

Peak value transmission rate is LTE downlink link to solve the main problems, in order to realize system downlink peak value rate target, LTE downlink mainly adopts QPSK, 16QAM, 64QAM three modulation mode, and uplink mainly adopts BPSK, 8PSK and 16QAM QPSK,. Channel coding, LTE main consideration the mature Turbo yards, and if you can obtain significant gain will also consider other encoding.

Since there are difficult to solve the "synchronization problem", LTE on unicast business does not use the downlink macrodiversity, just in providing more village broadcasting business, they can be adopted to solve community prefix large circulation between synchronization problem. Moreover, considering that the network structure of LTE highly "flattening", "decentralization", LTE system will not adopt uplink macrodiversity technology.

4 LTE Network Architecture and the Hollow Agreement Structure

The whole framework of the 3GPP LTE the revolutionary change. E - UTRAN system structure as shown in Fig. 2 shows:
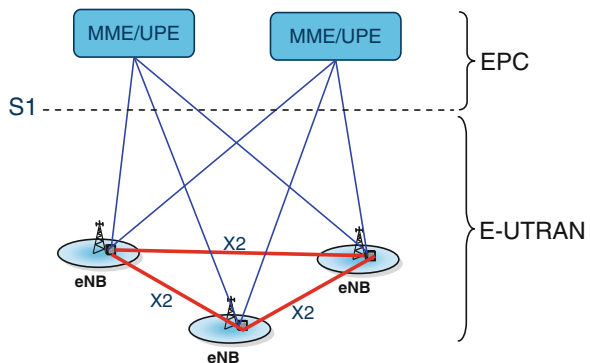

Fig. 2 E - UTRAN structure

LTE system core network using two layers of flat network architecture, and adopts full IP distributed structure and support VoIP, Mobile IP IMS, wait for all sorts of technology. Access eNodeB (mainly by eNB) and access gateway (aGW) constitute, aGW is actually a boundary nodes, if it as part of the core network, the access network mainly by eNB a layer of composition. In addition to its original of LTE eNB Node B's functions, but also bear the original RNC most of the function. LTE adopted this similar to typical IP broadband network structure to simplify network and reduce delay and achieve low delay, low complexity, low cost requirements, the system structure of the 3GPP far-reaching impact.

E - UTRA air interface protocols are divided into L1 layer, L2 layer and L3 layer.

Physical layer is responsible for coding, physical HARQ processing, modulation and multi-antenna processing and signal to the appropriate physical time-frequency resources mapping, etc.

L2 layer divided into the following several sub-systems: layer of media Access Control (MAC) Access Control, forward, wireless Link Control (Link Control, RLC) have grouped Data gathering and Packet Data Protocol (Convergence, PDCP) (The 3GPP TR 25.813 V7.1.0 (2006)[2]). Each child between layers use Service Access points (Service Access Point, SAP) as end-to-end communication interface.

PDCP, executive IP head compression, in order to reduce transmission through wireless interface bit quantity. PDCP is also responsible for was to send data encryption and integrity protection. In the receiving side PDCP agreement, the corresponding encryption and decryption operation.

RLC, responsible for segmentation/string receive, retransmission processing and the sequential delivery to senior etc.

MAC, processing downlink scheduling and HARQ retransmission. Schedule function located in eNodeB, every village have a MAC entity, and is responsible for ascending and descending links. HARQ agreement in part a MAC protocols of the receiver and the sender do exist.

Wireless Resource Control (have RRC) layer Control Resource, the main function is to broadcast with NAS (Non - Access Stratum, the Access layer) and AS (Access Stratum, Access layer) related system information, paging, establishing and maintaining and release UE and E - UTRAN between RRC news RRC connection, the integrity of protection, establishing and maintaining and release the point-to-point wireless between bearing, including UE measurement reports and village switching, selection and residential area to choose again, mobile sexual function.

References

3GPP TR 25.913 V8.0.0 (2008–12) Requirements for Evolved UTRA (E - UTRA) and melky Evolved UTRAN (E - UTRAN) (8) Release

ChuLinBo LiuShuXia, mobile communication and the Internet. HuangShengHua fusion technology. Zte technology, 2002.01

HuHaiNing lynch soldiers. The next generation wireless network LTE introduction. Neat

LinHui 3GPP LTE 3 G evolution (about) work. Telecommunication network technology, March 2006 period

NieXiao system performance evaluation LTE. Beijing university of posts and telecommunications, 2007

SunLiXin, YouXiao tiger, zhang ping. Third generation mobile communication technology. Beijing: People's wiley&sons press, 2000

The 3GPP LTE core technology. ShenJia and standardized progress. Mobile communications, 2006.4

The 3GPP TR 25.813 V7.1.0 (2006–09) Evolved Universal Terrestrial have Access (E - UTRA) and melky Universal Terrestrial have Evolved a Access (E); UTRAN - Have protocol margin and keep (with Release 7)

The 3GPP TR 25.912 V8.0.0 (2008–12) study of Universal Terrestrial evolved Feasibility seas have Access (Universal Terrestrial UTRA) and melky have Access UTRAN) (a (8) Release

XuZhiYu, HanWei, PuYingChun. HSDPA technology principle and network planning practice. People's wiley&sons press, 2007

Exchanging Data in Heterogeneous Relational Databases via XML Documents

Le Lei

Abstract With the strong ability of describing and transfer ring data, XML offers an applicable way for data exchange between different data sources. This paper introduces a model-drive method to exchange data in heterogeneous relational databases via XML documents. The relevant processes mainly involve establishing xml documents as a robust middle data source and realizing an effective bidirectional mapping between XML Schema and relational mode.

Keywords XML • Heterogeneous database • Relational database • Data exchange • Mode mapping

1 Introduction

The database technology is now widely used to store and manage data. But it's inconvenient to exchange information because databases are sometimes heterogeneous (Richard Wang and Madnick 1990), which means that data are structured in different ways and then the conversion has to be done before these data are exactly exchanged. Obviously, using a middle data source which can be easily converted to data of other structure is a possible way to deal with such situations. The eXtensible Markup Language (XML) is just an appropriate middle data source.

XML (Tim et al. 1998) is a subset of SGML, developed by W3C. XML documents usually contain a declaration about data type, which makes data in XML easily understood by both human and computer. New data types can be defined and added into XML anytime without a group of data structures defined previously. And there is little limit to data elements. Because of these qualities mentioned above, XML is an ideal middle format when data are exchanged.

L. Lei (✉)
School of Economy & Management, Tongji University, Shanghai, China
e-mail: lelei1988@126.com

The Relational Database (RDB) (Codd 1970) is now widely used and the relevant technology is quite mature. This paper focuses on data exchange between different RDBs and mainly deals with the mapping between XML and RDB in the two directions.

This paper is organized as follows. In Sect. 2, I briefly introduce the two different kinds of mapping method between XML and RDB. In Sect. 3, I present how to convert relational data to XML. In Sect. 4, the conversion in the opposite direction is shown. So Sects. 3 and 4 together give a complete account of the bidirectional mapping between XML Schema and relational mode.

2 The Mapping Methods

There are two kinds of mapping methods about conversion between RDB and XML (Liu Xiaoming 2009): the template-drive method and the model-drive method. The main difference between them is the way to establish the mapping relation.

2.1 The Template-Drive Mapping

The template-drive mapping does not previously give a clear definition of the mapping relation between XML and other data documents. Instead, SQL commands with parameters are embedded in a template. These commands are recognized and executed during the conversion. And then the results are placed where the commands expect. After that, the real XML documents are formed. For example, as for an XML template below:

```
<Book>
<SelectRes> Select Book_id, Title from Book</SelectRes>
</Book>
```

When the data transferring middleware begins to deal with the document, each "select" sentence will be replaced by the results from the execution of itself.

Thus there will be a new and expected XML document which may seem like this:

```
<Book>
<Row>
<Book_id> BookId</Book_id>
<Title> BookTittle</Title>
</Row>
</Book>
```

The advantage of this method is that it is convenient for transition. Only if the template is given, XML documents are to be formed fast. However, the template-drive mapping is a shallow one because this method cannot extract the relational mode from RDB in fact and thus has to ignore all constraints reflected by original

database. It can only put the data in RDB into an XML text and has no ability to realize the mapping in the opposite direction. In a word, this kind of mapping method cannot really realize the data exchange between heterogeneous databases.

2.2 The Model-Drive Mapping

The model-drive mapping does not only rely on embedded SQL commands as the template-drive mapping does, but also uses a concrete model to realize the real mapping. The model-drive mapping is a deep mapping because it involves models between different data modes. The essence of the method is to establish a bidirectional mapping between relational mode and XML Schema, and to design a flexible and intelligible mapping model which can express the various constraints in the original RDB well. There are usually two different methods about the model-drive mapping. One is the table-based mapping and the other is the object-relation mapping. The table-based mapping applies to RDB while the object-relation mapping applies to object database and hierarchical database. This paper discusses on heterogeneous RDBs, so the table-based mapping is exactly what I adopt. In the table-mapping, XML documents are considered as a (or a group of) table(s) during the transition. A very simple model about that is shown as below:

```
<Database>
<Table>
<Row>
<Column1> . . . </Column1>
. . .
</Row>
</Table>
</Database>
```

In the table-based mapping method, when the RDB is converted to XML documents, the query results will be filled in the expected places just as what is done in the template-based mapping. And when XML documents are converted to a relational database, the contents in the documents will be inserted into corresponding tables and records.

3 The Mapping of Relational Database to XML

In the model-drive mapping, two steps are indispensable when RDB is converted to XML documents. First, map the relational mode in RDB to the XML Schema document. This step is called "mode mapping". Second, extract data from the source database and then use these data to compose XML data documents according to the XML Schema document formed in the "mode mapping" step. This step is called "data embedding".

3.1   Mode Mapping

3.1.1   Establish the Data Type Mapping Table

The difference in data types is a primary aspect of heterogeneous data. A good middle data source must be able to bridge similar but different data types. So before the mode mapping, a data type mapping table must be established to associate data types of heterogeneous databases with those of XML.

The Data type mapping table is stored in an XML document. There is a simple example of a data type mapping table below (Shi Yu-jing et al. 2003):

<? xml version="1.0" ?>
<reflection xmlns="x-schema:c:\midxml\refschema.xml">
<r id="1" d="SMALLINT" x="int"/>
<r id="2" d="INTEGER" x=" r8"/>
<r id="3" d="VARCHAR" x="string"/>
. . . . . .
<r id="12" d="DATETIME" x="date"/>
</reflection>

In the example, "d" represents the data type of RDB, and "x" stands for the data type of XML.

When the mapping table is established, similar data types from different databases should be close to each other. That means their "id" values should be very close. It is because different types in databases may correspond to the same XML data type. If their id values close to each other, the conversion needs less calculation, which means better efficiency.

3.1.2   Generating the Structure Tree

In RDB, data are stored in the form of a 2-dimension table. They are structured data. XML is a self-description language which organizes data in a hierarchal and embedding way. The data are semi-structured. So realizing the structure mapping between XML and RDB is the first and a key step in the mode mapping. There is an easy but effective method to map the different structures: Generate a structure tree from the relational table, according to the table structure and reference constraints.

To deal with the structure of a table, the whole table, each line and each row is separately considered as an element. The whole database is considered as a root element as well. And then the simple structure tree is formed and it is shown in Fig. 1.

On the basis of the simple structure tree, let's think about various constraints of relational data. The constraints include the entity integrity constraint, the reference integrity constraint and the user defined integrity constraint. To realize the user defined constraint, corresponding attributes are added to the column elements.

Fig. 1 The simple structure
tree
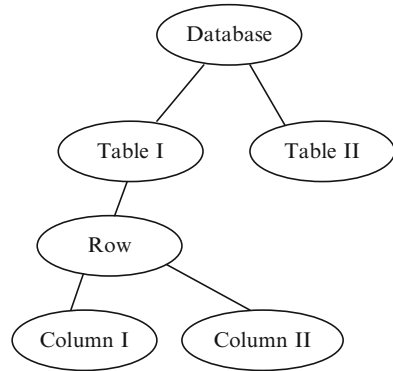
Fig. 2 The complete
structure tree
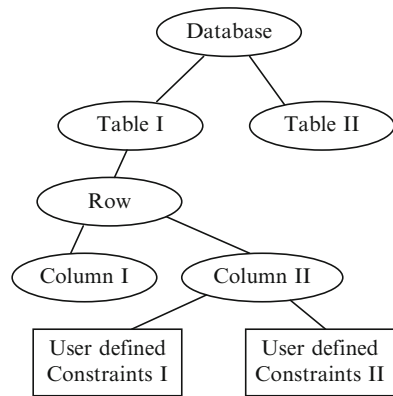
As for the entity integrity constraint, there are two ways to deal with it. In the first
method, the columns of the candidate keys are considered as the attributes of the
original columns rather than elements. For example, in the relational tree of Fig. 2,
if Column1 is the candidate keys of table 1, then Column1 should be the attribute of
Row rather than a single element. The advantage is that it highlights the candidate
keys. However, the reference constraint is difficult to be dealt with in this method
because foreign keys are usually elements rather than attributes. And if a candidate
key consists of more than one primary attributes, this method cannot work. What's
more, considering candidate keys as attributes makes documents hard understood
and the mapping in the opposite direction is not easy to be implemented. So the
second method is adopted in this paper. The candidate keys are still considered as
the sub elements of line elements and they will be dealt with when the Schema
document is written. It will be introduced soon in the paper. And the reference
integrity constraint will be dealt with as well in the Schema document.

The complete structure tree is shown in Fig. 2. The ovals in the figure represent
elements while the rectangles represent attributes.

3.1.3 Establishing the XML Schema Document

Establish the XML Schema document according to the structure tree. The main steps are listed as below:

Determine global variants about XML Schema: The variants are the attribute values of schema, which is the root element, such as namespace, target namespace and so on.

Establish the root element of database: The element is named after the database and the type is complexType.

Establish the table element: The elements are named after tables and the type is complexType.

Establish the line elements: Establish the element which stands for one record in the table. It's best to name it as "A_row". "A" is the name of the table. For example, if the element is set up on the basis of a line of table Book, then the line element is named as "Book_row". This way to name the elements can avoid name conflicts of different tables in the same database. The type of the line element is usually complexType.

Establish the column element: The elements are named after columns.

Determine the data type: According to the data type mapping table, determine the XML data type of column elements or the type to be inherited. If the type has to be intertied, column elements are in fact defined as simpleType or complexType.

Map the constraints of column elements: It is mainly about the user defined integrity constraints of column elements' attribute nodes. Most of such constraints can be dealt with in the way of restriction. Constrain data types of column elements with restriction tags and then delete the attribute nodes which are considered as constraints in the structure tree. As for some other constraints (such as default value and fixed value), the attribute nodes (such as default attribute and fixed attribute) may be kept in the tree. If there are user defined constraints which are unable to be dealt with by all the two ways mentioned above, retain their attribute nodes while do not actually constraint them in XML Schema.

Deal with the entity integrity constraints and the reference integrity constraints: Deal with candidate keys, primary keys and foreign keys. Candidate keys are dealt with by unique tag, primary keys by key tag and foreign keys by keyref tag. The selector tag, field tag and XPath attribute are also needed. Selector tag corresponds to the name of the table and the field tag to the name of the fields. For example, there are two tables in the database Bookstore, table Book and table Author. Book has two candidates Book_id and ISBN. Book_id is selected as the primary key. Book has a field called Author_id, which is a foreign key referring to the candidate key Author_id in table Author. It is assumed that all table elements and column elements about the two tables are established. Then the foreign keys, primary keys and foreign keys are processed as below:

As for the candidate but not the primary key:
<xs:unique name="Book_unique_1">
    <xs:selector xpath="Book"/>
    <xs:field xpath="ISBN"/>
</xs:unique>

As for the primary key:
<xs:key name="Book_key">
    <xs:selector xpath="Book"/>
    <xs:field xpath="Book_id"/>
</xs:key>
<xs:key name="Author_key">
<xs:selector xpath="Author"/>
    <xs:field xpath="Author_id"/>
</xs:key>

As for the foreign key:
<xs:keyref name="Author_id_ref" refer="Author_key">
    <xs:selector xpath="Author"/>
    <xs:field xpath="Author_id"/>
</xs:keyref>

If a candidate key consists of more than one primary attributes, embed a replacement element in the line element and then refer these primary attributes to the replace element. Deal with the replace element when candidate keys need to be handled.

3.1.4 Embedding Data

Once XML Schema document is established, the mode mapping from RDB to XML document is finished. Embedding data is what to be done next. Extract data from the source database and fill in the contents and attribute values of corresponding elements according to the XML Schema document. It is similar to the template-drive mapping. Embed select sentences into the template and fill in data from top to bottom.

3.1.5 Documents Verification

After the data are embedded, the mapping from RDB to XML is almost done. But it's best to verify the XML documents to make sure that the middle data sources are strong enough. There are at least 3 kinds of documents: the data type mapping table, the XML Schema documents and the XML data documents. And the verification includes 2 aspects. The first one is to check the format of all the 3 kinds of XML documents. The second one is to check the matching of XML data documents and XML Schema. There are some tools which can realize the verification.

4 The Mapping of XML to Relational Database

Like the mapping in the opposite direction, the mapping from XML to RDB includes two major steps. The first step is called the mode mapping and the second step is called the data embedding.

4.1 The Mode Mapping

In the mode mapping, the task is to map XML Schema to the relational mode. So objective database and tables will be created and constraints will be set up.

4.1.1 Creating the Database and Tables

With the interface of the Document Object Model (DOM) offered by XML, resolve the XML data document. Extract the root node which is named after the database, the child nodes of the root node, which are named after tables, and the child nodes of the table nodes' child node, which are named after fields in the tables. Create the database and tables in the objective database according to the names in those nodes' tags. Take use of the data type mapping table when defining types of the fields in tables, to convert XML data types to those of a specific RDB.

4.1.2 Set Up All Constraints

By the technology of DOM and Xpath, extract attribute nodes which are created as user defined constraints and extract the contents of the <restriction> tags which are embedded in the column elements. And then the user defined integrity constraints can be established. Similarly, set up the candidate keys, primary keys and foreign keys according to the contents of the <unique>, <key> and <keyref> tags.

4.2 Data Embedding

After the mode mapping, embed all the original data stored in the XML data documents into the new database. Traverse data documents through the interface of DOM, extract the text contents of every column nodes, and insert the contents into the tables as a record by means of some database object technology.

5 Conclusion

In this paper, I try to introduce a general method to exchange data between heterogeneous RDBs via XML. This method is model-drive and can not only exchange data but also convert modes.

However, because some user defined constraints are not strongly structured as other constraints, it is not easy to deal with them since we cannot know them exactly before the exchange. Maybe a document which contains a list of formatted common user defined constraints can be created and there is a mechanism to expand the list when a new kind of constraints occurs. How to organize this table and how to integrate it to the method in this paper is worth further researching.

This paper just introduces a method but refers little to the implementation. Literature proposes the architecture of data integration through XML (Pendyala et al. 2003). A similar architecture can be built to realize this method. So establishing the architecture is another interesting work for the future.

References

Bray T, Paoli J, Sperberg-McQueen CM (1998) Extensible Markup Language (XML) 1.0., World Wide Web Consortium (W3C)

Codd EF (1970) A relational model of data for large shared data banks. Commun ACM 13(6):377–387

Liu Xiaoming (2009) Integration of heterogeneous database based on XML [J]. Inf Technol (2): 48–51

Pendyala VS, Shim SSY, Gao Z (2003) An Xml based framework for enterprise application integration [J]. IEEE international conference on E-commerce, 2003–06, pp 128–135

Richard Wang Y, Madnick SE (1990) A polygen model for heterogeneous database systems: the source tagging perspective. In: VLDB. Morgan Kaufmann Publishers Inc., San Francisco, pp 519–538

Shi Yu-jing, Niu Chun—liang, Ma Xin—na (2003) Transmitting data between heterogeneous data sources by XML[J]. Mod Comput 2003(11):79–81

A Study on the Defining Model and Packaging Method of Legacy Systems Based on the Survey and Design Integration Platform

Xiang-Qing Wang and Ding-Fang Chen

Abstract Based on the analysis of the legacy systems in large railway design institutes, a defining model, which features railway system packaging method, is defined for the legacy systems. At the same time one of packaging method for the legacy systems based on web service technology has been proposed by using SOA. The legacy systems has been packaged and incorporated into the survey and design integration platform.

Keywords Survey and design integration platform • Legacy systems • Packaging methods • SOA • Web Service

1 Introduction

In the process of designing high-speed railway, large railway design institutes have a high degree of dependence on the specialized software and survey and design integration environment with a wide range of professionals involved and complex coordination. In an effort to improve the design efficiency, a large number of specialized software, which play a significant role in the railway design, have been set up in these design institutes after the accumulation of more than a decade.

