ACTA TECHNICA

Volume 61 (2016), Number 4B

ISSN 0001-7043

Institute of Thermomechanics AS CR, v. v. i.

Acta Technica Volume 61 (2016), Number 4B ISSN 0001-7043

 MK ČR E4662

Publisher

Institute of Thermomechanics CAS, v.v.i.
Czech Academy of Sciences

Jiří Plešek Head of the Publisher

Editorial Office

Address: Acta Technica
Institute of Thermomechanics CAS, v.v.i.
Dolejškova 5
182 00 Praha 8, Czech Republic

E-mail: journal@it.cas.cz *Phone:* +420 26605 2058 *Fax:* +420 28689 0433

Home Page: http://journal.it.cas.cz/

Subscription Information

Address: Ivo Doležel
Kolínská 19/1666
130 00 Praha 3

E-mail: dolezel@it.cas.cz *Phone:* +420 72324 8626

Continued on inside back cover

CONTENTS

Acta Technica *61*, No. 4B/2016, 1–8

Design of a ZigBee-based intelligent household electricity monitoring and control information system

Xueqin Lu[1], Yinyan Li[1], Yan Tang[1]

Abstract. A ZigBee-Based Intelligent Household Electricity Monitoring and Control Information System (ZBEMCIS) is proposed. The system is composed of hardware and software. The hardware consists of sensors, a ZigBee-based gateway, and a client, while the system software is divided into three layers: the data service layer, the logic layer and the user layer. The sensors collect the data from electric equipment and household environment. Through the ZigBee-based home gateway with routing and external network, the data reach the client, which makes all kinds of home appliances connected into an intelligent system. While the software householders can realize monitoring and metering management of household electrical appliances through browsers, meanwhile, the system can automatically report to users the abnormality warnings by various communication modes. The ZBEMCIS can easily monitor electricity usage with a high level of accuracy, therefore, it can be used for electricity monitoring and controlling for household applications, and also it may make intelligent home appliances more diversified.

Key words. ZigBee, household electricity, monitoring and control, information system..

1. Introduction

With the emergence of innovative technologies, living standards and quality of life have reached an all-time high. A significant part of the modern lifestyle is intertwined with the usage of electronic and electrical devices. In traditional way, if users want to know the household electricity information, they can acquire only the total amount of electricity by bills, internet or telephone, rather than the usage of each electrical terminal. And users have no idea of how to monitor the usage of the terminals and to optimize the electrical consumption performances, not to mention the intelligent household electricity. On the other hand, if there is something wrong with some terminal, or users forget turning off the lamp, or the gas is leaking out, given these conditions, the users cannot get in time.

[1]Department of Electronic Commerce, Ningbo Dahongying University, Ningbo City, 315100, China

With advancement in wireless technologies and through the implementation of distributed sensor networks, intelligent residential systems are concerned more and more. By eliminating the need to run wires in an existing facility, wireless technologies can help reduce the cost of construction in "intelligent" household electricity. Due to their small footprints, wireless nodes can easily be mounted without interruption of current usage and without inconveniencing householders for renovations and changes. Another benefit of wireless technologies that makes them appropriate for residential use is their low energy consumption, because they can be powered by batteries with long service lives.

In order to monitor and control household electrical terminals more effectively, and safeguard in the household, we propose to design this ZigBee-based system combined with wireless sensor nodes to meet the need.

2. System architecture

The ZBEMCIS consists of a ZigBee-based gateway, a client, and sensors. The sensors monitor and control the electricity usage of the electrical terminals and transmit data to the gateway. The gateway's purpose is to connect sensor nodes to an existing home network. The client accept the datum and storage them into database. Users can access the status of electrical terminals by computer or smart phone.

2.1. System hardware architecture

System hardware architecture consists of sensor nodes and a ZigBee-bases gateway, which is the basis of ZBEMCIS. The sensor nodes are divided into electricity measure nodes and environmental monitoring nodes. The measure nodes are based on CC2530 chip with its development kits and can do the following things:

1. Measurement of power parameters, such as the voltage, current, and power of the outlet.
2. Control of the power output of the outlet.
3. Security protection.
4. Transmission of information from each node to the client through ZigBee.

The sensor node is shown in Fig. 2.

The environmental monitoring nodes are mainly used to monitor the home environment information, including temperature, humidity, smoke concentration and gas concentration. Their principle function is to collect the environmental parameters in real time, and get the actual environment parameters by comparison with the reference parameters. When the parameters of the acquisition exceed the threshold range, the alarm signal can be inspired.

The ZigBee-based gateway (its architecture being depicted in Fig. 3) is to encode and process the data from the nodes in accordance with certain rules of encoding

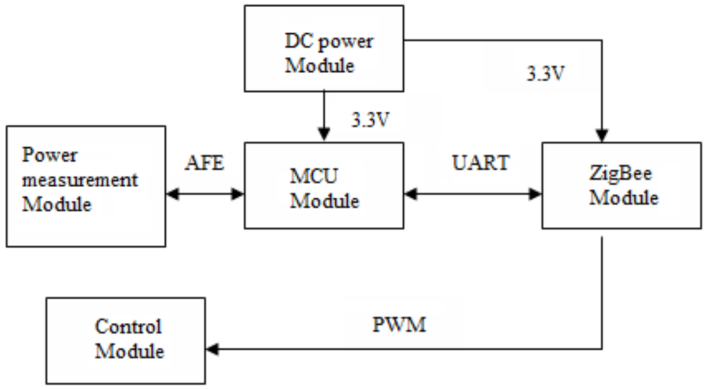

Fig. 1. Block diagram of sensor node

and processing, and then transmit the data through the RF communication module, which is connected with the home network to the client. When it receives instruction from the users, it can decode in the same way.

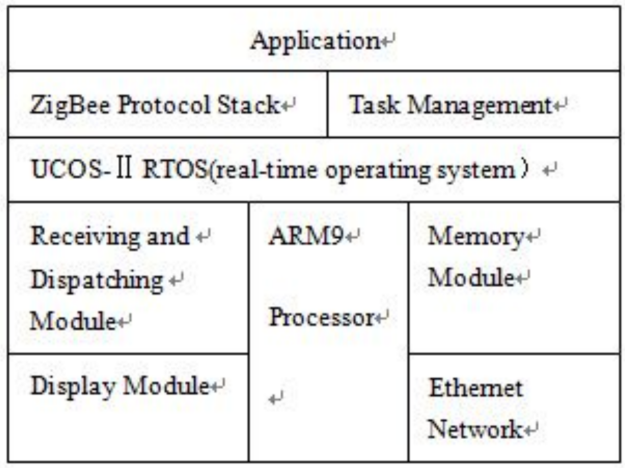

Fig. 2. ZigBee-based gateway architecture

As a part of data exchange, the ZigBee-based gateway will transmit the data from the home environment to the back-end computer network, which realizes the data acquisition of wireless sensor nodes; through the computer network, the client receives the data, and can store, manage and control the data.

2.2. System software architecture

According to the modular design principle, the system software is divided into three layers: the data service layer, the logic layer and the user layer. The data service layer stores the data from the nodes to the database; the logic processing layer provides a series of data processing interface to merge, eliminate redundancy and do other operation according to the requirements of the system; the user layer interact directly with users, so it accepts the user's request and reports to users if there is any exception.

2.2.1. The data service layer. ZBEMCIS uses relational database MySQL to storage the data. The table structure is shown in Fig. 4.

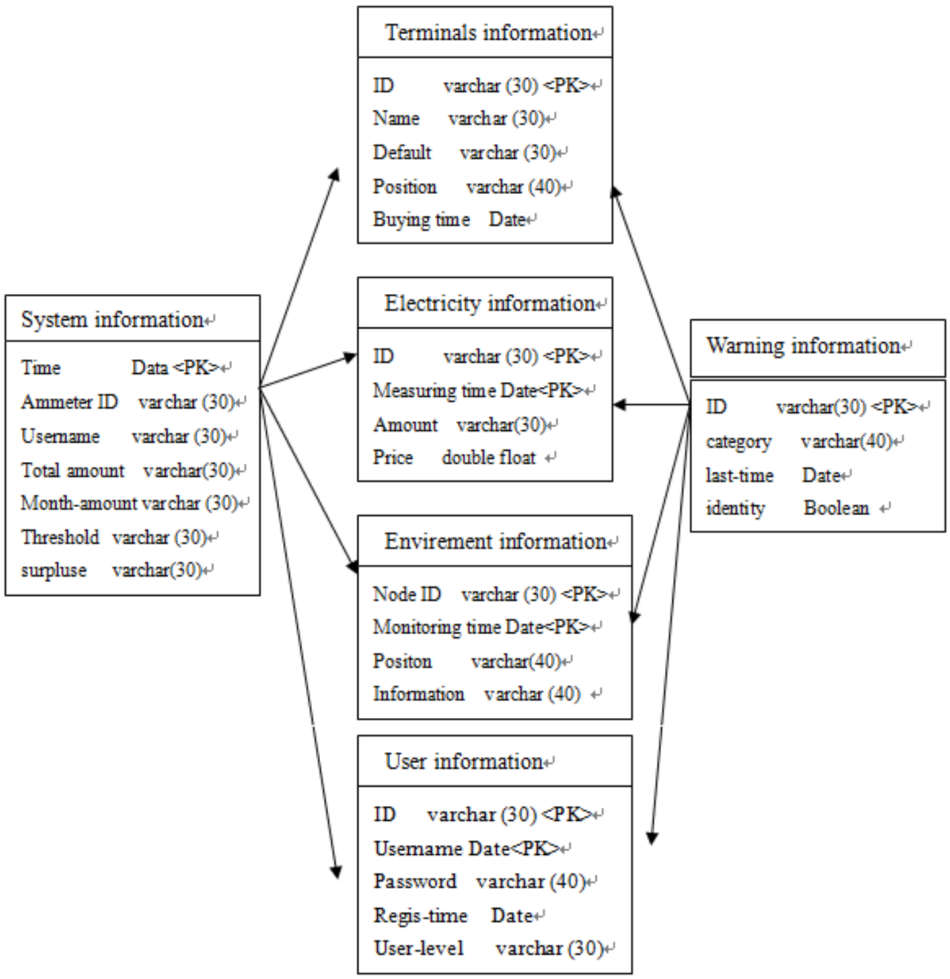

Fig. 3. Database table structures

The database of the system contains six tables used to store all data, and carry out a unified numbering system for the user's meter, electricity terminals, environmental monitoring position and user information. And the unified numbering system is used as the unique identifier to connect each entity.

In order to make the system high cohesion, low coupling characteristics, and easy to operate, we design to integrate Hibernate framework in the system because Hibernate is closed to JDBC, which makes Java programmer use Access Object DAO to operate the database. The system also has a unified interface to access database, so that the database can be realized and hide concrete realization.

2.2.2. Logic layer. The logic layer provides data access interface for the system, which is divided into four business sub modules: the business interface module, the business processing module, the data interface module and the data buffer module, shown in Fig. 5.

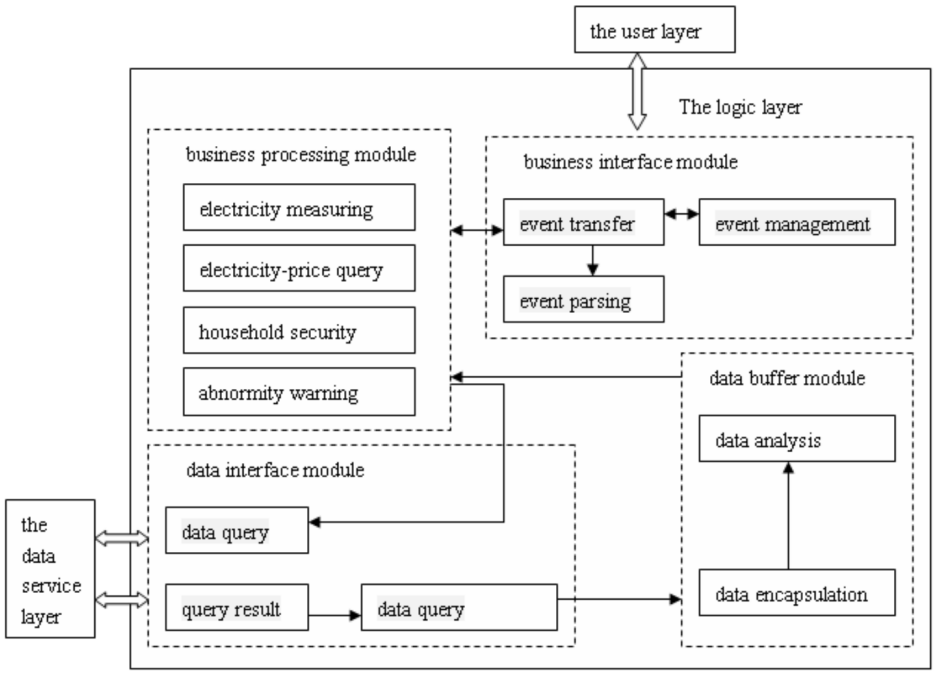

Fig. 4. Logic layer

When the program starts running, the status of each sensor is obtained, which includes the voltage, electricity of the household terminals, the indoor and outdoor environment, the smoke and gas concentration, and then determines whether these conditions are abnormal. If they are, the system will write the relevant information into the database. Otherwise, through the Internet, the system can connect to power network of administration of power supply to get the price. The program can count

the electric charge for users.

2.2.3. User layer. The system uses B/S framework, combining the PHP language and the CSS to develop the user layer. In order to ensure the system has a good response mechanism, the Ajax technology is used in the process of implementation.

3. System functions

The main functions of the system are as follows:

1. Checking usage of the power of the electric terminals in real time, and realize the unified control and management of all household appliances in the home.
2. Change of the parameters of the household terminals through the browser, thereby remotely controlling the switch of the terminals.
3. Automatic generation of electricity bills containing all of the equipment used, including the total number of electricity, the power consumption of the device.
4. Setting the power and energy consumption threshold for each household terminal, when the terminal is used or the total energy consumption exceeds the set default threshold, by sending e-mail to remind, users can decide whether or not to remotely close the power equipment.
5. Monitoring of home environment information system real-time, automatic control of indoor temperature and humidity information.
6. The smoke and gas detection sensors can monitor the indoor gas concentration, if abnormal events occur, there will be timely warning.

4. Conclusions

ZigBee has become a more and more important application of the Internet of Things (IoT) technology in intelligent household. In this paper, the authors propose a cost-effective ZBEMCIS, which can monitor and control the household electricity terminals, and the users can easily access the status and information from sensor nodes, which has a very important role in realizing savings in electricity consumption, monitoring, overload managing and remote control. And also it may make intelligent home appliances more diversified.

References

[1] Changhai Peng, Kun Qian: *Development and application of a ZigBee-based building energy monitoring and control system.* The Scientific World Journal *8* (2014), 1–13.

[2] S. Ahuja: *ZigBee based web enabled power monitor*. Masters-Project Report, California State University, Sacramento, CA, USA, (2010), 92 p.

[3] A. Nedelcu, F. Sandu, M. MacHedon-Pisu, M. Alexandru, P. Ogrutan: *Wireless-based remote monitoring and control of intelligent buildings*. Proc. IEEE International Workshop on Robotic and Sensors Environments, 6–7 Nov. 2009, Milano, Italy, 47–52.

[4] F. Shariff, N. A. Rahim, W. P. Hew: *Zigbee-based data acquisition system for online monitoring of grid-connected photovoltaic system*. Expert Systems with Applications *42* (2015), No. 3, 1730–1742.

[5] Liu Pan, Ding Rui: *Network nodes design of the internet of things based on ZigBee technology*. J Anhui Agricultural Sciences *20* (2015), 356–359.

[6] Wang Fuwen, Zhang Shengchao, Li Jianqi, Li Xiaofeng, Li Zhenyu: *Wireless nodes monitoring environment in vineyard based on ZigBee*. Computer and Digital Engineering *3* (2014), 535–538.

[7] C. Wilson, T. Hargreaves, R. Hauxwell-Baldwin: *Smart homes and their users: A systematic analysis and key challenges*. Personal and Ubiquitous Computing *9* (2015), No. 2, 463–476.

[8] Y. Yamagata, H. Seya: *Proposal for a local electricity-sharing system: A case study of Yokohama city, Japan*. IET Intelligent Transport Systems. *9* (2015), No. 1, 38–49.

[9] Tian Hang Zhang, Jia Yan, Shao Qing Xu: *An intelligent household electricity load control method based on demand response*. Applied Mechanics and Materials. *740* (2015), 307–310.

[10] P. Baronti, P. Pillai, V. W. C. Chooka, S. Chessa, A. Gottab, Y. Fun Hua: *Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards*. Computer Communications *30*, (2007), No. 7, 1655–1695.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 9–18

Vibration analysis of a shaft-bladed system by using solid models[1]

Chunlong Zheng[2], Xueshi Yao[2,3]

Abstract. To study the vibration characteristics of a rotating shaft-bladed system, the spin softening effect of finite element method (FEM) was incorporated with prestress effect to analyze the shaking, swing, torsion and their coupled vibration. The mathematical formula was derived from the Coriolis effects and gyroscopic inertia. The study shows that the centrifugal force exerts influence on the different order vibration frequencies. The effect of the spin softening has significant influence on the first order swing vibration frequency of the blades, but has little influence on the shaking and torsion vibration frequencies. With the variation of rotating velocity, the coupled vibration among shaking, swing and torsion could happen because their vibration frequencies change with the different magnitude of the increase. The coupled vibration is more destructive. The shaft's bending vibration split into forward whirl and backward whirl modes and the torsional vibration frequency does not change with the increase of the speed. The accuracy of the solid model and the solution techniques have been demonstrated by comparison with beam model results of a commercial software. The results further show the complexity of the dynamic characteristics of the shaft-bladed system. The work is helpful to improve the dynamic stability of the shaft-bladed system.

Key words. Shaft-bladed system, Coriolis effect, spin softening, swing vibration, coupled vibration..

1. Introduction

Resonant vibrations of a rotating shaft-bladed system can occur in many engineering structures such as aircraft propellers, helicopter and wind turbine rotors. To

[1]The authors are in debt to professors J. Chen and L. Q. Chen, Drs. T. Jia and W. H. Huang for their fruitful discussions. The work was supported by the Natural Science Foundation of China under Grant No. 11172181, the Natural Science Foundation of Guangdong Province of China under Grant No. 10151200501000008, the Educational Foundation of Scientific and Technical Innovation of Guangdong Province of China under Grant No. 2013KJCX0167, and the Scientific Research Foundation of Key Discipline of Guangdong Province of China.

[2]School of Physics and Electromechanical Engineering, Shaoguan University, Shaoguan 512005, China

[3]College of Mechanical and Electrical Engineering, Guangdong Baiyun University,Baiyun 510450, China

design the system, we must think of undergoing both global motion and longitudinal deformation. The coupling between the stress stiffness and deformation makes the dynamic characteristics of the system more complex, showing a strong nonlinearity. Due to the rotating speed over the critical speed, the system could have a destructive and non-synchronous whirl components because of the effects of Coriolis, variable load and deformation. Strong vibration will cause the accident, and bring about huge economic losses. In order to meet the engineering requirements, it is in an urgent need to improve the design level and optimize the shaft-bladed system.

Research on rotor dynamics is continually and steadily expanded in recent years. The coupled nonlinear equation which takes into account the stiffening effect is derived by applying the Lagrange equation for the moving beam, and based on the Newmark direct integration method and the Newton-Raphson iteration method, the computational procedures of the numerical method for solving the nonlinear equation are given [1].The influence of the curving and twisting of an elongated blade on its vibrations during complex rotation is studied, and it is shown that these geometrical factors may cause additional resonant vibrations [2].The blades are modeled as discrete multi-degree-of-freedom systems using the finite element software code ANSYS, and ANSYS is used to obtain the stiffness matrices of the blades, allowing the free vibration characteristics of the rotating blades to be determined by analytical formulation [3].The cylindrical rotor modes are not influenced by gyroscopic effects and remain at a fairly constant frequency versus rotor speed. Conversely, conical rotor modes are indeed influenced and caused to split into forward and backward whirl components that respectively increase and decrease in frequency with increased rotor speed [4]. Using solid models for the rotor dynamic analysis, the backward whirl is dominated by spin softening effect compared to the forward whirl frequencies where the stress stiffening plays a significant role [5–7].The study results on rotor dynamics provide a necessary foundation for further study on the shaft-bladed system [8–9].

2. Dynamic equation

2.1. Dynamic equation of beam in rotating reference frame

Using the rotating beam model to simulate a blade, the dynamic equation [8] is

$$[M]\{\ddot{u}_r\} + [C_{\mathrm{cor}}]\{\dot{u}_r\} + \{[K_{\mathrm{e}}]^{\mathrm{e}} + [K_{\mathrm{g}}]^{\mathrm{e}} + [K_{\mathrm{v}}]^{\mathrm{e}}\}\{u_r\} = \{F\}\ . \tag{1}$$

Here, $[M]$ is the mass matrix, $[C_{\mathrm{cor}}]$ denotes the Coriolis matrix, $[K_{\mathrm{e}}]^{\mathrm{e}}$ stands for the axial-lateral stiffness matrix of Euler-Bernoulli beam element model, $[K_{\mathrm{g}}]^{\mathrm{e}}$ represents the stiffness matrix of the centrifugal load corresponding to the angular velocity, $[K_{\mathrm{v}}]^{\mathrm{e}}$ is the spin softening stiffness matrix and F represents the load vector. Finally, symbol u_r stands for the vector of displacements.

The Coriolis matrix $[C_{\mathrm{cor}}]$ is given by the formula

$$[C_{\mathrm{cor}}] = 2\int_v \rho[N]^{\mathrm{T}}\varpi[N]\,\mathrm{d}v\,, \tag{2}$$

where $[N]$ is the shape function matrix,

$$[C_{\mathrm{cor}}] = 2\int_v \rho[N]^{\mathrm{T}}\varpi[N]\,\mathrm{d}v\,. \tag{3}$$

Here

$$\varpi = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_y \\ -\omega_y & \omega_z & 0 \end{bmatrix}, \tag{4}$$

where $\omega_x = \omega_y = 0$ and $\omega_z = \dot{\psi}$ is the rotating angular velocity ρ is the density and v is the volume of element. Finally,

$$u_r = [u_1, v_1, \beta_1, u_2, v_2, \beta_2]\ , \quad (\beta_1, \beta_2) = (v_1', v_2')\ .$$

In our case

$$[C_{\mathrm{cor}}] = \begin{bmatrix} 0 & \frac{-7}{10} & \frac{-1}{10} & 0 & \frac{-3}{10} & \frac{l}{15} \\ & 0 & 0 & \frac{3}{10} & 0 & 0 \\ & & 0 & \frac{l}{15} & 0 & 0 \\ & \text{antisym} & & 0 & \frac{-7}{10} & \frac{l}{10} \\ & & & & 0 & 0 \\ & & & & & 0 \end{bmatrix}, \tag{5}$$

Further,

$$[K_{\mathrm{e}}]^{\mathrm{e}} = \begin{bmatrix} A & 0 & 0 & -A & 0 & 0 \\ & \frac{12I}{l^2} & \frac{6I}{l} & 0 & -\frac{12I}{l^2} & \frac{6I}{l} \\ & & 4I & 0 & -\frac{6I}{l} & 2I \\ & & & A & 0 & 0 \\ & \text{sym} & & & \frac{12I}{l^2} & -\frac{6I}{l} \\ & & & & & 4I \end{bmatrix}, \tag{6}$$

$$[K_{\mathrm{g}}]^{\mathrm{e}} = \frac{EA}{60l^2}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ & 72\tilde{u} & 6\tilde{u}l & 0 & -72\tilde{u} & 6\tilde{u}l \\ & & 8\tilde{u}l^2 & 0 & -6\tilde{u}l & -2\tilde{u}l^2 \\ & & & 0 & 0 & 0 \\ & \text{sym} & & & 72\tilde{u} & -6\tilde{u}l \\ & & & & & 8\tilde{u}l^2 \end{bmatrix}, \tag{7}$$

where l is the beam element length and $\tilde{u} = u_2 - u_1$. The spin softening stiffness matrix is:

$$[K_{\mathrm{v}}]^{\mathrm{e}} = -\dot{\psi}^2\left([M_{\mathrm{t}}]^{\mathrm{e}} + [M_{\mathrm{r}}]^{\mathrm{e}}\right)\ , \tag{8}$$

where

$$[M_{\mathrm{t}}]^{e} = \frac{\rho Al}{420}\begin{bmatrix} 140 & 0 & 0 & 70 & 0 & 0 \\ & 156 & 22l & 0 & 54 & -13l \\ & & 4l^2 & 0 & 13l & -3l^2 \\ & & & 140 & 0 & 0 \\ & \text{sym} & & & 156 & -22l \\ & & & & & 4l^2 \end{bmatrix}, \tag{9}$$

and

$$[M_{\mathrm{r}}]^e = \frac{\rho I}{l}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ & \frac{6}{5} & \frac{l}{10} & 0 & -\frac{6}{5} & \frac{l}{10} \\ & & \frac{2l^2}{15} & 0 & -\frac{l}{10} & -\frac{l^2}{30} \\ & & & 0 & 0 & 0 \\ & \text{sym} & & & \frac{6}{5} & -\frac{l}{10} \\ & & & & & \frac{2l^2}{15} \end{bmatrix}. \tag{10}$$

2.2. Dynamic equation of beam in stationary reference frame

Using the rotating beam model to simulate a shaft, the dynamic equation is

$$\{[M]_{\mathrm{t}} + [M]_{\mathrm{r}}\}\,\ddot{u} + [G]\dot{u} + [K]_{\mathrm{e}}u = F\,. \tag{11}$$

Here, $u^{\mathrm{T}} = [v_i, w_i, \theta_{iy}, \theta_{iz}, v_j, w_j, \theta_{jy}, \theta_{jz}]$, $[G]$ is the gyroscopic matrix, $[K]_{\mathrm{e}}$ is the elastic stiffness matrix and $\{u\}, \{\dot{u}\}, \{\ddot{u}\}$ are the displacement, velocity and acceleration vectors, respectively. The matrix $[G]$ may be expressed as

$$[G] = 2\int_0^l \rho I\,[N']^{\mathrm{T}}\,[\omega]\,[N']\,\mathrm{d}x = 2\int_0^l \rho I \begin{bmatrix} N_1' & 0 \\ 0 & N_1' \\ 0 & N_3' \\ N_3' & 0 \\ N_2' & 0 \\ 0 & N_2' \\ 0 & N_4' \\ N_4' & 0 \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} 0 & -\dot{\psi} \\ \dot{\psi} & 0 \end{bmatrix} \cdot \begin{bmatrix} N_1' & 0 \\ 0 & N_1' \\ 0 & N_3' \\ N_3' & 0 \\ N_2' & 0 \\ 0 & N_2' \\ 0 & N_4' \\ N_4' & 0 \end{bmatrix}^{\mathrm{T}} \mathrm{d}x =$$

$$= 2\dot{\psi}\rho l \begin{bmatrix} 0 & \frac{6I}{5l^2} & -\frac{I}{10l} & 0 & 0 & -\frac{6I}{5l^2} & -\frac{I}{10l} & 0 \\ & 0 & 0 & -\frac{I}{10l} & \frac{6I}{5l^2} & 0 & 0 & -\frac{I}{10l} \\ & & 0 & \frac{2I}{15} & -\frac{I}{10l} & 0 & 0 & -\frac{I}{30} \\ & & & 0 & 0 & -\frac{I}{10l} & \frac{I}{30} & 0 \\ & & & & 0 & \frac{6I}{5l^2} & \frac{I}{10l} & 0 \\ & & & & & 0 & 0 & \frac{I}{10l} \\ & \text{antisym} & & & & & 0 & \frac{2I}{15} \\ & & & & & & & 0 \end{bmatrix}. \tag{12}$$

Symbols $M_{\mathrm{t}}, M_{\mathrm{r}}$ denote the translational and rotational inertial mass matrices and $\dot{\psi}$ is the angular velocity and $I = I_y = I_z$ is the bending moment of inertia. The one-dimensional beam models require good modeling techniques to approximate the three dimensional shaft. The effects of stress stiffening and spin softening are not included in the beam models as there is no cross-sectional dimension in the analysis. It is difficult to accurately obtain the calculation results in the rotor dynamics beam model.

2.3. Dynamic equations of solid in rotating reference frame

Using the solid model in rotating reference frame to simulate a blade, the dynamic equation is

$$[M]\{\ddot{u}_r\} + \{[C] + [C_{\mathrm{cor}}]\}\{\dot{u}_r\} + \left\{[K] + [K_\sigma] - \dot{\psi}^2[M]\right\}\{u_r\} = \{F\}\ . \tag{13}$$

Here $[M]$ is the global mass matrix, $[C]$ is the global damping matrix, $[K]$ denotes the global stiffness matrix, $[K_\sigma]$ stands for the global prestress stiffening matrix and $-\Omega^2[M] = -\dot{\psi}^2[M]$ is the stiffness matrix corresponding to the spin softening.

Matrix $[K_\sigma]$ can be expressed as

$$[[K_\sigma] = \begin{bmatrix} S_0 & 0 & 0 \\ 0 & S_0 & 0 \\ 0 & 0 & S_0 \end{bmatrix}, \tag{14}$$

$$[S_0] = \int_v [S_{\mathrm{g}}]^{\mathrm{T}}[S_{\mathrm{m}}][S_{\mathrm{g}}]\ \mathrm{d}v\ , \quad [S_m] = \begin{bmatrix} \sigma_x & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z \end{bmatrix} \tag{15}$$

and

$$[S_{\mathrm{g}}] = \begin{bmatrix} \frac{\partial N_1}{\partial x} & \frac{\partial N_2}{\partial x} & \cdots & \frac{\partial N_8}{\partial x} \\ \frac{\partial N_1}{\partial y} & \frac{\partial N_2}{\partial y} & \cdots & \frac{\partial N_8}{\partial y} \\ \frac{\partial N_1}{\partial z} & \frac{\partial N_2}{\partial z} & \cdots & \frac{\partial N_8}{\partial z} \end{bmatrix}. \tag{16}$$

2.4. Dynamics equation of solid in stationary reference frame

Using the solid model in stationary reference frame to simulate a shaft, the dynamic equation is

$$[M]\{\ddot{u}\} + \{[C] + [G_{\mathrm{gyr}}]\}\{\dot{u}\} + \left\{[K] + [K_\sigma] - \dot{\psi}^2[M]\right\}\{u\} = (F)\ . \tag{17}$$

The rotating shaft-bladed system to be modeled must be axisymmetric and the gyroscopic matrix generated is valid only for the linear analysis. The centrifugal effects of distributed shafts and mounted blades are included. Stress stiffening and spin softening are included, which have significant effect on the critical speeds and unbalance response. The rotor dynamics analysis can now predict the whirl amplitudes

more accurately, rather than just estimating the critical speeds and unstable regimes, thus improving the design capabilities of the shaft-bladed system [6].By setting the effects of gyroscopes, prestress and spin softening, formula (17) ($\{f_{\mathrm{g}}\} = [G_{\mathrm{gyr}}]\{\dot{u}\}$ being the gyroscopic moment) could correctly complete the dynamic analysis of the rotating shaft-bladed system by using the solid model of Ansys version 10.0.

3. Numerical results

3.1. Illustrative example

A helicopter blade radius is 10 m (along y axis), its width is 0.23 m. The blade cross-sectional area does not change along the radius. Based on a single blade as an example, its cross-sectional model is depicted in Fig. 1 and particular points have the following coordinates:

$$1:\ (0,0,-0.14)\,,\quad 2:\ (0,0,0.06)\,,\quad 3:\ (0.0133,0,0.09)\,,$$
$$4:\ (0.03,0,0.05)\,,\quad 5:\ (0.0167,0,-0.04)\,.$$

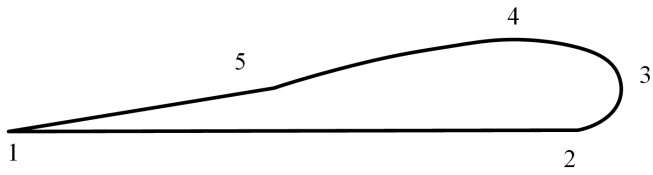

Fig. 1. Model of blade

The shaft length is 2.5 m with its outer diameter 0.06 m and inner diameter 0.04 m. The disk is placed in the axial direction of 1.74 m from the rotating blade center, of the disk thickness 0.08 m and diameter 0.4 m. The four blades are far from the disk, mounted on the outside of the bearing. Bearings of length 0.1 m are located on the center of shaft 0.1 m and 2.5 m. The actual shaft-bladed system is connected into a whole solid by the four blades and a shaft. The material performance parameter is

$$E = 42.6\cdot 10^{9}\,\mathrm{Pa}\,,\ \mu = 0.28\,,\ \rho = 1950\,\mathrm{kg\,m^{-3}}\,.$$

3.2. Calculation results

On the platform of Ansys 10.0, the blade is discretized by a mesh with 6362 nodes and 6478 solid46 elements. Based on different rotating speed, the vibration model of the blades is calculated in the shaft-bladed system. The Campbell diagram is given in Fig, 2, and the first swing vibration model as shown in Fig. 3. The results show that the centrifugal force exerts influence on the different order vibration

frequencies. Spin softening has great influence on the first swing frequency, but only a little influence on shaking and torsion vibration frequencies. With the variation of rotating velocity, the coupled vibration among shaking, swing and torsion can occur because their vibration frequencies change with the different magnitude of the increase. The coupled vibration is more destructive. The accuracy of the solid model and the solution techniques have been demonstrated by comparison with the beam element model.

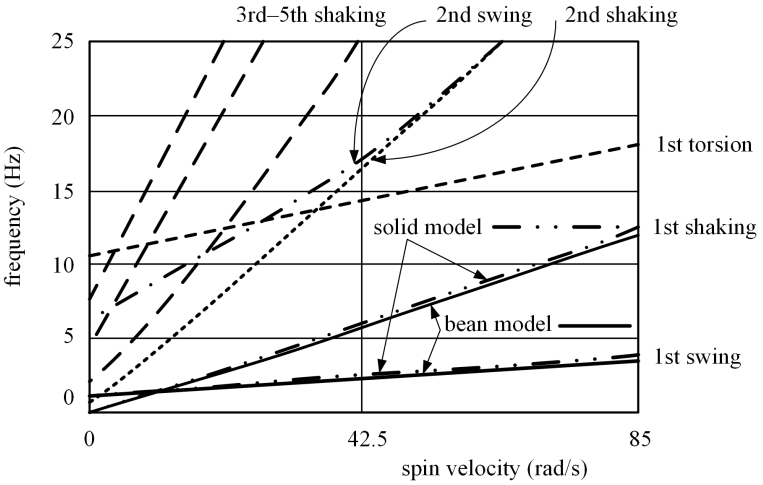

Fig. 2. Vibration model of blades

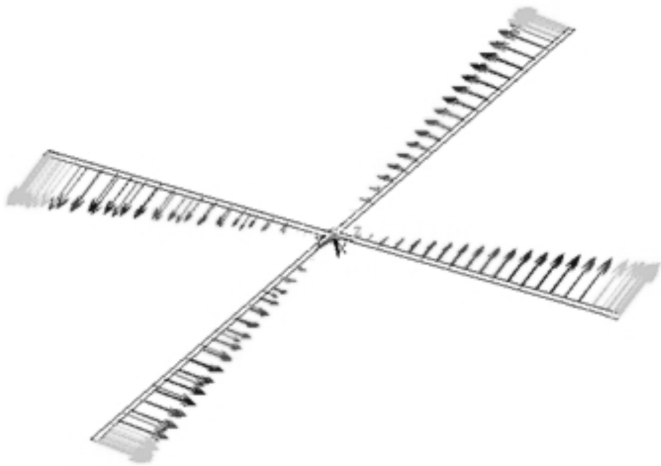

Fig. 3. First swing vibration

The shaft's Campbell diagram in the shaft-bladed system is given in Fig. 4. The

vibration modes are compared between the solid and beam models. The bending vibration split into forward whirl mode and backward whirl mode.The bending vibration modes are related with the shaft structure. The torsional vibration frequency does not change with the increasing speed. The beam model of the shaft does not take account the stress stiffness and spin softening effects that play an important role in the solid model. The solid model rotor dynamics provides an accurate solution for such problems.

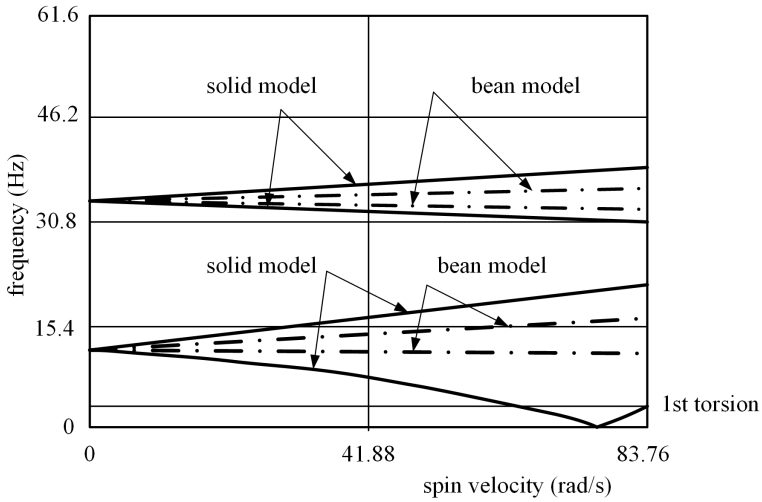

Fig. 4. Vibration model of shaft

The shaking, swing, torsion and their coupled vibration of the shaft-bladed system are calculated by solid and beam modes. The calculation shows that the spin softening has great influence on the first swing frequencies of the blades, but only a little influence on shaking and torsion vibration frequencies. The shaft bending vibration split into forward whirl mode and backward whirl mode and the torsional vibration frequency does not change with the increase of speed. The bending vibration models are related to the shaft structure. With the variation of rotating velocity, the coupled vibration between the blades and shaft can be induced easily.

A solid model shaft is used here, where centrifugal effects can play significant role in stress stiffening and spin softening effects, which are not considered in beam models. A specific advantage of the solid model lies in the fact that the whole model of the shaft-bladed system can be accounted in one analysis, which is impractical in beam models. Solid model rotor dynamics provides an accurate solution for such problems. The work lays a basic foundation for improving the dynamic stability of the shaft-bladed system.

4. Conclusion

Shaking, swing, torsion and coupled vibration of the shaft-bladed system are calculated by solid and beam modes. Calculation shows that the spin softening has great influence on the first swing frequencies of the blades, but has only a little influence on shaking and torsion vibration frequencies. The shaft bending vibration splits into forward whirl and backward whirl modes and torsional vibration frequency does not change with the increase of the speed. The bending vibration models are related to the shaft structure. With the variation of rotating velocity, the coupled vibration between the blades and shaft can be induced easily.

Solid model shaft is used here, where centrifugal effects can play significant role in stress stiffening and spin softening effects, which are not considered in the beam models. A specific advantage of the solid model lies in the fact that the whole model of the shaft-bladed system can be calculated in one analysis, which is impractical in the beam models. Solid model rotor dynamics provides an accurate solution for such problems. The work lays a basic foundation for improving the dynamic stability of the shaft-bladed system.

References

[1] A. Y. T. Leung, G. R. Wu, W. F. Zhong: *Vibration analysis of flexible beam undergoing both global motion and elastic deformation.* IJ Str. Stab. Dyn. *4* (2004), No. 4, 589–598.

[2] V. I. Gulyaev, S. N. Khudolii: *Vibrations of curved and twisted blades during complex rotation.* Int. App. Mech. *41* (2005), No. 4, 449–454.

[3] P. J. Murtagh, B. Basu, B. M. Broderick: *Mode acceleration approach for rotating wind turbine blades.* J Multi-body Dynam. *218* (2004), No. 3, 159–167.

[4] E. Swanson, C. D. Powell, S. Weissman: *A practical review of rotating machinery critical speeds and modes.* Sound and Vibration *39* (2005), No. 5, 10–17.

[5] J. S. Rao: *Rotor dynamics of aircraft gas turbine engines.* J Aerospace Sci. Technol. *60* (2008), No. 3, 169–182.

[6] J. S. Rao, R. Sreenivas: *Dynamics of a three level rotor system using solid elements.* Proc. IC Joint Power Generation, 16–19 June 2003, Atlanta, USA, pp. 601–606.

[7] X. S. Yao, C. L. Zheng, Y. F. Liao: *Dynamics of turbocharger rotor with spin softening effect.* Appl. Mechanics and Materials *397–400* (2014), 90–94.

[8] M. A. Trinidade, R. Sampaio: *Dynamics of beams undergoing large rotations accounting for arbitrary axial deformation.* J Guidance Control and Dynamics *25* (2002), No. 4, 634–643.

[9] Z. L. Mahri, M. S. Rouabah: *Calculation of dynamic stresses using finite element method and prediction of fatigue failure for wind turbine rotor.* WSEAS Trans. Appl. Theor. Mechanics *3* (2008), No. 1, 28–41.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 19–28

Thermal oxidation kinetics analysis of ferrous sulfide under different heating rates[1]

Man Yang[2], Xianfeng Chen[2], Yujie Wang[2], Xiaoguang Yue[3]

Abstract. Ferrous sulfide (FeS) samples of the same mass were investigated by STA449F3 simultaneous thermal analyzer in order to analyze their thermal oxidation kinetics parameters. The reaction processes at different heating rates of 5 °C/min, 10 °C/min and 15 °C/min in air from 30 °C to 900 °C were comprehensively studied by thermogravimetry (TG), differential scanning calorimetry (DSC) and derivative thermogravimetry (DTG) methods. FeS kinetics parameters and reaction mechanism function were obtained with the combination of Satava-Sestak equation and Flynn–Wal–Ozawa (FWO) method. The results show that heating rates have significant effects on the characteristic temperatures of samples. With increase of heating rates, the characteristic temperatures of FeS samples gradually raise; the secondary reaction model acquired by the Satava–Sestak equation conforms to mechanism function; activation energy E obtained by FWO method is 153.83 kJ/mol, the pre-exponential factor A being $2.625 \cdot 10^8$.

Key words. Thermal oxidation kinetics, ferrous sulfide, mechanism function, characteristic temperatures, activation energy, pre-exponential factor..

1. Introduction

In storage and transportation of crude oil, active sulfur and organic sulfur in petroleum products together with the storage tank walls generate FeS. When FeS reacts with oxygen under some conditions coincident with exothermic reaction, continuous accumulation of heat can lead to high temperatures, even fire and explosion [1]. Meanwhile, oil tank fire and explosion seriously affect people's physical security and results in great losses in national economy [2–3].

Massive researches were made by many scholars for exploring the reasons of these accidents and thermal analysis kinetics was applied in the study of FeS ther-

[1]This work was supported by National Natural Science Foundation of China (projects No. 51374164, and No. 51174153).

[2]School of Resources and Environmental Engineering, Wuhan University of Technology, Wuhan 430070, China

[3]School of Civil Engineering, Wuhan University, Wuhan 430072, China

mal spontaneous mechanism in recent years [4–8]. The kinetic triplets [9], including activation energy A, pre-exponential factor E and mechanism function in differential form play an important role in evaluating energetic material stabilization in exothermic decomposition reaction. The activation energy regarded as the index of evaluating the spontaneous combustion tendency of FeS was usually calculated by the FWO, Kissinger, Coats–Redfern or Friedman method, respectively. To obtain the activation energy and mechanism function, the Iso-conversional method and Satava–Sestak equation were jointly applied in some references [10–12]. The activation energy and pre-exponential factor were acquired by the Iso-conversional methods, then the most probable mechanism function $f(\alpha)$ was inferred by combination with the Malek method [13]. However, the above methods could only obtain a part of the kinetic factors. Besides, according to our knowledge, there is no report regarding the application of a method to acquire FeS kinetic triplets in an easy and scalable way.

In this work, we report an easy way of obtaining activation energy, pre-exponential factor and mechanism at the same time, which is with the combination of Satava–Sestak equation and FWO method on the basis of TG, DSC and DTG analysis at different heating rates.

2. Experimental section

FeS ore ($\geq$85.0 % purity) was purchased from Tianjin Guangfu Fine Chemical Co., Ltd. (Tianjin, China). Then it was dryed and dewatered in the electrothermal drying oven, grinded and screened after constant weight was attained, and finally, 200–240 mesh FeS samples were selected.

STA449F3 simultaneous thermal analyzer, produced by NETZSCH Company in Germany was purchased, where air was used as the purge gas with an air flow rate of 30 ml/min during the whole experiment. The samples with mass of 5 mg were measured at the heating rates of 5, 10 and 15 °C/min, respectively, from 30 ° to 900 °C.

Figures 1–3 shows the typical TG, DSC and DTG curves under different conditions. Generally for FeS, weight loss and weight gain phenomenon appear obviously with the increase of temperature from 30 °C to 300 °C. The initial weight loss could be ascribed to the removal of absorbed water, and weight gain indicated the absorption of oxygen. FeS exhibited a constant increase from 300 °C to 600 °C, which could be attributed to the weakening of physical absorption, slight beginning of chemisorption. The steep decrease in the weight of FeS at around 600 °C could be attributed to drastic combustion and release of SO_2. As shown in the Table 1, FeS ignition temperatures were 594, 615 and 628 °C under the heating rates of 5, 10 and 15 °C/min respectively. DSC curves (Fig. 1) had a small phase transformation absorption peak at around 140 °C. FeS had an obvious exothermic peak under the different heating rates. The initial temperatures of exothermic peaks were 570, 589 and 603 °C (Table 1), the corresponding peak temperatures were 620, 643 and 665 °C (Table 1). As shown in Fig. 3, although an obvious weight loss peak was presented under the different heating rates, the heating rates had less influence on

DTG curve. The maximum weight loss rate had slight increase, and the corresponding temperatures were 618, 646 and 665 °C (Table 1). TG, DSC and DTG curves all move to higher temperature direction with the increasing heating rate, which can be ascribed to the heat delivered from crucible to sample resulting in differential temperature between furnace and sample. Once temperature gradient occurred, the differential temperatures were also increasing with the heating rates.

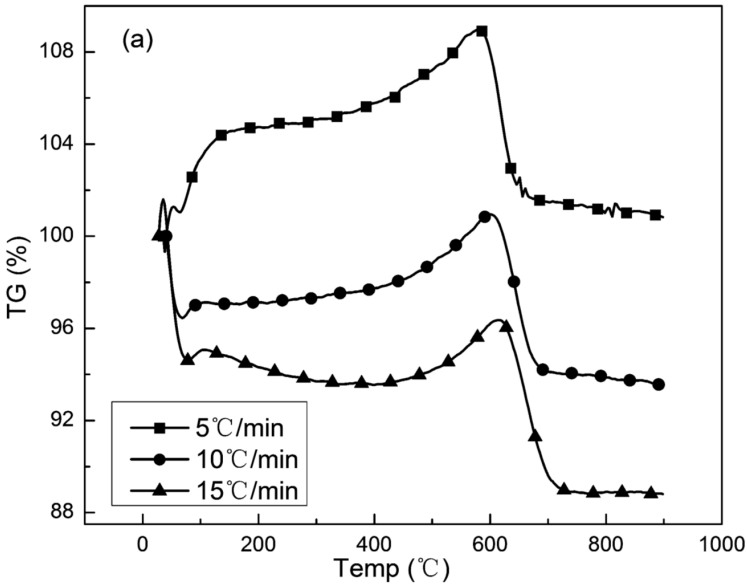

Fig. 1. TG curves of FeS under different heating rates

3. Theoretical analysis

The methods of thermal kinetics analysis [14, 15] can be classified as the application of thermal analysis techniques to study physical properties and chemical reaction rates and mechanism in materials, and to obtain the reaction dynamics parameters and mechanism functions. In order to determine the mechanism of solid state reaction and corresponding dynamics parameters, the study can start from the thermal kinetics equation.

FeS oxidation is a traditional gas-solid two-phase reaction, whose dynamics integration equation can be directly showed as follows

$$G(a) = \frac{A}{\beta} \int_0^T e^{-E/RT} \, \mathrm{d}T \,, \tag{1}$$

where α is the conversion degree, β is the heating rate (K/min), E is the activation energy (kJ/mol) A is the pre-exponential factor (K/s), $G(\alpha)$ is the integral expression

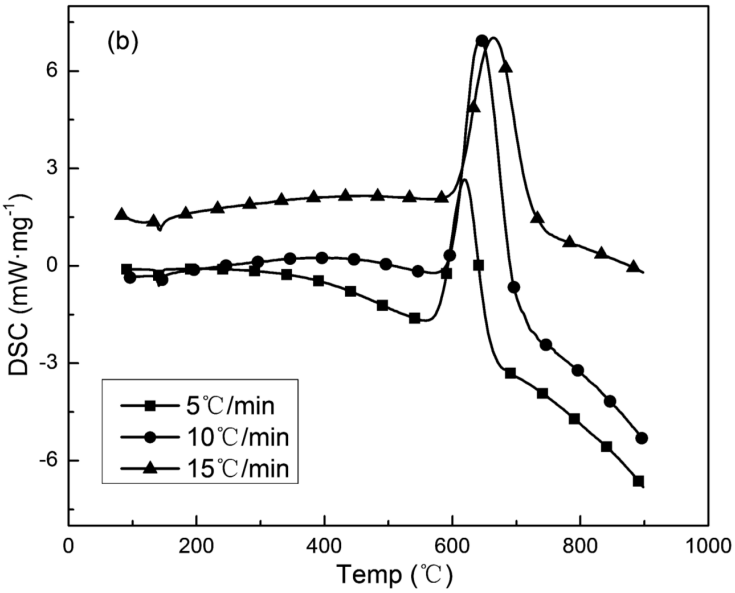

Fig. 2. DSC curves of FeS under different heating rates

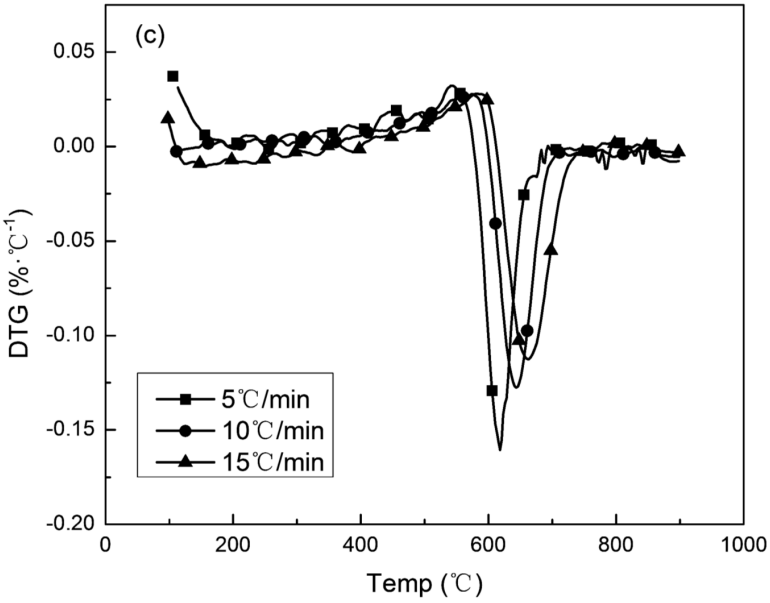

Fig. 3. DTG curves of FeS under different heating rates

of FeS oxidation reaction kinetics model, T is the reaction temperature, and R is the gas constant.

Table 1. Mode functions of commonly used gas solid reaction mechanism

Reaction mechanism	Symbol	Integral form $g(\alpha)$	Differential form $f(\alpha)$
One-dimensional diffusion	D1	α^2	$(2a)^{-1}$
Two-dimensional diffusion	D2	$\alpha + (1-\alpha)\ln(1-\alpha)$	$[-\ln(1-\alpha)]^{-1}$
Three-dimensional diffusion	D3	$\left[1-(1-\alpha)^{1/3}\right]^2$	$\frac{3}{2}(1-\alpha)^{\frac{2}{3}}\left[1-(1-\alpha)^{1/3}\right]^{-1}$
Four-dimensional diffusion	D4	$(1-2a/3)-(1-a)^{2/3}$	$3/\left\{2\left[(1-a)^{-1/3}-1\right]\right\}$
Random nucleation theory and subsequent growth	A2	$[-\ln(1-a)]^{1/2}$	$2(1-a)\,[-\ln(1-a)]^{1/2}$
Random nucleation theory and subsequent growth	A3	$[-\ln(1-a)]^{1/3}$	$3(1-a)\,[-\ln(1-a)]^{2/3}$
Interface reaction	R1	α	1
Interface reaction	R2	$1-(1-\alpha)^{1/2}$	$2(1-\alpha)^{1/2}$
Interface reaction	R3	$1-(1-\alpha)^{1/3}$	$3(1-\alpha)^{2/3}$
First-order reaction	F1	$-\ln(1-\alpha)$	$1-\alpha$
Second order reaction	F2	$(1-\alpha)^{-1}-1$	$(1-\alpha)^2$

By integral transform of equation (1), FWO equation can be written as follows:

$$\lg\beta = \left[\frac{AE}{RG(\alpha)}\right] - 2.315 - 0.4567\frac{E}{RT} \tag{2}$$

With different β and the same α, $G\alpha$ is constant. Although FWO method is a common method to obtain the activation energy, E is directly obtained avoiding the problem of selecting suitable reaction mechanism functions, but the pre-exponential factor and reaction mechanism functions cannot be obtained.

After transformation of equation (1), Satava–Sestak equation can be attained in

the following form

$$\lg G(\alpha) = \left[\frac{A_{\mathrm{s}}E_{\mathrm{s}}}{R\beta}\right] - 2.315 - 0.4567\frac{E_{\mathrm{s}}}{RT}\,, \tag{3}$$

where A_{s} is a pre-exponential factor (in K/s) and E_{s} is the activation energy (kJ/mol). Both these quantities are obtained from the deformation of formula (2).

When β_i is fixed, T_{ij} and α_{ij} are substituted into the expression (3), then k equations can be expressed as follows

$$\lg G(\alpha_{ij}) = \left[\frac{A_{\mathrm{s}}E_{\mathrm{s}}}{R\beta_i}\right] - 2.315 - 0.4567\frac{E_{\mathrm{s}}}{RT_{ij}}\,, \tag{4}$$

where $i = \text{const}$ and $j = 1, 2, \ldots, k$. As β_i is fixed, then $\lg(A_sE_s/(R\beta_i)$ is a constant, thus the equation (4) is a linear equation and the linear least square method can be applied to solve these equations.

Substituting the mode functions (Table 1) into the equation (4) under different heating rates, and plotting $\lg G(\alpha)$ versus $1/T$ (see Figs. 4–6), one can obtain a series of curves, in which the best linear correlation curve represents FeS reaction mechanism.

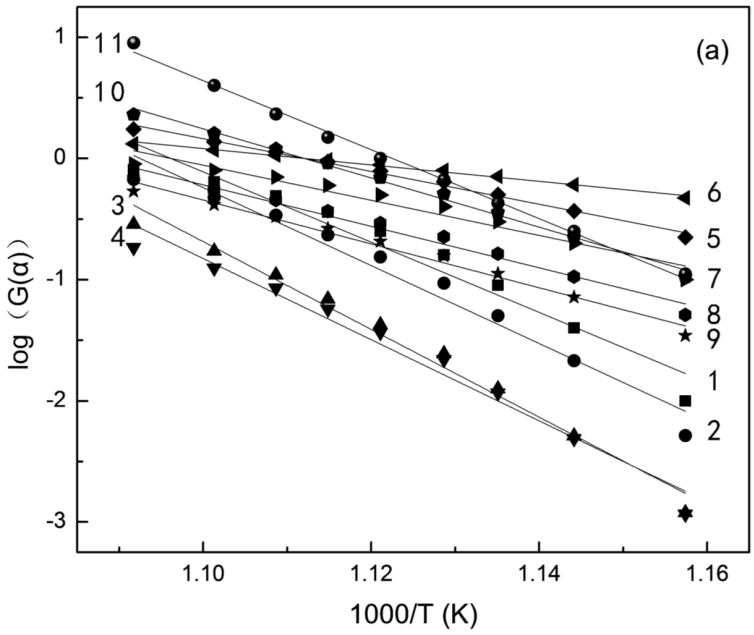

Fig. 4. Plots of $\lg G(\alpha)$ versus $1000/T$ at heating rate 5 °C/min: 1–D1, 2–D2, 3–D3, 4–D4, 5–A2, 6–A3, 7–R1, 8–R2, 9–R3, 10–F1, 11–F2

Then the determined reaction function of FeS is substituted into the expression (2), and after plotting $\lg G(\alpha)$ versus $1/T$, a curve can be obtained, and the

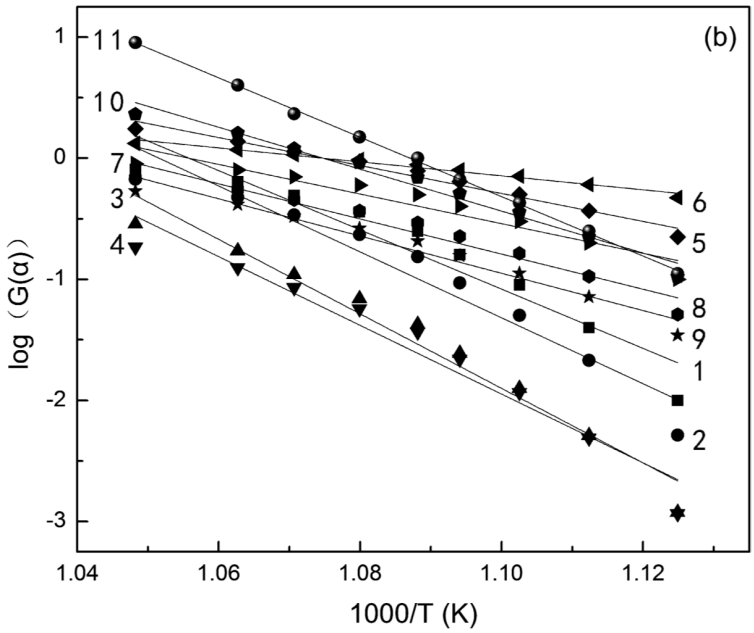

Fig. 5. Plots of $\lg G(\alpha)$ versus $1000/T$ at heating rate 10 °C/min: 1–D1, 2–D2, 3–D3, 4–D4, 5–A2, 6–A3, 7–R1, 8–R2, 9–R3, 10–F1, 11–F2

corresponding activation energy and pre-exponential factor of FeS with different conversion degrees can be calculated by slope intercept form.

Table 2. Correlation coefficients of various mechanism functions with temperature rate

Temperature rate (K/min)	5	10	15
D1	−0.9747	−0.956 89	−0.962 79
D2	−0.983 35	−0.968 38	−0.973 47
D3	−0.991 66	−0.980 32	−0.984 33
D4	−0.986 57	−0.972 85	−0.977 57
A2	−0.996 78	−0.988 85	−0.991 84
A3	−0.996 78	−0.988 85	−0.991 84
R1	−0.9747	−0.956 89	−0.962 79
R2	−0.988 18	−0.975 16	−0.979 67
R3	−0.991 66	−0.980 32	−0.984 33
F1	−0.996 78	−0.988 85	−0.991 84
F2	−0.997 91	−0.999 46	−0.999 54

Table 2 shows results of different kinds of functions substituted into equation (4) under different heating rates (5, 10 and 15 °C/min). It can be found that F2 mode has the best correlation, which can be taken as reaction mechanism function

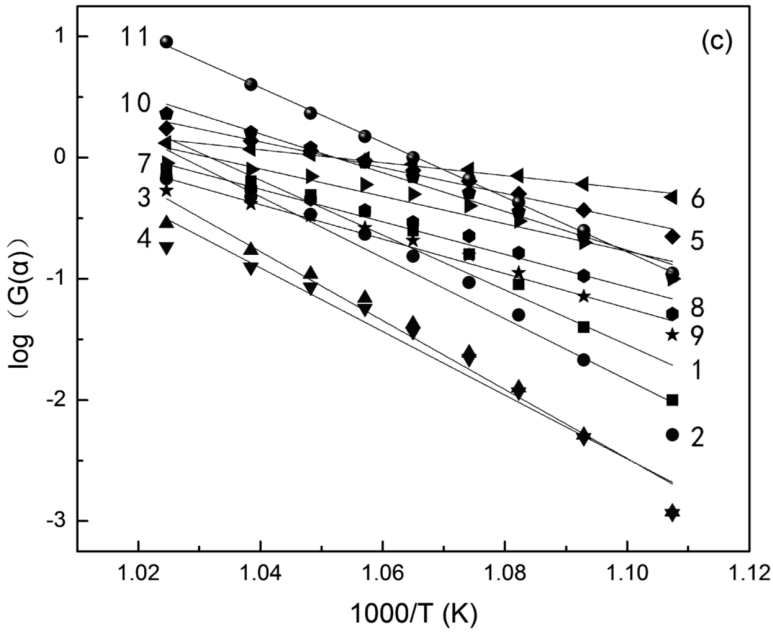

Fig. 6. Plots of $\lg G(\alpha)$ versus $1000/T$ at heating rate 15 °C/min: 1–D1, 2–D2, 3–D3, 4–D4, 5–A2, 6–A3, 7–R1, 8–R2, 9–R3, 10–F1, 11–F2

in the corresponding temperature interval. In other words, $G(\alpha) = (1-\alpha)^{-1}$, $f(\alpha) = (1-\alpha)^2$ represents the second order reaction. Besides, $E_{s5} = 167.46\,\mathrm{kJ/mol}$, $E_{s10} = 158.52\,\mathrm{kJ/mol}$ and $E_{s15} = 139.11\,\mathrm{kJ/mol}$.

Then substituting F2 mode into the expression (2), kinetic parameters can be obtained by the same conversion degree. The average conversion degree of each kinetic parameter is shown in Table 3, where the apparent activation energy $E_0 = 153.83\,\mathrm{kJ/mol}$ and $\lg A = 8.42$.

Table 3. Kinetic parameters of FeS conversion degree rate

α	E (kJ/mol)	A (K/s)	$R2$ (−)
0.1	173.071 920 9	552 981 875.2	0.998 97
0.2	169.692 617 6	586 904 014.2	0.999 73
0.3	164.805 616 7	423 117 732.1	0.999 43
0.4	159.133 637 3	264 105 090.2	0.999 98
0.5	155.570 286 4	212 791 055.6	0.995 01
0.6	150.960 540 2	150 989 879.3	0.998 22
0.7	143.791 057 8	79 637 006.5	0.999 99
0.8	138.437 293 3	59 113 532.42	0.999 27
0.9	128.998 254 5	32 468 236.13	0.999 27
average	153.829 025	262 456 491.3	

Finally

$$\left|\frac{E_0 - E_{s15}}{E_0}\right| = 0.00957 < 0.1\,, \tag{5}$$

$$\left|\frac{E_0 - E_{s15}}{E_0}\right| = 0.03049 < 0.1\,, \tag{6}$$

$$\left|\frac{E_0 - E_{s15}}{E_0}\right| = 0.0886 < 0.1\,. \tag{7}$$

According the above criterion, the results of expressions (5), (6) and (7) are smaller than 0.1, thus FeS thermo-oxidation is in accordance with the secondary reaction model (F2): integral form $G(\alpha) = (1-\alpha)^{-1}$, the corresponding activation energy E is 153.83 kJ/mol, and the pre-exponential factor is $A = 2.625 \times 10^8$.

4. Conclusion

- According to the TG, DSC and DTG curves, the heating rates have significant effects on the peak temperatures of samples. Under the heating rate of 5, 10 and 15 °C/min, respectively, FeS ignition temperatures were 594, 615 and 628 °C, the initial temperatures of exothermic peak were 570, 589 and 603 °C, the corresponding peak temperatures were 620, 643 and 665 °C, the corresponding temperatures of the maximum weight loss rate were 618, 646 and 665 °C. The results show that the characteristic temperatures of FeS samples will gradually raise with the increase of heating rate.

- FeS dynamic parameters and reaction mechanism functions of thermo-oxidation were solved with combination of Satava–Sestak method and FWO method. FeS thermo-oxidation determined by Satava–Sestak equation is in accordance with the secondary reaction model (F2): integral form $G(\alpha) = (1-\alpha)^{-1}$, activation energy E of FeS thermo-oxidation obtained by FWO method is 153.83 kJ/mol and the pre-exponential factor $A = 2.625 \times 10^8$.

References

[1] Z. H. Zhang, S. L. Zhao, P. Li, Y. Han: *Spontaneous combustion process of hydrogen sulfide corrosion products formed at room temperature.* Acta Petrolei Sinica (Petroleum Processing Section) *28* (2012), No. 1, 122–126.