However, we can find some problems in these software, such as long span of developing time, considerable difference in software structures, different data storage formats, etc. On the other hand, the survey and design integration platform has been established in the cloud, and all the work has been incorporated into its collaborative design platform, but in these so-called legacy systems(Aminian 2003), there's no access to the cloud computing environment if not packed (Bian et al. 2004).

X.-Q. Wang (✉) • D.-F. Chen
College of Logistic Engineering, Wuhan University of Technology, Wuhan 430063, China
e-mail: wxq178@163.com

To solve this problem, the paper aims to propose a defining model and package the legacy systems by using SOA architecture and Web Service interface technology. Then the legacy systems are incorporated into the survey and design integration platform as a part of cloud.

2 Analysis of the Legacy Systems in the Survey and Design Integration Platform

2.1 Introduction to the Survey and Design Integration Platform

With the rapid development of railway survey market, design process required to achieve the integration of design and management, which means the three personnel in survey, design and management will operate on the same platform. The survey and design integration system's ultimate goal is to create a intelligently integrated design system that combines a number of functions (such as project's design and calculation, technical management, production management, results archiving, etc.). Besides, this system, which achieves the integration of survey, design and management, is also the connection between various survey and design stages, individual designers, and departments of designing, technical management and production management (National Enterprise Management Innovation Approval Committee 2011).

2.2 Relationship Between the Survey and Design Integration Platform and Legacy Systems

The survey and design Integration platform is a collaborative one built within the enterprise's internal cloud computing environment. To save IT resources, legacy systems need to be packaged so that they can be served as a part of the cloud and exchange services with the survey and design integration platform system. The cloud is primarily to resolve the integration problem among multiple heterogeneous systems, which packages the enterprise's professional computing capacity and data resources into a service. With different Web Service interfaces as the main forms of service, it provides on-demand to internal users through the network, satisfying their large amount of data computation and storage requirements when designing the railways.

Being the professional tool software used within different departments, legacy systems fulfill their functions independently. In addition, it provides service outside of the profession as well. Since it is the platform for the professional co-design,
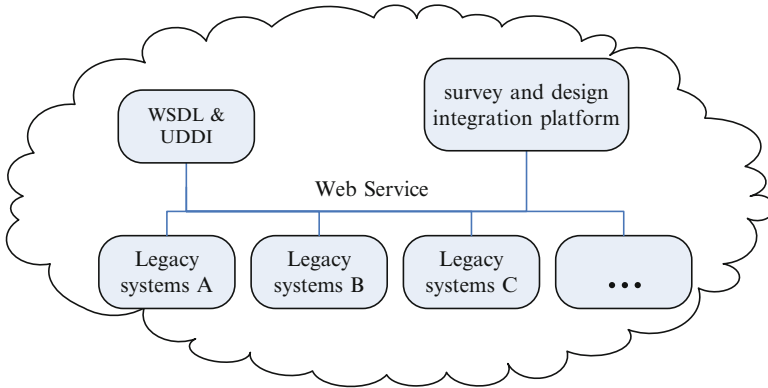
Fig. 1 Relationship between the survey and design integration platform and legacy systems

integration is not only the user of legacy systems' services, but also provides coordinated services to the legacy systems, which we call the loosely coupled relationship (Fig. 1).

2.3   Current Problems in Legacy Systems

Legacy systems primarily include such categories as various management software, professional software used by railway designing departments, CAD-based design software and commercial software purchased from outside, among which most are specialized software and design software with a leading structure of stand-alone version and C / S. Their data are mainly stored in the stand-alone version.

After years of informatization, the large railway design institutes have built a great number of different types of systems within their enterprises. But they can't be integrated into the enterprise's cloud computing environment if without being packed and then will probably become the "information islands" with the following problems:

1. the long span of construction time. With no access to network interfaces, some systems are even the stand-alone ones written in C language that appeared during the 1990s. Because some legacy systems have been used for a very long time, it's meaningless to modify their packages again. So they can be directly redeveloped or not incorporated into the collaborative environment.
2. different system structures. Most legacy systems are stand-alone ones, as well as C / S and B / S structure. The disaccord in system structures will result in non-standard data access interfaces.
3. inconsistent data storage. Data structures and data access interfaces are not standard, namely typical heterogeneous system. For example, most professional

design software use AutoCAD as a secondary development platform. Data are not separated for centralized storage and have no network access interfaces, instead, they are stored directly in the drawings during the design process.

3 Study on Defining Models for Legacy Systems

3.1 Basic Concept for Defining Models of Legacy Systems

Package defining models are the control ones based on properties, which determine the packaging methods by using discriminant function to analyze the attributes of legacy systems. The properties of legacy systems include the system's construction period, software architecture and data storage. Whereas the packaging methods of legacy systems contain redeveloping under the survey and design integration platform, changing local data into those released by Web Service, and transforming data interface directly.

1. the system's construction period indicates the one from the year when it is put into use to the present. It mainly affects the system's language, scope and the need of redeveloping. For instance, if the system's construction period is over 5 years, we can directly define the need to redevelop in accordance with the collaborative design standard.
2. software architecture is the architecture used during system development, complete with stand-alone system, B / S architecture, C / S framework, etc.
3. aiming at the data produced by the legacy systems, data storage refers to storage location, sharing method and external interfaces, which mainly involves no data storage, local storage, drawing storage, network text storage, network database storage, etc.

3.2 Definition of Defining Models

1. Let P be the system construction period, S be software architecture, D be data storage, B be packing method (Bian et al. 2004).
2. PAk (k ∈ [1, K]), SAm (m ∈ [1, M]), DAn (n ∈ [1, N]), represent respectively the construction period of legacy systems, software architecture and data storage properties.
3. ATTR (p), ATTR (s), ATTR (d) indicate the assignment relationship among construction period P, software architecture S and data storage property D. So we have:

Table 1 Defining results of legacy systems in the survey and design integration platform

Legacy system	Period	System structure	Data storage	Result	Action
A bridge design App based on AutoCAD	10Y	Standalone app	Local storage	0	Redevelop
Human resources system	3Y	Client/server	Web-based database	55	Partial transformation
Electrification design resources database	2Y	Web/server	Databased&Ftp	75	Building interface
Design task decomposition & Plan system	4Y	Client/server	Network database	45	Partial transformation
File management system	4Y	Web/server	Web-based database	70	Building interface
Highway subgrade design CAD	1Y	Client/server	Network database	35	Partial transformation
Contract management	4Y	Web/server	Web-based database	85	Building interface

$$ATTR(p)\ PA1 \times PA2 \times \ldots \times PAK$$

$$ATTR(s)\ SA1 \times SA2 \times \ldots \times SAM$$

$$ATTR(d)\ DA1 \times DA2 \times \ldots \times DAN$$

Assigning the subject's single property, for example, Year (y) = 5, Structure (s) = C / S, Data (d) = Local storage and so on.

4. In one rule, the packaging methods for model B are determined by a f a function based on p, s, d. Rule: Choice (s, r, e) ← f (ATTR (p), ATTR (s), ATTR (d)) function, according to property value of p, s, d and its return value to determine the types of packing method. For example, if returning 0, we need to redevelop the system or not to incorporate it into the integration platform. If returning between 0 and 60, we need to transform locally while build interfaces if returning between 60 and 100.

5. A strategy can be composed of multiple rules, with each consisting of the construction period Y, the software architecture S and data storage D.

3.3 Applications of Defining Models

Based on the defining models above, we can define and classify the legacy systems under the survey and design integration platform. Shown in the following Table 1, definition and classification results of some legacy systems are as follows:

Giving the way of definition and classification, the defining models of legacy systems also provide the basis for the selection of packaging method. As shown in the above table, most legacy systems in the survey and design integration platform can be transformed, needing no redevelopment. Otherwise, to meet the requirements, we can simply package them by local adaptation and establishing data interfaces.

4 Study on the Packaging Method of Legacy Systems Based on web Service

4.1 SOA Architecture Analysis

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

The essence of Service-oriented architecture (SOA) is all the business objects or business models. As for the "service", it's a form exposed by these business models. Because of this uniform service, it's more convenient for different providers and customers to exchange information (Jun-hao et al. 2010). Using SOA architecture and Web Service to solve enterprise's integration problems shows the following advantages: (1) simplicity: with a unified coding format, Web Service is easy to design, develop, maintain and use. (2) open standards: Web Service is based on a set of open standards such as SOAP, WSDL, UDDI, HT-TP and so on. (3) Flexibility: integration based on Web Service is very flexible. Being the loose coupling between service publishing applications and services using applications, Web Service can pass through the firewall. (4) cross-platform: no matter what language Web Service uses and what platform to operate, you can call/visit Web Service if just knowing its WSDL. Therefore, the integration of enterprise resources based on SOA and Web Service captures a lot of attention among research institutions and enterprises, and reaps certain application (Jiang et al. 2010).

4.2 Packaging Method of the Legacy Systems in the Survey and Design Integration Platform

To package the legacy systems in the survey and design integration platform, it brings all the legacy systems into SOA framework for a unified management. Integration systems and legacy systems share data via Web Service. In this paper, steps of the web service-based packaging for legacy systems are as follows:
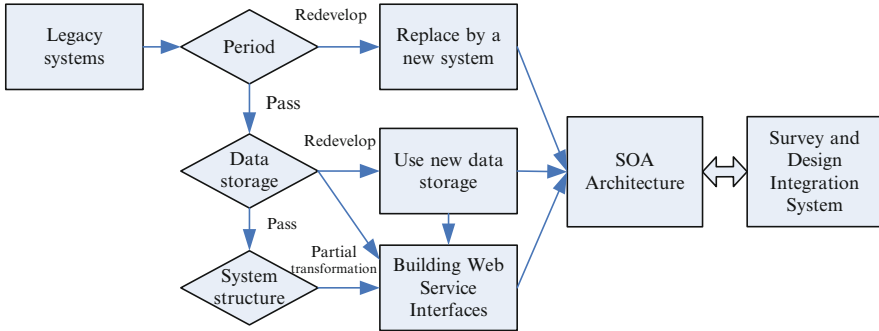
Fig. 2 Legacy system development based on SOA architecture

Step1: analyze and judge the legacy systems in accordance with their defining models. For the one, of which the construction period is over 5 years, we can directly redevelop or not incorporate it into the integration platform. While for that which is less than 5 years, we proceed to Step2.

Step2: For the stand-alone version, that is, the system's local storage data fails in data sharing, data storage needs to be changed. First, we can change from the stand-alone storage into network storage, and then build Web Service interface to meet the requirements of SOA architecture. Continue to Step3.

Step3: For the data of C / S and B / S structure, build Web Service interfaces at the data storage side (background process).

Step4: Build WSDL and UDDL, register all Web Service interfaces for legacy systems, and incorporate them into the process of integration synergies according to the requirements of SOA architecture (Fig. 2).

4.3 Packaging Implementation of Legacy Systems Based on Web Service

There are also many C / S system-based desktop systems within the enterprise. The major development technologies contain NET platform related technologies, J2EE platform related technology and C++ technology. And most application software systems are two-tier structures in the form of C / S, mainly including the client application and database server programs, which are respectively defined as the front and background programs.

If not considering that the front and background programs need specific software applications to support desktop systems, we use the following methods to package the general systems into Web Service: the way to package the system under the. NET platform into Web Service is similar to that used in the. NET platform, which directly packs Web Service-oriented Web Resource. This means to build a new Web

Service project which induces the whole system's DLL files in front and background programs and adds [Web Method] to NET Act. This method can be released into the Web Service, that is to say, packaging the system's related methods into Web Service.

The way to package the system developed under the J2EE platform into Web Service is similar to that used in the. NET platform, which directly packs Web Service-oriented Web Resource. This time, we can induce the JAR package of front and background programs and create a new implementation class. Then we can import the relevant implementation method with the help of JAR package, compile the class which contains the method, and release WSDD document into the Web Service by modifying the relevant business (Zhou et al. 2010).

5   Examples to Apply the Packaging Method

Combining the research on key technologies and some railway design applications, we have selected a number of typical legacy systems to verify their packing methods discussed in this paper, including the railway station design system in the stand-alone version, bridge design system in C / S structure, B / S Human Resource Management Systems and other technologies for Web systems and desktop systems.

Take the railway station design system in the stand-alone version as one example, this system is a typical professional design software based on AutoCAD. Its system construction period is 3 years and data are stored in the drawing. After judging by a defining model, we need to transform the data storage, modify software architecture and build Web Service interfaces. With the help of packaging method, the system's architecture transformation makes data stored in the database according to corresponding projects. Moreover, the construction of Web Service interfaces provides a unified standard for data design, which does achieve the collaboration throughout the whole railway station's design process.

6   Conclusion

All in a word, based on the study of the packaging method for the legacy systems in the survey and design integration platform, this paper has proposed a defining model suitable for the large railway design institutes, with which a number of legacy systems have been modified. After that, the legacy systems have been incorporated into the survey and design integration platform, achieving the collaborative purpose.

References

Aminian HA (2003) The legacy system dilemma, making the right choices, Technical report, USA, 36(10):153–176

Bian Y-G, Xu H, Wang Q (2004) A study on the legacy systems packaging method based on the WebServices, Computer Era, 12.26–30

Jiang Z, Lu Y, Ruizhi S (2010) On integration of heterogeneous enterprise resources based on SOA. Comput ApplSoftw 27(1):53–59

Wen Jun-hao, Zeng Jun, Zhang Zhi-hong (2010) Security policy of attribute-based access control in SOA. Comput Sci 37(9):147–150

National Enterprise Management Innovation Approval Committee (2011) The 17th national enterprise management modernization innovation achievement, Management Press, 3

Jiang Zhou, Lu Yang, Rui-Zhi Sun (2010) On integration of heterogeneous enterprise resources based on SOA. Comput Appl Softw 27(1):51–53

The Research of Tank Control System Based on AT89S52

Yong Shi, Kuan Huang, Chang tao Wang, and Bin Ma

Abstract AT89S52 microcontroller of the control system for the control of the core, combined with sensor technology, amulticontrol functions in one, including temperature, automatic lighting, automatic water level control, to achieve the control of fish tank.

Keywords SCM • Tank • Water level control

1 Introduction

Currently, many families and hotels, shopping malls and other public places are still using an artificial hand for viewing the water tank test, level control, water cycle, feeding and other operations. For the aquarium's conservation issues, the market also has been found in a variety of fish tank water temperature control, drainage (Xuehai Hu 2005), aeration and lighting equipment, such as filters, heaters, pumps and oxygen equipment to improve water quality. However, as to many products, function is not uniform, and mostly non-intelligent, single, constant temperature control, oxygenation, or lighting systems(Atmel corporation 2010). Ornamental fish tanks for manual operation bring a lot of inconvenience, thus the focus of this paper is to watch the fish tank on the current status of control equipment, analytical and research applications, a fish tank control system design(Hailan Zhao et al. 2004).

Y. Shi (✉) • K. Huang • C.t. Wang • B. Ma
School of Information, Shenyang Jianzhu University, Shenyang, Liaoning, China
e-mail: shiyong13@163.com; huangkuan@sjzu.neu.cn; wct@sia.cn; mabin5511@163.com

2 Programs and functions of the overall system

A. *system functional components*

The system was developed with the full knowledge and analysis of various types of aquarium controllers currently carried out under the premise of the whole system is divided into the following functional subsystems: the automatic heating/cooling temperature control, automatic/timed aeration system, automatic Water level control system, automatic lighting systems. These subsystems have their own input and control output signal detection capabilities, combined with the system clock circuits, data storage circuit/watchdog/reset circuit, and the function of each subsystem parameter settings and the LCD display circuit, a common integration set of functional control system (Yong Quanyu 2004).

B. *system control parameters*

The control object for the tank system, the control aim is to make the system automatically adjust to providing optimum aquarium water quality and living environment. Note that the water quality design and environmental parameters: water temperature, dissolved oxygen of water, the water level is high or low ambient lighting and other parameters. Based on the same design principles and methods for other environmental requirements, different environmental parameters can be controlled to achieve a unified design to improve scalability. Table 1 shows the parameters of the control system processing.

2.1 The hardware structure diagram and the overall structure

A. *System's hardware*

The system microcontroller core CPU, the formation of a tank outside can be placed in stand-alone, separate ways to achieve water quality and environmental testing signal input and output control signals from various quarters, the system hardware design shown in Fig. 1.

Central control module: The main microcontroller as the core, including the crystal oscillator, reset circuit, extended memory, and so on. Interactive button

Table 1 Control parameters

Project	Control parameters	Corresponding measures
1	Water temperature	According to the system settings control the heating temperature range of bars to start, stop
2	Dissolved Oxygen	Oxygen content of oxygen under control pump start and stop Time to time or according to the system to start, stop
3	Water level	Inlet valve under the control of water level start, stop
4	Lighting	Set the time according to the system timing control lights turn on and off
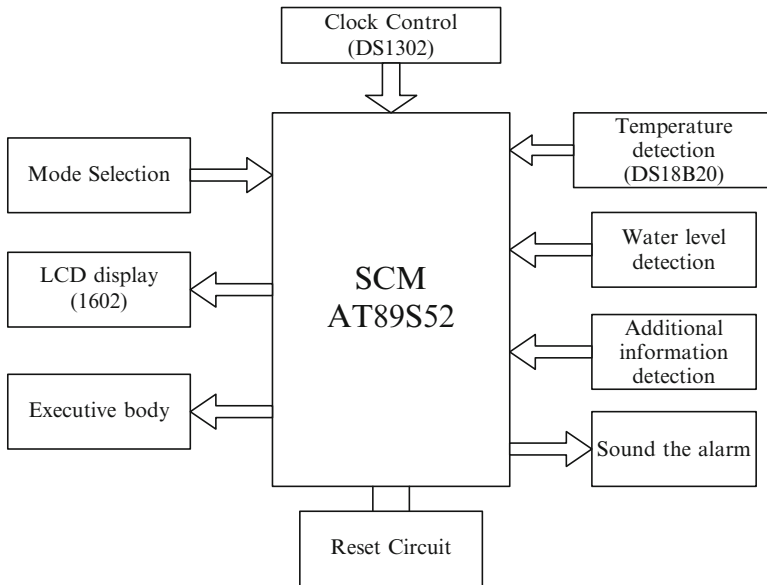
Fig. 1 System hardware block diagram

control module: the function keys for system settings, and set the appropriate course of LCD liquid crystal display control to achieve a good human-computer interaction function.

1. **in real time through the LCD liquid crystal display parameter values.**
2. **generated by LCD gauge tube and the buzzer sound and light alarm.**

Clock control circuit module: Implementing the system clock and the absolute control of the relative clock synchronization is time for environmental parameters The number of detection and the control device start and stop operation of the key. Data Storage/watchdog/reset the module: the achievement of key parameters of the storage, system working process of monitoring and restart the reset exception.

B. *The overall structure of the hardware*

Microcontroller hardware and software design are two important aspects of the system. The system hardware design, the main departure from the following principles:

1. **hardware design and software design optimized combination of hardware circuitry.**
   Some of the hardware features are available for software to implement, in turn, the function of some of the software can also be used to complete the hardware. Hardware goes with the software to function. The software's response time is not only longer than hardware, but also takes CPU time.

However, the function of the hardware with software is to simplify the hardware structure and improve the reliability of hardware, simplifying system upgrades, etc. It can also reduce costs, and therefore the system design process is in real time to meet the feasibility and possible under the premise Possible hardware functions implemented in software.

2. **reliability and anti-jamming design.**
   According to the reliability design theory, the system is less than the number of chips used in the system longer MTBF, and the less number of chips used, the address data bus on the circuit board, the less likelihood of interference, so the basic system Microcontroller. The idea is to meet the function of the number of cases and strive to use fewer chips and circuits, the system selected DS18B20 digital temperature sensor is also based on this consideration. The system mostly used features advanced DIP devices, so from the component count, board space, power consumption, interference and system costs have been significantly improved.

3. **flexible and feature upgrades and system expansion.**
   A design is often not fully taken into account all aspects of the system. The system needs to constantly improve, the need for functionality upgrade; and, the system should be considered in the design of applications in the future extension of the convenience. Extensions of the original design when the system should not need to conduct in case of big changes, modify the software and hardware, even a small amount can be done without modifying the hardware. Extensions are flexible in a measure of the importance of the merits of a system pointer. For example, the system can control a total of 8 outputs, the main control section takes up only 4, if you need to increase multi-channel control, system hardware extension is very convenient, and only minor modifications to the software can be completed. According to the system hardware requirements and some design principles above, determine the system hardware schematic diagram. AT89S52 chip microcomputer system for the central processing unit, the temperature sensing element, water level sensors, liquid crystal display, sound and light alarm, signal output unit circuit. The following major circuit designs are described in detail.