[2] S. Lei, S. Jian, X. Kui: *Fuzzy fault tree assessment based on improved AHP for fire and explosion accidents for steel oil storage tanks.* J Hazardous Materials *278* (2014), 529–538.

[3] D. Y. Zhang, A. Z. Ding, S. C. Cui, C. Hu, S. F. Thornton, J. Dou, Y. Sun, W. E. Huang: *Whole cell bioreporter application for rapid detection and evaluation of crude oil spill in seawater caused by Dalian oil tank explosion.* Water Research *47* (2013), No. 3, 1191–1200.

[4] S. R. Niranjan, N. Datta, K. Suresh: *Effect of Fe modification on H_2S sensing properties of rheotaxially grown and thermally oxidized SnO_2 thin films.* Mat. Chem. Phys. *156* (2015), 227–237.

[5] N. Y. Dzade, A. Roldan, N. H. de Leeuw: *Adsorption of methylamine on mackinawite (FeS) surfaces: A density functional theory study.* J Chem. Phys. *139* (2013), No. 12, 28–36.

[6] W. H. Wang, Z. S. Xu, B. J. Sun, D. Liang: *Experiment investigation on spontaneous combustion of iron sulfides in oil tank.* Advanced Materials Research *750–752* (2013), 1758–1764.

[7] F. Q. Yang, C. Wu, H. Liu, W. Pan, Y. Cui: *Thermal analysis kinetics of sulfide ores for spontaneous combustion.* J Central South University (Science and Technology) *42* (2011), No. 8, 2469–2474.

[8] M. J. Glenn, J. A. Allen, S. W. Donne: *Thermal investigation of a doped alkali-metal carbonate ternary eutectic for direct carbon fuel cell applications.* Energy & Fuels *29* (2015), No. 8, 5423–5433.

[9] M. Malow, U. Krause: *The overall activation energy of the exothermic reactions of thermally unstable materials.* J Loss Prevention in the Process Industries *17* (2004), No. 1, 51–58.

[10] M. Yang, X. F. Chen, Y. J. Shang, R. D. Bao: *Particle size effect on FeS spontaneous combustion characters in petroleum.* Appl. Mech. Mater. *608–609*, (2014), 971 to 975.

[11] S. P. Zhao, J. C. Jiang, Y. X. Yang: *Thermal kinetic analysis on oxidation of FeS based on thermal gravimetric experiments.* Acta Petrolei Sinica (Petroleum Processing Section) *26* (2010), No. 6, 972–976.

[12] Huang Lei, Chen Yucheng, Liu Geng, Li Shengnan, Liu Yun, Gao Xu: *Non-isothermal pyrolysis characteristics of giant reed (Arundo donax L.) using thermogravimetric analysis.* Energy *87* (2015), 31–40.

[13] J. Z. Liang, J. Z. Wang, G. C. P. Tsui, C, Y. Tang: *Thermal decomposition kinetics of polypropylene composites filled with graphene nanoplatelets.* Polymer Testing *48* (2015), 97–103.

[14] M. Montazeri-Gh, M. Mahmoodi-k: *Optimized predictive energy management of plug-in hybrid electric vehicle based on traffic condition.* J Cleaner Production *149* (2016), 935–948.

[15] A. N. D'yachenko, R. I. Kraidenko: *Processing oxide-sulfide copper ores using ammonium chloride.* Russian Journal on Non-ferrous Metals *51* (2010), No. 5, 377–381.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 29–36

Study on influence of the NO_x emissions characteristics entering the steady point in different initial catalyst temperature

Bi Shiying[1]

Abstract. A bench test was carried out in order to study the influence of the emissions characteristics of NO_x entering the steady point in different initial catalyst temperatures. The results showed that firstly the NO_x emissions gradually reduced, then started rising when reduced to about 10 ppm, and finally tended to be stable. The temperature of entering the point was lower, the time when NO_x reaches the equilibrium state was longer. The ammonia nitrogen ratio was higher, the minimum value NO_x that could be achieved was lower, the ammonia storage cushioning effect for NO_x was stronger, and the time in which emissions achieved steady state was also longer.

Key words. Diesel engine, nitrogen oxides, elective catalytic reduction, ammonia.

1. Introduction

At present, Phase IV Emission Standard has been fully implemented in emissions requirements of China's heavy-duty automotive diesel engine. SCR (Selective Catalytic Reduction) technology which uses urea as a reducing agent has been widely used because of its advantages such as improved fuel consumption by optimized in-cylinder combustion and strong anti-sulfur poisoning, etc., which plays an important role in alleviating air pollution caused by diesel exhaust. SCR upstream temperature has a direct impact on catalyst performance. During the test, presence of hot car stability time inconsistency, piping layout as well as original machine status will result in SCR upstream temperature difference in steady state cycle, thereby affecting NO_x emission results. In this paper, through A100 operating point of upstream temperature of different catalysts to be switched to steady circulation, effect of initial temperature of different catalysts on NO_x emissions and ammonia leakage of the steady-state operating point is studied.

The main technical parameters of the diesel engine are listed in Table 1.

[1]Weifang University Shandong, Weifang, 261061, China

Table 1. Main technical parameters of diesel engine

Parameter	Unit	Value
Rated power	kW	110
Rated speed	r/min	2300
Maximum torque	N·m	550
Maximum no-load stabilized speed	r/min	2650
Minimum no-load stabilized speed	r/min	650

2. SCR reaction

2.1. SCR reaction fundamentals

Basic operating principle of SCR system consists of several steps. First, exhaust enters exhaust pipe after flowing out of supercharger turbo. Meanwhile, urea injection unit installed in SCR system upstream sprays quantitative aqueous solution of urea into exhaust pipe in the form of mist. Urea solution drop has hydrolysis and pyrolysis reactions under the role of high temperature exhaust gas, generating required reducing agent NH_3 which selectively restores NO_x to nitrogen N_2 under catalyst. The main reaction is as follows:

Urea formula is $(NH_2)_2CO$. This agent is decomposed into NH_3 and CO_2 at high temperature, after adding H_2O. This reaction can be completed without catalysis conditions:

$$(NH_2)_2CO + H_2O \rightarrow 2NH_3 + CO_2\,. \tag{1}$$

Ammonia NH_3 in SCR catalyst reacts with NO and NO_2 in exhaust gas, generating N_2 and H_2O [1–3]

$$4NH_3 + 4NO + O_2 \rightarrow 4N_2 + 6H_2O\,, \tag{2}$$

$$4NH_3 + 2NO + 2NO_2 \rightarrow 4N_2 + 6H_2O \tag{3}$$

and

$$4NH_3 + 2NO_2 + O_2 \rightarrow 3N_2 + 6H_2O\,. \tag{4}$$

2.2. Ammonia storage capacity

Adjust fixed point hot car to idling after stabilization, then cool catalyst temperature, so that inlet temperature begins to decrease. When inlet temperature reaches 400 °C, 350 °C, 300 °C, 250 °C, 200 °C, respectively, rapidly switch to A100 operating point, open urea switch and record each parameter variation of NO_x, NH_3, SCR second by second. The relevant parameters of A100 operating point under steady state are shown in Table 2.

As can be seen from Fig. 1, as the temperature rises, ammonia storage capacity of the catalyst continues to decrease. When the temperature reaches 450 °C, there is almost no ammonia storage capacity.

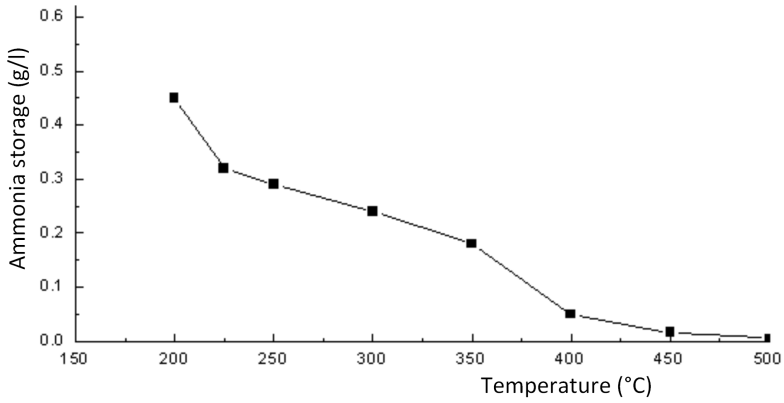

Fig. 1. Ammonia storage amount of catalyst obtained in sample test

3. Bench test system

A 4l diesel engine was chosen for bench test. The main parameters of diesel engine are shown in Table 1, and test bench system layout is shown in Fig. 2.

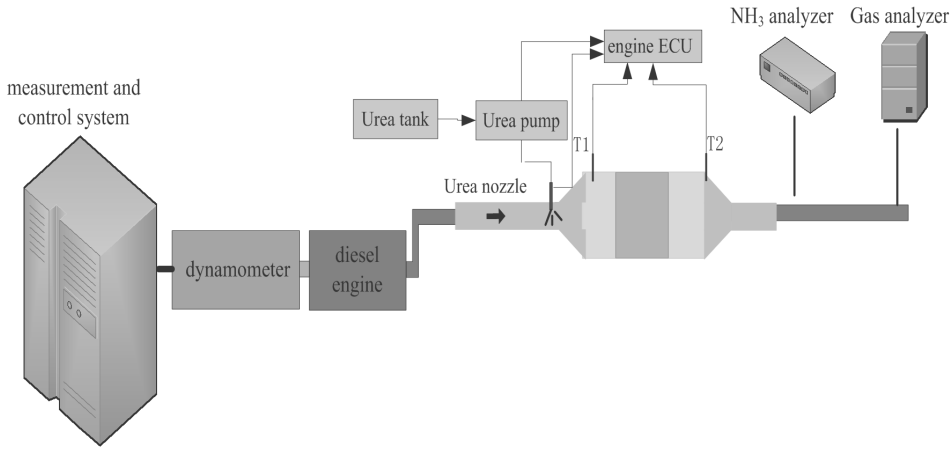

Fig. 2. Bench test system

Table 2. Parameters of A100 operating point under steady state

Parameters	Unit	Value
Rotate speed	r/min	1417
Torque	N·m	536
Power	kW	79.7
Airspeed	h^{-1}	27000
Set ammonia nitrogen ratio	-	0. 83

4. Test results and analysis

When initial temperature was lowered to 250 °C, NO_x, NH_3 and SCR inlet temperature variations with time are shown in Fig. 3. NO_x and NH_3 emission results at different initial temperatures are shown in Figs. 4 and 5.

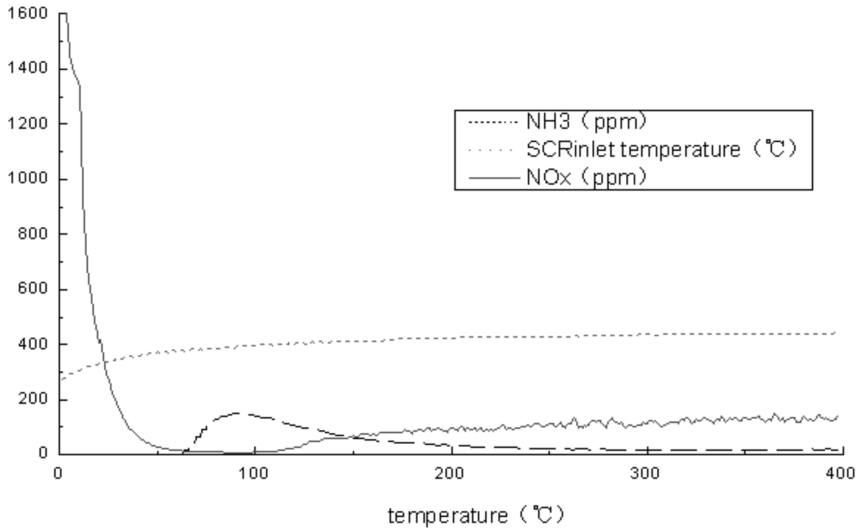

Fig. 3. NO_x, NH_3 and SCR upstream temperature variations vs. time

As can be seen in Fig. 3, after start of urea injection, NO_x first decreases, then begins to rise after reaching a low point, to be gradually stabilized after a period of time; ammonia leakage first gradually increases, and begins to decrease after a peak to be stabilized after a period of time. This is because in initial injection, temperature of catalyst inside SCR is low, catalyst activity is weak while ammonia storage capacity is strong, catalyst saturated ammonia storage capacity is large, then a large amount of NH_3 is adsorbed on active catalyst sites [4]. As the temperature increases, catalyst activity gradually improves, with NO_x conversion efficiency constantly improved and NO_x emissions decreased. Since the temperature rises rapidly, when NH_3 is massively released at low temperature, it reacts with NO_x, so NO_x emissions decreases rapidly. When released NH_3 cannot be completely consumed, ammonia leakage is caused. As the temperature continues to rise, catalyst activity continues to improve, ammonia storage capacity gradually weakens, with increasingly less storage of NH_3. NO_x rises slowly, and when the time reaches about 400 s, NO_x emission is stabilized.

Define that NO_x concentration corresponding to maximum theoretical conversion efficiency is A, then

$$A = A_0 - A_1 \,. \tag{5}$$

Here, A_0 represents the original machine NO_x emission concentration at oper-

ating point, A_1 represents theoretically consumed NO_x volume concentration at ammonia injection complete reaction at the operating point(NH_3 and NO_x are calculated in 1: 1 ratio). According to calculation of actual urea injection quantity and mass of exhaust gas, NO_x concentration corresponding to point A is 107 ppm. As can be seen in Fig. 3, at about 230 s, NO_x concentration reaches 107 ppm, indicating that ammonia stored at low temperature before the time is involved in the reaction, but the time can be approximately considered as the moment when ammonia storage is almost exhausted.

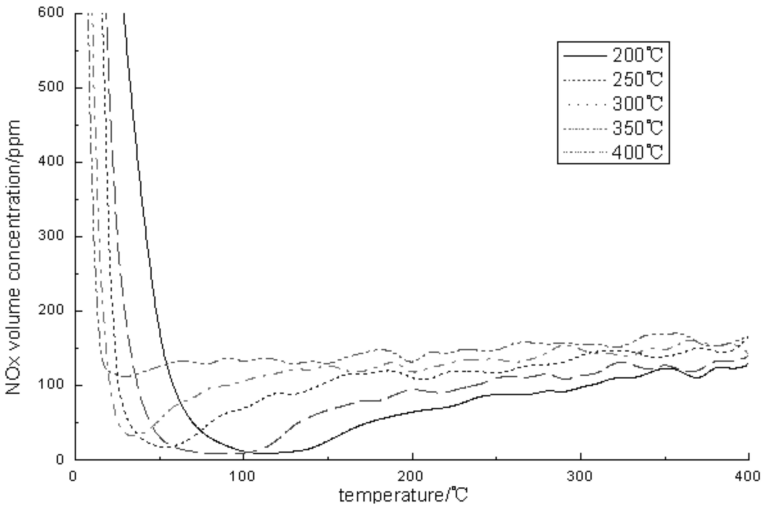

Fig. 4. NO_x volume concentration variation under different initial temperatures vs. time

As can be seen from Fig. 4 and Fig. 5, the lower is initial temperature at operating point, the lower is minimum NO_x value, the longer time is required to reach the lowest point and the higher is ammonia leakage peak. This is because the lower is temperature at the operating point, the stronger is catalyst ammonia storage capacity, the longer is the storage time and the greater is ammonia storage amount. As the temperature rises, more NH_3 can be released, so minimum value of NO_x emissions is lower, corresponding leakage of ammonia will be higher, and the time needed for increasing NO_x to reach equilibrium will be longer.

As can be seen from Fig. 4, when the temperature is 200 °C, the time for NO_x emissions to reach the lowest point is longest, about 130 s. Also, the time for ammonia storage exhaust is longest, about 310 s; at 400 °C, after circulation, the time required for NO_x to touch the bottom is the shortest, about 30 s, indicating that there is basically no ammonia storage at 400 °C,.

Each operating point emission result recorded by ESC cycle is the average of measured values after the point runs 30 s (what is shown is about 90–120 s region). As can be seen from Fig. 4, under the condition that urea injection amount is equal, different cycle temperatures will lead to quite different collected NO_x emission results. To ensure consistency of emission results, temperature into circulation should

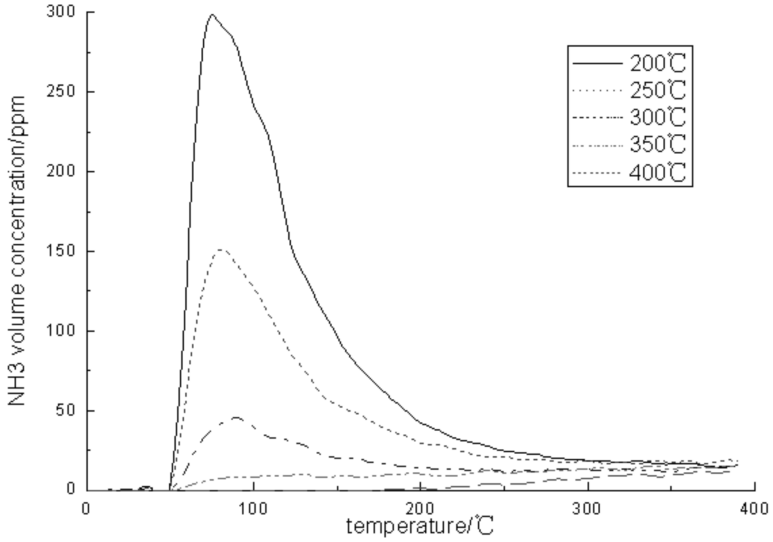

Fig. 5. NH_3 volume concentration variation under different initial temperatures vs. time

be consistent as far as possible.

Figures 6 and 7 show over-time variation of NO_x and NH_3 volume concentration corresponding to urea injection in accordance with different ammonia nitrogen ratios under the same conditions after switched to A100 operating point. As shown in the Figure, with increase of ammonia nitrogen ratio, minimum value that NO_x can achieve is lower, and NO_x duration seen from the chart in lower emission region is also longer. This is because the larger ammonia nitrogen ratio, the more ammonia stored in low-temperature period after spray, the greater ammonia emissions with increasing temperature, and the stronger NO_x buffering effect. But when ammonia nitrogen ratio reaches 1, NO_x emission reaches the lowest value almost without further rise. This is because urea injection at this time is large, and all the NO_x can be completely consumed in theory.

5. Conclusion

1. The lower temperature after starting driving cycle, the lower minimum value of NO_x that can be achieved, the longer time needed to reach the lowest point, the longer time for exhaust of ammonia storage, and the higher ammonia leakage peak. Therefore, to ensure consistency of the test results, stabilization of hot car time and other measures should be taken to ensure consistency of each time entering test cycle temperature.

2. The higher ammonia nitrogen ratio, the lower minimum value of NO_x that can be achieved, the stronger buffering effect of ammonia storage on NO_x, and the

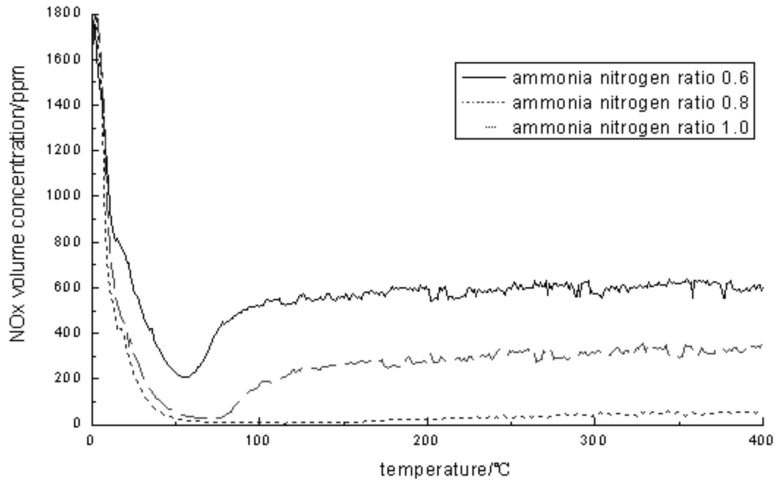

Fig. 6. NO_x volume concentration variation vs. time under different ammonia nitrogen ratios

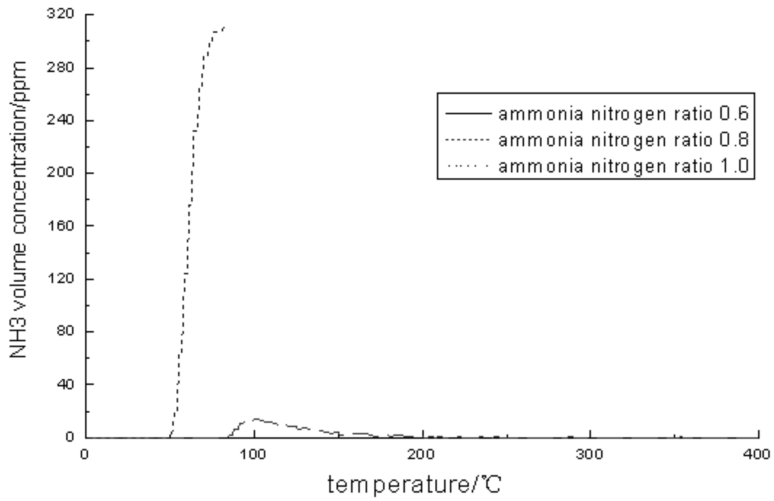

Fig. 7. NH_3 volume concentration variation vs. time under different ammonia nitrogen ratios

longer time for emission results to be stabilized. Therefore, when ammonia nitrogen ratio is relatively high, operating point entry should be maintained at a higher temperature to reduce NH_3 overflow.

References

[1] D. L. Donoho: *De-noising by soft-thresholding.* Trans. Info. Theory *41* (1995), No. 3, 613–627.

[2] R. B. W. Heng, M. J. M. Nor: *Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition.* Applied Acoustics *53* (1998), Nos. 1–3, 211–226.

[3] R. F. M. Marcal, M. Negreiros, A. A. Susin, J. L. Kovaleski: *Detecting faults in rotating machines.* IEEE Instr. Meas. Magazine *3* (2000), No. 4, 24–26.

[4] F. A. Andrade, M. N. M. Badi, I. Esat: *A new approach to time-domain vibration condition monitoring: gear tooth fatigue crack detection and identification by the Kolmogorov–Smirnov test.* J Sound Vibr. *240* (2001), No. 5, 909–919.

[5] L. Peretto, R. Sasdelli, R. Tinarelli: *On uncertainty in wavelet-based signal analysis.* IEEE Trans. Instrum. Measurement *54* (2005), No. 4, 1593–1599.

[6] Munoz-Cobo, C. Perez *Hilbert-Huang analysis of BWR neutron detector signals: application to DR calculation and to corrupted signal analysis.* Annals Nucl. Energy *30* (2003) No. 6, 715–727.

[7] B. Liu, S. F. Ling: *On the selection of informative wavelets for machinery diagnosis.* Mech. Systems Signal Proc. *13* (1999), No. 1, 145–162.

[8] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, R. O. Wells: *Noise reduction using an undecimated discrete wavelet transform.* IEEE Signal Proc. Letters *3* (1996), No. 1, 10–12.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 37–46

A comparative study of virtual power plant operation optimization strategy[1]

Zejing Qiu[2], Jie Xiang[2], Zhao Shuang[3], Qiqi Qian[2], Chunxue Li[2], Ming Zeng[3]

Abstract. A strategy of a virtual power plant is proposed, consisting of a cogeneration system and a solar power system. First, an optimization model is built based on day-ahead market and real-time optimization. Then, different optimal operation strategies of real-time optimization and reference situation are compared, which is illustrated with an examples.

Key words. Virtual power plant, operation optimization strategy, total operation cost, imbalance error reduction.

1. Introduction

In recent years, the rapid development of world's economy results in an increasing demands on energy [1-2]. However, with increasingly serious energy shortage and environmental issues, people started to realize that the economy development relied on fossil fuels is unsustainable [3]. How to improve energy utilization efficiency and develop renewable energy has becoming a problem urgently waiting to be solved in the process of energy development [4]. Renewable energy power generation technology mainly includes wind power, photovoltaic power generation, biomass power, tidal power, etc. In the existing renewable power generation technology, wind power and photovoltaic power generation are the most mature and most widely used technologies. However, the output of renewable energy is uncontrollable, which exhibits a great deal of volatility and uncertainty. And it will influence the safe and stable operation of power grid in the process of grid connection, and this also becomes the

[1]This work is supported in part by technology project of State Grid Corporation of China (title: Key technology and operation mode of efficiency power plant with typical high energy consumption customers' research).

[2]Nari (Wuhan) Electrical Equipment & Engineering Efficiency Evaluation Center, Wuhan, 430074, China

[3]School of Economics and Management, North China Electric Power University, Beijing, 102206, China

obstacle in the way of large-scale renewable energy development [5-7].

In order to solve the above problem and increase the renewable energy generation, now we can use virtual power plant (VPP) combined by different distributed generation technology to compensate the output volatility of renewable energy. The output instability of renewable energy can be covered by other distributed generation system, thus the energy efficiency of the whole system will be improved, and we can achieve the purpose of energy saving and emission reduction. Nowadays, there exists no clear definition of VPP, and one of the typical concept follows: VPP is an organic combination of traditional power plant, distributed power, controllable load and energy storage system in a certain area, participated in the power grid as a whole system through a control center management. The typical structure of a VPP is shown in Fig. 1.

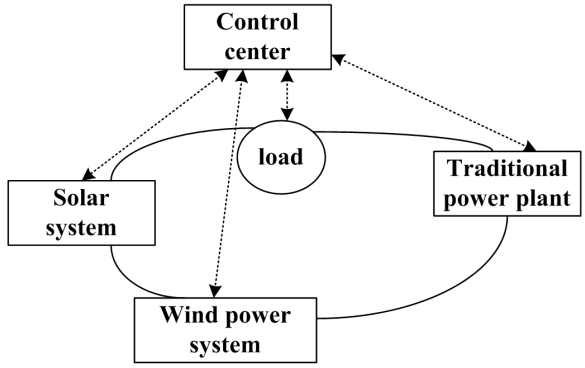

Fig. 1. Typical structure of VPP

The figure above typically contains tow networks, which are energy network and information network. The solid line represents energy network, that consists of the electricity transmission network, connected with the solar system, wind power plant and traditional power plant. The dotted line represents the two-way communication between the control center and other cells in VPP.

At present there is no large-scale application of virtual power plant, and the relevant research is still in its infancy. There are many researches related to virtual power plant, ranging from concept to the joint optimal operation. Ruiz et al. [8] provide an optimization algorithm to manage a VPP composed of a large number of customers with thermostatically controlled appliances. Masuta et al. [9] evaluate the frequency and trend of outages and surpluses of power due to the forecast error of PV power generation by numerical simulations conducted using the power system model of Kanto area in Japan. Eto et al. [10] propose a method deciding a schedule of node movement to cover the target agricultural field from plant to harvest by the minimal number of nodes in order to reduce the node deployment cost. Pandžić et al. [11] studied the operation optimization of a virtual power plant including wind power generation, energy storage system and conventional power plants. El-Sayed and Obara [12] propose a prediction algorithm based on a neural network (NN) to

predict the electricity production from a solar cell. Also, Pandžić et al. [13] consider a weekly self-scheduling of a virtual power plant composed of intermittent renewable sources, storage system and a conventional power plant. Ajabshifizadeh et al. [14] consider the predictions of both the amplitude and timing of the next solar cycle will assist in estimating the various consequences of Space Weather.

The structure of this paper is as follows: in section I we introduce the background of virtual power plant, while in section II the optimization models are built up. In section III, we use a case to verify the validity of the model. In section IV, the article is summarized.

2. Material and methods

2.1. Formulation of the problem

Considering the uncertainty of renewable energy output, in order to simplify the model. In this paper, we study the optimal operation of a virtual power plant consisting of a cogeneration system and distributed solar power system. The cogeneration system includes a generator, a waste heat utilization device and an auxiliary boiler, which is used to meet the electricity and heat demands of the users and smooth the solar output.

In order to study the optimal operation strategy of virtual power plant, we will build models based on day-ahead and real-time scenarios, respectively.

2.2. Day-ahead optimization algorithm

The objective function of day-ahead optimization aims at the minimum of the operation cost of the whole system, and may be written in the form

$$\operatorname{Min}\sum_{t=1}^{T}(C-S)=\operatorname{Min}\sum_{t=1}^{T}\left[(fc_{\mathrm{CHP}}+fc_{\mathrm{boiler}})-(r_{\mathrm{sold}}+r_{\mathrm{con}})\right], \tag{1}$$

where fc_{CHP} and fc_{boiler} represent the fuel costs of the CHP system and boiler, respectively, r_{sold} represents the revenue from selling electricity to the grid, and r_{con} represents the saving electricity expense.

The relationship between the fuel consumption and electricity output of the CHP system is given by the equation

$$Q_{\mathrm{CHP}}=\alpha\cdot\sum_{t=1}^{T}\beta(t)\cdot E_{\mathrm{CHP}}(t), \tag{2}$$

in which α is the fuel consumption rate related to the type of equipment and operation environment, β is the running state variable of the CHP system: $\beta=0$ represents its shut down and $\beta=1$ represents on.

In this paper, we assume that the thermal efficiency of boiler γ is constant, thus,

the fuel consumed by boiler can be calculated as

$$Q_{\text{boiler}} = \frac{\sum_{t=1}^{T} H_{\text{boiler}}(t)}{\gamma \cdot Q}\,, \tag{3}$$

where Q is the calorific value of natural gas.

Therefore, the total fuel cost of whole system can be obtained by the total fuel cost multiplied with the fuel cost p_{fuel}, i.e.

$$fc_{\text{CHP}} + fc_{\text{boiler}} = (Q_{CHP} + Q_{\text{boiler}}) \cdot p_{\text{fuel}}\,. \tag{4}$$

The total production of electricity E_{CHP} consists of the part generated by CHP system used to satisfy the users' demand E_{con}, and the excess part E_{sold} that can be sold to the grid,

$$E_{\text{CHP}} = E_{\text{con}} + E_{\text{sold}}\,. \tag{5}$$

Therefore, the revenues can be obtained by adding the power usage multiplied with corresponding price

$$r_{sold} + r_{con} = p_{\text{sold}} \cdot E_{\text{sold}} + p_{\text{con}} \cdot E_{\text{con}}\,. \tag{6}$$

In the process of optimization, the system mainly has two aspects of constraints, that is, operation constraints and technical constraints.

The operation constraints ensuring the heat demands of the users can always be satisfied,

$$D_{\text{heat}}(t) = H_{\text{CHP}}(t) + H_{\text{boiler}}(t) + H_{\text{storage}}(t)\,, \tag{7}$$

where H_{CHP} and H_{boiler} represent the heat generated by the CHP system and the boiler, respectively. Symbol H_{storage} is the heat obtained from the heat storage device.

The status of the heat storage device is calculated on the condition that the efficiency η is constant, and its value is 90% in this paper.

$$H_{\text{storage}}(t) = \eta H_{\text{storage}}(t-1) - H_{\text{storage}}(t) \cdot \triangle t\,, \tag{8}$$

where $\triangle t$ is the time interval.

The following technical constraints are used to ensure the safe and stable operation of devices,

$$0 \le H_{\text{storage}} \le H_{\text{s\_ max}}\,, \tag{9}$$

$$0 \le H_{\text{boiler}} \le H_{b\_ \text{max}}\,, \tag{10}$$

$$H_{\text{CHP\_ min}} \le H_{\text{CHP}}(t) \le H_{\text{CHP\_ max}}\,, \tag{11}$$

$$E_{\text{CHP\_ min}} \le E_{\text{CHP}}(t) \le E_{\text{CHP\_ max}}\,. \tag{12}$$

2.3. Real-time optimization algorithm (Strategy 1)

The imbalance of the system mainly comes from the gap between the electric forecast and actual outputs. In this paper, the imbalance of the VPP system is composed from two parts. One of them is the solar output prediction error and the other is the output error of CHP system due to the demand changes.

The optimization objective of Strategy 1 is to minimize the total imbalance error. The economic problems are not considered here and the objective function is given as

$$\mathrm{Min}\sum_{t=1}^{T}[E_{\mathrm{real}}(t)-E_{\mathrm{forecast}}(t)]^2=\mathrm{Min}\sum_{t=1}^{T}[E_{\mathrm{error}}(t)]^2\,, \tag{13}$$

where R_{real} and E_{forecast} represent the actual and forecast power outputs of the VPP.