3   Host controller hardware design

As the core to form the main controller of AT89S52 the hardware circuit, shown in Fig. 2.

Hardware components:

1. real-time clock circuit: the 7,6,5 ft from the DS1302 and the CPU of the P2.4, P2.5, P2.3 connected to form (Huang Zhi 2003 and Tao Chen 2008);
2. reset circuit: the key S4 the RST pin CPU connection with the composition;
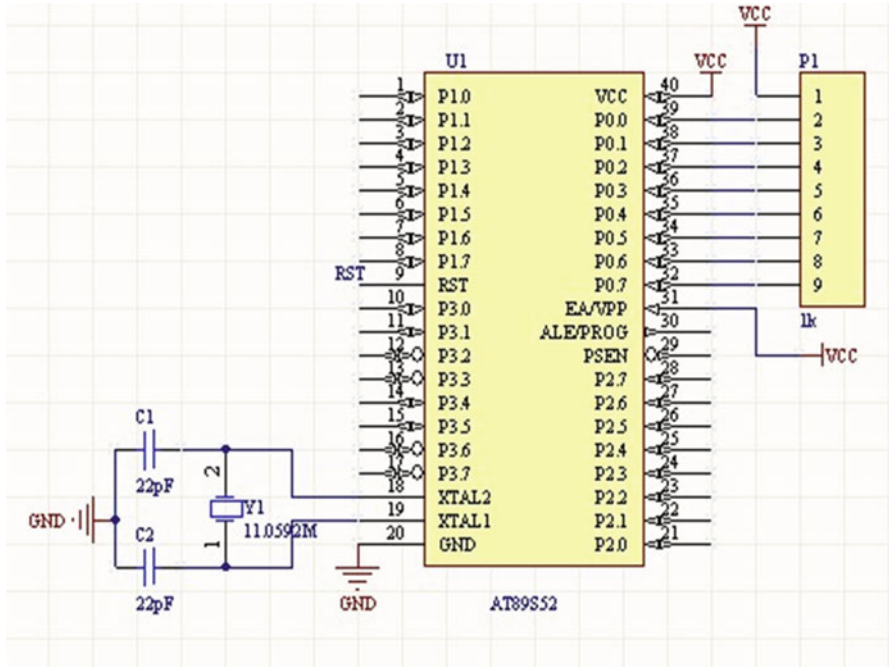
Fig. 2 The host controller hardware circuit diagram

3. keyboard interface circuit: S1, S2, S3 button pin and CPU of theP1.0∼P1.2 connected to form;
4. LCD 1602 interface circuit: DB0∼DB7 and the CPU of the P0.0∼P0.7, RS, R/W, E and P2.2, P2.1, P2.0 connected to form;
5. temperature detection circuit: the DS18B20 the I/O pins and the CPU's P2.6 connected to form;
6. buzzer circuit: and the CPU's P2.7 connected to form;
7. actuator drive circuit: the relay and the CPU of the P3.0, P3.1, P3.2 connected to form.

4 Software design

Control system software design is the design of the tank's main content and the focus of the system should be based on functional requirements to the system hardware as the basis for system software design. To make software features designed specifically to read and convenient debugging, reliability, and general use of structured programming. SCM system program is used to develop the C programming language, the SCM system (Jiacai Fu 2004 and Xin Zhang 2005), the

Fig. 3 System main program flow chart
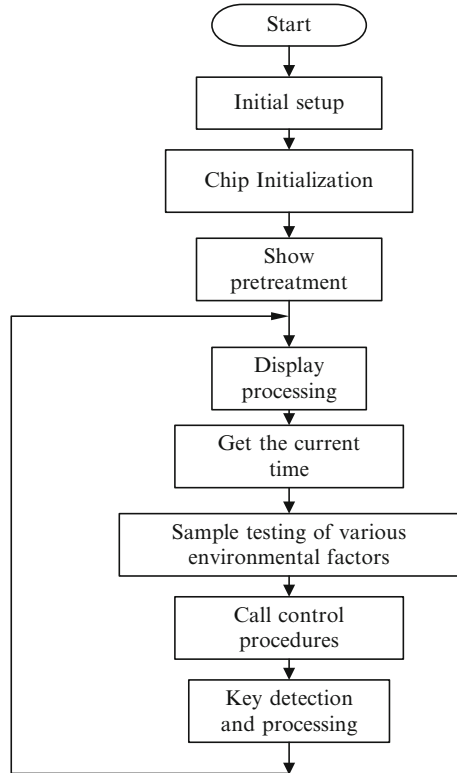
program is divided into six modules, namely the main program module, time and temperature acquisition module, detection module collection tank environmental parameters, keyboard control and display processing module, and Data storage, power protection, power on reset and watchdog module. Each module has certain features, some of which also contain a number of sub-modules, both independent and interconnected, low-level module can be called high-level module. System software described in this article includes: the main program, initialization routines, interrupt service routine, sampling and testing procedures, keyboard and display processing, sampling and testing procedures, state control procedures, and data storage, power-down protection, Power-on reset and watchdog program. The following description of each of the program modules will be the design and preparation. System main program flow chart is shown in Fig. 3.

Main program code is as follows:

```
        void delay (unsigned int delay _i);// delay procedures
void display1 ();// display program
void Init ();// initialize
void chuli1302 ();// time handler
void an jian ();// function keys
```

```
void wendu18b20 ();// function of temperature treatment
void xian wei kai guan ();// motor control function electric bar lamp
void main ()
{
 Init ();// initialize including the 1302 18b20 1602 int1 initialization function, etc.
 while (1)
   {
     xian wei kai guan ();// motor control function electric bar lamp
     an jian ();// function keys
     chuli1302 ();// time handler
     wendu18b20 ();// function of temperature treatment
     display1 ();// display program
       }
   }
```

5 Conclusion

The results show that through the practice of the control system stable and reliable, easy operation, while the system is designed to be flexible, simple, low cost, which has some use and reference.

References

Atmel corporation (2010) 8-bit Microcontroller with 8K Bytes In-System Programmable Flash AT89S52. www.atmel.com/atmel/acrobat/doc1919.pdf

Guoqin Zheng, Zhelei Xia (2002) Integrated sensor applications started. Zhejiang Science and Technology Press, Hangzhou, pp 27–33

Hailan Zhao, Jian Zhu, Xiangwei Zhao (2004) DS1302 real-time display of time and application of the principle. Electronics, pp 43–46

Huang Zhi (2003) Start with the real time clock chip DS1302 data acquisition system. Microcontroller and embedded system applications (1–6):365–367

Jiacai Fu (2004) SCM control engineering practice technology. Chemical Industry Press, Beijing, pp 155–160

Tao Chen (2008) Microcontroller applications and C51 programming. Mechanical Industry Press, Beijing, pp 6–18

Xin Zhang (2005) Microcontroller theory and applications. Electronic Industry Press, Beijing, pp 6–13

Xuehai Hu (2005) Microcontroller theory and applications design. Electronic Industry Press, Beijing, pp 3–6

Yong Quanyu (2004) World popular MCU technical manual –U.S. series. BeijingAeronautics and Astronautics Press, Beijing, pp 3–5

Study on Human-Simulated Intelligence Controller of–Based on SCM the Boiler

Yong Shi, Chang tao Wang, Kuan Huang, and Bin Ma

Abstract This article take the boiler water temperature controller for research object and investigate the elevator control system based on taking the SCM a microcontroller as the core as a boiler controller, sensor technology, the basis for the system design to achieve the temperature and water level of the collection. AT89S52 microcontroller as the core of the system, combined with simulated intelligent PID control algorithm performance analysis, and hardware design, but also for the actual boiler research to meet the energy saving boiler control and safety requirements.

Keywords SCM • PID control • Human-simulated intelligent control • Sensors

1 Introduction

Heating boiler drum water level and water temperature protection control in the whole system plays a very important role in the central heating boiler drum run, maintaining water level and water temperature within a certain range to ensure the necessary conditions for the normal operation of the boiler (Lanjun Li et al. 2007; Songyue Peng and Liu Fruit 2007). Water protection function should satisfy the time when the boiler water protection from Ganguo and burn water wall, when there is full of water valve opens automatically; when the water level reached the limits of water level will shutdown, shut down, turn off the main Valve to prevent equipment damage (Zhang and Peng 2004)

Y. Shi (✉) • C.t. Wang • K. Huang • B. Ma
School of Information, Shenyang Jianzhu University, Shenyang, Liaoning, China
e-mail: shiyong13@163.com; wct@sia.cn; huangkuan@sjzu.neu.cn; mabin5511@163.com

2 Circuit Design of the Boiler Master Controller

A. Temperature Acquisition Circuit

Temperature Acquisition using the DS18B20 digital temperature sensor, the sensor if required it can be accurate temperature conversion, I/O line must ensure power supply during the conversion. Since the work of current to 1 mA of DS18B20, so the pull-up resistor 5 K alone is not enough to provide power, and when there are a few DS18B20 hanging in the same I/O lines and also to the temperature conversion, the problem is more acute (Ai-Ghazzawi et al. 2001).

Way is to DS18B20 supply VDD pin is access from an external power supply, shown in Fig. 1. The advantage is I/O line does not need to strengthen the pull, and bus controller not always remain high during the conversion temperature. This can allow the conversion during the single data bus for other exchanges. In addition, single bus can carry as many pieces DS18B20 temperature sensor, and if they all use external power, then it can first send a Skip ROM command, followed by a Convert T command, so that they at the same time the temperature conversion. Note that when coupled with an external power supply, GND pins can not be left floating. When the temperature is higher than 100 °C, the parasitic power use is not recommended, because in this temperature DS18B20 show relatively large leakage current, communications may not be possible. In this case the temperature should be used DS18B20 the VDD pin.

B. Water level monitoring circuit

Hall liquid level sensor installed in the container outside the permanent magnet in the float on the branch, with the level changes, the role of the Hall devices on the changes in the magnetic flux density, which measured level. Structure shown in Fig. 2.

Hall effect sensor is achieved by Hall converted a magnetic sensor. It has high sensitivity and wide application characteristics. Its working principle is: a semiconductor thin, its length L, width B, thickness D, magnetic induction placed
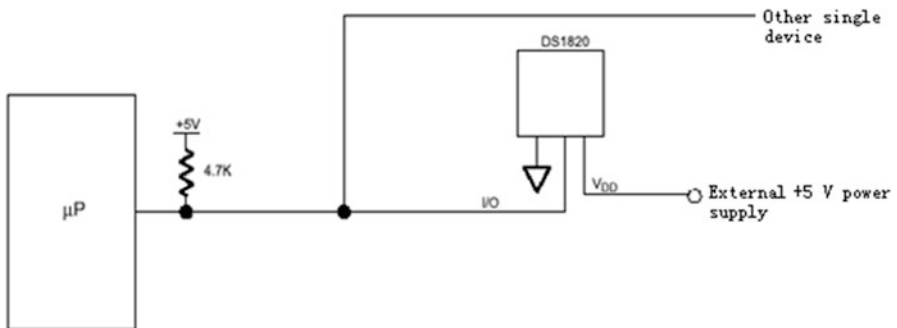

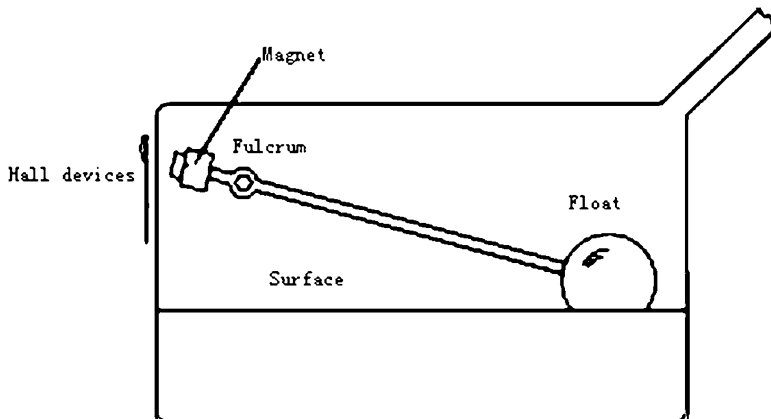**Fig. 1** Reference to an external power supply

Fig. 2 Hall devices measuring the water level

in the magnetic field strength B, in opposite sides of the pass to control the current I, and orthogonal magnetic field direction and current direction, Then both sides will produce a semiconductor with the control current and the magnetic induction is proportional to the product of the electric potential U, is the Hall voltage of the electric potential, with the U_H, said that $U_H = K_H IB$, where K_H is the sensitivity of the Hall element, the semiconductor thin Is the Hall element

Similarly there are two permanent magnet magnetic field the same, with the polarity of the relative placement. When the surface area far greater than the distance between the two when the magnetic field strength of the middle O, along the z axis in the gap between the formation of a uniform magnetic field gradient $dB/dx = K$ (K is a constant.) $B = 0$ Service as a reference to the origin of the displacement x, then $x = O$ when, $B = O$, $U_H = O$. When they move to the middle of the Hall element x at the time, U_H Department of B size determined by the x. $U_H = K_H IB$ known by the formula: I remain unchanged, $dU_H/dx = IK_H dB/dx = K_H I = K$, points after that $U_H = Kx$, the Hall potential is proportional to the displacement. The greater the magnetic field gradient, the higher the sensitivity, the more uniform magnetic field, U_H, and x the linear as possible.

The connection diagram of the Hall element and the microcontroller is shown in Fig. 3.

Hall element with the induced magnetic field intensity voltage signals, signal amplification by the differential amplifier, and then after a filter is filtering to get a clearer signal, the signal again after A/D converter to convert the spread to the microcontroller for processing. The design is by the microcontroller as a master controller, microcontroller to control the thyristor heating, lower machine PID regulator, to the signal converter and solenoid valve, add water by adjusting the inverter control pump and control solenoid valve
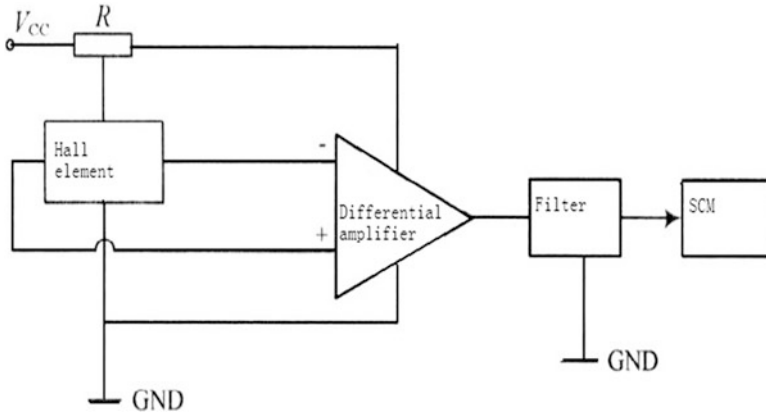
Fig. 3 Hall element connected with the microcontroller

3 Human-Simulated Intelligent PID Control Algorithm and the Capability Analysis of Controlling Algorithm

Human-simulated intelligent PID control combines the traditional PID control algorithm and human-simulated intelligent control algorithm together and uses proportion (Mansour et al. 2005), integral and differential control function to realize selecting feature information and combined control scheme according to dynamic process. In human-simulated intelligent control, the solving information space is made up of system error e (t), error rate e' (t) and time t. e (t) = r (t)−y (t), r (t) is input, y (t) is output. Apparently whether fixed value control or servo control, the aim is when t tends to infinity, e (t) = e' (t) = 0. The movement trajectory of system in the information space not only reflects all dynamic information but also reflects the control function in the system.

It can be seen from the static characteristics, the system deviation from the equilibrium state, the error increases when it is the direction of proportional control, error reaches extreme point, the control volume reaches the maximum, then decreases the error began, Control the amount of time the most extreme point with a lot more than a small amount of error and no longer change with the change, but "wait", "observation" error changes. When the error is zero when the system has the opposite direction to increase the percentage of control. Have roughly the same repetition and control process, so that after one or two oscillations, the system quickly stabilize the new equilibrium point (Lee 1990).

From the automatic control system performance indexes, the concrete analysis of the human-simulated intelligent PID is as the followings:

1. The Stability
   The method of combining open-loop and closed-loop is good for stability of large delay system. From energy perspective, when the error increases,

Fig. 4 PID control program flow chart
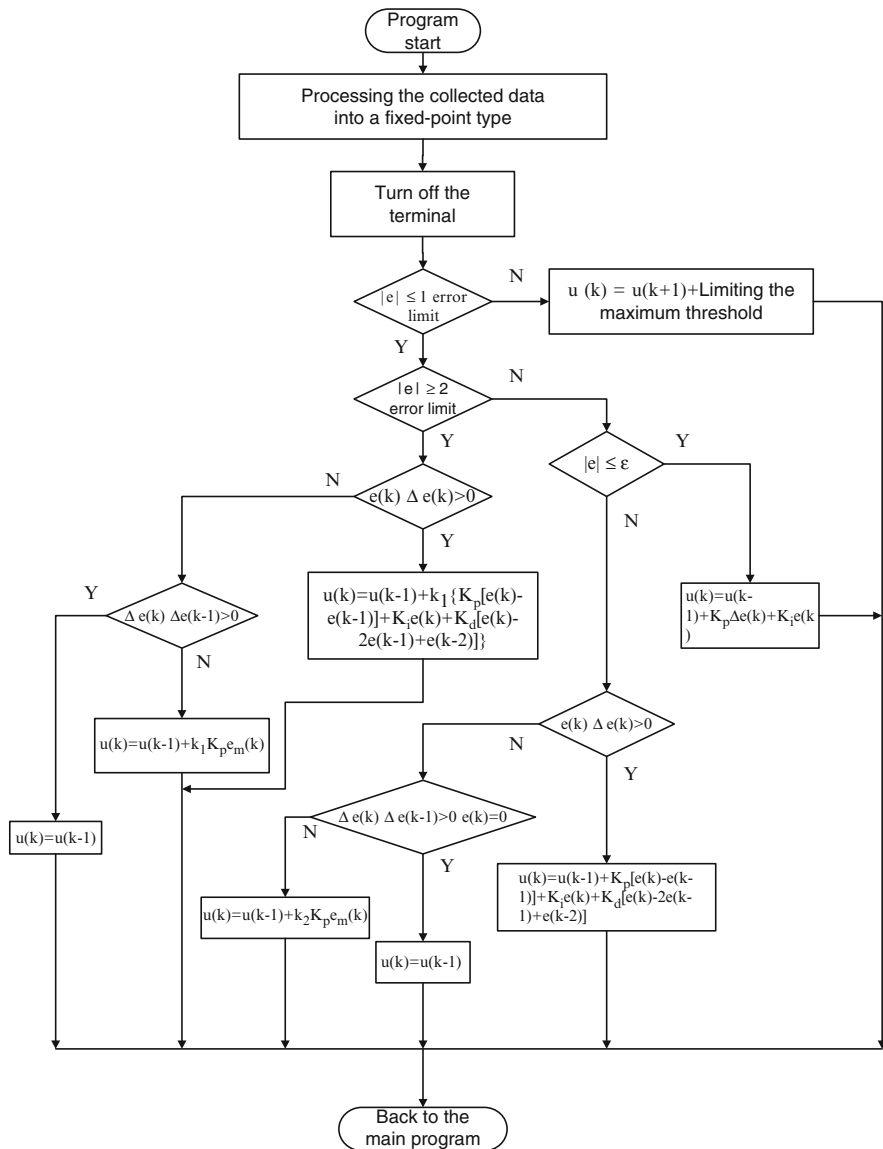
human-simulated intelligent PID controller inputs larger control energy by using the proportional control to offset disturbance input energy. When the error decreases, the intelligent control algorithm switches to the open-loop holding mode to guard its change. Such strategy avoids system oscillation and improves the system stability effectively (Guangzheng Peng 2004).

2. Quickness

For multi-model control, because the system will automatically switch control mode, so when the error becomes larger or has increasing trend, it can input larger control quantity and the system reflects quickly. In addition, the system can stabilize after one switch mode by setting appropriate parameters and it also shortens the setting time.

3. Accuracy

Because of using the open-loop hold mode and control depending on memory characteristic, multi-model control have integration function and the system stable error is zero. Overall, relying on the flexible recognition and decision control method, the multi-model control mode solves the problem of the control quality (power, stability, accuracy) in contradiction.

4   AT89S52 Microcontroller Programming

PID control program flow chart in Fig. 4

5   Conclusion

In this paper, SCM boiler control system, combined with practical requirements of the boiler control operation designed to SCM (AT89S52) control the main line, with simulated intelligent PID control design can be achieved out of the boiler water level, temperature control of the hardware circuit, the system SCM control system with a series of advantages, less selection, and simple circuit structure, software function quite perfect, cost-effective. Boilers can control some of the more simple, with a certain value.

References

Ai-Ghazzawi A, Ali E, Nouh A et al (2001) On-line tuning strategy for model predictive controllers. J Process Control 11:265–284

Lanjun Li, Shouyi Yu, Song Gao (2007) Fuzzy PID control of micro-boiler level [J]. Computer 11(1):33

Lee CC (1990) Fuzzy logic in control systems: fuzzy logic controller [J]. IEEE Trans Syst Man Cybern 20(2):419–435

Mansour SE, Kember GC, Dubay R et al (2005) Onlineoptimization of fuzzy2PID control of a thermal process [J]. ISA Trans 44(4):305–314

Mingli, Guangzheng Peng (2004) Fuzzy PID control algorithm in cylinder position servo control application. J Mach Hydraul 10

Songyue Peng, Liu Fruit (2007) PLC based control system of heating furnace temperature [J]. Autom Appl 10:32

Zhang L, Peng G (2004) PLC-based central heating boiler feed water protection control. Computer Information 13:25–28

Based on the Control of P2P Networks Data Transmission and the Model of Data Sharing Research

Zhang Yizhi and Liu Jiangfeng

Abstract Recently, P2P's applications develop very fast and are becoming very important in the network. As gradually increasing, P2P applications definitely cause the network's immensely consuming and network congestion. Therefore this article gives a method of distinguishing P2P's flow based on flow and payload in order to improve this situation. At the same time, in order to establish a better campus network for sharing resources, this article also gives a model of PRNM data sharing based on P2P technique, which is using P2P half distributed architecture to put the resources on each peer.

Keywords P2P network • Control of data transmission • Control of flow • Model of files sharing

1 P2P Operating Principles

P2P is Peer-to-Peer which is also called "point to point". The Peer technique is new for network, which is relay on both capability and bandwidth of participants in the network, not few servers. It allows a client to connect another client's computer directly to share and exchange data, and changes the ways of browsing and downloading through servers. It reveals the basic goal of the network, that is, connected with each other and point to point, and gives a good further view in distributed and shared network (Xiren Xie 2008). In some opinions, P2P is kind of traditional network technique because the basic designed plan of network is that makes computers can communicate with each other without the third party.

Z. Yizhi (✉) • L. Jiangfeng
Network Center, Chongqing Three Gorges University, Wanzhou, Chongqing, China
e-mail: zyz38@163.com; ljfl38@163.com

Table 1 Traffic identification technology is relatively tables

Name	Description	Pros	Cons
Deep Packet Inspection (DPI) technique	Detect application layer's payload to find applications of P2P in a packet	Easy implementation; convenient maintain; no restrict on the dynamic port changes	Weak on catching the encrypt P2P flow; need update to catch unknown P2P flow
Depth of flow behavior Inspection and recognition technique	Using flow characters to inspect P2P applications, it is developing	Strong on catching the encrypt P2P flow and unknown P2P flow	Accurate is hard to improve; Complicated implementation; Weak on sorting of P2P applications

2 The Research of the Control Technique of P2P Data Transmission Flow

Classic P2P flow recognition technique is developing from Central architecture and Distributed architecture to mixed architecture, and the characters of port develops from fixed port and random dynamic port to camouflage port. With describing the recognition technique, this article gives the method of P2P flow recognition based on both flow and payload (Zuev and Moore 2005; Lixia Chen 2005).

The method of P2P flow recognition based on flow and payload is compared and researched with the methods of P2P flow's recognition below. The technique of P2P flow's recognition can be categorized: Deep Packet Inspection (DPI) technique and Depth of flow behavior Inspection and recognition technique. They are compared as following (Table 1):

3 Construct the Model of P2P Data transmission's Flow and Data Sharing

With the scalar of campus network and applications of network increased, campus network has become the major method of obtaining information between teachers and students. Campus network's range is increased from few miles to couples ten of miles even different branches in a city, the scalar is expanded from hundreds of clients to couples thousand of clients. More and more campus networks are started to implement and the resources are becoming abundant (Xingjia Xu et al. 2006; Hongyi Yu 2005). P2P as a new data sharing technique is developing strong enough to make the network easier to share and exchange.

In order to realize the construction of network information better and let clients use network resources better, this article presents a model named P2P file sharing (PRNM) which is used for campus network. The key technique for realizing P2P network is discovering services that are used for any resources. The principle of

peer discovery is: firstly, a signal is sent by peer; then, the other peers receive this signal, update the list of peers and send back a response to the previous peer (Hui Tang and Guojie Zhang 2005; Yan Lin and Qing Yao 2003).

Flow characters analysis in P2P network, computer S is server and also client, so characters are as following: a lot of inflow connections(the other ports connect theirs' monitoring ports) and also a lot of outflow connections(which connect the other points' monitoring ports). When S is not a part of P2P, S must be a server or client, so the connection of client S must not be balance. Up flow and down flow is approximately balance. For traditional network's applications, such as HHTP, FTP, generally client sends a request and waits the data needed from a server, which up flow and down flow is not balance.

Based on the above characters, we can present this model: according to the result of in connections divide out connections in a period time, and comparing this result with experience observed value of traditional network's clients, we can judge whether or not client S is a part of P2P. Meanwhile, S_{in} means the in connections of client S, and S_{out} means the out connections of client S.θ is the observed value.

$$\frac{1}{\theta} \leq \frac{S_{in}}{S_{out}} \leq \theta \tag{1}$$

According to the divide value between up flow and down flow in a connection, when this value is in a particular range, we can decide that this connection is of P2P. Meanwhile, T_{up} is the up flow, T_{down} is the down flow, and θ is the observed value.

$$\frac{1}{\theta} \leq \frac{S_{up}}{S_{down}} \leq \theta \left(T_{up} > T_{min}, T_{down} > T_{min} \right) \tag{2}$$

According to the divide value between all of the up flow and down flow, when this value is in some particular range, we can decide that this client is a part of P2P. ΣT_{up} is the up flow of client S, $\sum T_{down}$ is down flow of client S, and θ is the observed value.

$$\frac{1}{\theta} \leq \frac{\sum T_{up}}{\sum T_{down}} \leq \theta \left(\sum T_{up} > T_{min}, \sum T_{down} > T_{min} \right) \tag{3}$$

We will group the peers in order to increase the capability of resource sharing and the efficiency of resource searching. The principles of grouping are: according to the peers' IP addresses, we should separate most of network transmissions in the same group by grouping them into many small groups, in order to decreasing the use of network bandwidth (Xiaohui Su 2007). For example, in the campus network, all teachers are separated in a single peer group and all students are separated in a single peer group. So, in every peer group, the information user capture is in the same format. Because the interests and problems they concern have similarity, if they share the resources in the peer group, this will improve the search efficiency, decrease the count of routing, and give a better network bandwidth, also solve the problem of controlling of security of resource (Yahua Yu 2006).
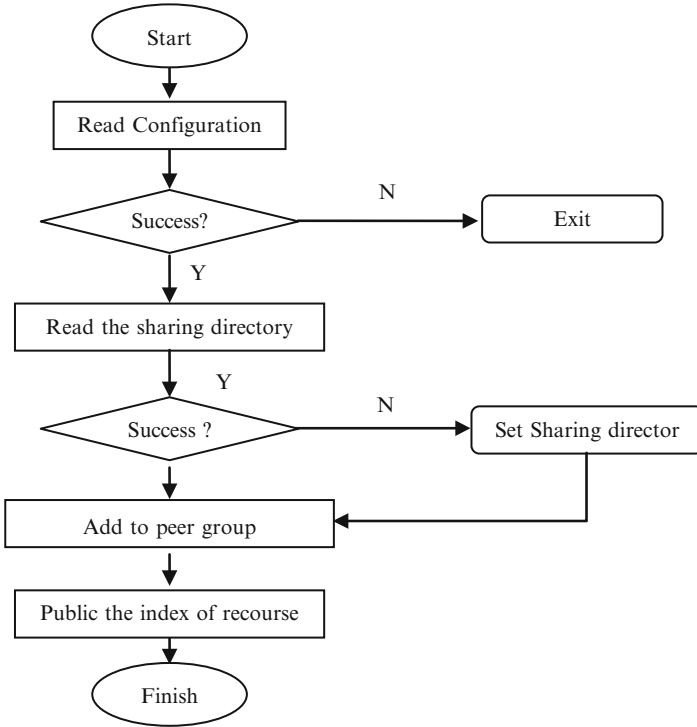
Fig. 1 File search work flow model

We can use the IP multicast technique to realize that files searching which means the process of searching the sharing resources in the network. Firstly, users need to files' names they want to search, and then system will pack them into searching info and send them to super peer groups in the peer group. After traversing the hash of resource lists and searching the file in the peer group, system will return that file if search successful, if not, the searching information will be sent to FTP Server. The procedure of hot point backup resources on FTP Server will search this file in the backup resource after receiving the search request. If search is successful, system return that file, if not, the hot resource backup routine will send the search information to the other peer groups and return a result. The model of process is following in Fig. 1:

The model of searching file includes two functionalities. One is sending the information of searching and then displays the result. The other one is monitoring the information of searching by super peer, and progress the information monitored. Firstly, sending the searching information, system return the results if searching is successful. The peer sends the searching information which is the name of resource and then attaches the multicast addresses of the peers. Secondly, system sends the searching information to the super peer in the peer group, and displays the results on the sharing resource lists (Liang Zhang and Futai Zou 2004).

Fig. 2 Model of Resource backup file
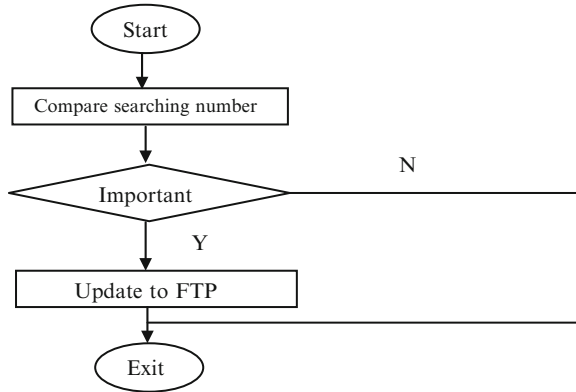
Important resource backup mechanism which is that setting the routine of backup important resource on the FTP Servers. This function maintains a list of the rank of searching resource. This routine will monitor the searching information. If a peer sends a searching request, the routine will monitor this information and update the list of the rank of searching resource. See the Fig. 2 for the model of resource backup.

There are two functionalities of important resource back: one is backup the important resource, another is when the peers start searching and transferring the information to FTP servers, the routine of important resource backup will progress the searching information, which is searching in the backup resource and return the results (Xinbao Zhang 2003).

4   Conclusion

With the increasing of network users, especially when BT and Xulei software are prevalently, depress under broad band becomes heavier and heavier, the problems of serious congestion in network are easy to prop up. The paper aims to strengthen management and monitor on network flow, optimize network quality so as to provide a method to recognize P2P flow based on flow and payload, in order to make P2P locally and minimize the cost of flow. Also, this article designs a model of controlling of P2P network flow to solve these problems.

This article compares the techniques of different kinds of P2P and analyzes the campus network. Based on these studies, this article provides model named P2P resource network (PRNM) which improves traditional one. Based on the peer technique and FTP servers, the author provides the mechanism of important resource backup which improves the ability of sharing in the network. Also, the model of PRNM cannot solve all the questions, it needs more research.

References

Hongyi Yu (2005) New self-organizing P2P network and routing algorithm. Posts & Telecom Press, Beijing

Hui Tang, Guojie Zhang (2005) A hybrid P2P network model research and design. Comput Appl [J] (3)

Liang Zhang, Futai Zou (2004) A survey and prospects of information retrieval on the peer-to-peer Network. Comput Sci [J] (4)

Lixia Chen (2005) P2P-based model of knowledge sharing, forum contemporary education [J] (15)

Yahua Yu (2006) Study on P2P software's impact on campus network's data traffic. J Jianghan Univy (Soc Sci) [J] (2)

Xingjia Xu, Houxin Wang, Yan Zhou, Weidong Jin (2006) Discovery implementation by IP multicast in peer to peer network. The development of computer technolgy [J], June

Xiaohui Su (2007) Collaborative work model and Application based on P2P. J Pingdingshan Inst of Technol [J] (4)

Xinbao Zhang (2003) Research infringement on the internet. China Renmin University Press [M], Beijing, the 11th issue of journal

Xiren Xie (2008) Computer network. Publishing House of Electronics Industry, Beijing, pp 240–242

Yan Lin, Qing Yao (2003) New type of hierarchical P2P model for research. Comput Appl [J] (12)

Zuev D, Moore A (2005) Traffic classification using a statistical approach. In: Passive & active measurement workshop, Boston, USA, March 2005

Design and Implementation of the Anti-spam System Based on AIS

Jue Huang and Changwu Liao

Abstract With the wide use of Email on the internet, it is urgent to resolve the severe problem of spam. Inspired by the biological immune system, it proposes an algorithm that can classify Spam without re-training. The result of experiments showed that the algorithm, compared with Bayesian algorithm, can improve recall rate and has a higher dynamic adaptability and stability.

Keywords Artificial immune system • Spam • Bayesian • Algorithm

1 Introduction

With the growing popularity of email, it plays a more and more role for people to communicate with each other. However, as the carrier of commercial advertising, viruses or content-sensitive, the spam occupies space and network bandwidth, people have to spend much time to pick normal mail from lots of spam. Spam even makes serious damage for our life and network. Anti-Spam has become a global problem. As spam is dynamic and various, we designed and implemented an anti-spam system based on the artificial immune.

2 Current Status

Because spam has became serious, there are many anti-spam system has been developed. They can be classified to based on email server or based on email client, and the email detect technical can be classified to based on rules

J. Huang (✉) • C. Liao
Department of Computer and Software, Nanjing Institute
of Industry Technology, Nanjing, China
e-mail: huangj@niit.edu.cn; liaocw@niit.edu.cn

(Carreras and Marquez 2001), based on statistic(Sahami et al. 1998), based on connection(Cao et al. 2004).

The detect technical based on rules, such as decision tree, interact rule, they are used mainly in some special field, they need domain knowledge of the field and rule lib to support and the rule's design, update, self learning is very difficult. So they are not used very popular and the performance is not very good.

The detect technical based on connection, such as Artificial Neural Network. The method is good in some aspect, such as self learning. But it needs lots of sample to train system and the train need lots of time. So the method is hard to be used.

The detect technical based on statistic, it adopt pure statistic principle and system self learning method. It builds up characteristic set by learning known sample to detect new sample such as Bayesian algorithm. Most of the current anti-spam systems use the technical.

But the Bayesian algorithm is not perfect. First, it based on independence assumption. The technical doesn't consider the connection between words or phrase, so it can't analyze email content especial Chinese. Second, it can't continuously self learning. If the mail content has much change, the correct rate will downgrade. And the technical has worse anti-interference performance. The Bayesian classify system classify email based on known sample, It has high detect rate. But the detect performance depends on training sample. So for the unknown sample especially mutation sample, its detect result is not very good. Unfortunately, more and more spam sender use the method to escape detect system.

For that reason, the paper adapt to the protect mechanism of biology immune, use artificial immune system to build up anti-spam system. Give a new anti-spam system module based on artificial immune system.

3 Artificial Immune System Introduction

3.1 Biology Immune System VS Anti-spam System

Artificial immune system is a new research direction after artificial neural network and evolvement calculate. Recently, artificial immune system has become a research hot spot for message security and network security.

Biology immune system is a complex system consist of immune cell, organ. The major function of Immune system is distinguishing self and Non-self. When antigens begin to invade, immune system will generate lymph cells to resist antigens. Immune system can self learning, recognize antigens, immune remember and so on (Mo 2003).

Biology immune system and anti-spam system are similar. The following Table 1 is their compare relation.

Table 1 Natural system and email system

Natural system	Email system
Self cells	Normal emails
Non-self cells	The emails to be detected
Native Cells	Native detector
Memory Cells	Memory detector
Antigens	Spam
Co-stimulation	Feedback message from user
Affinity	Similarity between emails
Cells lifecycle	Detector lifecycle

3.2 Immune Remember

When immune system meet antigens at first time, It will generate immune response, when it meet the same antigens once more, because associate member, the member cells will be wake up, immune system will give response fast then last time and generate lots of immune body to destroy antigens. During anti-spam system running, remember detect set will be build up according to immune cells' remember characteristic.

3.3 Affinity

Affinity was defined to match rate between antigens and immune body. The spam detect process just is match process between detect set and unknown characteristic and judge whether the unknown characteristic is self or not.

The vector length of detect set Ta, The vector length of unknown mail Tb, statistic the same word between two vector and assign to Tsame, so we can get.

Tsame = Ta ∩ Tb, Tmin = MIN(Ta, Tb), Affinity = Tsame/Tmin.

When affinity greater than threshold, the mail will be judge to spam and require customer response for the detect result.

3.4 Detect Set Generate Algorithm

Detect set is the collection of immune body. Extract cell section randomly to generate un-mature detect set, then un-mature detect set match with normal email, if match can't success, the un-mature detect set will evolve to mature detect set. If the count of match spam in limit time exceed algorithm, mature detect set will evolve to remember detect set. If mature detect set can't match in a long time, it will be delete from mature detect set.

4 Chinese Spam Filter System Design

4.1 Mail Pre-Process

When Chinese email incoming, first parse email content, then split words from mail content and filter stop words.

4.2 Pick Characteristic

The popular characteristic select method is word appear frequency, message gain, mutual information, statistic, expect cross entropy. They all calculate statistic message for a Chinese word, then set a threshold, delete the words their statistic message less than the threshold. Others are valid characteristics. The system will use mutual information as evaluate function.

$$difference(W) = \left| \left| \log \frac{P\left(W|C_1\right)}{P(W)} \right| - \left| \log \frac{P\left(W|C_2\right)}{P(W)} \right| \right|$$ (1)

P(W):the word W appeared probability in all documents;
P(W|Cj):the word W appeared probability in document j.
We sort the difference according to the calculation result, select some words as characteristics.

4.3 Build Up Vector Space Module

Phase email, then split words, select characteristics, generate words set, then describe email as a method which fit to calculate, and classify them.

4.4 The Judgment of Unknown Email

The following flow will imitate the characteristic of Biology, use member detect set, mature detect set analyze unknown email.

```
//unknown email detect algorithm
Read mail vector
{Read remember detect set;
    Match email vector with each remember detect set member;
If(affinity>MT) //if it can match, memory detector regard the mail as spam
User confirm the detect result
```

```
    If (user think it is a spam)//email.flag=1
    {Put EM into spam set;
        Put all cells of EM to cell set;
        Ab1.count=Ab1.count+1;}//the match count of immune body +1
    Else delete remember detect set;
  Else {Read mature detect set;
    Match mail vector with each member of mature detect set;
        If (affinity>MT) //if it can match, native detector regard the mail as spam
            User confirm the detect result;
            If(user think it is a spam)//email.flag=1
                {Put EM into spam set;
                Put all gene from EM into gene set;
                Ab2. count=Ab2. count+1;
                Ab2. age=Ab2. age+1;}
        Else delete mature detect;}
  Else User confirm the email;
    If(user think it is a spam)//email.flag=1
        {Put EM into spam set;
            Put all gene from EM into gene set;}
    Else Put EM into self set;}
```

5 Experiment and Result

The purpose of the experiment is compare artificial immune system and Bayesian algorithm. The operate system is Windows XP Professional, and language is C++. We select 2,000 Chinese emails as the sample.

5.1 Accuracy

Accuracy is the correct rate of system detect. From the Fig. 1, the accuracy of AIS is 85–90%, the accuracy of Bayesian is 82–90%, so we can see that AIS has a better stability robustness than Bayesian.