Although the objective functions are different, the operation and technology constraints are the same, so here we do not discuss the same constraints again. While there still is a unique constraint, the predict power output of VPP has to be satisfied by the actual power output and electricity E from auxiliary market

$$E_{\mathrm{forecast}}(t)=E_{\mathrm{real}}(t)+E(t)\,. \tag{14}$$

2.4. Real-time optimization algorithm (Strategy 2)

The optimization objective of Strategy 2 is to minimize the total operation cost of the whole system, including the fuel cost of CHP system, boiler, and expense for covering the imbalance error

$$E_{\mathrm{error}}=\mathrm{Min}\sum_{t=1}^{T}(fc_{\mathrm{CHP}}+fc_{\mathrm{boiler}}+c_{\mathrm{error}})\,. \tag{15}$$

As we can see from the equation above, when $E_{\mathrm{error}}>0$, the VPP system can sell extra electricity and get revenues. In order to avoid opportunistic phenomenon, we set a constraint that the total imbalance error cannot excess the imbalance error caused by the solar instability, that is,

$$\sum_{t=1}^{T}E_{\mathrm{error}}(t)\le\sum_{t=1}^{T}E_{\mathrm{error\_PV}}(t)\,. \tag{16}$$

2.5. Prediction of solar output

The output of the solar system depends largely on the real-time light intensity. Figure 2 shows the solar system output curve under different weather conditions in summer in China, and we can see that the output of solar system has a great deal of uncertainty and unpredictability. But with the permanent development of prediction technology, it will be more and more accurate and reliable.

According to the difference of the methods, the existing prediction methods

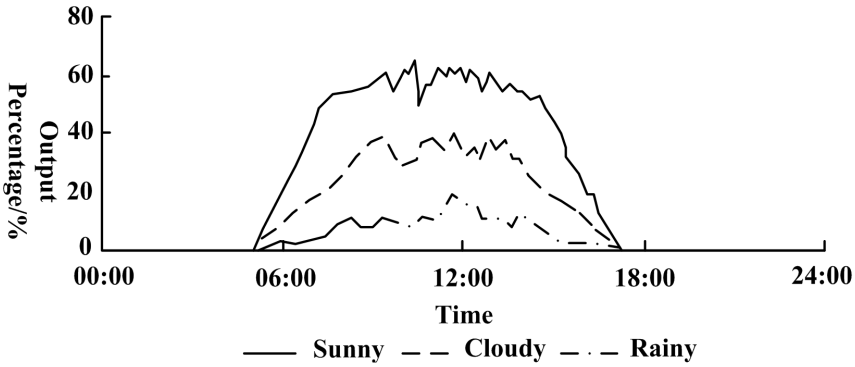

Fig. 2. Output percentage of PV under different weather conditions in summer

mainly can be divided into two groups: statistical methods and physical methods. The former one using the existing physical model obtains the prediction value by input the meteorological forecast data directly. The latter one is based on some statistical methods, analysis of historical data in order to find the internal rule by establishing the relationship between numerical weather prediction and historical output data to predict, mainly including support vector machine (SVM), grey prediction, regression model and so on.

With the continuous development of prediction technology, the accuracy of photovoltaic output prediction in the future will continue to improve, and this will lead to the improvement of the utilization of solar system.

3. Results

In this section, we use the actual output of VPP system without reschedule strategy as reference, compared the two real-time strategies by different indexes. The first index is the capacity for reducing the imbalance error, namely the difference with the day-ahead optimization. The second index is the total operation cost.

3.1. Capacity to reduce imbalance error

Figures 3 and 4 show the output of different scenarios under typical winter and summer days.

In the figures above, the black solid line represents the electric output in day-ahead optimization, and the shadow part represents the output in three different scenarios. As we can see from the preamble, Strategy 1 provides the maximum reduction of imbalance error in theory. We can also draw the conclusion that no matter what the optimization objective is, using optimization strategy can reduce the imbalance error efficiently. However, in Fig. 4 we can find that when using Strategy 2, there still remains larger imbalance error compared to the reference,

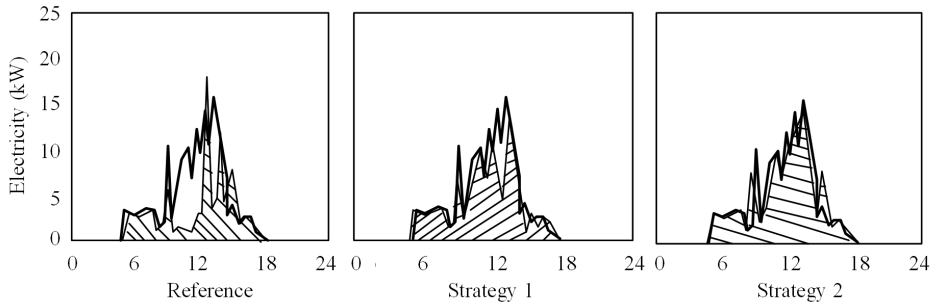

Fig. 3. Electricity outputs of different scenarios in typical winter day

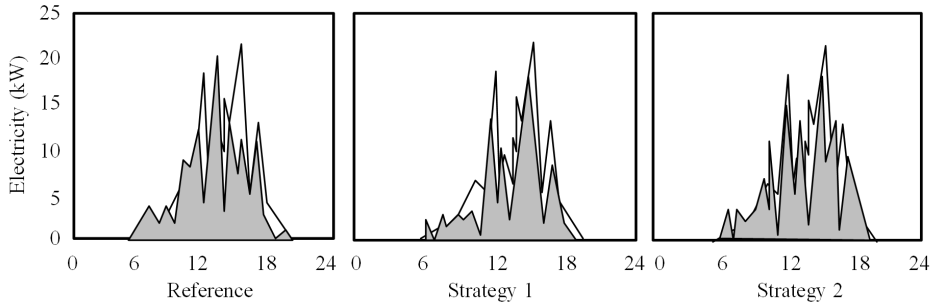

Fig. 4. Electricity outputs of different scenarios in typical summer day

which means that reducing imbalance error in this situation is uneconomical, and the latter analysis will also show the same conclusion.

Figures 5 and 6 reveal the remaining imbalance error in typical winter, summer and general situations. We can easily find that the imbalance error in summer is larger than that in winter, and the positive imbalance error is particularly significant.

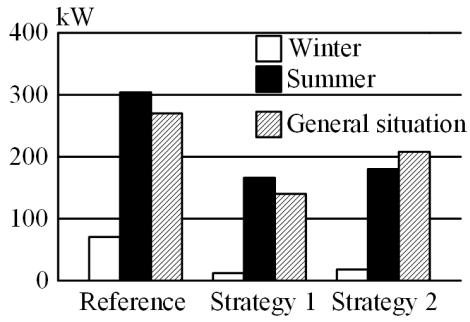

Fig. 5. Positive imbalances in winter, summer and in general situation

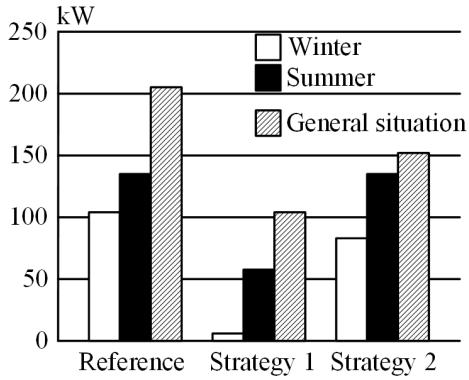

Fig. 6. Negative imbalances in winter, summer and in general situation

3.2. Total operation cost

The total operation cost of VPP system mainly consist of the fuel cost of CHP system, solar power system and the cost for reducing imbalance error (see Fig. 7).

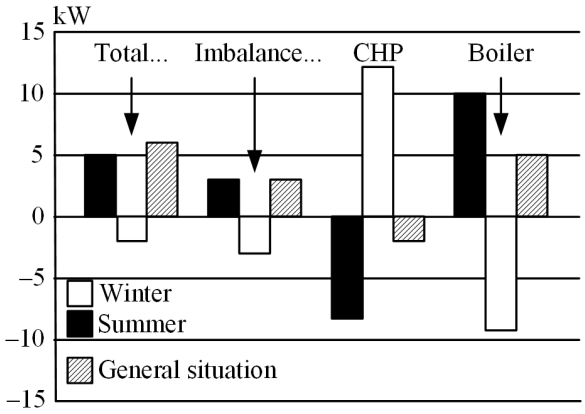

Fig. 7. Cost comparison between "Strategy 1" and "Reference" (yuan/week)

Compared with the reference, the fuel cost of CHP system in winter increases remarkable. This is because the illumination intensity in winter decreases, thus the CHP system has to compensate the shortage of solar generation. Meanwhile, the fuel cost of boiler is reduced, also because of the excessive output of the CHP system (see Fig. 8).

Combined with the above figure we can see that there is a larger positive imbalance error in summer and in general situation, so that the CHP system reduces the output and the fuel cost is lower as well. But this also leads to a great increase of the cost of reducing imbalance error, which proves to be non-economical, as also

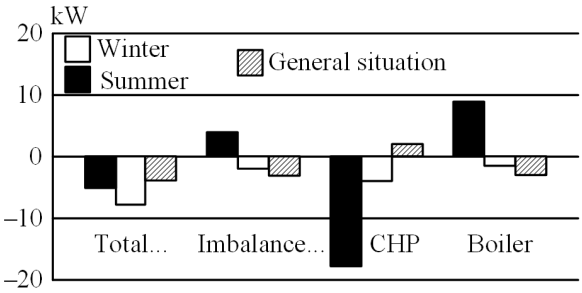

Fig. 8. Cost comparison between "Strategy 2" and "Reference" (yuan/week)

shown above.

4. Conclusion

In this paper, we designed optimized operation models of VPP system under different scenarios, and then compared the different strategies by factual data. As the results showed, no matter whether we want to maximize the imbalance error reduction or minimize the total operation cost, compared to the reference, the imbalance errors are largely reduced by reschedule. However, the results also tell us that the strategy used for reducing the imbalance errors is not always economical, especially in summer.

The application of virtual power plant can effectively improve the utilization efficiency of renewable energy under the condition of ensuring the safe and stable operation of power system. In the meantime, as one of the ways of demand side management, the use of virtual power plants has a great significance in strengthening the demand side management, and realizing the energy saving. While the development and application of virtual power plant still exhibit certain problems, it needs cooperation among technology, management and policy.

Limited by space and the depth of research, the research in this paper about the optimization operation of virtual power plant still has many deficiencies. So, in the further study about VPP, we can carry on the researches on system capacity configuration, price mechanism, the related policy and so on. In order to build a resource-conserving and environment-friendly society, the application of VPP will be surely wider and wider.

References

[1] A. Arnette, C. W. Zobel: *An optimization model for regional renewable energy development.* Renewable & Sustainable Energy Reviews *16* (2012), No. 7, 4606–4615.

[2] P. D. Wright: *Local aspects of UK renewable energy development: Exploring public beliefs and policy implications.* Local Environment *10* (2005), No. 1, 57–69.

[3] X. Tan, B. Rose: *Renewable energy-the path to sustainability.* Ecological Economy *4* (2008), No. 1, 15–23.

[4] A. Bhattacharyya: *A strategic policy for sustainable development.* Renewable Energy, Policy Brief *10* 2010, 1–7.

[5] Ž. B. Rejc, M. Čepin: *Estimating the additional operating reserve in power systems with installed renewable energy sources.* IJ Electrical Power and Energy Systems *62* (2014), No. 2, 654–664.

[6] C. Binding, D. Gantenbein, B. Jansen, O. Sundström, P. B. Andersen, F. Marra, B. Poulsen, Ch. Træholt: *Electric vehicle fleet integration in the Danish EDISON project - A virtual power plant on the island of Bornholm.* Proc. Power and Energy Society General Meeting IEEE. 25-59 July 2010, Minneapolis, MN, USA, 1–8.

[7] D. Pudjianto, C. Ramsay, G. Strbac: *Virtual power plant and system integration of distributed energy resources.* IET Renewable Power Generation *1* (2007), No. 1, 10–16.

[8] N. Ruiz, I. Cobelo, J. Oyarzabal: *A direct load control model for virtual power plant management.* IEEE Trans. Power Systems *24* (2009), No. 2, 959–966.

[9] T. Masuta, T. Oozeki, J. G. da Silva Fonseca Junior, A. Murata: *Evaluation of outage and surplus power due to forecast error of photovoltaic generation output.* IEEJ Trans. Power & Energy *134* (2014), No. 4, 286–295.

[10] M. Eto, R. Katsuma, M. Tamai, K. Yasumoto: *Efficient coverage of agricultural field with mobile sensors by predicting solar power generation.* Proc. IEEE Advanced Information Networking and Applications (AINA), 24–27 March 2015, Guwangiu, South Korea, 62–69.

[11] H. Pandžić, J. M. Morales, A. J. Conejo, I. Kuzle: *Offering model for a virtual power plant based on stochastic programming.* Applied Energy *105* (2013), No. 5, 282 to 292.

[12] A. G. El-Sayed, S. Obara: *Energy supply characteristics of a combined solar cell and diesel engine system with a prediction algorithm for solar power generation.* J Power & Energy Systems *4* (2010), No. 1, 27–38.

[13] H. Pandžić, I. Kuzle, T. Capuder: *Virtual power plant mid-term dispatch optimization.* Applied Energy *101* (2013), No. 1, 134–141.

[14] A. Ajabshirizadeh, N. M. Jouzdani, S. Abbassi: *Neural network prediction of solar cycle 24.* Research in Astronomy and Astrophysics *4* (2011), 491–496.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 47–56

Evaluation and optimization for virtual power plants under different operational strategies[1]

Zejing Qiu[2], Jie Xiang[2], Qiqi Qian[2], Chunxue Li[2], Ming Zeng[3]

Abstract. A Reliability and Operation Model (ROM) and unit commitment model are used for modeling and accessing virtual power plants under self-supply maximization and market profits maximization for two different operational strategies. Furthermore, a numerical example is presented to verify the scientific and validity of the model and provide certain decision support for the operation of virtual power plants.

Key words. Virtual power plant, operational strategies, ROM model.

1. Introduction

Virtual power plants, which are considered as a way of demand response, Virtual power plants which are considered as a way of demand response effectively combine distributed generators, controllable load and a variety of distributed energy storage device together, and by the coordinated regulation technology and communications technology to integrate and module all types of distributed energy [1]. And they have great significance to balance the power supply and demand, ensure the economy and security of the operation of power system, and help alleviate energy shortage problem and environmental deterioration problem.

Current research on virtual power plants has exhibited continuous increase. Zdrilić et al. [2] maximize the profit of the virtual power plant by a mixed-integer linear

[1]This work is supported in part by technology project of State Grid Corporation of China (title: Key technology and operation mode of efficiency power plant with typical high energy consumption customers' research).

[2]Nari (Wuhan) Electrical Equipment & Engineering Efficiency Evaluation Center, Wuhan, 430074, China

[3]School of Economics and Management, North China Electric Power University, Beijing, 102206, China

programming model which incorporates long-term bilateral contracts with weekly forecast hourly market prices. Shabanzadeh et al. [3] introduce an efficient MILP model based on robust optimization approach proposed to enable informed decision making under different levels of uncertainty. To manage a VPP, Ruiz et al. [4] provide an optimization algorithm composed of a large number of customers with thermostatically controlled appliances, and Kieny et al. [5] consider two types of aggregation EU project FENIX: the commercial virtual power plant (CVPP) that tackles the aggregation of small generating units with respect to market integration and technical virtual power plant (TVPP) that tackles aggregation of these units with respect to services that can be offered to the grid. Kuzle et al. [6] and Sučić et al. [7] consider the dispatch of VPP and come up with different programming model to overcome it. Tai-Her et al. [8] present a novel approach based on Weibull distribution to determine the capacity of wind turbine generators using capacity factor, normalized average power and product of CF and Punder different values of tower height and rated wind speed. Kuntschke et al. [9] present a fundamental system architecture and a concept enabling VPPs and DSOs to negotiate their positions to satisfy the needs and requirements of both sides. Robu et al. [10] design a payment mechanism that encourages distributed energy resources to join "cooperative" VPPs with large overall production. In ref. [11] by Caldon et al., an optimization algorithm is proposed to integrate many DG into a VPP, which will be able to generate and sell both thermal and electrical energy. Vale et al. [12] present a multi-level negotiation mechanism for Smart Grids optimal operation and negotiation in the electricity markets, considering the advantages of VPPs' management.

The main structure of this paper is as follows: Section 2 establishes the Reliability and Operation Model (ROM) and unit commitment model to evaluate virtual power plants on the basis of maximizing self-supply and market profits two different strategies. Section 3 selects one Chinese province as a numerical example to verify the scientific and validity of the model established before. Finally, analyzing and summarizing the influence of virtual power plants is performed on the power system under different operating strategies.

2. Model construction

In this paper, we use ROM model to simulate the overall operation of power system. ROM model was proposed at Comillas Pontifical University and is widely used for scheduling optimization in day-ahead power system. Meanwhile it is subject to restrictions related to units' operating conditions, including: unit output's upper and lower limits, units' ramp constraints, availability of intermittent units and regulatory reserve requirements. When some unexpected events happen, for example, the power units become faulty or units are deviated from the preset level, the day ahead optimal scheduling results should be initialized. Meanwhile, the target year for this study is 2020, thus before 2020, the daily operation of power system all need to be solved.

2.1. Unit commitment model

In ROM model, we use a unit commitment model to simulate the daily operation of power system before the target year. The specific objective function has the form

$$\min C_{\text{total}} = \sum_{i,t}(C_i^{\text{fd}} \cdot \phi_{i,t} + C_i^{\text{vb}} \cdot P_{i,t}^{\min} \cdot \phi_{i,t} + C_i^{\text{vb}} \cdot P_{i,t}) + \sum_t (C_{\text{ne}} \cdot N_t)\,, \tag{1}$$

where C_{total} are the total variable thermal costs, C_i^{fd} are the fixed costs, C_i^{vb} are the variable costs, C_i^{st} are the start-up costs, and C_{ne} are the unit costs of non-supplied energy. The variable $\phi_{i,t}$ represents the unit commitment results of thermal power unit i in the period t. Symbol $P_{i,t}^{\min}$ denotes the lower limit of the unit output, $P_{i,t}$ stands for the power output over the smallest output, and finally, N_t stands for non-supplied energy quantity in period t. As for the quantities $\phi_{i,t}$, they are binary variables, and their value is 0 if the thermal power plant i is shut down in period t.

Equation (1) is used to calculate the daily operating costs. Meanwhile, the values of relevant variables in the previous day are used as the initial values to calculate the following day's operating costs. Furthermore, the annual operating costs are cumulative results of daily operating costs.

This paper takes into account a series of technical restrictions when establishing the model, which are shown in the following relations.

$$D_t - D_t^{\text{gen}} - N_t = \sum_i P_{i,t}^{\min} \cdot \phi_{i,t} + P_{i,t}\,, \tag{2}$$

$$\sum_i (P_{i,t}^{\max} - P_{i,t}^{\min}) \cdot \phi_{i,t} - P_{i,t} \geq I_t^{\text{up}}\,, \tag{3}$$

$$\sum_i P_{i,t} \geq I_t^{\text{down}}\,, \tag{4}$$

$$P_{i,t} \leq (P_{i,t}^{\max} - P_{i,t}^{\min}) \cdot \phi_{i,t}\,, \tag{5}$$

$$P_{i,t} - P_{i,t-1} \leq R_i^{\text{up}}\,, \tag{6}$$

$$P_{i,t-1} - P_{i,t} \leq R_i^{\text{down}}\,, \tag{7}$$

$$\phi_{i,t} - \begin{pmatrix} U_i, & p=1 \\ \phi_{i,t-1} & 1 \end{pmatrix} = \lambda_{i,t} - \mu_{i,t}\,. \tag{8}$$

Here, D_t is the customers' power demand, D_t^{gen} is the output of the distribution network, N_t is the quantity of non-supplied energy, I_t^{up} and I_t^{down} stand for the upper and lower limits of regulatory reserve requirements, respectively, $P_{i,t}^{\max}$ is the largest output of the single unit i, R_i^{up} and R_i^{down} represent the upper and lower unit

ramp rate, respectively, while U_i stands for the initial state of the units. Finally, $\lambda_{i,t}$ is the parameter for start-up of units in period t, $\mu_{i,t}$ is the decision-making parameter for shut-down of units in period t. Both these parameters are binary variables and their values are either 0 or 1; $\lambda_{i,t} = 0$ if thermal power plant i is shut down in period t, while $\mu_{i,t} = 1$ in the same case, and vice versa.

Equations (3) and (4) show the regulatory reserve requirements, equations (5) and (6) show the power output limits and equation (8) ensures the logical sequence of starting up, unit commitment and shutting down.

2.2. Elasticity demands of virtual power plants

In this paper, we use the demand variable d^t to describe the constraints related to energy conservation, and then consider both the original demand D^t and the demand variable d^t to obtain the final power demand. The demand variable d^t is given as

$$d_t = D_t + v_t^{\text{up}} - v_t^{\text{down}} \tag{9}$$

where v_t^{up} and v_t^{down} stand for the upper and lower limits of demand response, respectively.

In this paper, we use the changes of demand as the only available demand response, so that during system's routine operation the increasing and decreasing amounts of power demand are balanced. This means that the two amounts of power demand during daily operation are equal, and the constraint can be expressed as follows

$$\sum_t v_t^{\text{up}} = \sum_t v_t^{\text{down}} . \tag{10}$$

Logically speaking, for the elasticity of demand some constraints need to be met, and the variable of load must be limited between values ε_{up} and $\varepsilon_{\text{down}}$ denoting certain percentages of the total demand.

$$\sum_t \varepsilon_{\text{up}} \cdot D_t \geq \sum_t v_t^{\text{up}} , \tag{11}$$

$$\sum_t \varepsilon_{\text{down}} \cdot D_t \geq \sum_t v_t^{\text{down}} . \tag{12}$$

2.3. Power supply and demand in virtual power plants

We assume that in virtual power plants, about 10% is the distributed power generation system. Then we need to estimate the power demand in virtual power plants. The specific process is described by equations

$$\min \sum_t (G_t - \alpha D_t) . \tag{13}$$

As

$$0 < \alpha < 1 , \tag{14}$$

$$\sum_t G_t - \alpha D_t > 0\,. \tag{15}$$

Here G_t is the generated energy of virtual power plants in period t and α is a scale parameter.

2.4. Operation strategies of virtual power plants

2.4.1. Self-supply strategy. Self-supply strategy is related to virtual power plants' internal supply and demand. With the help of the model stated above, we can simulate virtual power plants' operation process which can be divided into two phases. In phase one, due to virtual power plants having the ability of self-supply to some extent, the possibility of purchase and sale power at this stage does not need to be considered. Based on this, it needs to minimize the power supply costs of local load, including the cost of non-supplied energy in virtual power plants. Meanwhile, in this phase, non-supplied energy in virtual power plants is avoided by a penalty factor, and all the local power generation has the priority to meet the internal power demand of virtual power plants, details is shown in equation (9) to (12). In addition, in this phase, the power demand which is met by self-supply of virtual power plants needs to be estimated, and when the power demand is higher than the corresponding power supply, it is necessary to go for the next phase, and that means meeting the power demand by purchasing power from the market.

Additionally, in phase two, we allow virtual power plants to purchase and sale power in the power market, and by means of equation (1)–(11) to minimize the operating costs of power system. At the same time, we need to ensure that the power purchased from the market in phase two less than the shortage power which cannot be met by self-supply in phase one and then the elasticity of demand used to minimize system costs in virtual power plants can be guaranteed.

2.4.2. Market profits maximization strategy. Virtual power plants mainly maximize the market profits by purchase power in a low price when power is needed and sale power in a high price when power is surplus in virtual power plants. At the same time, the elasticity of power demand in virtual power plants also helps to increase the market profits to some extent.

Similarly to self-supply strategy of virtual power plants, the implementation and simulation of market profits maximization is also divided into two phases, while the market profits of daily operation before the target year are all required to be calculated, and then accumulated to obtain the total market profits. Based on the variable of power demand in virtual power plants, the calculation of daily market profits is divided into two phases: in the first phase, we assume that the market price is known, and virtual power plants are designed to maximize their market profits, and then on the basis of this to calculate the demand variable. The second phase, we must iteratively ensure the consistency between demand variable calculated in the first phase and the market price. When the differences of two successive iterations of the market price fall below a certain threshold, we consider that this iterative

process is converged.

During each iteration, the obtained solutions of virtual power plants' market profits all need to be compared with the solution in the former iteration, and if the solution is lower than the former solution, it is proved that the solution is not the optimal solution, and iteration is needed to obtain the optimal solution; meanwhile, if the solution is not lower than the former solution, then we stop the iterative process and get the optimal solution. If the maximum number of iterations is reached, we also stop iteration and recognize the obtained solution as the final optimal solution. Through the continuous iterations we can progressively approach to the optimal solution, until we finally find the optimal solution. Specific solving process is shown in the equation

$$\max M_j = \sum_t -D_{t,j}^{\text{net}} \cdot [Q_{t,j-1} + \frac{Q_{t,j-1} - Q_{t,j-2}}{D_{t,j-1}^{\text{net}} - D_{t,j-2}^{\text{net}}} \cdot (D_{t,j}^{\text{net}} - D_{t,j-1}^{\text{net}})] , \tag{16}$$

here M_j is the market profit, $D_{t,j}^{\text{net}}$ is the net demand in virtual power plants, and $Q_{t,j-1}$ and $Q_{t,j-2}$ are the reference prices in iteration $j-1$ and $j-2$, respectively.

Meanwhile, when the net demand $D_{t,j-1}^{\text{net}}$ and $D_{t,j-2}^{\text{net}}$ in the previous two iterations are the same, the demand variable is set to zero; in equation (16), variable $D_{t,j}^{\text{net}}$ and $Q_{t,j-1}$ are consistent with the market profits of virtual power plants. If power demand is greater than power supply, that is, the net demand $D_{t,j}^{\text{net}}$ is positive, the surplus power will be sold in the power market at price $Q_{t,j-1}$, which represents a cost of virtual power plants. If power supply is greater than the power demand, that is, the net demand $D_{t,j}^{\text{net}}$ is negative, then the market profits of virtual power plants will increase. The other profit part of the equation is intended to reflect possible influence that the changing demand may have on the market price. In addition, when using this equation to obtain the optimal solution, equation (9) and (10) also need to be considered.

3. Case study

The principal data of the study are summarized in Table 1 and Table 2.

Table 1. Supply and demand data of one Chinese province

Annual demand (TWH)	60.8
Peak demand (GW)	10.1
Nuclear power (GW)	1.4
Thermal power (GW)	7.2
Hydropower (GW)	3.1
Wind Power (GW)	3.6
PV (GW)	1.4
Cogeneration (GW)	0.9
Biomass power (GW)	0.5

Table 1 displays the specific power supply and demand data of one specific province in China. From Table 2 it can be seen that the technical parameters are very de-

tailed, including scheduled outage rate, forced outage rate, minimum and maximum power outputs, ramping rate and other related parameters. In addition, there is some data related to costs and emissions. These are all the input parameters of the model we established before. Meanwhile, this paper assumes that from 2020 the distributed power system will account for 10% of virtual power plants.

Table 2. Average technical parameters of thermal plants in one Chinese province

	Thermal power	Closed-cycle gas turbines with one axis	Closed-cycle gas turbines with more axes	Gas turbines	Nuclear power
Scheduled outage rate (%)	0.05	0.015	0.015	0.023	0.08
Forced outage rate (%)	0.05	0.04	0.05	0.09	0.05
Minimum output (MW)	71	59	124	16	500
Variable heat rate (te/MWh)	1280	760	760	1200	1250
Non-load heat rate (te/h)	29,000	140,000	260,000	60,000	
Fuel costs (EURO/te)	0.17	0.34	0.34	0.34	0.019
Maximum output (MW)	176	196	398	153	500
Upper limit of ramping (MW/h)	51	68	139	65	
CO_2 costs (Euro/t CO_2)	18	18	18	18	
Emissions (t CO_2/MWh)	0.876	0.312	0.312	0.312	
Startup consumption (te/str)	710,000	180,000	390,000	250,000	
Startup time (h)	18	6	6	13	
Number of power plants	15	10	28	6	6

This paper's study of virtual power plants is representative to some extent. So far, the size estimation for the virtual power plant which greatly organizing and utilizing distributed energy barely has been studied except for a European report. This paper demonstrates single virtual power plants, and then expands it to all virtual power plants existing in China's large power system. Therefore, on the basis of virtual power plants' power supply situation, we expand our research based on the above 10% of cases. In addition, we assume that in virtual power plants wind power, photovoltaic power cogeneration account for the same percentage. At the

same time, we assume that the total power consumption of one Chinese province in 2020 is 60.8 Terawatthours. Then we select the demand data in 2011 as the basis, and the average annual growth rate is 2 percent until 2020, so that we can estimate the power consumption in 2020. The final calculation results obtained from the model are shown in Table 3 and Table 4.

Table 3. Solutions under different operation strategies from perspective of virtual power plant

	Self-supply strategy	Profits maximization	Costs minimization
Purchases of power (GW)	313	683	689
Sales of power (GW)	459	752	759
Net sales of power (GW)	146	69	70
Costs (Mio. Euro)	21	41	41
Incomes (Mio. Euro)	25	49	48
Net profits (Mio. Euro)	4	8	7
Demand coverage rate (%)	91.6	83.2	83.9
Variations of demand (GWh)	441	607	671

Table 4. Solutions under different operation strategies from the perspective of power system

	Self-supply strategy	Profits maximization	Costs minimization
Thermal cost (Mio. Euro)	1442	1445	1430
Emissions (Mio. Euro)	7.3	7.5	7.5
Spillage (GWh)	1186	1183	1191

It can be seen from the results that in terms of power demand coverage the self-supply strategy obviously surpasses the market profits maximization strategy, while market profits maximization strategy is relatively beneficial to virtual power plants' profits. And the differences between self-supply strategy and market profits maximization strategy are not obvious in terms of thermal costs, emissions and power spillage.

4. Conclusion

With the development of China's economy, problems of energy shortages and environmental pollution have become increasingly prominent, and this has made the integrated and controllable virtual power plants become more and more important. This paper utilizes the ROM model and unit commitment model for modeling and accessing virtual power plants under self-supply maximization and market profits maximization with two different operational strategies. The illustrative example used for verification of the scientific value, practicality and validity of the model proves that it can be used for modeling and evaluating virtual power plants' operation to promote the healthy and orderly development of China's power system.

References

[1] M. Franke, D. Rolli, A. Kamper, A. Dietrich, A. Geyer-Schulz, H. Schmeck, Ch. Weinhardt: *Impacts of distributed generation from virtual power plants*. Proc. 11th Annual IC Sustainable Development Research, 6–8 June 2005 Helsinki, Finland, CD-ROM.

[2] M. Zdrilić, H. Pandžić, I. Kuzle: *The mixed-integer linear optimization model of virtual power plant operation*. 8th IC Energy Market, 25–27 May 2011, Zagreb, Croatia, 5–9.

[3] M. Shabanzadeh, M. K. Sheikh-El-Eslami, M. R. Haghifam: *The design of a risk-hedging tool for virtual power plants via robust optimization approach*. Applied Energy *155* (2015), 766–777.

[4] N. Ruiz, I. Cobelo, J. Oyarzabal: *A direct load control model for virtual power plant management*. IEEE Trans. Power Systems. *24* (2009), No. 2, 959–966.

[5] C. Kieny, B. Berseneff, N. Hadjsaid, Y. Besanger, J. Maire: *On the concept and the interest of virtual power plant: Some results from the European project Fenix*. Proc. Power & Energy Society General Meeting, 26–30 July 2009, Calgary, Canada, CD-ROM.

[6] I. Kuzle, M. Zdrilić, H. Pandžić: *Virtual power plant dispatch optimization using linear programming*. Proc. 10th IC Environment and Electrical Engineering (EEEIC), 8–11 May 2011, Rome, Italy, CD-ROM.

[7] S. Sučić, T. Dragičević, T. Capuder, M. Delimar: *Economic dispatch of virtual power plants in an event-driven service-oriented framework using standards-based communications*. Electric Power Systems Research *81* (2011), No. 12, 2108–2119.

[8] Y. Tai-Her, W. Li: *A study on generator capacity for wind turbines under various tower heights and rated wind speeds using Weibull distribution*. Proc. IEEE Trans. Energy Conv. *23* (2008), No. 2, 592–602.