5.2 Recall

Recall stands for call back rate. The high recall rate mean less spam leak by system. From the Fig. 2, the recall rate of AIS is better than Bayesian.
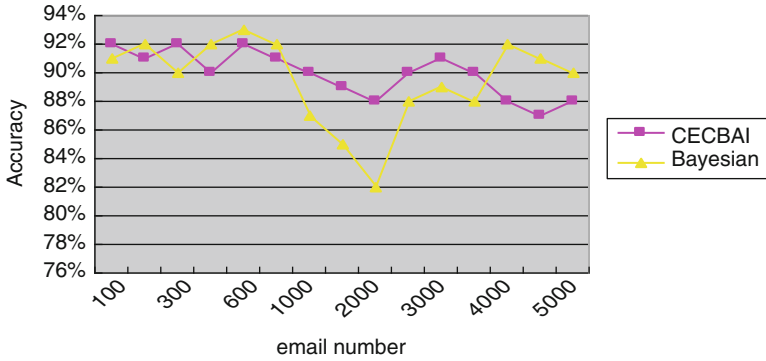
Fig. 1 The accuracy of AIS and Bayesian
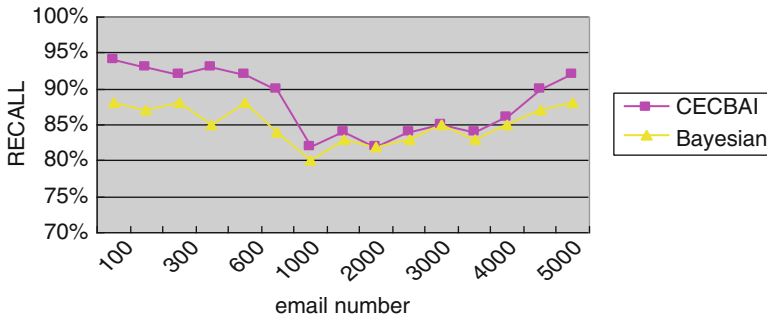
Fig. 2 The recall of AIS and Bayesian

6   Conclusion

The paper implements anti-spam system based on artificial immune algorithm. The system can handle mutagenesis and has better performance. In recall rate, it is better than Bayesian algorithm.

References

Cao YK, Liao XF, Li YF (2004) An email filtering approach using neural network, Lecture notes in computer science. Springer-VerlagGmbH, Heidelberg, pp 688–694

Carreras X, Marquez L (2001) Boosting trees for anti-spam email filtering. In: Proceedings of Euro conference Recent Advances in NLP (RANLP – 2001), pp 58–64

Mo HW (2003) The principles and application of artificial immune system. Harbin industry university publishing house, Harbin, pp 38–48

Sahami M, Dumais S, Heckerman D (1998) A Bayesian approach to filtering junk email. In: Proceedings of AAA'I workshop on learning for text categorization, pp 55–62

A Study of Design Technique of Dynamic Data Report in PowerBuilder11.5

Liu Huacheng

Abstract The data report is one of the frequently designed softwares. According to the changing feature of data, data report has two forms: static statement and dynamic statement. This paper discusses the method of designing dynamic data report in PowerBuilder11.5 and the technical points of realizing dynamic data statement. This method applies to all types of dynamic report forms.

Keywords Report forms • Dynamic • Datawindow object • Field attribute

1 Introduction

PowerBuilder11.5 is the forehead developing tools of database with formidable function, popular with software developer. But the format of report is unitary and the styles are a few provided by PowerBuilder11.5. If only simply using the report object provided by Power Builder to make report, it cannot satisfy manufacturing requirement of some specialized report, especially the manufacture of dynamic report. Therefore, after completing the design of data report, according to the practical situation of application management system and uses' demand, it can realize dynamic manufacture of data report by dynamically control various attributive setting of data report object.

L. Huacheng (✉)
Chongqing Three Gorges University, Wanzhou Chongqing 404000, China
e-mail: bbagg2004@sina.com

2 Design Idea and Solution

Regarding to A4 size paper, after fixing good title and footnote position, the content of report can contain different numbers in accordance with different choices of designing report. How many records can be held on longitudinal and how many field projects can be held on the deflection in a report, which is invisible for developer to design report object and only visible only when program runs. If the number of rows of each page can be gained in the process of running procedure, dynamic controlled lines are high; it can realize the function of each page of dynamic alignment report. Similarly, if depositing the attribute of field projects of report object into data sheet, we can modify the attributes value regards to practice after operating applicative procedure, for example: Whether the field demonstrates, width occupied by setting, setting of demonstration coordinate and so on, thus realizes the dynamical manufacture of data report (Sha Jin and Jia Yufeng 2003).

2.1 Demonstrations and Printed Number of Controlling Data Report

Generally speaking, each page of data report will demonstrate the number of lines of data, when just used to print and output. However, in the process of development, it is invisible. In fact, the key point of realizing control of printed number of rows is Height value of Detailed Band of controlling DataWindow object.

There are two ways of controlling Detail the Band high method, the one method is the computation method: Namely attribute of Units of DataWindow object is changed into 1/1,000 cm, and then the height (centimeter) of printed paper is gained. Using this height to subtract Header Band, Summary Band, the Footer Band, then divide printed number of lines of each page to get height of Detail Band. The formulas are following (Ye Zhijun):

Detail.Height=[Paper.Height–(Header.Height+Summary.
Height+Footer.Height)]/(Number of lines/each page)

Another method is: not to change the attribute of Units of DataWindow object, Units is still PBU, then reading the number of lines of each page in the process of procedure, comparing with number of rows input by users. If current number of rows are larger than that input by users, we can increase the height of Detail Band, if less, we can decrease.

As the systematical unit of PB definition, according to the system, PBU calculates its PBU unit according to system front information installed by machine, which is better than previous one in the light of portability; Next, the second method is use PBU unit, generally speaking, the integer arithmetic is more simply than the previous one in respect of controlling part of procedure.

In order to realize controlling of printed number of row relying on the second one, we first of all need to know some attributes relevant to data window (Jung and Hong 2000).

2.1.1 LastRowOnPage Property

There is LastRowOnPage attribute in the PB data window; its function is records the line number of last record in each page of data window surrounded in current printing model. For example: When the data window is in the non-printing pattern, each page has 5 records, then the LastRowOnPage value is 5, 10, 15 . . . ; Similarly the same data window under printing pattern, each page may hold 30 records, then the LastRowOnPage value is 30, 60, 90 . . . , obviously the LastRowOnPage value has something to do with the current printing pattern and current page (Ahuactzin and Gupta 1999).

The PB data window has two kind of patterns: Non-printing pattern and printing pattern. The switch between two kinds of patterns is realized by modifying, setting the attribute of Print Preview as Yes or No. In case of default, the value of data window object, PrintPreview, is demonstrated as No, by the data window control of window. Meanwhile, the size of page in data window object is the same as data window control. That means the data window control in the window can show the lines of data, the page in the data window only can show the same data. We should pay attention to a key point when using LastRow Page that there are should be data in current data. If there is no, the value is 0; if the data in data window are less than a page, the value is just the line number of last record (Sha Shixuan and Wang Shan 2001).

2.1.2 Detail Band

Detail Band is the place where upload data window object, demonstrate and compile data. The record taken out from the back database will be loaded completely in the detailed band. Therefore, it is also the important part to be adjusted in the printed report. Each row of Detail Band shows a record of database. Consequently, the number of row that each page can contain has something to do with the row height of Detail Band when printed. After making definition of the height of Detail Band, Footer Band, Summary Band, the height of Detail Band is the key point to decide the number of printed row in each page (Xie Wei et al. 2002).

The attribute of Band of data window are various, we should use one attribute of Height. When procedure runs, we can use Describe sentence to obtain the Height attribute of Detail Band promptly, then according to the user's needs, Height value can be adjusted by Modify sentence to realize the control of printed row (Ke Jianxun et al. 2003).

Its core code is as follows:
// Obtains the Detail Band height
str_detail=dw_print_modify.Describe("DataWindow.&
Detail.Height")
// If the number of lines input is larger than the current number of rows, height is reduced.
If li_row>li_pagerow_user THEN

```
li_page_modify=li_row
If li_page_modify>= 1 THEN
DO
// will transform the height to integer
li_detail=long (str_detail)
// the height is decreased
li_detail=li_detail - 1
// Will transform the height into the string of characters
str_detail=string(li_detail)
// Revises the height of Detail Band
dw_print_modify.Modify("DataWindow.Detail.Height=&
"+str_detail)
li_row=long(dw_print_modify.Describe("DataWindow.&
LastRowOnPage"))
// If the number of adjusted row is larger than or equal to that input by users, it
will stop.
LOOP WHILE li_row<li_page_modify
li_pagerow_user=li_row - 1
sle_pagerow.Text=string(li_pagerow_user)
else
MessageBox("alarming information","each page should&
not be less than a row",Information!,OK!)
End if
else// If the number of row input by user is smaller than the current number of
row, its height will be increased
li_page_modify=li_row
If li_page_modify>= 1 THEN
DO
li_detail=long (str_detail)
li_detail=li_detail+1
str_detail=string(li_detail)
dw_print_modify.Modify("DataWindow.Detail.Height=&
"+str_detail)
li_row=long(dw_print_modify.Describe("DataWindow.&
LastRowOnPage"))
// If the number of adjusted row is smaller than or equals to that input by users,
it will stop.
LOOP WHILE li_row>li_page_modify
li_pagerow_user=li_row - 1
sle_pagerow.Text=string(li_pagerow_user)
else
MessageBox("alarming information","each page is not&
less than a row",Information!,OK!)
End if
End if
```

Table 1 Field attribute memory table(t_name)

Field name	Field type	Field width	Whether to permit NULL	Default value
Serial number	Integer	Default	No	Automatically
Tabulation title name	Char	60	Yes	None
Data window title	Char	60	Yes	None
Title quotation name	Char	60	Yes	None
Data quotation name	Char	60	Yes	None
Width value	Decimal	12(2)	Yes	None
X value	Decimal	12(2)	Yes	None
Y value	Decimal	12(2)	Yes	None
Whether to demonstrate	Char	20	Yes	None

2.2 Dynamic Control Data Report Form Field Attribute

When designing applicative procedure, we may design a data sheet t_name first, and stores the attribute name of each field in data report into this form.

The Table 1 shows Field attribute memory table(t_name)

In the application procedure, we can set attribute of each field in accordance with users' practical need. For example: Actual demonstration titles, demonstration position of data field, demonstration width as well as whether to demonstrate and so on. In the application system, after searching records of data in the database, we can modify attribute of each field to realize data report by calling Modify () function dynamic to realize dynamic manufacture of data report. The code of attribute of each field in these dynamic control data can be packed by defined function so as to call initially in the procedure. The core code is following (Cui Wei 2002):

// read each field of its attribute in the data sheet t- name, then Reads various fields and the attribute value one by one from data sheet t_name, then dynamically control data sheet through Modify() function.

```
sum=parent.dw_1.rowcount()
for li=1 to sum
// The control title demonstrated the name so to sees the attribute
ls_xsbt=parent.dw_1.object.data.primary.current[li,3]
ls_yybt=parent.dw_1.object.data.primary.current[li,4]
ls=ls_yybt+".text="+"~""+ls_xsbt+"~""
parent.dw_print.modify(ls)
ls_yysj=parent.dw_1.object.data.primary.current[li,5]
xs_flag=parent.dw_1.object.data.primary.current[li,9]
// Control displays attribute
choose case xs_flag
case "YES"
ls=ls_yybt+".visible=true"
parent.dw_print.modify(ls)
ls=ls_yysj+".visible=true"
parent.dw_print.modify(ls)
```

```
case "NO"
ls=ls_yybt+".visible=false"
parent.dw_print.modify(ls)
ls=ls_yysj+.visible=false"
parent.dw_print.modify(ls)
end choose
// Control displays width
li_w=parent.dw_1.object.data.primary[li,6]
ts_w=ls_yybt+".width="+"~""+string(li_w)+"~
parent.dw_print.modify(ts_w)
ds_w=ls_yysj+".width="+"~""+string(li_w)+"~""
parent.dw_print.modify(ds_w)
next

parent.dw_print.settransobject(sqlca)
parent.dw_print.retrieve()
end if
```

3   Conclusions

When designing the PowerBuilder11.5 database application procedure, we should design the data window object as the Group style to realize the function of classification and collection so as to take shape of database report, then dynamically control data report in the application procedure to demonstrate the number of row and attribute value of field so as to realize dynamical manufacture of data report, which satisfies users' desire to the full with advanced portability. The above codes of procedure have passed the test under the pattern of PowerBuilder11.5 C/S.

References

Ahuactzin JM, Gupta KK (1999) The kinematic roadmap: a motion planning-based global approach for inverse kinematics of redundant robots. IEEE Trans Robot Autom, USA 15(4):653–669

Cui Wei (2002) Tsinghua University publishing house. Tsinghua University publishing house, USA, CA

Jung HK, Hong SK (2000) Path planning and obstacle avoiding of robot using new wall Followinn algorithm. In: Proceedings of the Asian control conference, pp 1628–1633

Ke Jianxun, Zhang Tao, Shao Liang (2003) Enters step development. Tsinghua University publishing house, CA, pp 168–170

Sha Jin, Jia Yufeng (2003) PowerBuilder8.0 programs hundred examples. Scientific Publishing house, CA, pp 152–153

Sha Shixuan, Wang Shan (2001) Database system introduction, 3rd edn. Higher education publishing house, CA, pp 163–164

Xie Wei, Jiang Shiqing, Yu Zhenhong (2002) The enterprise information management data controls concurrently. CA: Micro Comput Appl 18(3):18–20

Ye Zhijun. PowerBuilder distribution network applied technology. Chinese Water conservation Water and electricity Publishing house, CA, pp 158–159

Research on Network Integration of BACnet and 6LoWPAN

Bo Wang and Yi Sun

Abstract It is necessary to introduce WSN (Wireless Sensor Network) into IB (Intelligent Building), so as to promote the intelligent monitoring and control level of building environment. As one of ISO standards, BACnet(a data communication protocol for Building Automation and Control NETworks) is a leading protocol of automation control networks in IB. BACnet defines a device type named B-SS (BACnet Smart Sensor) for WSN, but there is no implementation program on B-SS and no scheme of B-SS connected with IPv6 till now. In order to achieve a kind of WSN based on B-SS which can interconnect with IPv6, this paper analyzed the 6LoWPAN (IPv6-based Low-power Wireless Personal Area Networks) protocol put forward by 6LoWPAN WG in IETF and designed an integrated model which integrates BACnet and 6LoWPAN in TinyOS2.0. The integrated model was tested in TOSSIM and the result shows that the integration technology is reasonable and feasible.

1 Introduction

To provide a safe, comfortable, environmentally-friendly and energy-efficient living environment is the main goal of IB. Automatically monitoring and control the building environment is an effective means to achieve the goal. As a kind of communication network which is composed of wireless techniques, embedded system and micro-electro-mechanical systems, WSN is widely used to monitor environment in the industrial and military area. In order to promote the intelligent

B. Wang (✉) • Y. Sun
College of Computer Science, Chongqing University, Chongqing, China
e-mail: wangbo@cqu.edu.cn; 13983365240@139.com

degree, it is extremely necessary to introduce WSN to IB. BACnet is one of ISO standards of data communication in intelligent building. BACnet has defined six types of standard devices which could be applied to different application areas. As one of the six standard devices, B-SS was defined by BACnet protocol, but there is still no application scheme in IB. Thus, how to realize the WSN base on B-SS applied to IB becomes a problem which is needed to be solved.

All-IP is one of development trends of various communication networks. In order to achieve the convenient interconnection between WSN and IPv6, IETF set up a working group named 6LoWPAN in late 2004. To achieve the interconnection between BACnet and IP network is the key issue needed to be solved in BACnet standard. BACnet standard has defined BACnet/IP (ANSI/ASHRAE 135-2001 2003), however, BACnet/IP is not adapted to the B-SS with limited capacity and data-processing resources. So how to use 6LoWPAN protocol to realize the interconnection between WSN base on B-SS and IPv6 network becomes the main problem studied by this paper.

This paper introduced the architecture of 6LoWPAN protocol and its effect on the I B, analyzed the BACnet standard and B-SS functional requirements, designed and implemented a BACnet protocol for B-SS, realized an integrated model between BACnet and 6LoWPAN (hereinafter referred to as B-6 model).

2 6LoWPAN and IB

6LoWPAN is a protocol definition to enable IPv6 packets to be carried on top of low power wireless networks, specifically IEEE 802.15.4 (Mulligan 2007). Figure 1 describes the architecture of 6LoWPAN protocol.

The Maximum Transmission Unit (MTU) of IPv6 is 1,280 octets, but the largest package size supported by IEEE 802.15.4 physical layer is 127 octets, removing other overhead of MAC layer in IEEE 802.15.4, there are only 81 octets left for network layer in 6LoWPAN.Therefore, the adapter layer above MAC layer

Application layer	
Transport layer (UDP)	
Network layer (IPv6)	
Adapter layer	
IEEE 802.15.4	MAC
	PHY

Fig. 1 Architecture of 6LoWPAN

is necessary. How to compress the package and redesign router algorithm for 6LoWPAN is the key issues studied by 6LoWPAN WG now (Bormann 2008; Kim et al. 2008a).

The application problem of 6LoWPAN is also one of key issues needed solving. One draft concerned with 6LoWPAN application space was published by 6LoWPAN GW in 2008 (Kim et al. 2008b), which described several application scenarios. This paper mainly focused on the application scenarios of how to apply 6LoWPAN to IB. The draft only shows that 6LoWPAN could be applied to IB, but does not instruct how to realize this application.

3 BACnet and B-SS

BACnet protocol was issued by ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) in 1995. In 2003, BACnet protocol became ISO 16484-5 and EN/ISO 16484-5. Now, BACnet has become one of the most widely used building control network communication protocols. The architecture of BACnet protocol consists of application layer, network layer, virtual data link layer and physical layer. The application layer provides several BACnet services, while the network layer interconnects each BACnet networks. BACnet standard did not design a new data link layer for itself. The data link layer of BACnet includes several common data link layer protocols such as Ethernet and so on. Therefore, between network layer and data link layer, BACnet designed a new layer named BVLL (BACnet Virtual Link Layer), which provides an interface for higher layer.

The devices in BACnet are named Objects. These objects contain properties and functions. In BACnet protocol, there are six kinds of standard devices (ANSI/ASHRAE Standard 135-2001 2001); B-SS (BACnet Smart Sensor) is one of these. One B-SS device common identified with Device_Object and Analog_Input_Object. Device_Object is used to recognize one B-SS device in BACnet networks, Analog_Input_Object provides some data for B-SS functions. These two objects were designed as Table 1 in detail.

Table 1 Definition of objects in B-SS

Device_Object	
Object_Identifier	Device_1
Object_Name	"Temp_1"
Object_Type	DEVICE
Analog_Input_Object	
Object_ Identifier	Analog_Input_1
Object_Name	"B6"
Present_Value	NULL
Object_Type	ANALOG_INPUT
Units	DEGREES_CELSIUS

Fig. 2 Sketch diagram of integrated B-6 model
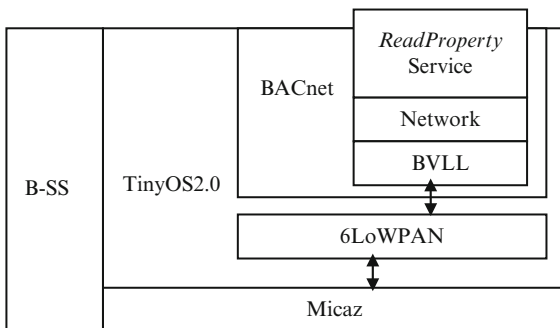
According to the BACnet standard, the necessary function of B-SS device is Data Sharing, which is realized by ReadProperty service in DataSharingReadProperty model of BIBB (BACnet Interoperability Building Blocks).

3.1 Integration of BACnet and 6LoWPAN

In order to achieve the interconnection between WSN based on B-SS and IPv6 networks, 6LoWPAN protocol and BACnet protocol in B-SS device were integrated, named B-6 model.

A collapsed integrated model for B-SS in TinyOS2.0 was designed (Levis 2006). As Fig. 2 shows, the redesigned BACnet protocol is composed of application layer, network layer and BVLL. Application layer mainly provides ReadProperty service to realize the DataSharing of B-SS. The network layer interconnected each WSN based on B-SS. BVLL provides an interface for network layer to link to a specific data link layer.

In B-6 model, BACnet protocol is a special application defined above the 6LoWPAN protocol. The arrow between BACnet and 6LoWPAN means the data exchange between these two protocol stacks.

The 6LoWPAN protocol in Fig. 2 was implemented by Matus Harvan (Harvan 2007). Micaz component in Fig. 2 mainly implements the IEEE 802.15.4 function for 6LoWPAN.

TinyOS2.0 is built up by components. Each function unit can be recognized as a component. Components in TinyOS2.0 are divided into system components and user components. Figure 3 is a mapping diagram of B-6 model in TinyOS2.0. AdcReadClientC component, TimerMilliC component and Cc2420ActiveMessageC component are system components.

AdcReadClientC component mainly provides the function of sensing data. TimerMilliC component provides the clock for whole system, and Cc2420ActiveMessageC component implements the function of IEEE 802.15.4.
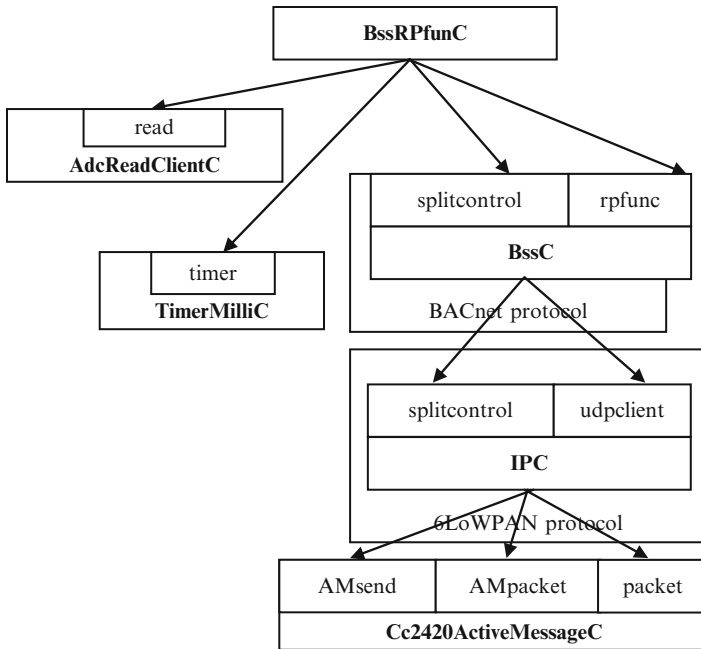
Fig. 3 Mapping diagram of B-6 model in TinyOS2.0

IPC component designed by Matus Harvan provides function of 6LoWPAN protocol. BssRPfunC component and BssC component defined by this paper mainly implements the function of BACnet protocol. A part of core program of the BssC component is:

```
configuration BssC {
/* provide interface */
    provides {
            interface Splitcontrol as BAControl;
            interface rpfunc[uint8_t i ];
                }
}
implementation{
/* Implement the wire between BssC and IPC */
components BssP ;
components IPC;
BssControl=BssP;
rpfunc=BssP.rpfunc;
BssP.IPControl ->IPP;
BssP.UDPclient ->IPP.UDPClient[ uint_16 , bpdu ];
}
}
```
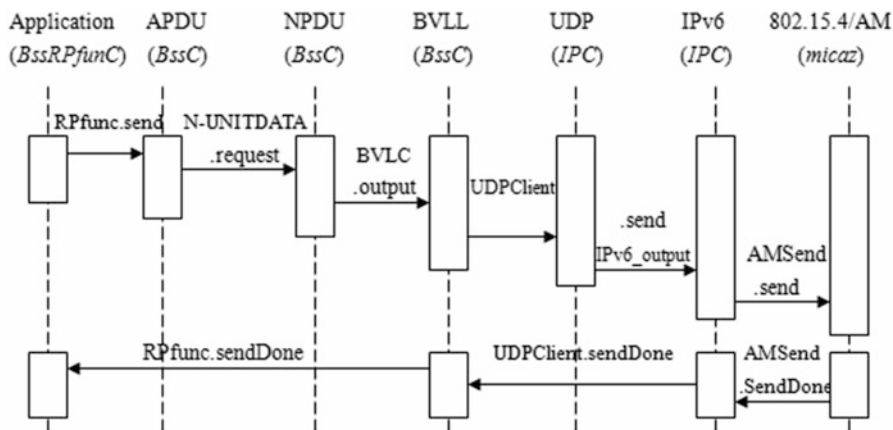
Fig. 4 Data stream of B-6 model

3.2 Process of Data Sending

In B-6 model, BssC component implements the function of application layer, network layer and BVLL of BACnet protocol, and BssRPfunC component mainly provides ReadProperty service for B-SS. Figure 4 shows the data stream between BACnet protocol and 6LoWPAN protocol in B-6 model.

In Fig. 4, BssRPfunC component first called the read interface of AdcRead-ClientC, and then assigned the data to Present_Value in Analog_Input_Object, and then sent the value including Present_Value to BssC component through rpfunc. In BssC component, the values from BssRPfunC component were encapsulated as APDU, NPDU and BPDU by corresponded functions. In the end, BssC component transferred the BPDU to IPC component through UDPClient.

3.3 Testing of B-6 Model

TOSSIM is a simulate software in TinyOS 2.0. The B-6 model with TOSSIM was tested. Firstly, in this section the function of BACnet protocol was tested in B-6 model with one single B-SS node, and then the function of B-6 model with one application scenarios according to draft "Design and Application Spaces for 6LoWPANs" issued by 6LoWPAN WG was test as well.

3.3.1 Testing BACnet Protocol

BssRPfunC component completed the function of the Data_Sharing by assign the value to Present_Value of Analog_Input_Object. And then BssRPfunC component called the BssC component by rpfunc.

Fig. 5 One typical
6LoWPAN application
scenario in IB
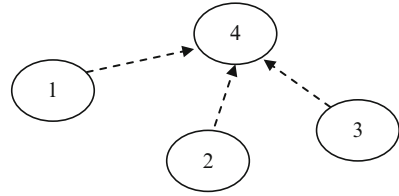
BssC component encapsulated the data from BssRPfunC component as APDU,
and then encapsulated the APDU as NPDU by network layer in BACnet protocol,
finally the BPDU was formed by BVLL. More details about the parameters of
APDU, NPDU and BPDU in testing result were mentioned in BACnet standard.

3.3.2 Testing Application Scenarios

One typical IB application scenarios of 6LoWPAN protocol was mentioned in
reference 5. The topology of this application scenario was a star topology network
as Fig. 5 shows, is composed by four smart sensor devices.

Node 1, node2 and node 3 respond for sensing data and sending data to node 4,
node 4 receives data from these three nodes. In this testing scenario, node 1, node 2
and node 3 were installed the B-6 model.

Node 1, node 2 and node 3 respectively generated the BPDU by BssC component
and then called the IPC component to encapsulate the BPDU as UDP, and finally
IPC component called the ActiveMessageC component to send the data to node 4
by IEEE 802.15.4.

4   Conclusions

The BACnet protocol for B-SS device was redesigned and the function of BACnet
protocol was implemented by defining the BssC component and BssRPfunC
component in TinyOS 2.0. The test shows that the redesigned BACnet protocol
met the needs of B-SS. The B-6 model can be achieved by calling the UDPClient
interface of IPC component in BssC component. It also can be concluded that B-6
model is reasonable and practicable.

References

American Society of Heating, Refrigerating and Air-Conditioning Engineers: Standard 135-2001
(2003) BACnet-a data communication protocol for building automation and control networks,
pp 526–537

American Society of Heating, Refrigerating and Air-Conditioning Engineers: Standard 135-2001 (2001) BACnet-a data communication protocol for building automation and control networks, pp 551–554

Bormann C (2008) Draft-bormann-6lowpan-cbhc-00.txt. Context-based header compression for 6lowpan. The 6LoWPAN working group, 8 July 2008

Harvan M (2007) Connecting wireless sensor networks to the internet – a 6lowpan implementation for tinyos 2.0. Master's thesis, Jacobs University Bremen, pp 28–42

Kim E, Kaspar D, Gomez C, Bormann C (2008) Draft-dokaspar-6lowpan-routreq-06. Problem statement and requirements for 6LoWPAN Mesh Routing. The 6LoWPAN working group, 14 July 2008

Kim E, Chevrollier N, Kaspar D, Vasseur JP (2008) Draft-ekim-6lowpan-scenarios-03. Design and application spaces for 6LoWPANs. The 6LoWPAN working group, 14 July 2008, pp 15–17

Levis P (2006) TinyOS programming. Cambridge University Press, New York, pp 27–46

Mulligan G (2007) EmNets '07 Proceedings of the 4th workshop on embedded networked sensors. The 6LoWPAN architecture. ACM Press, New York, pp 78–82

Discrete Core/Periphery Structure Characters of C2 Organizations Analysis Based on SNA

Shuchun Chang and Qiyi Zhang

Abstract The Command and Control Relationship Network Models are constructed. The structure characters of these models are analyzed using "discrete core/periphery model" in Social Network Analysis. The core/periphery structure characters of actual models can be quantitatively analyzed by calculating the relevancy degree between actual models and ideal models. It can be seen from research, if C2 organizations have clear core/periphery structure characters, the results analyzed with discrete model are relative ideal.

Keywords Social network analysis • C2 organization • Core/periphery structure character

1 Introduction

"Command and Control Relationships" research is an important content of Command and Control research fields. Most traditional researches are qualitative studies about basic conception, principle, mechanism and system design, while quantitative researches are relative few. The quantitative research on "Command and Control Relationships" is more and more important along with the development of modern war. The familiar methods such as Petri-net, SEA and Network theory, are used for constructing system models and efficiency assessment. Added to this, recently Social Network Analysis (SNA) was introduced to the field. For example, Anthony Dekker (Defence Science and Technology Organization, Australia) and his group exhibited traditional SNA and combined the new theory of complex network, and they have achieved material research findings on C4ISR architecture analysis

S. Chang (✉) • Q. Zhang
Bengbu Automobile NCO Academy Transportation Command Department, Bengbu, China
e-mail: changshuchun@163.com

(Dekker 2001, 2002, 2003). The researchers in USA DOD and British DOD also studied the command and control (C2) organization with SNA (Renfro 2001; Jarvis 2005). These studies make it clear that SNA is applicable to quantitative research on Command and Control Relationships.

Social Network Analysis is a series of techniques applied to depict and analyse the formal and informal relationships among organization members. It is generally agreed that SNA came into being as a method to study anthropology in British. In recent years, the advancements of complex network theory, computer technology and modern mathematics made it possible to study big and complex social networks with SNA. Now SNA is applied in the fields of management, economics, sociology, anthropology, physic, and military. There is a premise to study command and control relationships of C2 organizations with SNA: C2 organizations can be regarded as social organizations, the relative stable relationships are in existence among organization members, and the relationships can be studied with methods and theory of sociology field. With the premise, the conceptions of command and control relationships can be defined, the network models of relationships can be constructed, and some idiographic questions can be studied from sociological point. The attributes of C2 organization network models studied with SNA can quantificationally and qualitatively illuminate the structure characters of command and control relationship networks, and illuminate the influence of organization and their members. Then the relation between command and control relationship types and efficacy of C2 organizations can be discovered.

In this paper, with the theory of social network analysis, we will construct the Command and Control Relationship Network Models of C2 organizations, analyze the "core/periphery" structure characters of these models, and endeavor to discover the relationships between structure characters and organizations efficiency. The conclusions can be applied to the C4ISR architecture analysis and C2 organizations planning.

2 Core/Periphery Structure Analysis in SNA

"Core/periphery structure" analysis is an important part of SNA, it can make effectively quantitative research about the Core/periphery structure characters on certain network. It has been applied in the fields such as collectivity activity, science magazine citation network and crime network (Xu and Chen 2005; Krebs 2002).

SNA considers that if a social network of one organization has core/periphery structure character, the nodes of network can be separated into two groups. The relation among the members of one group is compact, and the group can be regarded as "core". The relation among the members of another group is loose, and the group can be regarded as "periphery".

In 1999, a paper written by S.P. Borgatti and M.G. Everett were published on social network analysis trade press "Social Networks". In the paper, the authors advanced a viewpoint that core/periphery structures could be studied from models

point of view. They developed two kinds of models of the core/periphery structure: a discrete model and a continuous model (Borgatti and Everett 1999). The progress of the method is that using one kind of core/periphery structure model, combining with actual data, calculating the relevancy degree between actual model and ideal model, quantitatively analyzing the core/periphery structure characters of actual model, and then reasonably judging the position of organization members(core, periphery and semi-periphery).

3   Command and Control Relationship Network Models Investigated

There are three key steps to research command and control relationships with the method of SNA. The first step is to construct the relationship network models of C2 organizations, the second step is to calculate the characteristic parameters of the models, and the third step is to analyze the structure characters of these models.

C2 organizations can be regarded as a collectivity that are composed of some entity elements which are combined together to achieve a same mission concurrently in certain circumstance. It can be seen that the functions of entity elements and the relationships among the entity elements determine the efficacy of C2 organizations. Referring these researches, we consider that the entity elements of C2 organizations can be classed to three kind units: Scouters(S), Decision Makers(D) and Actors(A).

The objects investigated here are six C2 organizations which were designed by A.H. Dekker in his research (Dekker 2001). These organizations have different architectures including the centralized architecture without information sharing (C2_centralized), the split architecture without information sharing (C2_split), the distributed architecture without information sharing (C2_distributed), the negotiation architecture without information sharing (C2_ negotiation), the distributed architecture with information sharing (C2_distributed_share) and the negotiation architecture with information sharing (C2_negotiation_share). Each one of these organizations has similar type and amount entity elements and represents a typical command and control relationship pattern. With these objects, we can deeply research the characters of command and control relationship architecture from the perspective of comparison.

The Command and Control Relationships Network Models(C2RNMs) of these organizations are built as the follow figures (see Figs. 1, 2, 3, 4, 5, and 6). It can be seen from the figures that each organization has two or four Decision Makers(D_A, D_S, D1, D2, D3 and D4), five Scouters(Ssat, S1, S2, S3 and S4) and four Actors(A1, A2, A3 and A4). Scouters gather all kinds of information about organizational environment and organization objective and send them to decision makers. Decision Makers receive the information from scouters and other decision makers and then arrive at decisions about the state and action for the nodes in organizations. Actors receive the information from decision makers and perform all kinds of missions.

Fig. 1 C2_centralize
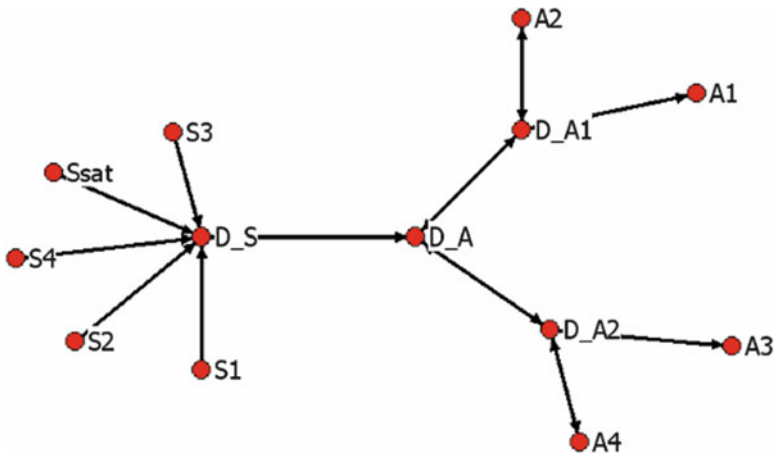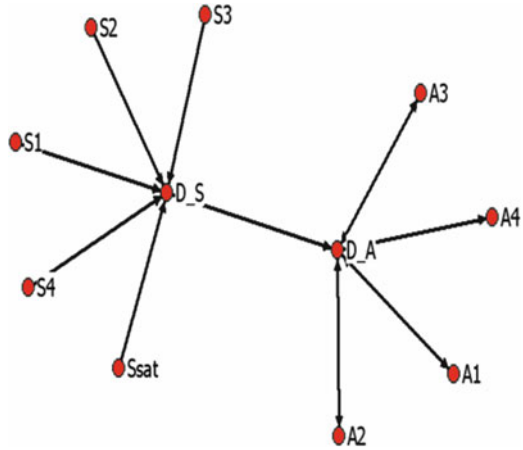
Fig. 2 C2_split

4 Core/Periphery Structure Characters of C2RNMS Analysis

As we know, there are two kinds of models of the core/periphery structure: a discrete model and a continuous model. In this paper, we will use the discrete one to study, and in subsequent research we will use the continuous one. With program UCINET (Borgatti et al. 2002), we can get the discrete core/periphery structure characteristic parameters of the six C2RNMs (such as Fig. 7).

In Fig. 7, "Starting measure" means the algorithm of rank correlation coefficient that used here is CORR, and there are other algorithms such as HAMMING, DENSITY, SXY, EMPTYPER. "Starting fitness" means the initial fitting value,

Fig. 3 C2_distributed
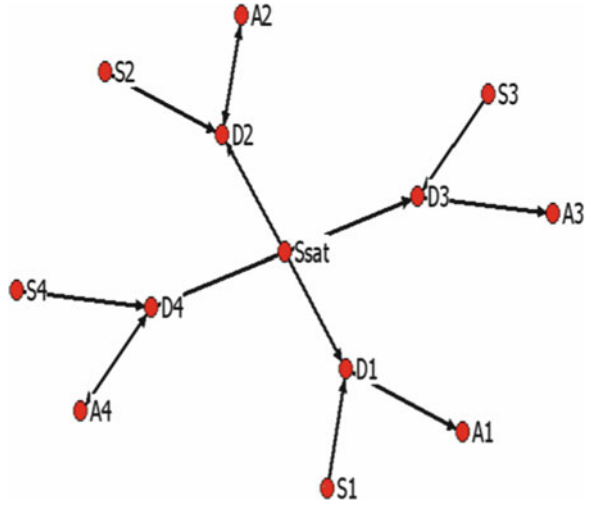
Fig. 4 C2_negotiation
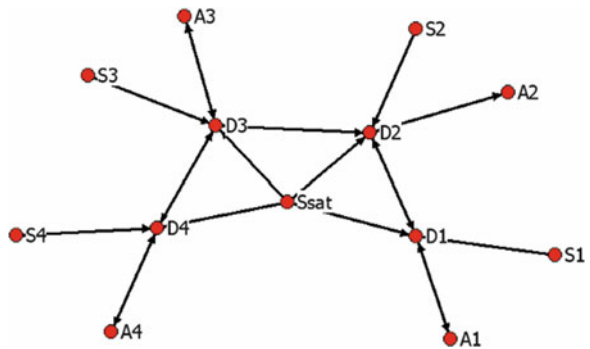
Fig. 5 C2_distributed_share

Fig. 6 C2_negotiation_share


```
SIMPLE CORE/PERIPHERY MODEL
--------------------------------------------------
Type of data:                    Positive
Fitness measure:                 CORR
Starting fitness: 0.408
Final fitness: 0.408

Core/Periphery Class Memberships:

    1:  A1 A2 A3 A4 D_A D_S
    2:  S1 S2 S3 S4 Ssat

Blocked Adjacency Matrix

                    1          1
            1 2 3 4 5 1   7 8 9 0 6
            A A A A D D   S S S S S
           ---------------------------
 1    A1 |           1  |             |
 2    A2 |           1  |             |
 3    A3 |           1  |             |
 4    A4 |           1  |             |
 5   D_A | 1 1 1 1   1  |             |
11   D_S |           1  | 1 1 1 1 1 |
           ---------------------------
 7    S2 |           1 |             |
 8    S3 |           1 |             |
 9    S4 |           1 |             |
10  Ssat |           1 |             |
 6    S1 |           1 |             |
           ---------------------------

Density matrix

          1      2
        -----  -----
 1    0.333  0.167
 2    0.167  0.000
```

Fig. 7 Characteristic parameters of C2_centralize

namely the correlation coefficient of initial matrix and ideal matrix. "Final fitness" means the final fitting value, namely the correlation coefficient of transposed matrix and ideal matrix. If the values of the two parameters are bigger, the core/periphery structure character degree of the network models will be clearer. "Core/Periphery Class Memberships" shows the core member group and periphery member group separated by UCNET. "Blocked Adjacency Matrix" shows the adjacency matrix of core and periphery, namely the adjacency relationships among the two group members. "Density Matrix" shows the relationship density matrix of the members. The relationship among core group members is closed, the relationship among periphery members is trifling or even is inexistence, and it is not required for each periphery member to relate to all core members.