[9] R. Kuntschke, M. Specht, M. van Amelsvoort, M. Wagler, M. Winter, R. Witzmann: *Economic optimization in Virtual Power Plants vs. stable grid operation—bridging the gap*. Proc. 20th IC Emerging Technologies & Factory Automation (ETFA), 8–11 Sept. 2015, Luxembourg City, Luxembourg, CD-ROM.

[10] V. Robu, R. Kota, G. Chalkiadakis, A. Rogers, N. R. Jennings: *Cooperative virtual power plant formation using scoring rules*. Proc. IC Autonomous Agents and Multiagent Systems, 4–8 June 2012, Valencia, Spain, 370–376.

[11] R. Caldon, A. R. Patria, R. Turri: *Optimisation algorithm for a virtual power plant operation*. Proc. 39th IC Universities Power Engineering (UPEC), 6–8 Sept. 2004, Bristol, Great Britain, 1058–1062.

[12] Z. Vale, T. Pinto, H. Morais, I. Praca, P. Faria: *VPP's Multi-level Negotiation in Smart Grids and Competitive Electricity Markets*. Proc. IC Power and Energy Society General Meeting, 24–29 July 2011, Detroit, USA, 1–8.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 57–74

Numerical analysis and simulation of liquid food temperature fluctuation in an open refrigerated display cabinet

Wensong Lin[1], Tianji Chen[1]

Abstract. Based on the numerical simulation and experimental study on temperature of vertical open refrigerated display cabinets (RDC), some methods of periodic functions and wave shape analysis theory were used for research of the food temperature fluctuation in such a cabinet and temperature fluctuating of chilled liquid food inside RDC was also simulated. The main factors influencing the food temperature fluctuations were analyzed and studied, and the mathematical models of cabinet temperature perturbation and corresponding food temperature response were built; the mathematical models were solved and analyzed by the CFD and wave characteristic functions. The characteristics of food temperature fluctuations were obtained from the results of numerical simulation, and four of them related to liquid foods were compared and analyzed under different on-off ratios and defrost conditions. As the results of simulations show, the temperature fluctuating time of the food with shorter defrost heating time and smaller on-off ratio is shorter and the the height of peak is lower. So, selecting suitable defrosting condition and on-off ratio will optimize the performance of RDC and improve the quality of displayed food.

Key words. Refrigerated display cabinets, food temperature fluctuation, wave shape analysis, numerical simulation, periodic defrost..

1. Introduction

As the terminal equipment of modern food cold chain, open display cabinets are widely used in supermarkets. They can not only guarantee the food quality and freshness, but also have the advantages of good displaying and easy taking the products stored. Compared with the closed display cabinets, the open display cabinets face more problems. The cabinet temperature is, for example, more sensitive to ambient temperature, and a slight temperature rise will cause microbial breeding dangerous to customers. As Schmidt et al. [1] studied, the temperature fluctuation during the defrost period will accelerate the chemical and enzymic changes in the tissue of meat and in dairy products even possibly speed up breeding of microorganism and may cause food spoilage and poisoning. Therefore, the control of RDCs,

[1]Department of refrigeration, Shanghai Ocean University, Shanghai 200090, China

especially their temperature, is very crucial for the chilled food storage industry with rapid development.

The fluctuations of the cabinet temperature are closely related with the characteristics of air supply and defrost of open refrigerated display cabinets. In this field, there is much research on the numerical simulation of the temperature field inside RDC based on computational fluid dynamics (CFD), mainly using the finite difference method, the finite element method, the transformation method, etc., as Cortella et al. [2] summarized. In this paper, based on the numerical simulation and experimental study on temperature characteristic of open vertical refrigerated display cabinets, some methods of periodic function and wave shape analysis theory were used in the research on the food temperature fluctuation. The temperature fluctuating of chilled liquid food inside RDC was simulated, the simulation of food temperature fluctuation was performed considering the disturbance of temperature of air supplied and periodic defrost, which is different from those recommended in literature. The temperature fluctuation of chilled liquid food inside RDC was controlled and suppressed by regulating the wave characteristic of the defrost cycle and on-off ratio of the refrigeration system in order to optimize the performance of RDC and improve the quality of foods on display.

2. Physical model of RDC and liquid food tank

According to the performance test standard of RDC [3], the ambient temperature and humidity was set as 25 °C and 60.0 %. The RDC simulated is an open vertical type with inner and outer dual air curtains, and it has four shelves and bottom display areas as shown in Fig. 1a. The dimension is $2430 \times 1100 \times 1940$ mm, inner volume is 1400 l and total display area is $5.9\,\mathrm{m}^2$. The cabinet temperature range is from -2 °C to 2 °C. Four kinds of liquid food, namely spring water, orange juice, milk, and yoghurt (plain) of same volume (396 ml) were filled fully into four cylindrical containers of same size (60 mm in diameter, 150 mm in height) and same material (paper of thickness 1 mm). The containers were put separately on the same position on the second shelf in RDC (two pieces of insulation material were pasted on the top and bottom of the container for thermal isolation, as shown in Fig. 1b). The influence of RDC horizontal direction was ignored. The standard M packages (oxy-ethyl-methyl-cellulose and water, referring to BS EN441-4 standard [4]) were laid on the other shelves and bottom area in RDC as a regular load in supermarkets.

3. Mathematical model

3.1. Theoretical analysis of numerical simulation

When an open RDC works, Laguerrea et al. [5] found that there is a heat exchange (between foods displayed and air flow inside the cabinet) by heat conduction, convection and radiation all the time. Moreover, food temperature fluctuates with the change in cabinet temperature. There are many factors affecting the change in

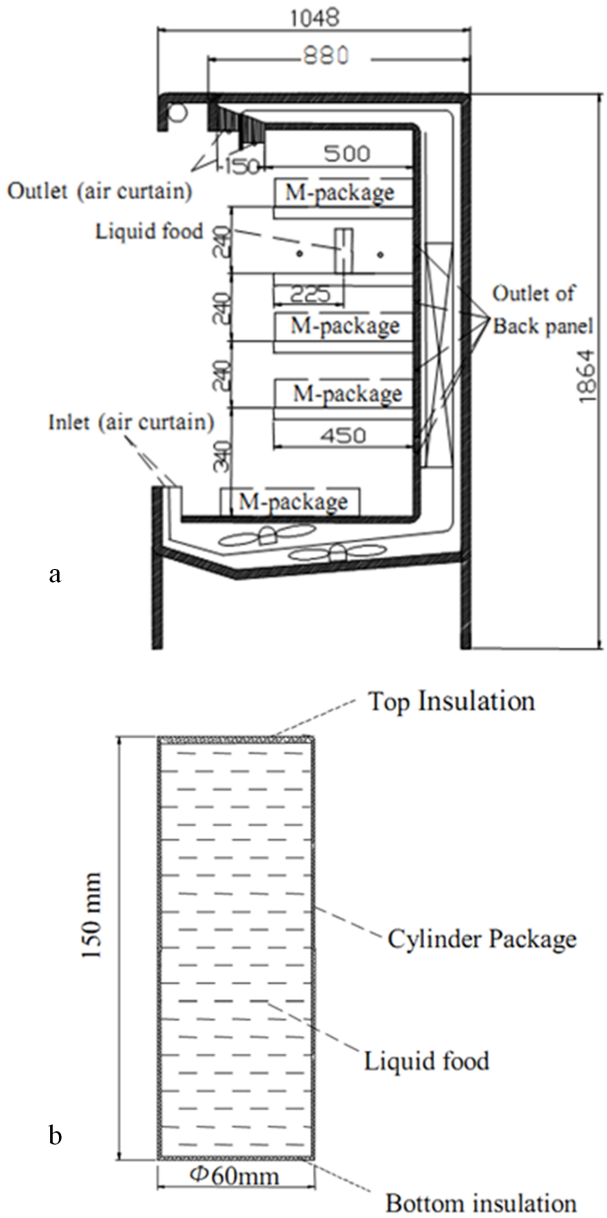

Fig. 1. Layout of tested RDC: a–arrangement (dimensions in mm), b–container with liquid food (dimensions in mm)

cabinet temperature. In general, they are divided into two main parts, periodical ones (the disturbances of temperature of air supplied etc.) and nonperiodical ones (such as the heat load of external air infiltration). The disturbance time of nonpe-

riodical ones is shorter and they are mostly irregular, but the periodical ones occur during the whole operation time of RDC, directly impacting the cabinet temperature. In this paper, three parameters, namely peak height H, fluctuating time λ and delay time t_{h} were used for wave shape analysis. Peak height H means the difference between the maximal and minimal temperatures in a temperature fluctuating cycle (half wave shape), and represents the amplitude of fluctuations. Its value is related to the operation condition of RDC and food category. The time from the start point to the end point of the fluctuating cycle was taken as the fluctuating time λ. It is half wavelength in a defrosting cycle representing the length of defrosting fluctuating time, and becomes full wavelength in a refrigeration cycle, indirectly reflecting the fluctuating frequency and the number of fluctuations. The absolute time different between the time of the peak point in a fluctuating cycle and that of the reference cycle was taken as the delay time t_{h}. It reflects the speed of temperature fluctuating transmission or response, its value is related with the food (package) category and has nothing to do with the operation conditions of RDC.

3.2. The disturbance model of air supply temperature of RDC

3.2.1. The disturbance model of periodic on-off control The start-stop of the refrigeration system and the defrost components are controlled by the control system to maintain the cabinet temperature at the value set. The neutral zone control method is used widely in the control system of RDC. When the cabinet temperature reaches the upper limit set, the controller sends out signals to open the liquid solenoid valve (or impulse electronic expansion valve) and start the compressor, and thus the refrigeration system works and the cabinet temperature decreases. When it reaches the lower limit set, the controller sends out signals again to shut off the liquid solenoid valve and stop the compressor, and thus cabinet temperature rises back. So the on-off control of a refrigeration system is periodic, and the ratio of uptime and downtime is almost constant under the steady operation condition. Thus the air supply temperature fluctuates periodically along with the periodic on-off control. Lazarin and Aprea [6–7] conclude that RDC with different cabinet temperature control system will have different control accuracy of temperature, and especially different on-off time ratio. If R is the on-off time ratio and P is one operation cycle, the uptime of the refrigeration system in one operation cycle is $R \cdot P(1+R)$ and the downtime is $P/(1+R)$. If $t(\tau)$ is the disturbance temperature function caused by on-off control, the periodic impulse τ in one operation cycle P is expressed as

$$t(\tau) = \begin{cases} t_{\mathrm{H}} & 0 < \tau < RP/(1+R) \\ t_{\mathrm{L}} & RP/(1+R) < \tau < P \end{cases}, \tag{1}$$

where t_{H} is the temperature of supplied air when the refrigeration system of RDC starts as the upper limit of cabinet temperature and t_{L} is the temperature of supplied air when the refrigeration system of RDC stops, as the lower limit of cabinet temperature.

By Fourier series expansion, equation (1) becomes

$$t_r(\tau') = \left(\frac{R}{R+1}t_{\mathrm{H}} + \frac{1}{R+1}t_{\mathrm{L}}\right) + \frac{2(t_{\mathrm{H}} - t_{\mathrm{L}})}{\pi}\sin\left(\frac{R}{R+1}\pi\right)\cos\frac{\pi}{P}\tau\,. \qquad (2)$$

3.2.2. The temperature disturbance model of periodic defrost Tasso et al. [8] found that frost forms and accumulates on the evaporator coil surface after a long period of operation, and the heat exchange performance of the evaporator degrades as the frost thickness increases, even blocking the flow passage and resulting in the failure of the compressor by liquid hammer. Therefore, periodic defrost is required to remove the accumulated frost on the evaporator coil in a setting time interval by the control system. When the control system switches to the defrost mode, the refrigeration system stops and the electrical heating runs. When the time or coil temperature reaches the upper limit set, the defrost stops and the cabinet temperature rises to some extent because of the effect of the defrost heating as Lawrence and Evans studied [9]. Because of the periodicity of defrosting process, the air supply temperature changes periodically, and the temperature change caused by defrost can be considered to be an individual disturbance factor. Symbol t_{d} denotes the air supply temperature when defrosting, τ_{d} is the heating time, and τ_{i} is the interval time between two defrosting process, the single defrosting cycle $P' = \tau_{\mathrm{d}}\tau_{\mathrm{i}}$. If $t(\tau')$ is the function of air supply temperature of RDC, the periodic impulse formula is

$$t(\tau') = \begin{cases} t_r & nP' < \tau < nP' + \tau_{\mathrm{i}} \\ t_d & nP' + \tau_{\mathrm{i}} < \tau < (n+1)P' \end{cases} \qquad n = 0, 1, 2, \ldots\,. \qquad (3)$$

In the above formula, t_{r} is the air supply temperature during refrigerated cycle, see equation (2). By Fourier series expansion, the formula (3) becomes

$$t(\tau") = \frac{1}{p'}(t_{\mathrm{r}}\tau_{\mathrm{i}} + t_{\mathrm{d}}\tau_{\mathrm{d}}) + \frac{2(t_{\mathrm{r}} - t_{\mathrm{d}})}{\pi}\sin(\frac{\tau_{\mathrm{i}}}{P'}\pi)\cos\frac{\pi}{P'}\tau\,. \qquad (4)$$

In (4), the defrosting time τ_{d} starts from the timing point when the air supply temperature is over the upper limit of the cabinet temperature t_{H} in the refrigerated cycle. The heating time τ_j generally is set as a fixed value by manual or depends on the time till the coil temperature reaches the fixed value set. When the heating time ends, the air supply temperature reaches the maximum value during the defrosting process, the refrigeration system starts, and the air supply temperature decreases gradually. The time of the defrosting process ends when the air supply temperature reaches the lower limit of the cabinet temperature. In the above process, the air supply temperature changes periodically and can be described as a half-wave sine function with peak of $t_{\mathrm{d,max}} - t_{\mathrm{H}}$ and fluctuating time τ_{d} as shown in Fig. 2.

Its Fourier progression expression is

$$t(\omega\tau'_{\mathrm{d}}) = \frac{t_{\mathrm{d,max}} - t_{\mathrm{H}}}{\pi}(1 + \frac{\pi}{2}\sin\tau_{\mathrm{d}} - \frac{2}{3}\cos 2\omega\tau_{\mathrm{d}}) \quad \text{and} \quad \omega = \frac{1}{2\tau_{\mathrm{d}}}\,. \qquad (5)$$

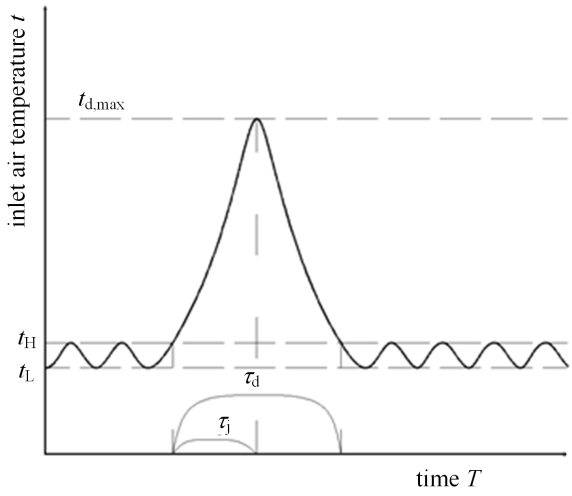

Fig. 2. Graph of half wave sine function of single defrost cycle

Here, quantities $t_{\mathrm{d,max}}$ and τ_{d} are calculated from the formulae

$$t_{\mathrm{d,max}} = t_{\mathrm{H}} + \eta_{\mathrm{e}} W_{\mathrm{e}} \left[\frac{\tau_{\mathrm{j}} - \frac{m_{\mathrm{w}}(C_{\mathrm{p,f}} \Delta T + h_{\mathrm{f}})}{Q_{\mathrm{h}}}}{c_{\mathrm{a1}} m_1} \right], \tag{6}$$

$$\tau_{\mathrm{d}} = \tau_{\mathrm{j}} + \left[\frac{m_1 (h_{1,t_{\mathrm{d,max}}} - h_{1,t_{\mathrm{L}}})}{q_{\mathrm{evp}}} \right]. \tag{7}$$

Here, symbol η_{e} denotes the defrost heating efficiency, W_{e} is the electric heater power in kW, m_{w} stands for mass of defrosting water in kg, $C_{\mathrm{p,f}}$ represents the specific heat capacity of frost in $\mathrm{J\,kg^{-1}\,K^{-1}}$, ΔT is the temperature difference between frost and electric heater in K, h_{f} denotes the fusion enthalpy of frost in J kg−1, Q_{h} represents the heat transfer from electric heater to evaporator in J, c_{a1} stands for specific heat capacity of air, $\mathrm{J\,kg^{-1}\,K^{-1}}$, $h_{1,t_{\mathrm{d,max}}}$ is the enthalpy of air with temperature of $t_{\mathrm{d,max}}$ in $\mathrm{J\,kg^{-1}}$, $h_{1,t_{\mathrm{L}}}$ denotes the enthalpy of air with temperature of t_{L} in $\mathrm{J\,kg^{-1}}$ and, finally, q_{evp} stands for the heat transfer flux from evaporator inside to air in W.

3.3. Simulation and solution of cabinet temperature field

3.3.1. Building of numerical model of temperature field Currently, as D'Agaro [10] concluded, turbulence models are mostly used for mathematical simulation of air curtain and RDC. The mainstream models are $k-\varepsilon$ model, Reynolds Stress Model (RSM) and large eddy simulation (LES) model. In this paper, an improved two-fluid model by Yu [11] is adopted, dividing the air flow into two fluids, the turbulent fluid from air curtain and the non-turbulent fluid outside the cabinet respectively, using the area coefficient of honeycomb as the volume fraction of air through air curtain.

Also a modification of mass transfer rate equation between two fluid is employed to improve the simulation accuracy of cabinet temperature field. In this improved two-fluid model, the food package is assumed as a hermetic cylinder, whose top and bottom are thermally insulated. The two-fluid model for the RDC can be established and the governing transport equations can be written as

$$\frac{\partial}{\partial x_j}(\rho_k r_k U_{kj} \Phi_k) = \frac{\partial}{\partial x_j}(r_k \Gamma_{\Phi k} \frac{\partial \Phi_k}{\partial x_j}) + \frac{\partial}{\partial x_j}(\Phi_k D_{\Phi k} \frac{\partial r_k}{\partial x_j}) + S_{\Phi k} + I_{\Phi k}\,. \qquad (8)$$

Here, the subscript k means the kind of fluid category, which is 1 or 2, the subscript j denotes the space coordinates, symbol Γ_Φ stands for the diffusion coefficient of fluid 1 or fluid 2, D represents the interphase diffusion coefficient between fluid 1 or fluid 2, S is the source term of fluid 1 or fluid 2, I is the interphase source term between fluid 1 and fluid 2, r stands for the volume fraction of fluid, and Φ denotes the dependent variable, corresponding to the continuity equation when taken as 1, momentum equation when taken as velocity, and energy equation when taken as temperature. Finally, ρ_k is the density of fluid k (k being 1 or 2), and U_{kj} is the velocity vector of fluid k (1 or 2) in the spatial y direction.

The disturbance temperature model of air supplied (2) and (4) is employed as the boundary function of the two fluid turbulence model in this paper, and this periodic disturbance boundary function is transferred and imported into user defined functions UDFs with the CFD software FLUENT, then the parameters of two fluids model are set and the computational domain is constructed and meshed for the simulation.

3.3.2. Computational case and grid The parameters of seven cases for simulation of RDC are listed in Table 1 and Table 2.

Table 1. List of input parameters

Parameter	Unit	Value
Ambient temperature T_a	°C	25
Ambient humidity R_a	%	60
Velocity of back air flow v_ba	$\mathrm{m\,s^{-1}}$	0.1
Velocity of outer air curtain* v_oa	$\mathrm{m\,s^{-1}}$	0.25
Velocity of inner air curtain v_{ia}	$\mathrm{m\,s^{-1}}$	0.3
Origin temp. of outer air curtain T_ao0	°C	10
Origin cabinet temperature T_c0	°C	2
Origin food temperature T_f0	°C	10
Food package surface (each side) S_f0	$\mathrm{m^2}$	0.014
Food package thermal conductivity k	$\mathrm{W\,m\,K^{-1}}$	0.9
Inner air curtain/back panel outlet temperature (Refrig. cycle) T_r	°C	$T_\mathrm{r}(\tau)$*
Inner air curtain/back panel outlet temperature (Defrost cycle) T_d	°C	$T_\mathrm{d}(\tau)$*

Note: * Defined temperature boundary function

The temperature change of spring water, orange juice, milk and yoghurt is investigated from case 1 to case 7, respectively. The initial temperature of the inner and outer air curtains is based on the experimental data for a better comparison. Thermophysical parameters of four liquid food samples are shown in Table 3 [12–16]. Neutral zone control is used for cabinet temperature control, and the initial cabinet temperature is set to be 2 °C. The on-off time ratio of the refrigerated system, defrost heating time and other parameters are shown in Table 2.

Table 2. Computational cases

Case	On-off ratio R	Defrost conditions	
No.	(-)	Heating time τ_d (min)	Interval time τ_i (min)
1	0.65	20	180
2	0.65	25	240
3	0.65	30	300
4	0.55	20	180
5	0.6	20	180
6	0.7	20	180
7	0.75	20	180

Table 3. Thermophysical properties parameters of liquid food samples [12–16]

Liquid food	Density ϱ ($\mathrm{kg\,dm^{-3}}$)	Specific heat capacity C ($\mathrm{kJ\,kg^{-1}\,K^{-1}}$)	Therm. conductivity k ($\mathrm{W\,m^{-1}\,K^{-1}}$)
Spring water	1.00	4.18	0.58
Orange juice	1.01	3.73	0.55
Milk	1.03	3.77	0.50
Yoghurt (plain)	1.05	3.52	0.48

As shown in Fig. 3, some parts or area of RDC as the evaporator, insulation layer and the propeller fan are removed in order to simplify the simulation. Considering that the open boundary conditions of the display cabinet are uncertain, the computational domain is extended until the effect of the display cabinet opening is negligible there. Then the computational domain becomes a irregular region of dimensions 1.07 m×1.41 m and the asymmetric grid consists of 44×36 cells, which were generated to mesh the computational domain. A denser grid is adopted in the air curtain inlet and outlet, back panel air outlet and surround of liquid food regions where the air flow is complex. For simplification of simulation, the plates of shelves are assumed to be insulated, see the details in Fig. 3, upper and bottom parts.

3.4. Response model of food temperature fluctuation

The heat transfer between the liquid food and airflow inside the cabinet occurs mainly by heat convection, and the cabinet temperature fluctuation gradually transfers into the interior of liquid food and causes the temperature fluctuated periodically

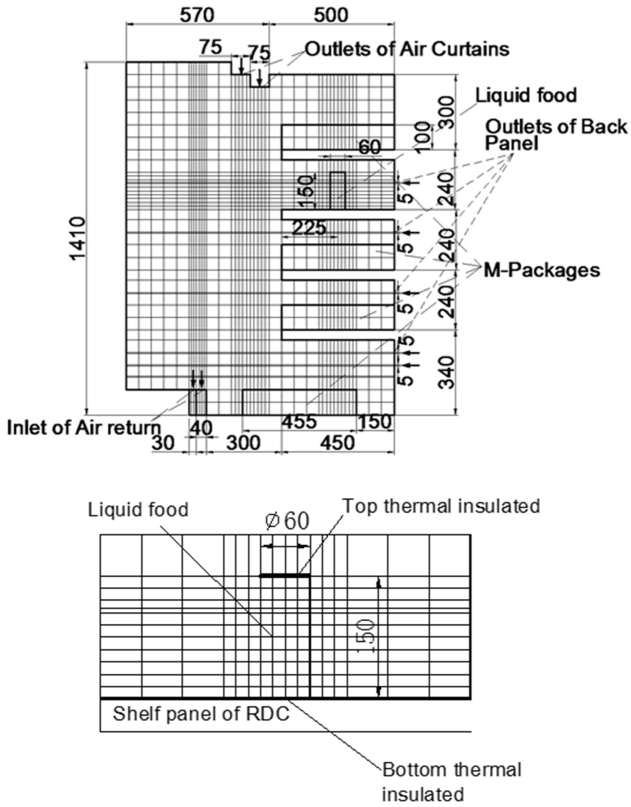

Fig. 3. Computational domain and grid of cabinet and liquid food: up–computing zone of RDC, bottom–computing zone around liquid food

with it. The food package is assumed as a hermetic cylinder with dry surfaces and insulated at top and bottom, so there is no mass transfer and the radiation is negligible. There is only convection between food package wall and airflow surrounded. A cylindrical coordinate system is built based on the structure of package, and a none-dimensional unsteady-state heat transfer model is established based on these assumptions and simplification. The instantaneous average temperature of 11 cells adjacent the wall of package on left and right sides (as shown in Fig. 3 bottom part) is used as the boundary air temperature in the calculation of this model, and the parameters related are set referring to the following equations [12–15].

According to the assumptions, the energy equation of liquid food is

$$\rho C_{\mathrm{p}} \frac{\partial T}{\partial t} = \frac{m}{r} k \frac{\partial T}{\partial r} + \frac{\partial k}{\partial r} \frac{\partial T}{\partial r} + k \frac{\partial^2 T}{\partial r^2} \,. \tag{9}$$

The mass equation has the form

$$\frac{\partial V}{\partial t} = \frac{m}{r} D \frac{\partial V}{\partial r} + \frac{\partial D}{\partial r} \frac{\partial V}{\partial r} + D \frac{\partial^2 V}{\partial r^2} \,. \tag{10}$$

The boundary condition between the food package and air is is given by the formula

$$-k\left[\frac{\partial T}{\partial r}\right]_{r=r_0} = \alpha F\left[T_{\mathrm{s}}(\tau,r) - \overline{T}_{\mathrm{air}}(\tau)\right] . \tag{11}$$

Here, $\overline{T}_{\mathrm{air}}(\tau)$ means the instantaneous average temperature of 11 cells adjacent the wall of package on both left and right sides.

The heat transfer equation between the liquid food and the food package is described by the formula

$$-k\left[\frac{\partial T}{\partial r}\right]_{r=r_1} = \omega\left[T_{\mathrm{s}}^{'}(\tau) - T_{\mathrm{i}}(\tau,r)\right] , \tag{12}$$

where $T_s^{'}(\tau)$ means the temperature of the inner surface of the package, and $T_i(\tau,r)$ denotes the temperature of the liquid food inside.

4. Results and discussion

4.1. Numerical simulation results of cabinet temperature fluctuation

The temperature field of the cabinet during the refrigeration cycle and the defrost cycle were simulated, respectively, as shown in Figs. 4 and 5. Figure 5 shows the distribution of temperature field at some point during the refrigeration cycle. From the figure, it can be seen that the distribution of the temperature field in each shelf has a certain difference, resulting in different distribution of temperature with M packages. The average temperatures in the 2nd and 3rd shelf are lower than those in other three shelves and the temperature of M packages in the 3rd shelf is also lower than those in other shelves. The significant temperature stratification is found inside the liquid food with package, and the temperature is high when the position is closer to the center of package.

Figure 5 shows the distribution of the temperature field at some point during the defrost cycle. From Fig. 5, it can be seen that the temperature distribution in each shelf is more uniform, and the distribution temperature of M packages in each shelf is very close, and the significant temperature stratification is also found inside the liquid food. The temperature is lower when the position is closer to center of the package.

4.2. Feature analysis of food temperature fluctuating response at different positions

The fluctuating response of the food temperature (in defrost cycles) in the cabinet in case 1 was calculated. The wave shape of average temperature curve in 2nd shelf was calculated and used as a reference wave shape, and then the characteristic numbers of temperature curve of different liquid food were calculated and analyzed basing on this reference wave shape. Figure 6 show the change in fluctuating char-

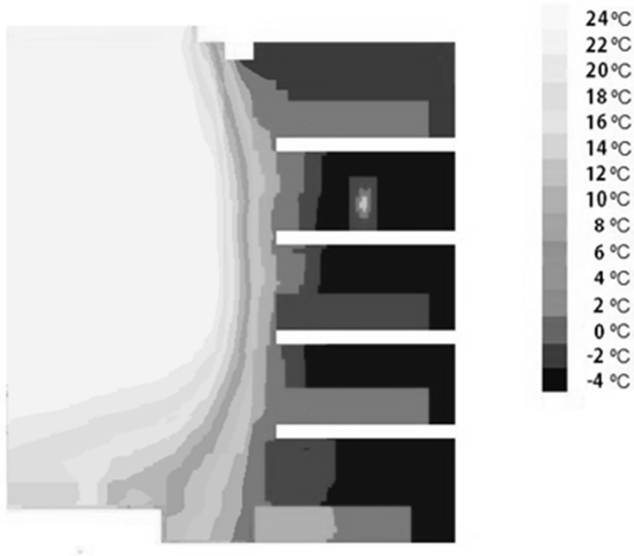

Fig. 4. Simulation of cabinet temperature field in refrigeration cycle (case 1)

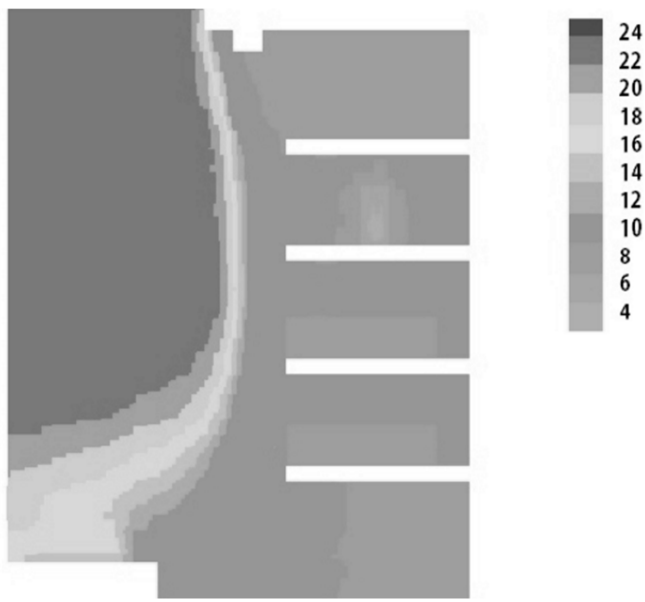

Fig. 5. Simulation of cabinet temperature field in defrost cycle (case 1)

acteristic numbers at different positions from outside to the center of the liquid food.

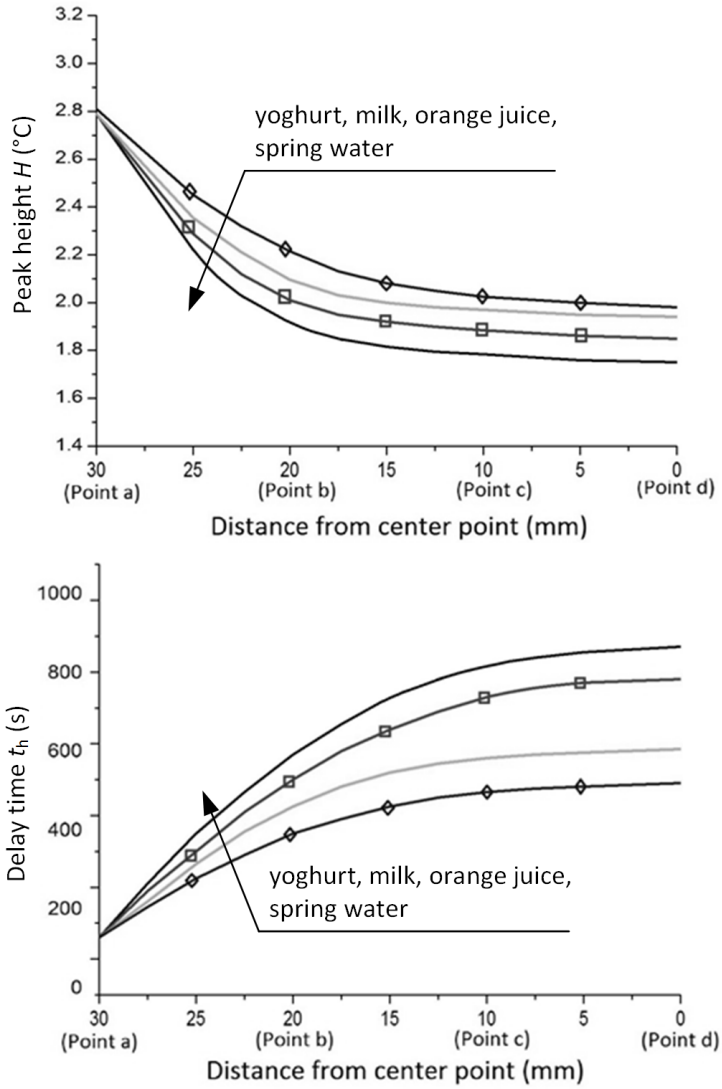

Fig. 6. Change in fluctuating characteristic numbers at different positions from outside to the center of the liquid food: up–peak height of each liquid food at different position, bottom–delay time of each liquid food at different position

It can be seen from Fig. 6 that some fluctuating characteristics of inner temperature of each liquid food correspond with the fluctuation of the cabinet temperature. The change of fluctuating characteristic of each liquid food from the outer points to inner points shows the obvious trend as follows: the peak height of temperature wave on the surface of each liquid food is higher around 2.8 °C, and drops gradually from the surface to center. On the contrary, the delay time of each liquid food increases. Taking spring water as an example, the peak height decreases by 1.04 °C

and delay time t_h increases 14.5 times. Furthermore, the fluctuating characteristic numbers of each liquid food shows a drastic change in the first 10 mm from outside (from the point a to point b), and gradually flattens in the final interval (from the point b to point d), such as the peak height of spring water decreases sharply by almost 1 °C (from 2.8 °C to 1.86 °C) in the first 10 mm and then decreases slightly by 0.1 °C (from 1.86 °C to 1.76 °C) in the final 20 mm. The delay time of spring water increases sharply almost 10 times (from 60 s to 590 s) in the first 10 mm and then it increases slightly 1.5 times (from 590 s to 87 s) in the final 20 mm.