The result (see Table 1) can be produced by calculating and gathering the core/periphery structure characteristic parameters of the six C2RNMs. It can be 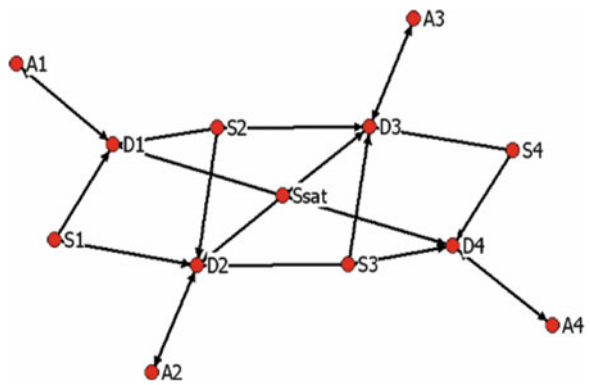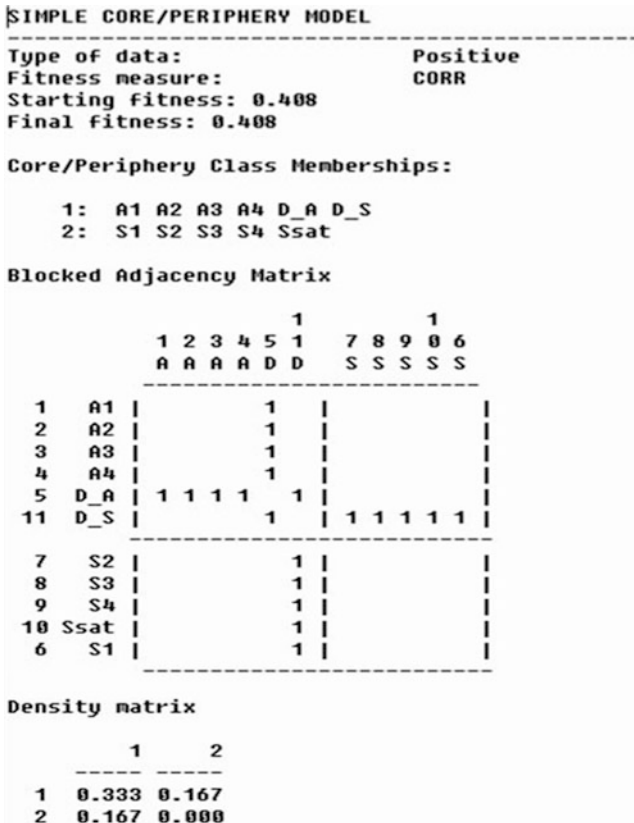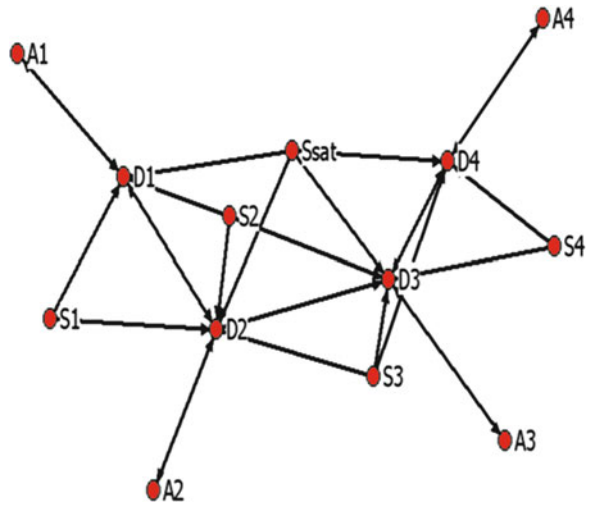seen from the table that the parameters "Final fitness" of three network models (C2_distributed_share, C2_split and C2_negotiation_share) are 0.581, 0.679 and 0.734. These values are higher indicating that the core/periphery structure characters of these three models are clearer. At this point, the result of core group and periphery group separated by UCNET can explain this kind of characters. For example, in C2_split network model, the core group includes D_A1, D_A2, D_A and D_S. In this organization (see Fig. 2), D_S handles intelligences gathered from all other Scouters, and send them to D_A. Then D_A reasonably allocates the tasks to D_A1 and D_A2, and D_A1 and D_A2 command the Actors that are affiliated with to perform their idiographic tasks. The units such as A1, A2, S1, S2 and so on are included in periphery group. These units gather intelligences from circumstance or act on circumstance directly, and they are in the periphery position. While the parameters "Final fitness" of other three network models (C2_centralized, C2_distributed and C2_negotiation) are 0.284, 0.357 and 0.408. Their values are lower indicating that the core/periphery structure characters of these three models are not clearer. At this point, the result of core group and periphery group separated is not ideal. For example, in C2_distributed (see Fig. 3), the members of core group are A3, A4 , D1, D2, D3, D4 and Ssat, the members of periphery group are A1, S1, S2, S3 and S4, while the structure character can not be identified easily.

The density matrix shows the relationship density among core group and periphery group. The "Core-in" parameters mean the relationship density among members of core group. The values of C2_distributed_share, C2_split and C2_negotiation_share are 0.467, 0.500 and 0.667, and these values are bigger. The "Periphery-in" parameters of these three models are 0.000, namely that there are no relationships among the members of periphery group. The "Core-Periphery" and "Periphery-Core" parameters of these three models are 0.262, 0.250 and 0.262 which are relative little. So, it can be seen that in these three models, the relation among core group is close, the relationship between core group members and periphery group members is loose, and the relation among periphery group members is absent. These results accord with core/periphery structure characters of ideal models. The clear characters make the units of C2 organizations cooperate with others for same goal. The density matrix values of other three models are relative littler that means their core/periphery structure characters are not clear.

Table 1 Characteristic parameters of six C2RNMs

	C2_centralized	C2_split	C2_distributed	C2_distributed_share	C2_negotiation	C2_negotiation_share
Final fitness	0.408	0.679	0.284	0.581	0.357	0.734
Core	A1 A2 A3 A4 D_A D_S	D_A1 D_A2 D_A D_S	A2 A3 A4 D1 D2 D3 D4 Ssat	D1 D2 D3 D4 Ssat	A2 A3 D1 D2 D3 D4 S2 Ssat	D1 D2 D3 D4 Ssat
Periphery	S1 S2 S3 S4 Ssat	A1 A2 A3 A4 S1 S2 S3 S4 Ssat	A1 S1 S2 S3 S4	A1 A2 A3 A4 S1 S2 S3 S4	A1 A4 S1 S3 S4	A1 A2 A3 A4 S1 S2 S3 S4
Core-in	0.333	0.500	0.250	0.467	0.357	0.667
Periphery-in	0.000	0.000	0.000	0.000	0.000	0.000
Core-Periphery	0.167	0.250	0.125	0.262	0.125	0.262
Periphery-Core	0.167	0.250	0.125	0.262	0.125	0.262

5   Conclusions

This paper analyzes the architecture characters of C2 organizations using "discrete core/periphery model" in Social Network Analysis. The progress of the method is that using discrete core/periphery structure model, combining with data of actual C2RNMs, calculating the relevancy degree between actual models and ideal model, quantitatively analyzing the core/periphery structure characters of actual models. If C2 organizations have clear core/periphery structure characters, the results analyzed with discrete model are relative ideal. On the other hand, if the results analyzed with discrete model are relative ideal, the core/periphery structure characters of C2 organizations are clear. Of course, our research is still preliminary, and there are still great deals of work to be done such as the structure characters analyzed with continuous core/periphery model, the relation analysis between structure characters and organization efficiency, and other contents that can be analyzed with social network analysis etc.

References

Borgatti SP, Everett MG (1999) Models of core/periphery structures. Soc Netw 21:375–395
Borgatti SP, Everett MG, Freeman LC (2002) UCINET for Windows: software for social network analysis. Analytic Technologies, Harvard
Dekker AH (2001) Applying social network analysis concepts to military C4ISR architectures. Connections 24(3):93–103
Dekker AH (2002) C4ISR architectures, social network analysis and the FINC methodology: an experiment in military organisational structure. DSTO report DSTO-GD-0313
Dekker AH (2003) Centralisation and decentralisation in network centric warfare. J Battlef Technol 6:23–28
Jarvis DA (2005) A methodology for analyzing complex military command and control (C2) networks. In: The 10th ICCRTS, McLean
Krebs VE (2002) Mapping networks of terrorist cells. Connections 24(3):43–52
Renfro RS (2001) Modeling and analysis of social networks. Dissertation, Department of the Air Force, Air University
Xu JJ, Chen H (2005) CrimeNet explorer: a framework for criminal network knowledge discovery. ACM Trans Inf Syst 23(2):201–226

System Architecture and Application of Wireless Sensor Network

Cai-jun Luo

Abstract Wireless Sensor Network technology is a kind of new data gathering and transmission technology, which combines with wireless communication, sensors, embedded computing and distributed information processing. It has the great potential with various applications in military defense, industrial automation and transportation. This paper mainly focuses on the developing present situation and trend of wireless sensor network routing protocol, its research hotspot and difficulty. It analyzes the application of wireless sensor network in various fields, pointing out the potential problems, and puts forward some improvement methods. Theoretical guidance is provided for further research and expanding the application range of wireless sensor network.

Keywords Wireless Sensor Network • Architecture • Application • Method

1 Introduction

In information society, the production, acquisition, transmission, storage and handling of information are the important parts of information science (Akylidiz et al. 2002). Various kinds of sensors are the major media which we access and use information resources. The progress of microelectronics technology, wireless communication technology, computer technology and internet technology, promoted the rapid development of sensors with low power consumption and multiple functions,

C. Luo (✉)

Department of Computer Science, Shaanxi Vocational and Technical College,
Xi'an, Shaanxi, China
e-mail: lcjsiny@sina.com

which have various integrated modules for data acquisition and processing, wireless communication, power supply, etc. The ways of data acquisition of traditional sensors have shifted from mono-style, wired transmission to intelligent, networked and wireless transmission, which have become the major technology of data acquisition.

2 Brief Introduction of Wireless Sensor Network

Wireless sensor network (WSN) is a self-organizing network in a monitored area, linked by wireless communication with many jumps formed from a large number of randomly distributed mini-nodes integrated with sensors, data processing modules and wireless communication modules (Sun Li-min et al. 2005). The main goal of WSN is collecting and dealing with information among the monitoring region which we are interested in by the cooperation of various sensors carried by nodes, such as temperature, humidity, light intensity, pressure, noise, size of moving object, speed and direction, etc. The transmission of the data to observers is wireless multihop through specified protocols (Zeng Peng et al. 2007).

Nowadays wireless sensor network can be roughly classified into three categories: wireless broadband network, wireless ad hoc network and WSN. WSN appliance is a highly relevant network system. It often has different network forms due to the different application and demand. Generally speaking, WSN has four features: self-organization, reliability, application of correlation and data-centric.

3 System Architecture of Wireless Sensor Network

3.1 Architecture of Sensor Node

Depending on applications, the WSN nodes have different configurations. But usually it composed of four parts: sensor module, processor module, wireless communication module and power module (Kottapalli et al. 2006).

Sensor module is responsible for collecting the required information within the monitoring area and data conversion and its types determined by the physical form of monitoring data signal; processor module is responsible for controlling the whole node, storage and processing the received data. It provides hardware support for the operating system and upper application software; wireless communication module is responsible for the wireless communication among nodes, transmission of data and controlling information; the power module provides energy for the node's normal work, usually by using micro batteries, solar battery, etc.
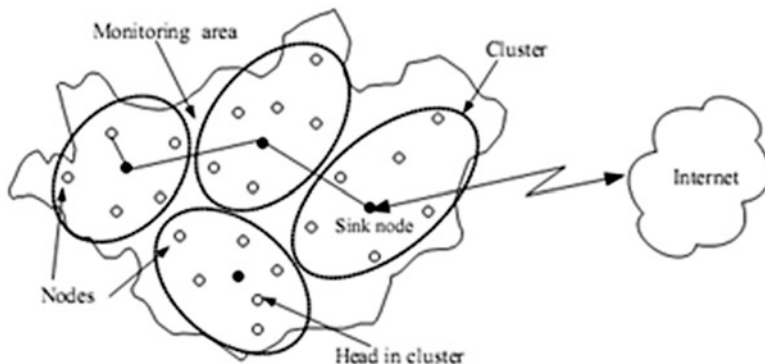
Fig. 1 Architecture of sensor network

3.2 Architecture of Sensor Network

Like the architecture of nodes, the architecture of WSN based on sensor network applications is also a variety of different. Basing on the function division, sensor network usually can be divided into the following three types of nodes: sensor node, sink node and management node, as shown in Fig. 1. The network is composed of massive sensor nodes random deployed in the monitoring region by the way of self-organizing. According to the certain protocol rules when monitoring data, finishing transmission by each sensor node hop by hop, and then gathering on the special node after the multi-hop routing. At last, arrive at the management node by internet and satellite. And the administrator achieving the configuration management to WSN by the node, issuing control command, etc. (Ruiz et al. 2003).

3.3 Protocol Stack Architecture of Sensor Network

Nowadays researchers have developed many types of protocol stacks for sensor network. Generally sensor network protocol stacks take the five-floor protocol in internet, which include physical layer, data link layer, network layer, the transport layer, the application layer and added the platforms with WSN characteristics, such as the energy management platform, mobile management platform and task management platform, etc. These platforms make sensor nodes work high efficiently under the priority of low power consumption.

Position layer and time synchronization layer were added in the improved module to provide the information support of the MAC protocol based on the time multiplex or the routing protocol based on the geographical location. Meanwhile network management, topology control, energy control and much other mechanism were presented to optimize and manage the agreement process.

3.4 Architecture Characteristics of Sensor Network System

System architecture means a lot for wireless sensor network. Although a WSN is a network system highly related to applications, which under different application requirements can be different in terms of the hardware and software of network nodes, network protocol as well as other relevant technologies. By considering some common features of WSN and providing a reference "system" and "framework" to developers, can greatly shorten the WSN development time and reduce its complexity. According to the different emphases, nowadays WSN system architecture includes dynamic protocol stack supporting system, stratified system, Self-adaption system, self-restore system, multitasking system and agent system, etc.

4 Application of Wireless Sensor Network

Sensor network has a broad prospect for applications. It can be widely used in many fields. Such as military, environmental supervising and prediction, medical, intelligent household, urban traffic, space exploration, safety supervising, etc.

4.1 Military Application of Wireless Sensor Network

In military field, sensor network will become a necessary part of system C4ISRT (Command control communication computing intelligence surveillance reconnaissance and targeting). Its goal is to design a battlefield command system combining the functions of commanding, controlling, computing, intelligence, surveillances, reconnaissance and positioning. This attracts the attention in military developed countries. Because sensor network is composed of intensive, low cost and random distributed nodes. Self-organizing and the ability of tolerating mistakes so that it will not lead the collapse of the whole system when some nodes are suffering malicious attack (Nogueira and Loureiro 2005). Just on this point, conducting real-time supervising on our army troops, equipment and supplies, Surveillance conflict zone, reconnoitering local terrain and laying out position attack target and access loss, etc.

4.2 Industrial Applications of Wireless Sensor Network

Development with computer, communication and network technology, the development of industrial communication technology has gone through the periods of analog instrumentation control systems in the 1960s and 1970s, distributed control systems (DCS) in the 1980s and 1990s, and the field bus control systems (FCS) at the end of the twentieth century, moving in the direction of intelligent

and networked. With the development of monitoring system, reduce the cost of investment and use of communication technology for industrial development of the urgent requirements. Under this circumstance, the low cost, using friendly and wide perception and other features of wireless sensor network catch people's attention. By applying these features, people can use the lower investment costs to monitor the whole industry progress, and using it as the foundation of implement optimized control to improve quality of products to realize the goal of reducing the leakage problems and increasing the energy efficiency.

4.3 Other Applications of Wireless Sensor Network

The environment science is used in a wide range. Traditional data acquisition methods cannot meet people's need any more. WSN provide convenience for data collection in a complex environment, such as tracking migration of birds, monitoring raw water, air and solid changes in composition. It also provides a great help for fire prevention, flash flood prediction (Steffan et al. 2005). Wireless sensor network plays an important role in medical and health caring field. Doctors can keep handling the patients progressing and deal with it in time when special sensor node be installed in patients body; it also can provide reference data for daily health research and new medicine development under the condition of without influencing the normal life.

5 Problems and Improving Methods of Wireless Sensor Network Application

5.1 Physical Manipulations of Sensor Node

Future sensor network will have hundreds and thousands of sensor nodes. It's difficult to monitor and protect each of them. Thus each node is a potential point to be attacked physically and by logic. In addition, sensors usually are deployed in no-human environment which make it more convenient for the attacker to capture sensor nodes. When captured the sensor node, the attacker can modify or get its information or code, modify its program code and load into sensor node. Obviously, the ordinary sensor node has great security holes. Since it's an unavoidable issue of security that sensor node can be easily physical manipulated, other technical solutions must be provided to improve the security performance of a WSN. An example is the node to node identity authentication, Another example is developing a new key agreement so that even a small amount of nodes are manipulated, attacker can not or difficult to obtain node information from other nodes is derived the key information. In addition, the sensing node can authenticate the legality of the software and other measures to improve the safety performance of the node itself.

5.2 *Information Wiretap*

According to the characteristics of wireless transmission and network distribution, it's easy for the attackers to get sensitive and private information through the transmission among nodes (Brannstrom and Jungert 2006). Encrypting transmit information can solve wiretaps, but it requires a flexible and strong key exchange and management scheme. What's more, the key management scheme should be easy to deploy and serve limited resources characteristics of sensor node. In addition, key management scheme has to ensure the security of the whole network to not be destroyed when section nodes out of work. Due to the limited internal storage resources of sensor node, it's impractical to realize the end-to-end security among nodes. While in sensor network, hop-hop information encryption can be realized in sensor network, thus sensor node only has to share the key with neighbor nodes.

5.3 *On Privacy*

Sensor network is designed to collect information. Attackers may get this sensitive information by the way of wiretap, adding forged illegal nodes, etc. An attacker can derive valid data from large amounts of acquired data if the attacker is able to obtain the relevant algorithm for acquiring limited information from completed information. Then the attacker can export the valid information through access a lot of useful information. The normal sensor private problem, which is unlikely obtain information through sensor network, but through remote monitoring of WSN by the attacker, so that to get large amount of information and work out the private problem with specific algorithm. Making sure the sensor information on network is visited only by trusted entity is the best way to solve this private problem, which can be achieve through data encryption and access control. Another way is to limit the granularity of transmit information on network, since the more detailed information, the more easily to leak the privacy, For example, a cluster node can achieve the data anonymous by dealing with the massive information received from the neighbor node and transmission results only.

6 Conclusion

Wireless sensor network have the characteristics of self-organization, miniaturization and external perception ability, etc. These features determine their vast application prospect in business field. For example, sensors embedded in home appliances and embedded systems can form wireless networks which with internet links are able to offer more intelligent and humanizing home environment. Sensor network can be used for space exploration, transmitting sensor nodes in outside

planet with spacecraft, the long-term monitoring of the interest environment, etc. It also can be used in disaster prediction and rescue, warehouse management, factory automation and other fields. Wireless sensor network would gestate new design and application mode.

References

Akylidiz F, Su W, Sanjarasubramaniam Y (2002) A survey on sensor networks. IEEE Commun Mag 40(3):12–14

Brannstrom M, Jungert E (2006) A scalable agent architecture for a dynamic sensor network. Electr Notes Theor Comput Sci 68(3):123–130

Kottapalli VA, Kiremidjian AS, Lynch Jerome P et al (2006) Two-tiered wireless sensor network architecture for structural health monitoring SPIE. In: 10th international symposium on smart structures and materials, San Diego, pp 2–6

Nogueira JM, Loureiro AAF (2005) On the design of a self- managed wireless sensor network. IEEE Commun Mag 43(8):95–102

Ruiz LB, Nogueira JM, Loureiro AAF (2003) MANNA: a management architecture for wireless sensor networks. IEEE Commun Mag 41:116–125

Steffan J, Fiege L, Cilia M, Buchmann A (2005) Towards multi-purpose wireless sensor net-works. In: Proceedings of conference on Sensor Networks (SENET'05), Italy, pp 342–345

Sun Li-min, Li Jian-zhong, Chen Yu et al (2005) Wireless sensor network theory, 1st edn. Tsinghua University Press, Beijing

Zeng Peng, Wang Zhong-feng et al (2007) Distributed communication protocol for wireless sensor networks. Communications 25(10):102–110

The Design and Implementation of Efficient SMS Query Software Based on Android

Jiao Guo, Chui Li, and Ying Xu

Abstract Analyzed the drawbacks of SMS query software in smartphone at present. In view of users requirements for query short messages in phones, and based on Android kernel and its architecture, highly effective SMS query software is presented. The implementations of the SMS query software is described, which provides a variety of modes of query, such as based on address book, based on contact time, based on current contact and based on content of message. Most current software does not yet have the above query modes for SMS, which should be adopted by many mobile phone software manufacturers in the future.

Keywords Android • Short message • Contact time • Fuzzy search

1 Introduction

The popularity of smart phones has brought people great convenience (Falaki et al. 2010). Smart phone not only has independent operating system, but also allows users to install various programs provided by third-party service providers. Except for phone calls, short message service is the most widely used as a communication service (Peersman et al. 2000). With a wide range of interpersonal communication, short messages in phone are more and more, but currently the short message query function of smart phones is not very perfect. At present, there are mainly two ways to search short message. One is the sequential search, another is to search based on contact's name. Sequential search asks users to traverse all messages to find the object message. Even if the query is based on contacts' name, the query must traverse all the messages related to the contacts. There are two problems in this

J. Guo (✉) • C. Li • Y. Xu
School of Computer Science and Engineering, Three Gorges University, Chongqing, China
e-mail: Legend_j@163.com

search method (Naor 2004). The first problem is that the way is a single query, and not friendly to use for user. The second problem is inefficient.