4.3. Influence of different on-off ratio on food temperature fluctuating

The change in fluctuating characteristic numbers at the center of each liquid food was simulated and calculated with different on-off ratio R from 0.55 to 0.75 (case 1, 4–7), and the results are shown in Fig. 7.

From Fig. 7, it can be seen that the peak height of each liquid food changes from 0.14 °C to 0.22 °C, and also increases to some extent, by 25.4 %, 25.6 %, 26.4 % and 28.6 % for spring water, orange juice, milk, and yoghurt, respectively. The fluctuating time of each liquid food during one single refrigeration cycle changes between 160 s to 195 s, and increases to some extent with an increase of the on-off ratio from 0.55 to 0.75 by 11.1 % for spring water, 12.8 % for orange juice, 14.9 % for milk, and 15.5 % for yoghurt, respectively. As the calculations show, when on-off ratio R of RDC is smaller, the on-off frequency is higher and the interval time between run and stop is shorter, then the fluctuating time λ and peak height H of temperature fluctuating inside food is shorter accordingly. Vice versa, when on-off ratio R is higher, the on-off frequency is lower and the interval time between run and stop is longer, then the fluctuating time λ and peak height H of temperature fluctuating inside food will be longer and higher accordingly. The calculations also show that the peak height of each liquid food at center position is shorter and between 0.14 °C to 0.22 °C with the change of R, but considering the temperature fluctuation is of high frequency (about 3 min per cycle) and long time (3–5 hours), its impact on food quality cannot be ignored, either.

4.4. Influence of different defrost condition on food temperature fluctuating

The simulation shows that the fluctuating characteristic of cabinet temperature is different under different defrost conditions and results in different fluctuating of liquid food temperature. The peak height of the average temperature fluctuating in the second shelf in three defrost cases (1–3) is listed in Table 4.

The temperature change at the center of four kinds of liquid food in three defrost cases was simulated, and the changes in peak height H and fluctuating time λ obtained by the method of the wave shape characteristic analysis are shown in Fig. 8.

From Fig. 8, it can be seen that the range of peak height H of each liquid food at the center point is between 1.8 °C to 2.3 °C and the range of fluctuating time λ is between 7500 s to 8600 s. In all the considered cases, with the increase of defrost

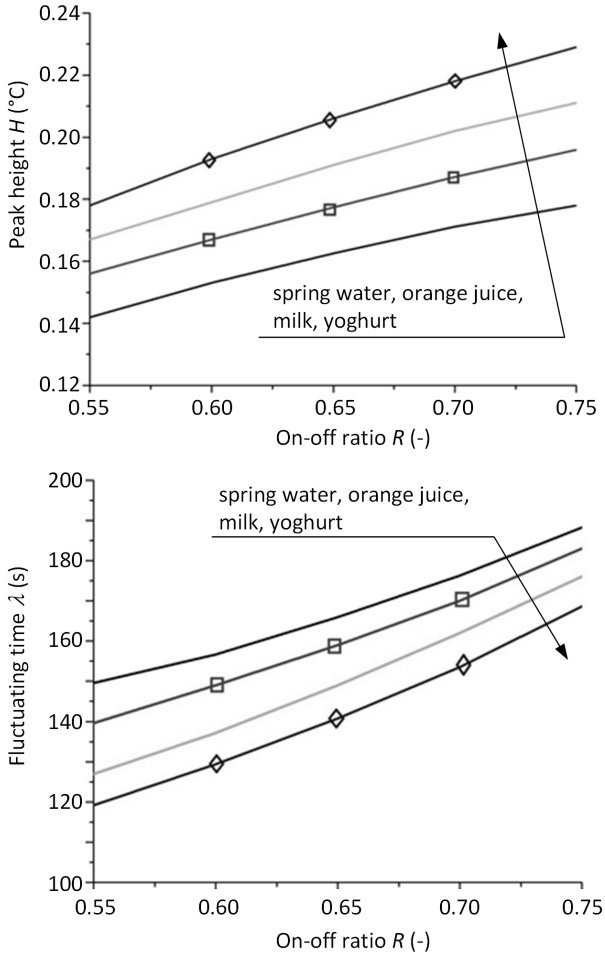

Fig. 7. Change in fluctuating characteristic numbers at center point of the liquid food with different on-off ratios R: up–peak height of considered liquid food with different on-off ratios, bottom–fluctuating time of considered liquid food with different on-off ratio

heating time and interval time, the fluctuating time in spring water during one single defrost cycle increases by 3.7 % and the peak height increases by 0.28 °C, the fluctuating time in orange juice increases by 5.1 % and the peak height increases by 0.31K; the fluctuating time in milk increases by 6.9% and the peak height increases by 0.33 °C, and finally, the fluctuating time in yoghurt increases by 8.1 % and the peak height increases by 0.35 °C. Thus, with the increases in the defrost interval time and heating time, the fluctuating time λ of each liquid food at the center point increases by 3.7 %–8.1 %, and peak height H also increases by 0.28–0.35 °C. It is likely that, in a certain range, if the defrosting interval time of RDC is shorter, the heating time will be also shorter and the temperature fluctuating will be smaller.

Thus, there will be more benefits to improve the quality of chilled food displayed.

Table 4. Peak height change of shelf temperature with various RDC defrosting cycle

Case	Interval time τ_{i} (min)	Heating time τ_{j} (min)	Average peak of second shelf temperature $\overline{T}_{\mathrm{air}}$ (°C)
1	3	20	9.25
2	4	25	10.85
3	5	30	12.35

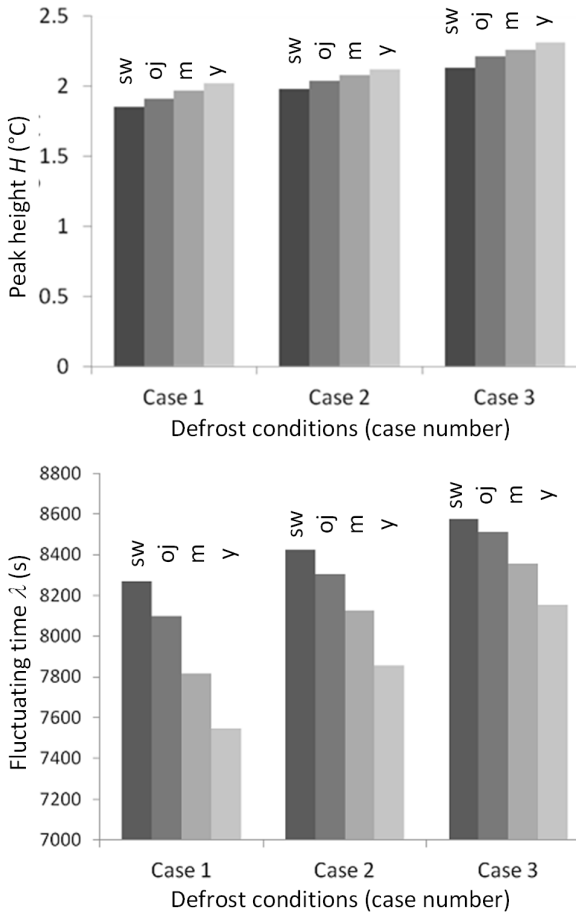

Fig. 8. Change in fluctuating characteristic numbers at center point of the liquid food with different defrost conditions: up–peak height of each liquid food in different defrost cases, bottom–fluctuating time of each liquid food with different defrost cases (sw-spring water, oj–orange juice, m–milk, y–yoghurt)

5. Conclusions

1. As theories and experiment indicated, periodic on-off control of refrigeration system and defrosting are two major interfering factors caused the temperature fluctuation in RDC, and different on-off ratio and defrosting conditions will cause different degrees of temperature rise and fluctuating.

2. A mathematic model of the temperature fluctuations inside the RDC and a fluctuating response model of food temperature corresponded have been built, and the disturbance temperature model of air supplied and periodic defrost has been employed to be the initial condition. The temperature field in RDC and the fluctuating characteristic of the liquid displayed have been simulated and calculated.

3. The temperature fluctuating characteristics of four liquid foods in RDC under refrigeration cycle and defrost cycle were simulated and calculated respectively. The results show that: under defrost cycle (case 1), the peak height of surface temperature wave of each liquid food is around 3 °C, and it decreases gradually from the surface to center. First, at 10 mm distance from the surface, the peak height decreases obviously from 3 °C to 2 °C, but it keeps around 2 °C with only a tiny change in the remaining 20 mm distance to the center. Under refrigeration cycle, the peak height of each liquid food at center position is shorter around 0.2 °C, but considering the temperature fluctuation of higher frequency (about 3 min per cycle) and long time (3–5 hours), its impact on food quality cannot be ignored.

4. As simulation results shown, when on-off ratio R of RDC is smaller, the on-off frequency is higher, the interval time between run and stop is shorter and fluctuating frequency of temperature inside food is relatively higher. Then the time λ and peak height H of temperature fluctuating inside food will be shorter accordingly. Conversely, when on-off ratio R is higher, the on-off frequency is lower, the interval time between run and stop is relatively longer, and the fluctuating frequency is lower. Then the time and peak height of temperature fluctuating inside food will be longer and higher accordingly. Such as, in a single refrigeration cycle, the fluctuating time of each liquid food increased by 11.1–15.5 % and peak height increased by 23.6–28.6 % with on-off ratio increased from 0.55 to 0.75.

5. This study shows that with the increase of defrosting interval time and heating time, the fluctuating time λ of each liquid food at center position increases by 3.7%–8.1%, and the peak height H also increased by 0.28–0.35K accordingly. So, in a certain range, if the defrosting interval time of RDC is shorter, the heating time will be shorter and the temperature fluctuating will be smaller, there will be more beneficial to improving the quality of chilled food displayed.

References

[1] V. J. S. Schmidt, V. Kaufmann, U. Kulozik, S. Scherer, M. Wenning: *Microbial biodiversity, quality and shelf life of micro filtered and pasteurized extended shelf life (ESL) milk from Germany, Austria and Switzerland.* IJ Food Microbiology *154* (2012), Nos. 1–2, 1–9.

[2] G. Cortella, M. Manzan, G. Comini: *CFD simulation of refrigerated display cabinets original research article.* IJ Refrigeration. *24* (2001), No. 3, 250–260.

[3] *ASHRAE Standard 72/1998, Methods of testing open refrigerators.* American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., Atlanta, 1998.

[4] *British Standard BS EN441-4: Refrigerated display cabinet, part 4*, 1995.

[5] O. Laguerre, M. H. Hoang, V. Osswald, D. Flick: *Experimental study of heat transfer and air flow in a refrigerated display cabinet.* J Food Eng. *113* (2012), No. 2, 310–321.

[6] R. Lazzarin, M. Noro: *Experimental comparison of electronic and thermostatic expansion valves performance in an air conditioning plant.* IJ Refrigeration *31* (2008), No. 1, 113–118.

[7] C. Aprea, R. Mastrullo, C. Renno: *Performance of thermostatic and electronic expansion valves controlling the compressor.* IJ Energy Research. *30* (2006), No. 15, 1313–1322.

[8] S. A. Tassou, D. Datta: *Influence of supermarket environmental parameters on the frosting and defrosting of vertical multi-deck display cabinets.* ASHRAE Transactions *105* (1999), No. 2, 491–496.

[9] J. M. W. Lawrence, J. A. Evans: *Refrigerant flow instability as a means to predict the need for defrosting the evaporator in a retail display freezer cabinet.* IJ Refrigeration *31* (2008), No. 1, 107–112.

[10] P. D'Agaro, G. Cortella, G. Croce : *Two- and three-dimensional CFD applied to vertical display cabinets simulation.* IJ Refrigeration *29* (2006), No. 2, 178–190.

[11] Yu Ke-zhi, Ding Guo-liang, Chen Tian-ji: *Modified two-fluid model for air curtains in open vertical display cabinets.* IJ Refrigeration *31* (2008), No. 3, 472–482.

[12] S. Peacock: *Predicting physical properties of factory juices and syrups.* International Sugar Journal *97* (1995), No. 1163, 571–572 and 575–577.

[13] R. Ibarz, V. Falguera, A. Garvin, S. Garza, J. Pagan, A. Ibarz: *Flow behavior of clarified orange juice at low temperatures.* J Texture Studies *40* (2009), No. 4, 445 to 456.

[14] I. M. Afonso, L. Hes, J. M. Maia, L. F. Melo: *Heat transfer and rheology of stirred yoghurt during cooling in plate heat exchangers.* J Food Eng. *57* (2003), No. 2, 179–187.

[15] I. H. Tavman, S. Tavman: *Measurement of thermal conductivity of dairy products.* J Food Eng. *41* (1999), No. 2, 109–114.

[16] R. Coquard, B. Panel: *Adaptation of the FLASH method to the measurement of the thermal conductivity of liquids or pasty materials.* IJ Thermal Sciences *48* (2008), No. 4, 747–760.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 75–86

Experimental and numerical analysis of I type crack blunt propagation mechanism[1]

S. L. Wang[2], W. Song[2], K. X. Dong[2], Y. B. Su[2],
D. S. Zhang[3], C. H. Li[3]

Abstract. A complex fracture network is investigated that may occur as a consequence of crossing morphological and hydraulic fractures. The study is based on weak interface bedding and natural fractures widely-distributed in unconventional reservoirs. By constructing three-point bending test of I-type crack extension, using digital speckle correlation experiments and numerical simulation method, change laws of surface strain field, stress field, crack tip opening displacement and stress intensity factor can be obtained. Through experimental and simulation results, we know that: When I-type crack intersects with weak bedding, the crack became blunt, interfacial shear strain increases rapidly, direction of crack tip maximum circumferential stress deflects, mode I crack transforms to a mixed mode I-II crack and ratio of K_{II}/K_I increases, then the crack extends deflected.

Key words. I-type crack, blunt, shear, deflect, weak bedding.

1. Introduction

Weak bedding development of unconventional reservoirs, hydraulic fracture intersects with weak bedding and extend along it, are the keys of forming fracture network [1]. Study of the fracture morphology after artificial fracture and weak bedding intersection has important significance for understanding fracture deflected propagation mechanism and revealing the fracture network formation mechanism.

By means of theoretical and experimental methods, many scholars have analyzed

[1]Supported jointly by National Basic Research Program of China (2015CB250900), Natural Science Foundation of China (51374074) and Northeast Petroleum University Innovation Foundation for Postgraduates (yjscx2015-024nepu).

[2]College of mechanical science and engineering, Northeast Petroleum University, Daqing, 163318, China

[3]Research Institute of Oil Production Engineering of Daqing Oilfield Company Ltd, Daqing, 163712, China

the intersections of artificial fractures and natural fractures, and put forward some extended criteria: Warpinski adopted linear friction theory to consider the shear slip failure caused by the fracture surface shear stress, and analyzed the tensile failure caused by the normal stress in the crack surface by using Mohr-Coulomb criterion, and proposed W–T criterion [2]. Rehshaw and Pollard studied the induced stress field of the hydraulic fracture tip, and proposed the evaluating criterion of the hydraulic fracture through the natural fracture (R–P criterion) [3]. Gu and Weng suggested the G–W criterion by extending the R–P criterion to non-orthogonal condition and obtained the critical curve of the hydraulic fracture through natural fracture [4]. Anderson carried out experiment research of hydraulic fracture through the bonding interface and discussed the influence of friction coefficient [5]. Zhou Jian et al. established a large size three-axial hydraulic fracturing experiment, and discussed the factors that affect the direction of the hydraulic fracture after intersection with natural fracture. The above analysis provided the weak bedding opening conditions qualitatively, it did not bring the displacement field and strain field of crack, then it could not obtain the crack deflected propagation mechanism. Using digital speckle correlation method and numerical simulation method can provide the variation of strain field and displacement field, so that the method can analyze the deflected propagation mechanism from microscopic point of view.

2. Experimental study of I-type crack propagation

As it is difficult to obtain a large number of unconventional shale cores, according to similarity criterion [7, 8] the experimental specimens were processed with gypsum and clay as similar materials. The experimental length, high and thickness of specimens are 300 mm, 150 mm and 50 mm, respectively. The specimens are divided into 3 layers, and in each layer of height 50 mm a crack of dimensions (1 mm×10 mm) was prefabricated in the middle of the specimen bottom, as shown in Fig, 1.

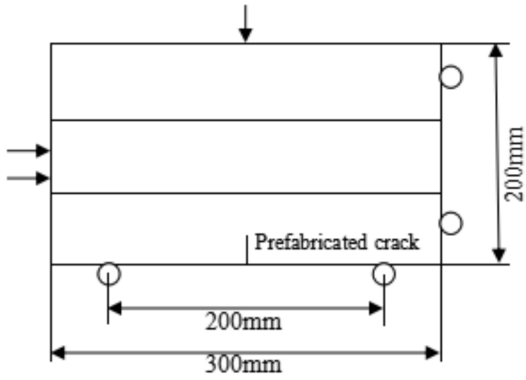

Fig. 1. Schematic diagram of material specimen

The specimens were loaded with servo loading test machine at the loading rate

of 0.03 mm/min (see Fig. 2). A high-speed video camera was used to capture the variation of the speckle field of the specimen surface. Using correlation algorithm, we can obtain the change law of displacement field and strain field in the process of the crack propagation.

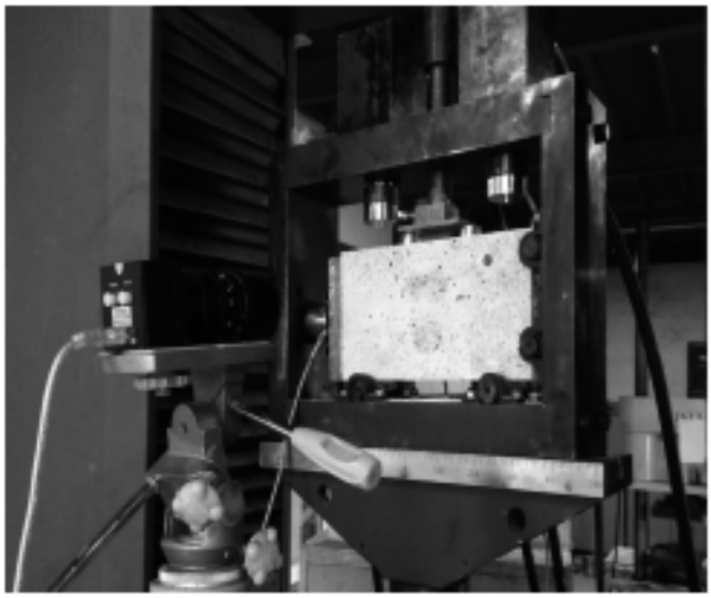

Fig. 2. Experimental loading of specimens

Figure 3 shows the evolution of horizontal strain field at different times. It can be seen from the trend of changes that with the gradual application of the load, the speckle field of the specimen surface gradually exhibits deformation localization band and extends. The final morphology of the deformation localization band is in agreement with the crack growth pattern, so that the displacement field and strain field can be used to analyze the crack propagation process.

In combination with curve in Fig. 4 we can see for loading time of 133 s that the local strain zone of the specimen surface extends to the interface position, and the relative value of the horizontal strain is 0.000116 mm at this time. For the loading time of 184 s, the local strain zone tip is still at the interface position and the length does not increase, but at this time the relative value of the horizontal strain increases rapidly to 0.001965 mm. The relative value of the horizontal strain at the crack tip is in a slow growth phase in the initial stage of loading, but from 133 s the relative value of the horizontal strain rapidly increases, and after 190 s it reaches the steady state. From this we can see that when the crack is extended to the interface, the crack tip loses the advantage of moving forward, crack growth reaches the steady state, but the crack tip opening displacement increases rapidly, and the cracks blunt phenomenon occurs here.

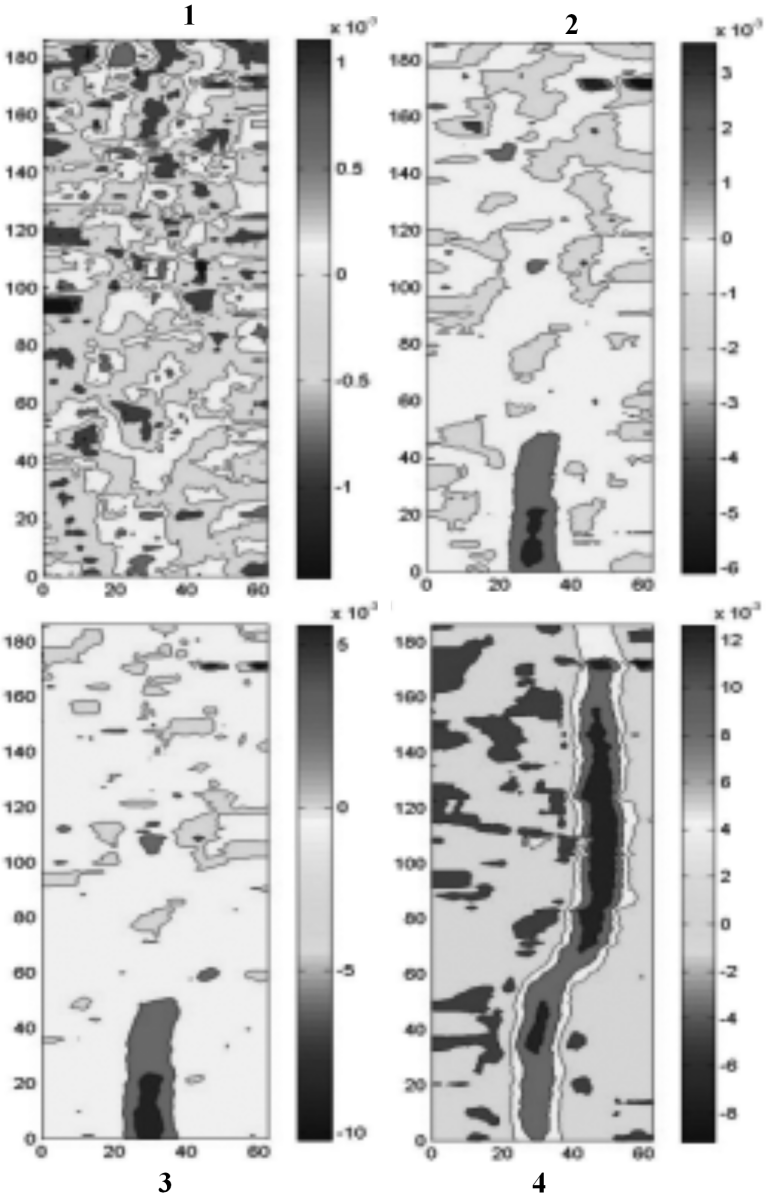

Fig. 3. Horizontal strain field changes of specimen in different times: 1–13.5 s, 2–133.6 s, 3–184 s, 4–190 s

Figure 5 shows the evolution process of shear strain field. It was shown that with the gradual application of the load, the shear strain of the crack surface also exhibits a deformation localization band. Comparison with horizontal strain field

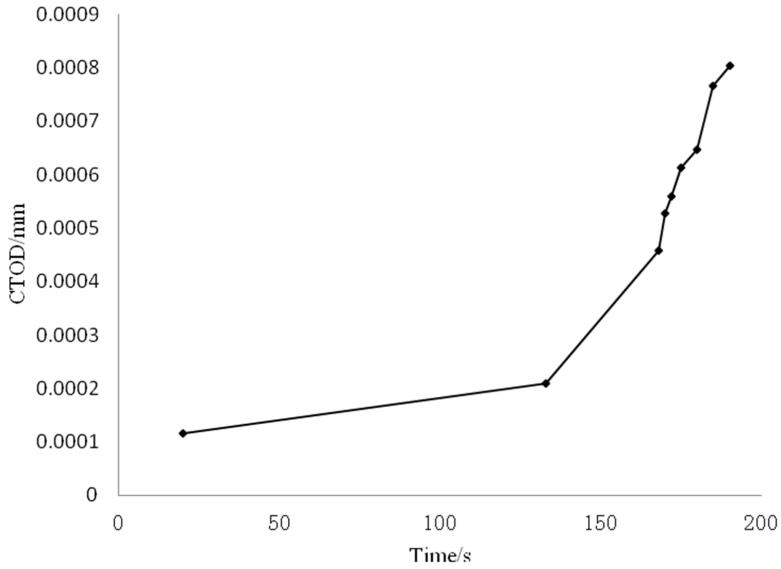

Fig. 4. Time evolution of horizontal strain field

shows that localization band appear at the interface layer. It explains that when crack extends to the interface layer where crack and interface layer join together, the interface shear concentrates there.

Figure 6 shows the distribution of shear strain in the interface layer at different times and different locations. It is obvious that shear strain of the interface increases gradually after crack extends to the interface layer. It reached the maximum 0.006848 mm after 190 s, which is much higher than in other places.

Curves in Fig. 7 show the distributions of shear strain at different horizontal positions for loading time of 190 s. The slice a was located at the interface layer, and points b, c and d were located at the upper part of the interface layer. By contrast, it is obvious that the shear strain of the interface layer is higher than the those in other horizontal positions. Upward from the interface layer, the shear strain decreases gradually and the maximum gradually decreases from 0.006848 mm to 0.005871 mm, then to 0.004126 mm and finally to 0.001644 mm. It can be seen from the above analysis that the crack gradually extends to the interface layer with the load application. After a period of time, the crack stops growing and the horizontal strain increases rapidly at the crack tip. It indicates that when the crack tip open displacement increases rapidly, it is at the blunt stage. Meanwhile, shear strain of the interface layer increases rapidly and is much higher than in other locations. Subsequently, the crack passes through the interface and extends inflected. Therefore, increase of shear strain which results in crack blunt at the interface layer, is the governing factor of inflected crack extension.

Based on the calculation method of stress intensity factor of I- and II-type crack, the crack tip opening displacement and sliding displacement are extracted, and the

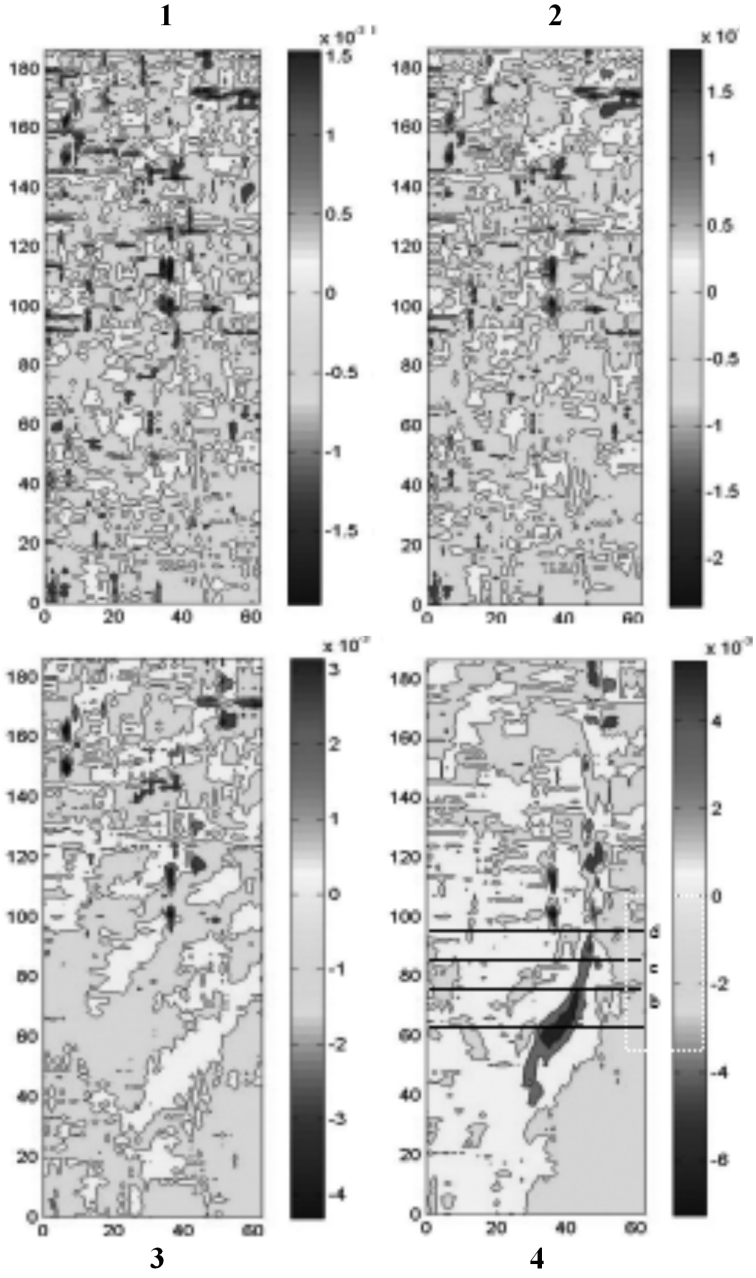

Fig. 5. Shear strain field changes of specimen in different times: 1–13.5 s, 2–133.6 s, 3–184 s, 4–190 s

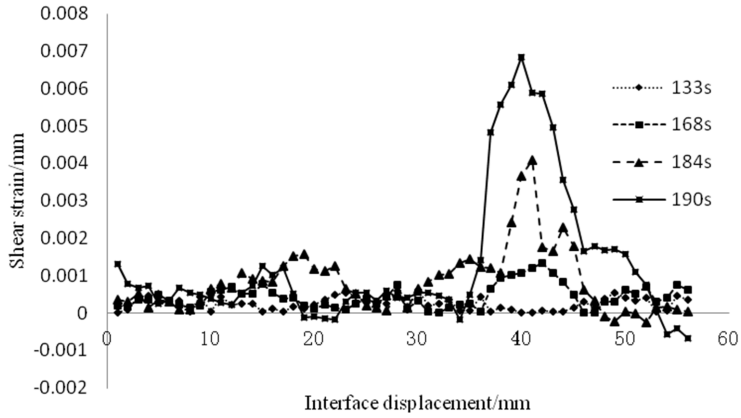

Fig. 6. Time evolution of shear strain field change curve

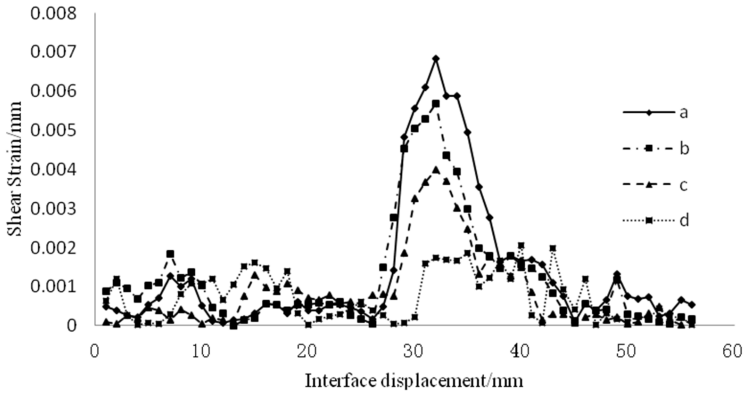

Fig. 7. Shear strain field change curve at different positions

crack tip K_I and K_{II} are calculated. At the time of 133 s, the crack tip K_I value is 0.12 MPa mm$^{0.5}$, K_{II} value is 0.0005 MPa mm$^{0.5}$. It is obvious that the K_I value is far greater than the K_{II} value, so that the latter one can be ignored. So before reaching the interface bedding, the crack is I-type open crack. With the load application, K_I gradually increases up to 5.01 MPa mm$^{0.5}$ at the time of 190 s, while the K_{II} value gradually increases to 2.04 MPa mm$^{0.5}$ and the value of K_{II}/K_I is 0.407. Coefficient K_{II} has an impact on the expansion of the crack. The crack changes from I-type to I-II composite crack, and then the crack extends deflected, the deflection angle being about 37 degrees.

3. Simulation study of I-type crack propagation

According to the three-point bending test method, the mechanical model of I-type crack propagation was established, and the geometrical size, force and boundary condition of the model are shown in Fig. 8. The finite element software ABAQUS is used to calculate the finite element model of three point bending with initial crack. The finite element model is shown in Fig. 9, where the fixed support model is applied on the left, the sliding bearing is applied on the right, while in the middle of the bottom span displacement load is applied, the speed of which being 0.03 mm/min. CPE4 elements are used for the discretization grid. In order to analyze the crack propagation, the grid encryption processing is adopted at the interface position.

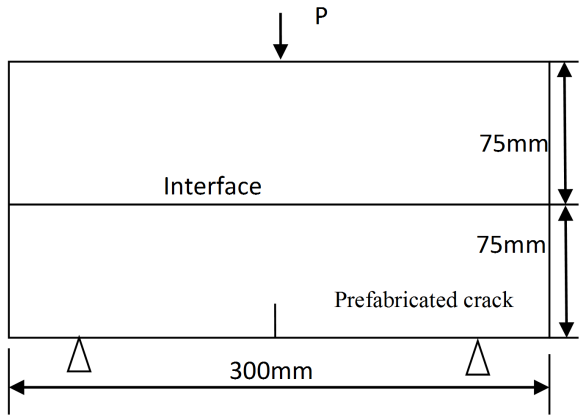

Fig. 8. Mechanical model of three-point bending

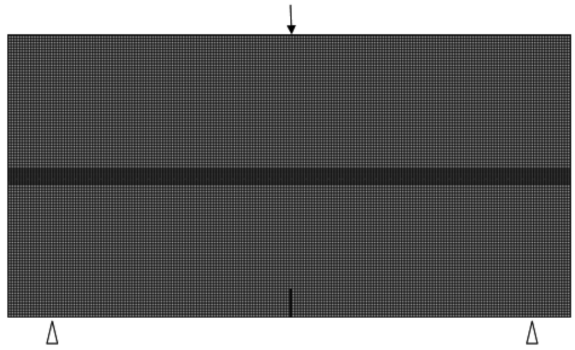

Fig. 9. Finite element model of three-point bending

Three following Figs. 10, 11 and 12 show evolutions of the crack. For better visibility, some parts are magnified.

Figure 13 shows that along with the load application, the crack length increases linearly with the load step, crack extends to the interface layer after 51 steps, crack

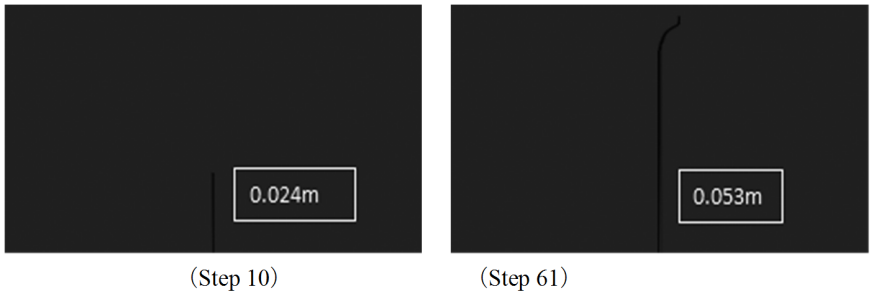

Fig. 10. Time evolution of crack propagation

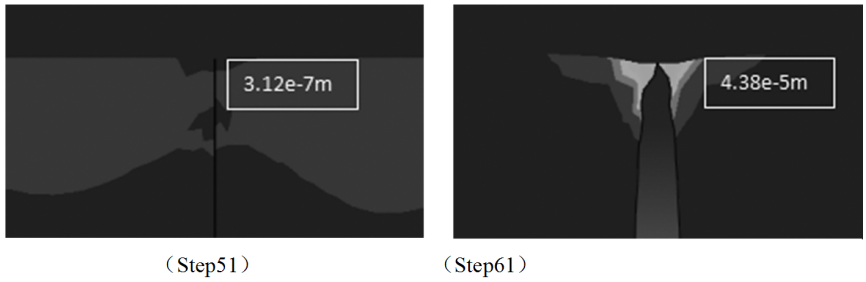

Fig. 11. Crack tip morphology in different times (magnified ten times)

Fig. 12. Time evolution of circumferential principal stress of crack tip

growth length is 46 mm, and crack length is the same after 61 steps, which says that between the 51st step to 61st step, the crack does not extend forward. From 70th step the crack continues to extend again, but the direction of propagation changes.

Figure 14 shows that after the crack extends to the interface layer, with the increase of the load step the crack tip opening displacement increases rapidly, from $3.12 \cdot 10^{-7}$ m (51st step) to $4.385 \cdot 10^{-5}$ m (61st step).

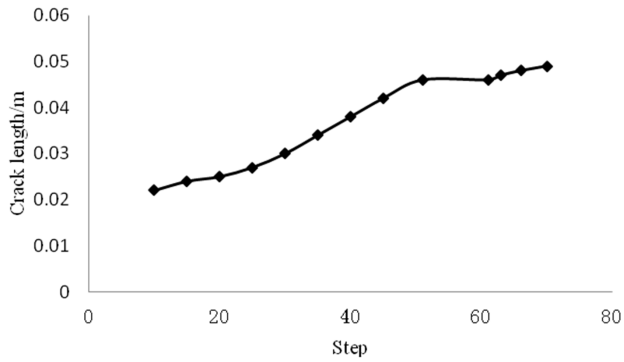

Fig. 13. Crack propagation length vs. number of steps

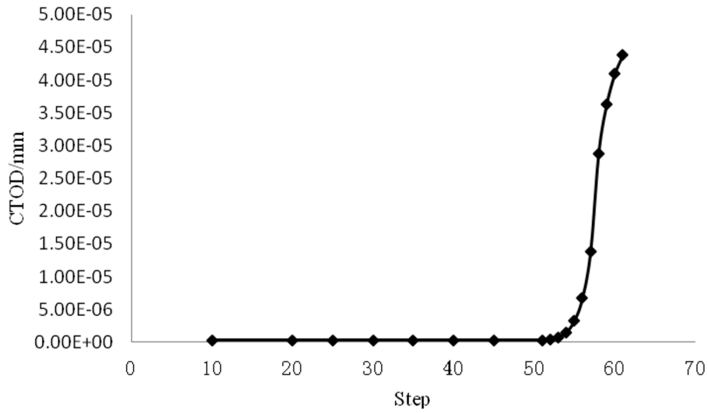

Fig. 14. Curve of crack tip opening displacement changes vs. loading step

4. Conclusion

In this paper, through the experimental and numerical simulation of I-type crack propagation, the following conclusions are obtained:

1. After I-type crack extends to the interface layer, as the interface strength is relatively low, the crack extend pauses, the crack tip opening displacement increases, and the I-crack tip decouples and blunts.

2. During the crack blunt process, the interfacial shear strain increases rapidly and interface shear strain is larger than other positions of shear strain. II-type stress intensity factor increases the crack tip to the maximum circumferential principal stress deflection, and I-type crack changes to I-II composite crack.

3. I-type crack tip caused by weak bedding, increased interfacial shear stress, maximum circumferential principal stress direction deflection, and crack tran-

sition from I-type to I-II type composite crack is are main control mechanisms of deflected crack extension.

References

[1] J. E. Olson, A. D. Taleghani: *Modeling simultaneous growth of multiple hydraulic fractures and their interaction with natural fractures.* SPE Hydraulic Fracturing Technology Conference, 19–21 Jan. 2009, The Woodlands, Texas, USA, paper SPE-119739-MS.

[2] N. R. Warpinski, L. W. Teufel: *Influence of geologic discontinuities on hydraulic fracture propagation.* J Petroleum Technology *39* (1987), No. 2, 209–220.

[3] C. E. Renshaw, D. D. Pollard: *An experimentally verified for propagation across unbounded frictional interfaces in brittle linear elastic materials.* IJ Rock Mechanics and Mining Sciences & Geomechanics Abstracts. *32* (1995), No. 3, 237–249.

[4] H. Gu, X. Weng, J. B. Lund, M. G. Mack, U. Ganguly, R. Suarez-Rivera: *Hydraulic fracture crossing natural fracture at no orthogonal angles: a criterion and its validation.* Proc. IC SPE Production and Operations, 24–26 Jan. 2012, The Woodlands, TexasUSA, 20–26.

[5] C. D. Anderson: *Effects of friction on hydraulic fracture growth near unbounded interfaces in rocks.* Society of Petroleum Engineers Journal *21* (1981), No. 1, 21–29.

[6] W. Burgert, M. Lippmann: *Models of translatory rock bursting in coal.* IJ Rock Mechanics and Mining Science & Geomechanics Abstracts *18* (1981), No. 4, 285–294.

[7] J. Marzbanrad, G. Soleimani, M. Mahmoodi-k, A. H. Rabiee: *Development of fuzzy anti roll bar controller for improving vehicle stability.* J Vibroengineering *17* (2015), No. 7, 3856–3864.

[8] H. Zhang, H. Song, Y. Kang, G, Huang, C. Qu: *Experimental analysis on deformation evolution and crack propagation of rock under cyclic indentation.* Rock Mech. Rock Eng. *46* (2013), No. 5, 1053–1059.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 87–98

Mathematical model to predict production of fractured horizontal well in tight gas reservoir and production analysis[1]

Yue Ming[2], Zhu Weiyao[2,3], Song Hongqing[2], Yang Lianzhi[2]

Abstract. According to different flow state around the fractured horizontal well associated with the threshold pressure gradient (TPG), combined with mass and momentum conservation equations, a mathematical model for steady gas flow in single fracture of fractured horizontal well was established. Considering the effect of interference between the fractures during the production period, production prediction model of fractured horizontal well was presented. Numerical analysis shows that the longer the fracture half-length, the bigger the productivity, but the productivity growth is reduced, and there is an optimal fracture half-length; the bigger the fracture conductivity, the bigger the productivity; the more the fracture quantity, the bigger productivity, but productivity growth is reduced due to the interference between fractures.

Key words. Tight gas reservoir, fractured horizontal well, fracture disturbance, elliptical flow.

1. Introduction

Tight sandstone gas resources in the global energy mix more and more important, is a strategic area at oil and gas exploration and development in future. In recent years, abundant tight sandstone gas resources were discovered in Ordos, Sichuan, Turpan-Hami basins, and its reserves, natural gas production accounts for one-third in China [1–4]. Compare to conventional gas reservoirs, it is difficult to develop

[1]This work is supported by National Program on Key Basic Research Project (973 Program) (Grant No. 2013CB228002), National Natural Science Foundation of China (Grant No. 11372033), and Fundamental Research Funds for Central Universities (No. FRF-TP-15-043A1).

[2]Civil & Resources Engineering School, University of Science & Technology Beijing, Beijing, 100083, China

[3]Corresponding Author

and achieve economic production less than expected using only horizontal wells due to complex geological occurrence of tight gas reservoir, dense reservoir, and small effective discharge area around the wells [5–6]. Therefore, multi-stage fractured horizontal wells are used to improve the well production and achieve cost-effective development. Numerical Simulation method is often used to study on production of fractured horizontal well, but complex data sources, long evaluation period, the application is not simple and easy [7–9]. Combined mass and momentum conservation equations, considering different flow characteristics [10–13], the fluid flow around the fracture is divided into two regions, a mathematical model of fractured horizontal wells in tight gas reservoirs was established, the analytic solution was obtained, productivity prediction equation was derived to evaluate development effect of fractured horizontal well intuitively.

2. Flow characteristics

The phenomena of non-Darcy flow in tight gas reservoir (TPG, slippage effect, and high-velocity non-Darcy flow) can be theoretically explained from the molecular force.

There are three forces during gas flow in the core: molecular adsorption between gas molecules and rock surface, interaction between gas molecules and differential pressure on both ends of cores.

Figure 1 shows the relationship between permeability and pressure gradient.

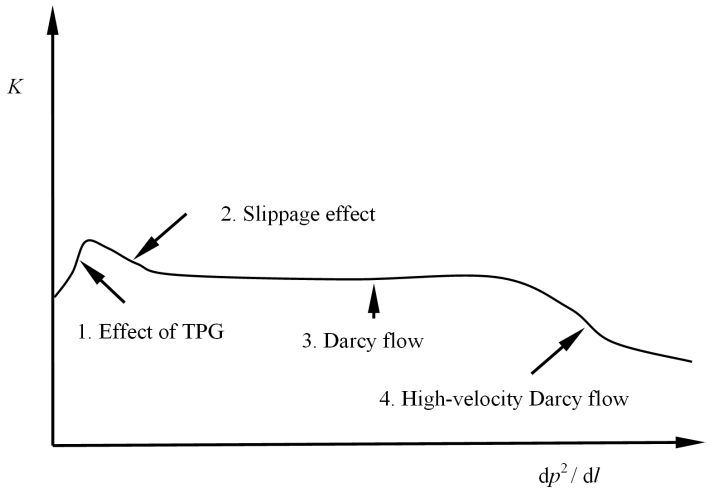

Fig. 1. Relationship between permeability and pressure gradient

As seen in the first stage of Fig. 1, the influence of adsorption force between gas molecules and rock surface is remarkable. When rock contains other polar fluid such as water or oil, the effect of gas TPG will be enhanced.

As seen in the second stage of Fig. 1, with the increase of gas pressure, gas molecule density increases, the interaction between gas molecules enhances, and the higher gas pressure, the lower the gas permeability. In this stage, the effect of interaction between gas molecules is remarkable.

As seen in the third stage of Fig. 1, slippage effect weakens gradually and gas enters the stage of stable Darcy flow, when the driving force is in the dominate status.

As seen in the fourth stage of Fig. 1, with more increase of air pressure, gas density becomes large, gas flow properties tend to be similar to the liquid flow, and gas flow is high velocity non-Darcy flow.

Due to the tiny pore throat, ultra-low permeability and high water saturation, the effect of TPG is the main influential factor of the production.

3. Conceptual model

3.1. Model assumption

Most hydraulic fracturing of horizontal wells are perforated after sealed with packer. In this case, fluid flow from matrix to wellbore is not considered.

Assumptions are given before establishing production forecast model of fractured horizontal as well as the reservoir is homogeneous and isotropic, gravity and capillary pressure are neglected, as well as wellbore storage and skin effect, fracture is symmetric around wellbore and located in the middle of layer with finite conductivity, fluid flows into fracture first and then flows into horizontal wellbore and fluid is only produced through perforation and fracture in horizontal well with case hole completion.

When the fractures of horizontal well are transverse, fluid flow could be divided into two parts: ellipse-shaped flow from stratum to cracks in horizontal plane, which is called external flow field, and linear and radial flow along cracks in vertical plane, which field is called internal flow field.

4. Mathematical model of fractured horizontal well

4.1. Productivity model of single fracture

4.1.1. Elliptical flow. According to Liu Ciqun, flow field of horizontal well could be recognized as isobaric ellipsoid and a family of hyperboloid streamline. Relationship between rectangular coordinate system and ellipsoid coordinate system is as follows:

$$x = a\cos\eta\,,\ \ y = b\sin\eta\,,$$

$$a = x_{\mathrm{f}}\cosh(\xi)\,,\ \ b = x_{\mathrm{f}}\sinh(\xi)\,, \tag{1}$$

where a and b are the major and minor axes of the ellipse, respectively, and x_{f} is the half-length of fracture.

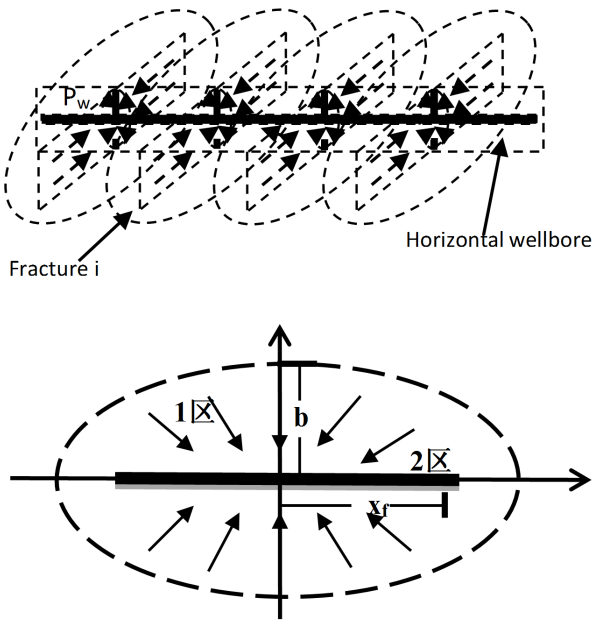

Fig. 2. Simplified schematic of flow field of fractured horizontal well

The mass conservation equation reads

$$\operatorname{div}\,(\rho_{\mathrm{g}}\boldsymbol{v}) = 0\,, \tag{2}$$

where ρ_{g} is the density of gas and $\boldsymbol{v}$ is its velocity.

The gas state equation may be expressed in the form:

$$\rho_{\mathrm{g}} = \frac{T_{\mathrm{sc}} Z_{\mathrm{sc}} \rho_{\mathrm{gsc}}}{p_{\mathrm{sc}}} \cdot \frac{p}{TZ}\,. \tag{3}$$

Here, T_{sc} is the standard state temperature, p_{sc} is the standard state pressure, Z_{sc} is the compressibility factor at the standard state and ρ_{gsc} is the density of gas at the standard state. Further, T, p and Z are the temperature, pressure and compressibility, respectively, in the real conditions.

The fluid flow follows the Darcy's law in elliptical flow region:

$$\boldsymbol{v} = \frac{k}{\mu}(\operatorname{grad} p - \boldsymbol{G})\,, \tag{4}$$

where k is the permeability of gas reservoir, μ is the gas viscosity, and $\boldsymbol{G}$ denotes the specific gravitational force acting on unit volume of gas.

Incorporating (3) and (4) into (2), and using a Cartesian coordinates to describe isobaric ellipse family, we get

$$\frac{\mathrm{d}\left(\rho v\right)}{\mathrm{d}y} = \frac{kT_{\mathrm{sc}}Z_{\mathrm{sc}}\rho_{\mathrm{gsc}}}{p_{\mathrm{sc}}T}\left[\frac{\partial}{\partial y}\left(\frac{p}{\mu Z}\frac{\partial p}{\partial y}\right) - C_\rho G\frac{p}{\mu Z}\frac{\partial p}{\partial y}\right] = 0\,, \tag{5}$$

where C_ρ is the gas compressibility coefficient.

Introducing the pseudo-pressure equation

$$m^* = 2\int_{p_{\mathrm{a}}}^{p}\frac{p}{\mu(p)Z(p)}\,\mathrm{d}p\,, \tag{6}$$

where p_{a} is a known pressure. Well-suited to engineering needs, $\mu(p)Z(p)$ can be simplified as μZ, the value of $\mu(p)Z(p)$ under constant temperature and average pressure of formation.

In combination with (5) and (6) we obtain the general governing equation in the form

$$\frac{\mathrm{d}^2 m^*}{\mathrm{d}y^2} - C_\rho G\frac{\mathrm{d}m^*}{\mathrm{d}y} = 0 \tag{7}$$

with constant-production inner boundary

$$\xi = \xi_{\mathrm{w}}\,,\quad y\frac{\mathrm{d}m^*}{\mathrm{d}y} = \frac{p_{\mathrm{sc}}Tq_{\mathrm{sc}}}{k\rho_{\mathrm{gsc}}T_{\mathrm{sc}}Z_{\mathrm{sc}}hx_{\mathrm{f}}\cosh(\xi)}\,. \tag{8}$$

The average minor axis of the inner boundary is

$$\bar{y}_{\mathrm{w}} = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} y\,\mathrm{d}\eta = \frac{2x_{\mathrm{f}}\sinh(\xi_{\mathrm{w}})}{\pi}\,. \tag{9}$$

The constant-pressure outer boundary is characterized by the conditions

$$\xi = \xi_{\mathrm{i}}\,,\quad p = p_{\mathrm{e}}\,,\quad m^* = m^*_{\mathrm{e}}\,, \tag{10}$$

Here, ξ_{w} is the elliptic coordinate near the wellbore, ξ_{i} is the elliptic coordinate of flow region and p_{e} is the initial reservoir pressure.

The average minor axis of the outer boundary is

$$\bar{y}_{\mathrm{i}} = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} y\,\mathrm{d}\eta = \frac{2x_{\mathrm{f}}\sinh(\xi_{\mathrm{i}})}{\pi}\,. \tag{11}$$

The analytical solution can be written in the form

$$-m_{\mathrm{xf}} + m_{\mathrm{e}} = \frac{p_{\mathrm{sc}}Tq_{\mathrm{sc}}}{2k\rho_{\mathrm{gsc}}T_{\mathrm{sc}}Z_{\mathrm{sc}}x_{\mathrm{f}}h\cosh(\xi_{\mathrm{w}})C_\rho G}\left[\mathrm{e}^{\frac{2x_{\mathrm{f}}C_\rho G}{\pi}(\sinh(\xi_{\mathrm{i}})-\sinh(\xi_{\mathrm{w}}))} - 1\right]\,. \tag{12}$$

4.2. High-velocity gas flow in fractures

The equation of motion in the fractures can be described as

$$-\text{grad } p = \frac{\mu}{k_\text{f}} v + \beta \rho v^2 \,, \tag{13}$$

where k_f is the permeability of fracture and β is the high-speed nonlinear flow coefficient. This coefficient can be determined from formula

$$\beta = \frac{0.005}{k_\text{f}^{0.5} \varphi^{0.5}} \,. \tag{14}$$

The mass flow rate dividing cross-sectional area yields velocity

$$v = \frac{q_\text{m}}{\rho_\text{g} A} = \frac{\rho_\text{gsc} q_\text{sc}}{2 \rho_\text{g} w h} \,. \tag{15}$$

Integrating (12) from 0 to x_f, the analytical productivity equation in fracture obtains the form

$$m_\text{xf} - m_\text{rw} = \frac{2 p_\text{sc} T x_\text{f}}{k_\text{f} T_\text{sc} Z_\text{sc} w h} q_\text{sc} + \frac{2 p_\text{sc} T \rho_\text{gsc} x_\text{f} \beta}{T_\text{sc} Z_\text{sc} w^2 h^2 \mu} q_\text{sc}^2 \,. \tag{16}$$

4.3. Productivity equation in single fracture

Adding (11) and (16) yields productivity equation in one single fracture,

$$m_\text{e} - m_\text{rw} = \frac{2 p_\text{sc} T \rho_\text{gsc} x_\text{f} \beta}{T_\text{sc} Z_\text{sc} w^2 h^2 \mu} q_\text{sc}^2 +$$

$$+ \left(\frac{p_\text{sc} T}{2 k \rho_\text{gsc} T_\text{sc} Z_\text{sc} x_\text{f} h \cosh(\xi_\text{w}) C_\rho G} \left[\text{e}^{\frac{2 x_\text{f} C_\rho G}{\pi} (\sinh(\xi_\text{i}) - \sinh(\xi_\text{w}))} - 1 \right] + \frac{2 p_\text{sc} T x_\text{f}}{k_\text{f} T_\text{sc} \text{Z}_\text{sc} w h} \right) q_\text{sc} . \tag{17}$$

As (17) is a quadratic polynomial, the productivity formula is derived as

$$q_\text{sc} = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \,, \tag{18}$$

where

$$a = \frac{2 p_\text{sc} T \rho_\text{gsc} x_\text{f} \beta}{T_\text{sc} Z_\text{sc} w^2 h^2 \mu} \,,$$

$$b = \frac{p_\text{sc} T}{2 k \rho_\text{gsc} T_\text{sc} Z_\text{sc} x_\text{f} h \cosh(\xi_\text{w}) C_\rho G} \left[\text{e}^{\frac{2 x_\text{f} C_\rho G}{\pi} (\sinh(\xi_\text{i}) - \sinh(\xi_\text{w}))} - 1 \right] + \frac{2 p_\text{sc} T x_\text{f}}{k_\text{f} T_\text{sc} \text{Z}_\text{sc} w h} \,,$$

and

$$c = m_\text{rw} - m_\text{e} \,.$$

4.4. Productivity equation of fractured horizontal well

When the horizontal well contains hydraulically caused several fractures, and some flow regions formed by fractures have interference, it is possible to reduce them to the one region. Assuming that every two elliptical flow regions intersect, the intersection area is S_i, and the production of single fracture considering the interference is

$$q_{\mathrm{f}i} = \left(1 - \frac{S_i}{\pi a_i b_i}\right) q_{\mathrm{sc}}\,, \quad i = 1, \cdots, n\,, \tag{19}$$

where S_i is the ith intersection area, $q_{\mathrm{f}i}$ is the gas production of the ith single fracture and n is the number of fractures. The ith intersection area is

$$S_i = 2 \cdot \left(\frac{1}{4}\pi a_i b_i - \frac{1}{2} a_i b_i \arccos\frac{y_i}{b_i}\right) - \frac{W_i}{2} \cdot y_i +$$

$$+ 2 \cdot \left(\frac{1}{4}\pi a_{i+1} b_{i+1} - \frac{1}{2} a_{i+1} b_{i+1} \arccos\frac{y_i}{b_{i+1}}\right) - \frac{W_{i+1}}{2} \cdot y_{i+1}\,, \tag{20}$$

where W_i is the width of the ith fracture ($i = 1, \cdots, n-1$) and

$$x_i = \frac{W_i}{2}\,, \quad y_i = \sqrt{\left[1 - \left(\frac{W_i}{2a_i}\right)^2\right] \cdot b_i^2}\,, \quad i = 1, 2, \cdots, n-1\,.$$

When there is no interference, $S_i = 0$.

The productivity prediction equation of fractured horizontal well in tight sandstone reservoir was presented considering the interference between fractures

$$Q_{\mathrm{sc}} = \sum_{i=1}^{n} q_{\mathrm{f}i}\,, \tag{21}$$

where Q_{sc} is the gas production of the horizontal well.

5. Analysis of influential factors

Numerical analysis is carried out using data of physical properties in Changling gas field. Its parameters are as follows: reservoir thickness is 8 m, initial formation pressure is 35 MPa, well bottom pressure is 25 MPa, TPG is 0.01 MPa·m^{-1}, gas viscosity at standard state is 0.027 mPa·s, gas density is 0.75 kg·m^{-3}, temperature in standard state is 293 K, reservoir temperature is 396 K, gas compressibility factor is 0.89, radius of wellbore is 0.1 m, porosity is 0.08, permeability is 0.1 mD, length of horizontal well is 1200 m, fracture quantity is 15, half-length of fracture is 150 m, fracture width is 3 mm, and permeability of fracture is 1 D.

Fig. 3 shows the relationship between production rate and pressure drawdown under different fracture half-length. As seen in the figure, production rate increases with the increase of pressure drawdown. Besides, production rate increases with

increasing fracture half-length under certain drawdown, and half-length of fracture that ranges from 150 m to 200 m have a good effect to productivity.

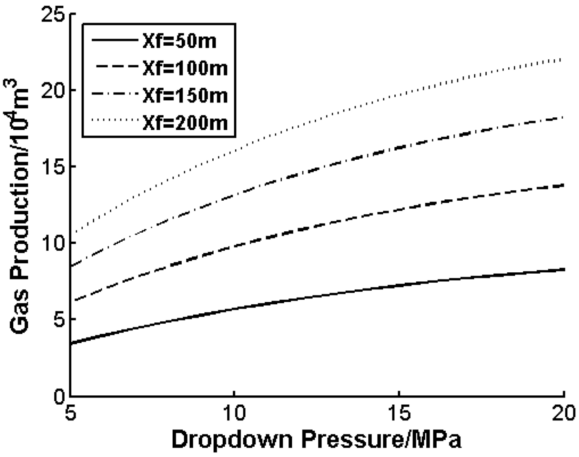

Fig. 3. Relationship between dropdown pressure and productivity under different half-lengths of fractures

Fig. 4 shows the relationship between production rate and pressure drawdown under different fracture conductivity. As seen in the figure, production rate increases with the increase of pressure drawdown. Besides, production rate increases with increasing fracture conductivity under certain drawdown. Improvement of the fracture conductivity have a good effect on productivity.

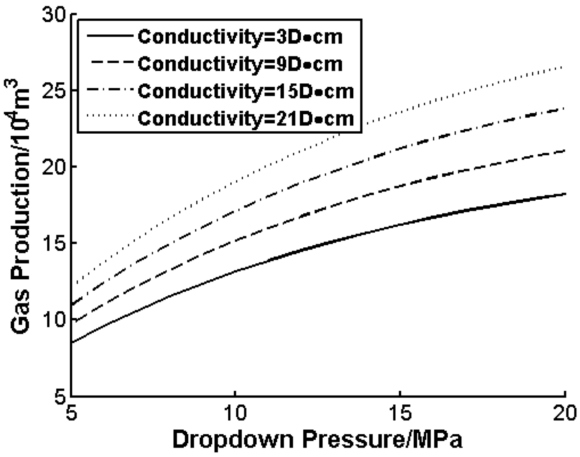

Fig. 4. Relationship between dropdown pressure and productivity under different conductivities of fractures

Fig. 5 shows the relationship between production rate and pressure drawdown

under different fracture quantity. As seen in the figure, production rate increases with the increase of pressure drawdown under certain fracture quantity. Besides, production rate increases with increasing fracture quantity under certain drawdown. With increasing fracture quantity, production of fractured horizontal well increases with decreasing increment and approaches a plateau due to the interference.

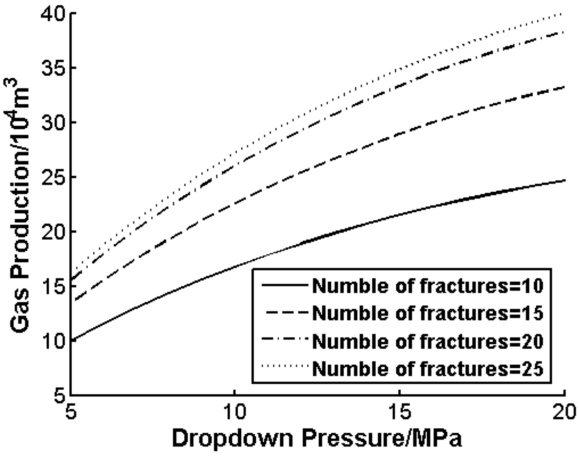

Fig. 5. Relationship between dropdown pressure and productivity under different fracture quantities

Fig. 6 shows the relationship between production rate and pressure drawdown under different TPG. As seen in the figure, production rate increases with the increase of pressure drawdown under certain TPG. Besides, production rate decreases with increasing TPG under certain drawdown. The fluid is hard to flow, and control area of well decreases due to the existence of TPG.

6. Conclusion

- The phenomena of non-Darcy flow in tight gas reservoir (TPG, slippage effect, and high-velocity non-Darcy flow) can be theoretically explained from molecular force, for high water-bearing tight gas reservoir, TPG being the main influential factor.
- According to different flow regions around the fractured horizontal well, considering the existence of threshold pressure gradient when the fluid flow in tight water-bearing gas reservoir, a mathematical model about fractured horizontal well coupled with two flow regions was established ,as the pressure at the junction of two regions equal, productivity equation of single fracture could be presented.
- The longer the fracture half-length, the bigger the productivity, but the productivity growth is reduced, and there is an optimal fracture half-length; the

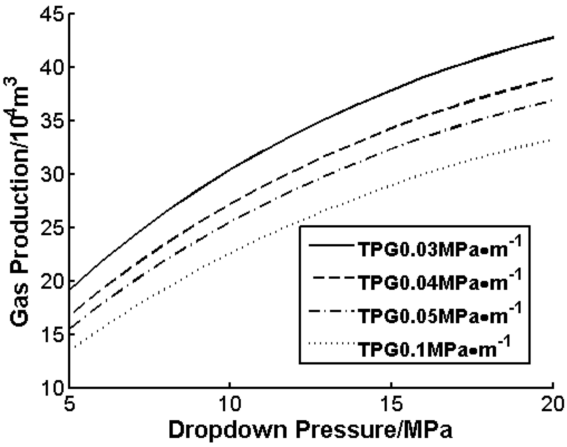

Fig. 6. Relationship between dropdown pressure and productivity under different values of TPG

bigger the fracture conductivity, the bigger the productivity; the more the fracture quantity, the bigger productivity, but productivity growth is reduced due to the interference between fractures.