In order to address the above issues, smart phones need third-party softwares (Short Message Service SMS Tutorial 2011), which could efficient query short messages. The solutions can be divided into two ways. One way is to increase the mode of query; the other is to improve the query efficiency. Query efficiency was mainly due to sequential search, so the first way is used mainly to solve the situation, in which the user can only do sequential search.

Now, Android platform is one of the more popular operating system for smart phone. This paper proposed an efficient message query system based Android platform, which will provide independent third-party software to efficiently query short messages for phone users.

2 Introductions to Android Platform

2.1 *Introduction to Android*

Android platform is the mobile phone operating system developed by the Google Company in November 5, 2007 (Xianhua et al. 2009). After the development for 2 years, Android platform holds the suitable share in the smart mobile field. Headed by Google, more than 40 leaders in the field of mobile communications form the Open Handset Alliance (OHA). Google and operators, equipment manufacturers, developers and other third parties form deep-level partnerships, and hope to establish a standardized, open software platform for mobile phones, in order to form an open ecosystem in the mobile industry. For this reason, Android is being accepted by more and more developers and users (Yuzhi et al. 2009). By the end of Mar 11, 2011, there are more than 250,000 applications in Android Market.

2.2 *Framework of Android*

The framework of Android platform consists of five parts totally. The core of underlying layer, which is an independent operating system based on the Linux 2.6 kernel, provides basic system services, including security mechanisms, memory management, process management, network stack and a series of driver modules. The middle layer of Android platform is run-environment, including system runtime library, Dalvik virtual machine, core libraries, and some various components supported by C/C++ libraries. In application layer of the platform, these API framework are used to develop new applications for developers. So the application layer could simplify application the architecture design for application development. Application layer program is written with the Java, which runs on virtual machine (Yaoyi and Liuwei 2008). The entire framework of Android platform is shown in Fig. 1.

Fig. 1 The entire framework
of Android platform

| Procedures of Application Layer |
| Framework of Application Layer |
| Other Libraries | Core Library |
| Core of Linux |
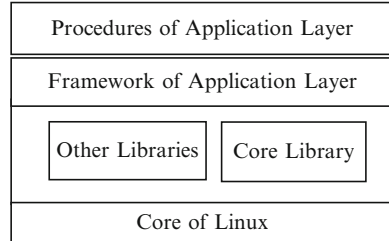
2.3 The Features of Android System

1. Openness
2. Open Applications
3. Applications are created under equal conditions
4. Applications can be easily embedded in the network
5. Program can run in parallel

3 Designs of the Solution

As the short message includes not only the text part, but also date, time, sender, recipient and other information, which are the query conditions, and corresponds to the various fields in SMS database. Therefore, increasing the mode of query can provide more available query conditions to find needed messages for users.

The additional modes of SMS query are as follows: query based on the address book information, query based on the contact time, query based on recent contact information, query based on message content. If all the above ways is alone, each way has its drawbacks. Therefore, the efficient SMS query software should integrate the above functions, and allow a variety of integrated ways to use.

3.1 Query Based on the Address Book

Query based on address book exits as essential function in many phones. The main principle is that users find the contact related to SMS message before they find short messages. The software lists all the messages related to the contact. And then users keep looking in this small range.

The mode of query is more flexible and fast. But the above range is not always small. For example, the user frequently contacts with one liaison person, so there are hundreds of short messages, it is very inefficient to sequentially search. The mode of query integrated with the following query mode, such as query based on the date, the efficiency will be improved a lot.

3.2 Query Based on the Contact Time

If users do not remember the contact's name when they query short messages, they can query the message based on general time. Query based on the contacting time can be divided into the precise search and fuzzy search. Precise search can be accurate to 1 day, for example, find the blessing of SMS received on Mid-Autumn Festival. Fuzzy Search can be used to searching in the general time period for use, such as searching in the last week. If this function and the query based on address book are combined, the search function will be more flexible, and users can more quickly find the target message.

3.3 Query Based on the Current Contact

Query based on the current contact is a new mode of query that combines the time with the contacts. The system lists the contacts based on the contacting time from near to far. If the contact corresponding to the short messages needed is in the first few persons, the user can directly view the contact's messages. Under normal circumstances, the recent incoming and outgoing short messages are used more probably, and the current contacts are generally linked, so the mode of query is easier to find the target messages.

3.4 Query Based on the Content of Message

The principle of query based on the content of message is similar to search engine, which allows the user to enter keywords, short message query software will find out all messages containing the keywords. The method relies on the content of the fuzzy search. When users cannot remember the specific time and contacts, they can easily find desired messages with some content of short messages. SMS query software integrated with this feature is similar to a micro-SMS search engine; its convenience can be greatly improved.

4 Implement Short Messages Query Based on Android Platform

The short messages and phone book in Android system are stored in a SQLite database. SQLite database is a small lightweight relational database. All inquiries about the short messages and contacts can be achieved through SQL statement. Android platform has packaged SQL statements into functions, so the query can
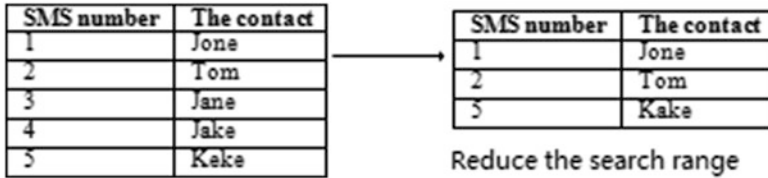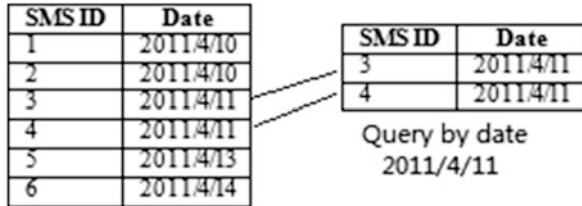
SMS number	The contact
1	Jone
2	Tom
3	Jane
4	Jake
5	Keke

SMS number	The contact
1	Jone
2	Tom
5	Kake

Reduce the search range

Fig. 2 A schematic of the query based on the address book

Fig. 3 A schematic of the query based on the accurate time

SMS ID	Date
1	2011/4/10
2	2011/4/10
3	2011/4/11
4	2011/4/11
5	2011/4/13
6	2011/4/14

SMS ID	Date
3	2011/4/11
4	2011/4/11

Query by date
2011/4/11

be achieved using the appropriate functions. On Android platform, query () is used to do database query, by setting different parameters for the query () method to construct various select statement, to achieve user's query.

4.1 Implement of Query Based on the Address Book

Query based on the address book involves two tables in the database, one is the contacts table, and a table is the short messages table. Firstly, obtain the contact number; secondly, according to the number retrieve the SMS message related to the number in the short messages table. The process is transparent to the user. Figure 2 is a schematic of the query based on the address book.

4.2 Implement of Query Based on the Contact Time

Query based on the contact time involves only the short messages table. If precise search is being done, let the date be equal to retrieve date, or let the date within a retrieve date range. Figure 3 is a schematic of the query based on the accurate time.

4.3 Implement of Query Based on the Current Contacts

Listing the current contacts need to sort all contacts based on the last contact time. After the user selects any of the contacts, the operation is as same as query based on the contact. Figure 4 is a schematic of the query based on the current contacts.
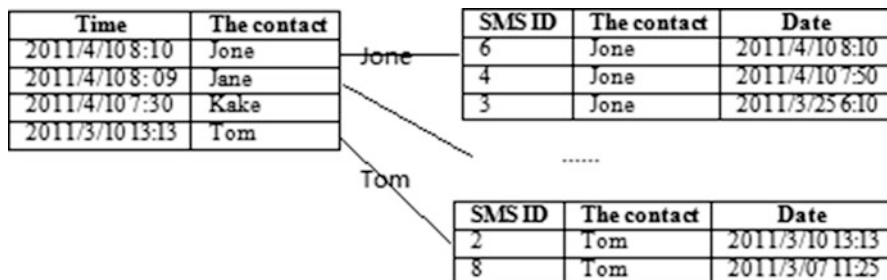
Time	The contact
2011/4/10 8:10	Jone
2011/4/10 8:09	Jane
2011/4/10 7:30	Kake
2011/3/10 13:13	Tom

Jone

SMS ID	The contact	Date
6	Jone	2011/4/10 8:10
4	Jone	2011/4/10 7:50
3	Jone	2011/3/25 6:10

....

Tom

SMS ID	The contact	Date
2	Tom	2011/3/10 13:13
8	Tom	2011/3/07 11:25

Fig. 4 A schematic of the query based on the current contacts

SMS ID	SMS Content
1	Hello!
2	Good Morning!
3	Good Afternoon!
4	Good Evening!

Good →

SMS ID	SMS Content
2	Good Morning!
3	Good Afternoon!
4	Good Evening!

Fig. 5 A schematic of the query based on keywords

4.4 Implement of Query Based on Keywords

Query based on keywords uses SQL statements:

Select * From sms Where body like %keyword%

By this method, all messages containing the keywords are retrieved, and then displayed to the user. Figure 5 is a schematic of the query based on keywords.

5 Conclusion

In this paper, based on Android core and its architecture, an efficient SMS query system is designed, and its basic features and technology adopted have been introduced. As third-party software for mobile phone users, it provides higher efficiency and more options compared to the short message query systems in most phones.

References

Falaki H, Mahajan R, Kandula S, Lymberopoulos D, Govindan R, Estrin D (2010) Diversity in smartphone usage. In: MobiSys 2010-6, San Francisco, CA, USA

Naor Z (2004) An efficient Short Messages Transmission in cellular networks. In: Proceedings of INFOCOM, Hong Kong

Peersman G, Cvetkovic S, Griffiths P, Spear H (2000) The global system for mobile communications Short Message Service. IEEE Pers Commun 7:15–23

Short Message Service/SMS Tutorial (2011) Technical report. http://www.developershome.com/sms/

Xianhua S, Zhenjun D, Rong C (2009) Research on mobile location service design based on Android. In: Proceedings of 5th international conference, WiCOM 2009, Beijing

Yaoyi M, Liuwei G (2008) Study of Android's architecture and its application development. Comput Syst Appl 11:110–113

Yuzhi L, Chenyu X, Zhenming J (2009) Examples of Google Android SDK development. Posts & Telecom Press, Beijing

Construction and Management of Virtual Experiment System

Liu Xu Yong

Abstract Virtual experiment system can effectively improve the teaching effect, stimulate student interest and enhance the technology content in teaching as well. This paper introduces the concept and basic characteristics of virtual reality technology and makes an insight into its advantage and the mode of administration, with the emphasis on its application to the computer and economic management courses.

Being essential for students' competence for knowledge acquisition and problem-solving ability, teaching Practice plays a very important role in higher vocational education. But the traditional teaching modes greatly hinder the students' innovation ability and therefore have some negative effect on their learning. The application of virtual reality technology can effectively solve the problem.

1 Introduction to Virtual Reality Technology

With three dimensional graph production technology, the multi-sensing interactive technology as well as the display technology, and the devices as keyboard, mouse, sensory equipment, hard helmet and data glove, the virtual reality

L.X. Yong (✉)
Electronic Equipment and Information Management Office, Shaanxi Academy
of Governance, Xi'an, Shaanxi, China
e-mail: liuxuyong@126.com

(Virtual Reality, VR) enables participants to enter and communicate in the virtual environment, in which they can sense and act on the various virtual objects (Hu Xiao Jiang 2005).

The basic characteristics of virtual reality technology include:

1.1 Multi-sensations

Employing the observation, to listen attentively to, touching, the kinesthesia sensing and the response installment, users may obtain various sensations in the virtual environment, and adopts the corresponding actual response and operates alternately.

1.2 Interaction

With kinds of equipments, users can sense virtual environment objects through their body movement and language in the virtual reality system. Such man-machine interaction may come very close to natural interaction through each, the user relies on comes in the sensation and the operation.

1.3 Immersion (Huang Ming ji 2005)

Based on the human physiological characteristic of vision, hearing and the computer three dimensional vectograph, the virtual reality technology get subjects immersed into the virtual setting.

2 The Characteristics of Virtual Experiment System

2.1 Support Cooperation

The virtual experiment system provides more than one way to communicate between the users.

2.2 Powerful Interaction

The virtual experiment system has the formidable interactive ability, which provides a lifelike experiment environment to the user, the subjects and the users can interact with each other simply by dragging the mouse Real-time observation on the experiment phenomenon and the result is also allowed in the setting.

2.3 High Sharing Degree

The virtual experiment system breaks the traditional regional and time limit, and enables wide-range subscribers' visit. Meanwhile, the laboratory technician may carry on the experiment in any time.

3 The Advantage of Virtual Experiment Teaching

Based on the present experiment teaching situation and the virtual experiment system's characteristic, the virtual experiment teaching has the following superiority:

3.1 The Open Object

In the experiment teaching, learners may be divided into the curriculum learner, the curriculum participant and amateurs. Rests on the hypothesized pilot system operation; they can be divided into the laboratory technician, the teacher and the technical personnel. Different objects carries on the study and the exchange under the virtual experiment platform according to their status respectively.

3.2 Open Organization

Closely related to the network, the experiment resource and the experiment project, the virtual experiment is not confined to time and place, which guarantees the users freedom to choose. Along with the experiment teaching's development and the request, there are more demands for the cross region, interdisciplinary and simultaneous experiment. The virtual experiment enlarges the space for the experiment teaching's development.

3.3 Open Resources

Teaching resources open completely to the experimental subjects in the virtual experiment, which not only includes the empirical datum, the curriculum analysis and summary, but also involves the virtual experiment project alternative's development, designs and operates (Xu Hong and Liu Yu 2009)? The student experimental design may use the system software program module and the experiment item design template; Empirical datum's analysis and processing may carry on through

the data analysis and the processing tool bag; the consolidated knowledge may be done through teacher's instruction, student's exchange information and experiment breakdown and error analysis.

4 Improve the Quality of Virtual Experimental Teaching Methods

To ensure a quality of teaching improved, so that effective teaching virtual experiments carried out smoothly, we must do a good job of teaching the management of experimental work, mainly in the following points:

4.1 Development of Realistic Experimental Syllabus

The experiment program of instruction is the university tests the teaching management the laws and regulations, it is organized and implemented teaching, check the fundamental basis of its quality.

4.2 The Preparation of Guidance Materials Virtual Experiment Class

According to the experimental curriculum, content and present laboratory experiments used in the virtual environment, to prepare for the professional use of the experimental tutorial series, available to the experimental class teachers and students to participate in the experiment. Tutorial should include the construction of experimental equipment, operating essentials, experimental project purpose, principles, experimental procedures, experimental results.

4.3 According to the Experimental Requirements of Syllabus Development of Assessment and the Scoring Criteria

Laboratory experiments to measure teaching quality assessment is an important indicator, is to check the effect of teaching to promote student learning, improve teaching, improve the quality of teaching methods, is to test the students master the content of experimental an important basis for the situation. Rating standards of impartiality, scientific and rational will directly affect student learning initiative.

4.4 Accurate Records of Experimental Teaching Log (Yang Yuhong and Huang Jinshan *2006*)

Experiment course the log is the actual record after Experiment, Is examines the experiment class entire process actual result the concrete management document.

5 Virtual Reality Technology Application

The application of virtual reality mainly focuses on the virtual laboratory, the virtual learning environment, virtual really teaches aspects virtual simulation campus, and virtual distance learning.

5.1 Virtual Reality in Application Economics and Management Curriculum

In economy teaching, problems like lack of practice but abundance in theory, disconnection between theory and practice, are common. Computer technology should be combined with the teaching media and the teaching method reform to construct a curriculum related Virtual pilot system, including the market marketing, the third party physical distribution, electronic commerce, the human resources development, the stock transaction, the finance, the finance, the traveling scene and so on. In this way, we can carry on the perfect development and the interactive operation. For example, the real estate marketing virtual system may formulate different user and make out different strategies to marketing. Through virtual experience, not only deepened to the theoretical knowledge understanding, moreover lets the laboratory technician realize the modern real estate marketing the interactive type, the experience type, the digitized marketing pattern. Moreover, the electronic commerce modeling system may establish the store, the bank, the commodity and so on corresponding model, the laboratory technician may simulate processes and so on commodity choice, payment, allocation. Moreover, the virtual reality technology mock tourism scenic spot, take real data and so on image, picture as a foundation, using the computer manufacture three dimensional image, depends upon the hard helmet type monitor or the three-dimensional eyeglasses, the data glove, the bodily spot tracking device and the special projected display equipment, "the depth" the information and the observation parallax produces the three dimensional three-dimensional synthesis results through the visual main center comprehensive object, and unifies the sense of hearing, the vision and the sense of touch realizes the scenic area in the virtual scene the geographical environment and the characteristic.

5.2 Virtual Reality Technologies in Computer Experiment's Application

The virtual experiment system includes the related test installation, the experimental subject, the teaching information resources and the laboratory environment, the virtual laboratory already may be the virtual conception laboratory, may also be the real laboratory reappearance. For example, to deepen WAN, the metropolitan area network concept understanding, it does not need to construct the real network, (Jiang Ping 2006). Moreover, in the computer assembling and the maintenance curriculum, the virtual machine software vmware foundation and the real computer same movement environment may be offered to ensure various experiment such as the processor, the memory, the graphic display device, the input-output device and so on, which may create a real-sense experiment without any influence to the computer (Li Wen Yu 2004).

The virtual reality technology may also experiment on the model devised by laboratory technician .the result and the effect is supposed to confirm the feasibility and the rationalization. For example, in the network experiment, risk and cost may get reduced through the virtual router, the switchboard, the server and the terminal carries on the network (Jiang Ping 2006)

6 The Virtual Reality Technology the Question Which Exists in the Application

With the economy and technology development, the virtual reality technology will increasingly widely used in education teaching. But the current of the virtual reality technology still has certain problem, mainly displays in:

6.1 The Hardware Environment Also Needs to Enhance

The virtual reality technology's related equipment universal existence use is not convenient, effect not good and so on situations, meets the requirements which with difficulty the virtual reality system needs, like central computer's processing speed has not been able to satisfy in the hypothesized world the huge timely data quantity processing need, based on the sense of smell, the sense of taste equipment did not have mature and the commercialization.

6.2   Virtual Reality's Effect Also Needs to Strengthen

Virtual reality's expression stresses in the geometry indicated that lacks the lifelike physics, the behavioral model, in the sensation, the visual synthesis aspect's research are many, but the sense of hearing, the sense of touch attention are few, its hypothesized effect also needs to further strengthen.

6.3   Realizes the Cost to Need to Reduce

The virtual reality system application's related equipment price is also quite expensive, and these equipment limitation is very big, needs to further develop its structure and the technique of manufacture, causes its dexterous and reduces the cost, can be more advantageous in the promotion applies.

6.4   Virtual Environment and a Real Environment of the Differences

Students learn in a virtual environment with existing in the environment in a real difference. If only consider interactive authenticity, so really operate machines is superior to any media teaching effect. Therefore, the virtual reality technology cannot completely replace the specific real operation; it can only to a certain extent instead of some experiments.

7   Conclusion

The real experiment environment is more advantageous than the Virtual experiment environment in deepening student's perceptual knowledge and beginning ability, therefore, not all experiments are suitable to Virtual ways (Zhu Min 2006). An experiment system should be comprehensive and the complete, which includes the true environment and the virtual environment. Only with the perfection of construction and the advance of experiment environment, can we improve the quality of experiment teaching greatly, promote the university's core competence, and fulfill the task to serve for the teaching, the scientific research, and hence have breakthrough in our science and technology innovation and our society's development and the progress.

References

Hu Xiao Jiang (2005) The virtual reality technology [M]. Machinery Industry Press, Beijing

Huang Ming ji (2005) Virtual numerical control technology and applications [M]. Chemical Industry Press, Beijing

Jiang Ping (2006) Virtual machine and its application in computer teaching [J]. Eng Inst Sichuan Univ 5:34–35

Li Wen Yu (2004) A web-based virtual laboratory environment [J]. Exp Technol Manag 21(3): 105–108

Xu Hong, Liu Yu (2009) Computer virtual experiment environment reform and practice [J]. Exp Technol Manag 2(2):90–92

Yang Yuhong, Huang Jinshan (2006) Reforming laboratory management system to improve teaching quality [J]. Northeast Agric Univ 4(1):77–78

Zhu Min (2006) Application of virtual experiment and teaching [D]. PhD thesis, East China Normal University