References

[1] Y. P. Ding, X. D. Wang, J. Xing: *A method of productivity calculation for fractured horizontal well.* Special Oil & Gas Reservoirs *15* (2008), No. 2, 64–68.

[2] Z. D. Lei, S. Q. Cheng, X. F. Li, H. P. Xiao: *A new method for prediction of productivity of fractured horizontal wells based on non-steady flow.* J Hydrodynamics *19* (2007) No. 4, 494–500.

[3] Y. Xu, T. Qi, F. Yang, H. Li, S. Zhou: *New model for productivity test of horizontal well after hydraulic fracturing.* Acta Petrolei Sinica *27* (2006), No. 1, 89–91, 96.

[4] F. H. Zeng, G. Zhao: *The optimal hydraulic fracture geometry under non-Darcy flow effects.* J Petroleum Sci. Eng. *72* (2010), Nos. 1–2, 143–157.

[5] T. L. Li, C. L. Li, Y. Wu, Y. D. Xu: *A new way to calculate fractured horizontal wells productivity in low permeability oil reservoirs.* J China University of Petroleum *30* (2006), No. 2, 48–56.

[6] D. G. Crosby, M. M. Rahman, M. K. Rahman, S. S. Rahman: *Single and multiple transverse fracture initiation from horizontal wells.* J Petroleum Sci. Eng. *35* (2002), No. 3–4, 191–204.

[7] S. Yuan, C. L. Li, B. Q. Zeng, X. Li, H. Su: *Experimental study on optimization of fracture parameters in horizontal well fracturing.* Petroleum Drilling Techniques *38* (2010), No. 4, 99–103.

[8] F. X. Zhang, X. D. Wu, X. F. Sui: *Study on fractured horizontal well productivity for low-permeability reservoirs based on electric analogy.* Special Oil and Gas Reservoirs. *16* (2009), No. 2, 90–93.

[9] X. B. Zhang, Y. Luo, S. P. Wang, G. Q. Zhang, G. X. Shi: *A new optimization design method of integral fracturing on the basis of the equivalent percolating resistance theory.* Inner Mongolia Petrochemical Industry (2007), No. 11, 280–285.

[10] C. Q. Liu: *The flow of fluid through porous media with a vertical fracture.* Petroleum Exploration & Development *3* (1987), 63–68.

[11] I. S. Nashawi, A. H. Malallah: *Well test analysis of finite-conductivity fractured wells producing at constant bottom hole pressure.* J Petroleum Sci. Eng. *57* (2007), No. 3, 303–320.

[12] Z. X. Lang, L. H. Zhang, L. S Cheng: *Investigation on productivity of fractured horizontal well.* J University of Petroleum Beijing *18* (1994), No. 2, 43–48.

[13] W. Y. Zhu, M. Yue, H. Q. Song, J. Yang, Y. Wang, M. Yu: *Productivity model of gas flow in CBM fractured reservoirs considering desorption and diffusion.* J Basic Sci. Eng. *21* (2013), No. 5, 953–960.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 99–112

Research on incentive mechanism of supply chain based on RFID technology[1]

Xin-yu Pan[2]

Abstract. The changes of the investment behaviors of supply chain members are firstly analyzed, when retailer uses the commitment contract as incentive mechanism. Secondly, based on the retailer's commitment contract and with the supplier's reverse mechanism of price discounts, the incentive effect and optimization function of the factors to supply chain, such as the order quantity of retailer commitment, the supplier price discount coefficient, the inventory loss rate and the identification success rate of RFID technology are studied.

Key words. RFID, applied research, supply chain, incentives..

1. Introduction

The supply chain is a system composed of a large number of enterprises whose goal is to maximize their own interests by cooperation and competition. In the process of cooperation, the decentralized decision of the main body often deviates from the optimal result under the centralized decision. Therefore, the enterprises need to establish a certain mutual binding mechanism to promote the supply chain members to take positive measures to achieve information coordination, mutual incentives, and reasonable distribution of profits and other objectives. Supply chain contract is to provide a reasonable and effective incentive measures to enhance the overall performance of the supply chain, so that the optimal decisions of each subject are the similar to the centralized decision under the optimal solution. The common supply chain contracts include Revenue-sharing Contract, Quantity-discount Contract, Price-discount Sharing Contract and Buyback Contract etc.

Revenue-sharing Contract refers to the contract form that the retailer will transfer

[1]This work was financially supported by National Natural Science Foundation of China: Research on Optimal Allocation and Operation of Supply Chain Resources under the Environment of Internet of Things (No.71472134).

[2]Department of Management and Economics, Tangshan University, Hebei, Tangshan, 063009, China; e.mail: xinyupan2014@163.com

to the supplier according to a certain proportion, which usually adopts the lower wholesale price when the supplier delivers the goods. In the actual situation, the franchise mode is a typical model of supply chain revenue sharing.

Quantity-discount Contract refers to the difference in the quantity of the buyer's order quantity in the process of the transaction. The greater the number of products ordered by the buyer, the greater the wholesale price discount and the lower the price is. It is a mean of promotion by the seller.

For the above supply chain contract, it has been already formed a more mature research system and application mechanism. However, due to the limited presence of rational people, contract design cannot satisfy all situations and solve all the problems. Then it will rely on the reputation of both sides or others as guarantee mechanism outside the negotiations between the two sides' trade relationship and legal constraints [1]. Thus it puts forward a new contract form of relational contract, and gets further research and development. Williamson combines relational contract and transaction cost theory, and points out that the relational contract has a significant improvement in the investment of proprietary trading partners [2].

Commitment Contract is one form of the contractual relationship. It is a kind of informal agreement or verbal agreement. On the basis of mutual trust between the two sides, it is bound by the credibility of the utility of an individual self-enforcing mechanism to promote the contract implementation. Cachon and Lariviere believe that the information asymmetry between supply chain nodes promotes the formation of the commitment contract. They simulated the game process of the supply chain enterprises: firstly after the manufacturer knows the products demand characteristics, it puts forward contractual commitments about the initial amount of related parts; secondly the supplier visits the income and opportunity cost of the commitment of the contract, in order to determine the production; then the manufacturer determines the final parts of the order quantity with the actual production requirements; finally the supplier organizes production. Although the actual demand for the product is not certain when the contract is concluded, the study shows that under certain conditions, the manufacturer's commitment contract stimulates the production of the supplier with higher production capacity [3]. The manufacturer and customer commitment contract was studied by Durango-Cohen and Yano. Considering the partnership between guarantee punishment mechanism for the implementation of the contract, namely if supplier's commitment to supply is less than the prediction of the customers, or suppliers do not meet the commitment amount in the final delivery, then the supplier needs to burden the linear penalty cost. The results show that the contractual commitments enhance the upstream and downstream information sharing level and reduce operating costs [4]. It can be found that the implementation of contractual commitments effectively promotes the revenue sharing and risk sharing, decrease operation costs and lay a good foundation for long-term cooperation for the partners, when there are many uncertain factors of supply chain decision-making situation. The combination of traditional contract and relational contract can enhance the overall competitiveness of the supply chain system.

2. Research on incentive mechanism of retailer's commitment contract

Considering the incentive measures for the downstream retailer to adopt the commitment contract, the decentralized decision-making supply chain system is introduced, and the decision-making process of the two sides of the game is as follows [5]. First of all, retailers promise the minimum quantity to upstream suppliers in advance. The order quantity is set as λ; secondly, according to the supplier's commitment to downstream, the retailers order quantity, combined with the prediction of market demand and determine the period of the optimal production Q_{c}. Finally, in the beginning of the season, retailers of the actual order quantity meet the $\max\{\lambda, x\}$, where x denotes the products that satisfy actual market demands. In this, the supplier continues to bear the retailer product cost caused by slow-moving inventories of actual order quantity, and has a salvage value of the product. Moreover, if retailer's actual sale is among $(0, \lambda)$, the residual value and inventory cost of unsold products are managed and enjoyed by the retailers. It is assumed that the unit inventory cost and the unit residual value of surplus products stored by retailers are the same as that in suppliers, which are still h and v. On the other hand, if the retailer sells the products, he usually cannot meet the actual needs of customers. Both retailers and suppliers will suffer from the loss of reputation and other aspects, and the unit out of stock losses are g_{R} and g_{S}, because products have loss in the process of suppliers managing inventory, and RFID technology cannot completely eliminate it. Thus, there is still $q_{\mathrm{c}} = (1 - \mathrm{e} + \theta\mathrm{e})Q_{\mathrm{c}}$, where q_{c} denotes the amount of products for the normal sales that suppliers can provide when retailers use commitment contract.

In the following text, the choice of different commitments for retailers, the game process and the optimal decision-making of the two sides are discussed.

2.1. Case 1: $\lambda < q_{\mathrm{c}}$

Under the model of suppliers managing inventory, when suppliers know that the committed order quantity of retailers is less than the amount of products that can be offered for normal sales, then after using RFID technology, the expected profit function can be expressed as:

$$\begin{aligned}\Pi_{\mathrm{S}}^{1} = \omega \left[\int_{0}^{\lambda_1} \lambda_1 f(x)\,\mathrm{d}x + \int_{\lambda_1}^{q_{\mathrm{c}}^{1}} x f(x)\,\mathrm{d}x + \int_{q_{\mathrm{c}}^{1}}^{\infty} q_{\mathrm{c}}^{1} f(x)\,\mathrm{d}x \right] - \\ - g_{\mathrm{S}} \int_{q_{\mathrm{c}}^{1}}^{\infty} \left(x - q_{\mathrm{c}}^{1}\right) f(x)\,\mathrm{d}x - (c+t)Q_{\mathrm{c}}^{1} - \\ - (h-v) \left[\int_{0}^{\lambda_1} \left(q_{\mathrm{c}}^{1} - \lambda_1\right) f(x)\,\mathrm{d}x + \int_{\lambda_1}^{q_{\mathrm{c}}^{1}} \left(q_{\mathrm{c}}^{1} - x\right) f(x)\,\mathrm{d}x \right], \end{aligned} \tag{1}$$

where λ_1 and q_{c}^1 denote the retailer's commitment to order quantity and the supplier can be used for the normal sales of the product. In the formula, the first retailers use a commitment contract and the supplier can obtain the income. The second, third and fourth retailers indicate the production cost of the supplier of out of stock losses, inventory cost, residual value of products and all products. The retailer's expected return function can be expressed as

$$\Pi_{\mathrm{R}}^1 = p\left[\int_0^{q_{\mathrm{c}}^1} xf(x)\,\mathrm{d}x + \int_{q_{\mathrm{c}}^1}^{\infty} q_{\mathrm{c}}^1 f(x)\,\mathrm{d}x\right] -$$

$$- g_{\mathrm{R}}\int_{q_{\mathrm{c}}^1}^{\infty}\left(x - q_c^1\right) f(x)\,\mathrm{d}x - (h-v)\left[\int_0^{\lambda_1}(\lambda_1 - x)\,f(x)\,\mathrm{d}x\right] -$$

$$- \omega\left[\int_0^{\lambda_1}\lambda_1 f(x)\,\mathrm{d}x + \int_{\lambda_1}^{q_{\mathrm{c}}^1} xf(x)\,\mathrm{d}x + \int_{q_{\mathrm{c}}^1}^{\infty} q_{\mathrm{c}}^1 f(x)\,\mathrm{d}x\right]. \tag{2}$$

Supplier gets retailer's commits order quantity, understands the relationship between the amount of commitment and amount of production that is available to the market demand. So the supplier's expected profit is given by the condition that the first derivative of the normal product sales volume is 0, in other words

$$\frac{\partial \Pi_{\mathrm{S}}^1}{\partial q_c^1} = -\left(\omega + g_{\mathrm{S}} + h - v\right)\int_0^{q_{\mathrm{c}}^1} f(x)\,\mathrm{d}x + (\omega + g_{\mathrm{S}}) - \frac{c+t}{1-\mathrm{e}+\theta\mathrm{e}} = 0\,.$$

The amount of products for normal sales, that makes the value of the equation (1) maximum under commitment contract, can be obtained, that is $q_{\mathrm{c}}^1 = q_1^*$. Thus, when $\lambda < q_c$, the optimal production of suppliers is still the production without using any incentive mechanism, that is $Q_{\mathrm{c}}^1 = Q_1^*$, which has nothing to do with the amount that retailers committed.

For term Π_{S}^1, its first derivative with respect to λ_1, i.e. $\frac{\partial \Pi_{\mathrm{S}}^1}{\partial \lambda_1} = (\omega + h - v)\int_0^{\lambda_1} f(x)\,\mathrm{d}x > 0$ means that under the promise of the contract, supplier returns, which is accompanied by an increase in the amount of upstream partners committed to increase in the amount. From (2) it can be obtained that when $\frac{\partial \Pi_{\mathrm{R}}^1}{\partial \lambda_1} = -(\omega + h - v)\int_0^{\lambda_1} f(x)\,\mathrm{d}x < 0$, it means that the retailer's revenue with the increase in the amount of commitment reduces, and from the derivative of Π_{R}^1 with respect to q_c^1, i.e. $\frac{\partial \Pi_{\mathrm{R}}^1}{\partial q_{\mathrm{c}}^1} = -(p + g_{\mathrm{R}} - \omega)\left(\int_0^{q_c} f(x)\,\mathrm{d}x - 1\right) > 0$ it follows that the retailer revenue is an increasing function of the supplier's sales volume. Therefore, for the retailers, when the committed quantity which is expected to be ordered to the suppliers meets the condition of $\lambda < q_c$, it can be expected that the product quantity that suppliers can sell is q_1^*, so the best decision is $\lambda_1^* = 0$. Because the amount of commitment in this case does not lead to the increase in the amount of suppliers available for normal sales, it will not bring the increase in the revenue of the retailer himself.

Therefore, for the case 1, under the consideration of their own profit to be maxi-

mize, the optimal committed amount of the retailer will be $\lambda_1^* = 0$, and the supplier's best output will not change, that is, $Q_c^{1*} = Q_{1*}$. So the maximum expected return of suppliers and retailers is still $\Pi_{\mathrm{S}}^{1*} = \pi_{\mathrm{S}}^{1*}$, $\Pi_{\mathrm{R}}^{1*} = \pi_{\mathrm{R}}^{1*}$.

2.2. Case 2: $\lambda > q_{\mathrm{c}}$

When $\lambda > q_c$, the supplier that should meet the retailer's order as the production target, can achieve its revenue maximization, so the optimal production amount is $Q_c^{2*} = \frac{\lambda_{2*}}{1-e+\theta e}$. Symbols Q_c^2 and λ_2, respectively, denote in this case the supplier's production and the retailer's commitment. If the retailer forecasts that his commitment will be greater than the amount of the product that is available for normal sales, that is, $\lambda > q_c$, the expected payoff function can be expressed as

$$\Pi_{\mathrm{R}}^2 = p\left[\int_0^{\lambda_2} xf(x)\,\mathrm{d}x + \int_{\lambda_2}^{\infty} \lambda_2 f(x)\,\mathrm{d}x\right] - \omega\lambda_2 -$$

$$- g_{\mathrm{R}}\int_{\lambda_2}^{\infty}(x-\lambda_2)f(x)\,\mathrm{d}x - (h-v)\int_0^{\lambda_2}(\lambda_2 - x)f(x)\,\mathrm{d}x\,. \tag{3}$$

Because Π_{R}^2 is the strict concave function about its committed amount λ_2, it can be obtained that the committed order amount λ_2^*, which makes the value of equation (3) the largest, satisfies $F(\lambda_2^*) = \frac{p+g_{\mathrm{R}}-\omega}{p+g_{\mathrm{R}}+h-v}$. Still assuming that the market demand is subject to the uniform distribution of $(0, D)$, then in the case 2, the retailer's best commitment to the amount of $\lambda_2^* = D\frac{p+g_{\mathrm{R}}-\omega}{p+g_{\mathrm{R}}+h-v}$. At this point, the supplier's expected return function is

$$\Pi_{\mathrm{S}}^2 = \omega\lambda_2^* - g_{\mathrm{S}}\int_{\lambda_2^*}^{+\infty}(x-\lambda_2^*)\,f(x)\,\mathrm{d}x - (c+t)\frac{\lambda_2^*}{1-\mathrm{e}+\theta\mathrm{e}}\,. \tag{4}$$

The maximum expected revenue functions of the retailer and supplier are then simplified, respectively:

$$\Pi_{\mathrm{R}}^{2*} = \frac{D}{2}(p+g_{\mathrm{R}}+h-v)\left(\frac{p+g_{\mathrm{R}}-\omega}{p+g_{\mathrm{R}}+h-v}\right)^2 - \frac{D}{2}g_{\mathrm{R}}\,, \tag{5}$$

and

$$\Pi_{\mathrm{S}}^{2*} = D\frac{p+g_{\mathrm{R}}-\omega}{p+g_{\mathrm{R}}+h-v}\left(\omega + g_{\mathrm{S}} - \frac{c+t}{1-\mathrm{e}+\theta\mathrm{e}} - \frac{g_{\mathrm{S}}}{2}\times\frac{p+g_{\mathrm{R}}-\omega}{p+g_{\mathrm{R}}+h-v}\right) - \frac{D}{2}g_{\mathrm{S}}\,. \tag{6}$$

Comprehending the analyses of Case 1 and Case 2, the range of the committed order quantity of the retailers is $\{\lambda_2^*, 0\}$. When the committed order quantity of the retailers is 0, the optimal production of the suppliers is $Q_c^{1*} = Q_{1*}$. When the committed order quantity of the retailers is positive, and ordering $\lambda = \lambda_{2*}$, the suppliers organize production to fulfill the orders, and the most optimal production is $Q_c^{2*} = \frac{\lambda_2^*}{1-\mathrm{e}+\theta\mathrm{e}}$. Only when $\lambda_2^* > q_1^*$ and $\Pi_{\mathrm{R}}^{2*} > \pi_{\mathrm{R}}^{1*}$ are both established, the

retailer will select the optimal commitment amount of λ_{2*}. To simplify these two inequalities, the condition of the optimal commitment amount of the retailers is λ_2^*

$$\frac{p+g_{\mathrm{R}}-\omega}{p+g_{\mathrm{R}}+h-v} > \frac{(1-\mathrm{e}+\theta\mathrm{e})(\omega+g_{\mathrm{S}})-c-t}{(1-\mathrm{e}+\theta\mathrm{e})(\omega+g_{\mathrm{S}}+h-v)} \cdot \left[1+\frac{(1-\mathrm{e}+\theta\mathrm{e})(h-v)+c+t}{(1-\mathrm{e}+\theta\mathrm{e})(\omega+g_{\mathrm{S}}+h-v)}\right]. \tag{7}$$

If the satisfying parameter set is recorded as $\Re_1$, the values of all variables are within the set. Retailers have the will to forwardly choose the positive commitment amount, and compared with the situation that commitment contract is not imported, the suppliers' expected profit has been improved. Moreover, the profit changes of suppliers increase in most cases.

3. Supplier price discount mechanism based on commitment contract

On the basis of retailers adopting commitment contract, considering the suppliers giving commitment amount a price discount of τ, then the wholesale price of advanced order part is $\tau\omega$. At this time, the members of the supply chain decision-making process [6] are as follows: first, the suppliers that need to determine the commitment ordering products can enjoy the discount coefficient τ; second, retailers reference price discount coefficient determined its commitments to the supplier minimum order quantity λ; then, suppliers according to the commitment ordering quantity to determine the optimal yield Q_{d} (at this time, for the amount of product sales and normal $q_{\mathrm{d}} = (1-\mathrm{e}+\theta\mathrm{e})Q_{\mathrm{d}}$). Finally, after the beginning of selling season, the actual order quantity of retailers should meet $\max\{\lambda, x\}$, x being the actual market demand. According to the following two cases, the expected return of the two parties in the supplier's price discount mechanism is discussed separately.

3.1. Case 3: $\lambda < q_{\mathrm{d}}$

Under VMI model [7], [8], when the supplier learned that the downstream partner's commitment order amount would be within the range, then by using RFID technology, its expected return function is

$$\Pi_{\mathrm{S}}^3 = \omega\left[\int_0^{\lambda_3}\lambda_3 f(x)\,\mathrm{d}x + \int_{\lambda_3}^{q_{\mathrm{d}}^1} x f(x)\,\mathrm{d}x + \int_{q_{\mathrm{d}}^1}^{+\infty} q_{\mathrm{d}}^1 f(x)\,\mathrm{d}x\right] - g_{\mathrm{S}}\int_{q_{\mathrm{d}}^1}^{+\infty}\left(x-q_{\mathrm{d}}^1\right) f(x)\,\mathrm{d}x - \frac{c+t}{1-\mathrm{e}+\theta\mathrm{e}}q_{\mathrm{d}}^1 - (1-\tau)\omega\lambda_3 - (h-v)\left[\int_0^{\lambda_3}(q_{\mathrm{d}}^1-\lambda_3)f(x)\,\mathrm{d}x + \int_{\lambda_3}^{q_{\mathrm{d}}^1}(q_{\mathrm{d}}^1-x)f(x)\,\mathrm{d}x\right]. \tag{8}$$

The q_{d}^1 and λ_3, respectively, refer to the amount of products available for sale and retailer's early commitment. For suppliers, because Π_{S}^3 is the strictly concave function of q_{d}^1, so it makes its derivative 0, that is $\frac{\partial \Pi_{\mathrm{S}}^3}{\partial q_{\mathrm{d}}^1} = 0$, it can be obtained that , under the mechanism that retailer using the commitment contract and the supplier using the price discount, the product quantity for the normal sales that making expect profit of the upstream suppliers maximum is $q_{\mathrm{d}}^{1*} = q_1^*$. So when $\lambda < q_{\mathrm{d}}$, the supplier will eventually still choose Q_1^* as the most excellent production of this period, the value of the retailer's commitment to order and supplier price discounts are not related.

On the other hand, for the retailer, when $\lambda < q_{\mathrm{d}}$, its revenue function can be expressed as

$$\Pi_{\mathrm{R}}^3 = p\left[\int_0^{q_{\mathrm{d}}^1} x f(x)\,\mathrm{d}x + \int_{q_{\mathrm{d}}^1}^{+\infty} q_{\mathrm{d}}^1 f(x)\,\mathrm{d}x\right] - g_{\mathrm{R}}\int_{q_{\mathrm{d}}^1}^{+\infty}\left(x - q_{\mathrm{d}}^1\right) f(x)\,\mathrm{d}x +$$

$$+ (1-\tau)\omega\lambda - (h-v)\int_0^{\lambda_3}(\lambda_3 - x)\, f(x)\,\mathrm{d}x -$$

$$- \omega\left[\int_0^{\lambda_3}\lambda_3 f(x)\,\mathrm{d}x + \int_{\lambda_3}^{q_{\mathrm{d}}^1} x f(x)\,\mathrm{d}x + \int_{q_{\mathrm{d}}^1}^{+\infty} q_{\mathrm{d}}^1 f(x)\,\mathrm{d}x\right]. \tag{9}$$

Calculating the first-order derivative of Π_{R}^3 with respect to q_{d}^{1*}, $\frac{\partial \Pi_{\mathrm{R}}^3}{\partial q_{\mathrm{d}}^1} > 0$ can be obtained. It can be seen that the retailer's profit is an increasing function of the product quantity available for sale of the suppliers. Ordering $\frac{\partial \Pi_{\mathrm{R}}^3}{\partial \lambda_3} = -(\omega + h - v)\int_0^{\lambda_3} f(x)\,\mathrm{d}x + (1-\tau)\omega = 0$ it can be got that, that in case 3, the retailer's optimal commitment amount is $\lambda_3^* = F^{-1}\left(\frac{(1-\tau)\omega}{\omega+h-v}\right)$, the commitment amount is also the commitment amount that making the suppliers' expected profit minimum in case 3.

On the basis of $x \sim U(0, D)$, equations (8) and (9) can be simplified as follows

$$\Pi_{\mathrm{R}}^{3*} = q_{1*}(p + g_{\mathrm{R}} - \omega)(1 - \frac{q_{1*}}{2D}) - g_{\mathrm{R}}\frac{D}{2} - (\omega + h - v)\frac{\lambda_{3*}^2}{2D} + (1-\tau)\omega\lambda_3^*, \tag{10}$$

$$\Pi_{\mathrm{S}}^{3*} = \frac{1}{2D}\left(\omega + g_{\mathrm{S}} + h - v\right) q_{1*}^2 - \frac{D}{2} g_{\mathrm{S}} + (\omega + h - v)\frac{\lambda_{3*}^2}{2D} - (1-\tau)\omega\lambda_{3*}. \tag{11}$$

When comparing the retailer's commitment to the quantity of products provided by the supplier can be used for the normal sales quantity, the increase of the supplier's price discount mechanism, the income of the members of the supply chain is changed, as follows

$$\Pi_{\mathrm{R}}^{3*} - \pi_{\mathrm{R}}^{1*} = \frac{\lambda_3^*}{2}(1-\tau)\omega = \frac{D\left[(1-\tau)\omega\right]^2}{2\left(\omega + h - v\right)} > 0,$$

$$\Pi_{\mathrm{S}}^{3*} - \pi_{\mathrm{S}}^{1*} = -\frac{\lambda_3^*}{2}(1-\tau)\omega = -\frac{D\left[(1-\tau)\omega\right]^2}{2\left(\omega + h - v\right)} < 0.$$

Relative to case 1, the incentive mechanism taken by the supplier can increase the retailer maximum expected profit, and decline its own expected revenue. The higher price discount τ takes, the smaller retailers' profit increment is. Therefore, when $\lambda < q_{\mathrm{d}}$, the supplier's optimal decision is not to provide a price discount; and the retailer's optimal decision is $\lambda_3^* = 0$. At the same time the supplier's most excellent production has not changed, is still Q_{1^*}.

3.2. Case 4: $\lambda > q_{\mathrm{d}}$

When the supplier knows the retailer's commitment order quantity is greater than its production which can be used for normal sales, the supplier will make to order, so the optimal production quantity is $Q_{\mathrm{d}}^2 = \lambda_4/(1 - \mathrm{e} + \theta\mathrm{e})$, Q_{d}^2 and λ_4 represent the supplier's production and retailer's commitment order quantity in case 4. When $\lambda > q_{\mathrm{d}}$, the upstream supplier's revenue function is

$$\Pi_{\mathrm{R}}^4 = p\left[\int_0^{\lambda_4} x f(x)\,\mathrm{d}x + \int_{\lambda_4}^{+\infty} \lambda_4 f(x)\,\mathrm{d}x\right] -$$

$$- g_{\mathrm{R}} \int_{\lambda_4}^{+\infty} (x - \lambda_4)\, f(x)\,\mathrm{d}x - (h - v)\int_0^{\lambda_4} (\lambda_4 - x) f(x)\,\mathrm{d}x - \tau\omega\lambda_4\,. \tag{12}$$

In the same way, when the commitment order quantity λ_4^* meets $F(\lambda_4^*) = (p + g_{\mathrm{R}} - \tau\omega)/(p + g_{\mathrm{R}} + h - v)$, the retailers expect the maximum revenue. Apparently $\lambda_4^* > \lambda_2^*$, which means that the supplier's price discount mechanism can improve the retailer's commitment order quantity relative to case 2. If it is still assumed that the market demand is subject to the uniform distribution of $(0, D)$, then the optimal commitment quantity of the retailer, the maximum expected return of the supplier, and the maximum expected profit of the retailer are, respectively

$$\lambda_4^* = D\frac{p + g_{\mathrm{R}} - \tau\omega}{p + g_{\mathrm{R}} + h - v}\,, \tag{13}$$

$$\Pi_{\mathrm{S}}^{4*} = D\frac{p + g_{\mathrm{R}} - \tau\omega}{p + g_{\mathrm{R}} + h - v}\left(\tau\omega + g_{\mathrm{S}} - \frac{c + t}{1 - \mathrm{e} + \theta\mathrm{e}} - \frac{g_{\mathrm{S}}}{2}\cdot\frac{p + g_{\mathrm{R}} - \tau\omega}{p + g_{\mathrm{R}} + h - v}\right) - \frac{D}{2}g_{\mathrm{S}}\,, \tag{14}$$

$$\Pi_{\mathrm{R}}^{4*} = \frac{D}{2}(p + g_{\mathrm{R}} + h - v)\left(\frac{p + g_{\mathrm{R}} - \tau\omega}{p + g_{\mathrm{R}} + h - v}\right)^2 - \frac{D}{2}g_{\mathrm{R}}\,. \tag{15}$$

3.3. Result analysis

3.3.1. Situation 1. Comparison of the price discount mechanism before and after the retailer's maximum expected return changes. From (5) and (15) we get

$$\Pi_{\mathrm{R}}^{4*} - \Pi_{\mathrm{R}}^{2*} = \frac{D\omega}{2(p + g_{\mathrm{R}} + h - v)}\left[\omega\tau^2 - 2(p + g_{\mathrm{R}})\tau - \omega + 2(p + g_{\mathrm{R}})\right]\,.$$

Obviously, when $(p+g_{\mathrm{R}})/\omega > 1$, the retailer's expected return decreases with the increase of the supplier's price discount coefficient, so when $\tau < 1$, the maximum expected return of the retailer is increased, which is similar to the conclusion of Case 3.

3.3.2. Situation 2. Compare the maximum expected return of the supplier before and after the use of the price discount mechanism.

By comparing (6) and (14), when the other factors are constant, Π_{S}^{4*} is the function about the price discount coefficient τ function, therefore, Π_{S}^{4*} about τ, and, therefore, its first-order derivative is

$$\frac{\partial \Pi_{\mathrm{S}}^{4*}}{\partial \tau} = \frac{D\omega}{p+g_{\mathrm{R}}+h-v} \cdot$$

$$\cdot \left[-\omega \left(2 + \frac{g_{\mathrm{S}}}{p+g_{\mathrm{R}}+h-v} \right) \tau + p + g_{\mathrm{R}} - g_{\mathrm{S}} + \frac{c+t}{1-\mathrm{e}+\theta \mathrm{e}} + \frac{g_{\mathrm{S}}\left(p+g_{\mathrm{R}}\right)}{p+g_{\mathrm{R}}+h-v} \right] .$$

As

$$\frac{\partial^2 \Pi_{\mathrm{S}}^{4*}}{\partial \tau^2} = \frac{-D\omega^2}{p+g_{\mathrm{R}}+h-v}(2 + \frac{g_{\mathrm{S}}}{p+g_{\mathrm{R}}+h-v}) < 0 \,,$$

put $\frac{\partial \Pi_{\mathrm{S}}^{4*}}{\partial \tau} = 0$, and then we get the maximum price discount below which can make the supplier's expected return Π_{S}^{4*} maximum

$$\tau^* = \frac{p+g_{\mathrm{R}}-g_{\mathrm{S}}+\frac{c+t}{1-\mathrm{e}+\theta \mathrm{e}}+\frac{g_{\mathrm{S}}(p+g_{\mathrm{R}})}{p+g_{\mathrm{R}}+h-v}}{\omega(2+\frac{g_{\mathrm{S}}}{p+g_{\mathrm{R}}+h-v})} .$$

When $0 < \tau^* < 1$, then compared to the situation two, the supplier adopts the price discount mechanism and the maximum expected return will be increased.

However, only when $\lambda_4^* > q_1^*$ and $\Pi_{\mathrm{R}}^{4*} > \pi_{\mathrm{R}}^*$ are both established, the retailer will choose λ_4^* for its commitment amount. By simplifying these two inequalities, the condition that the retailer chooses the optimal commitment amount for λ_4^* is:

$$\frac{\left(p+g_{\mathrm{R}}-\tau\omega\right)^2}{p+g_{\mathrm{R}}+h-v} > \frac{\left(1-\mathrm{e}+\theta \mathrm{e}\right)\left(\omega+g_{\mathrm{S}}\right)-c-t}{\left(1-\mathrm{e}+\theta \mathrm{e}\right)\left(\omega+g_{\mathrm{S}}+h-v\right)} \cdot$$

$$\cdot \left(p+g_{\mathrm{R}}-\omega\right)\left[1+\frac{\left(1-\mathrm{e}+\theta \mathrm{e}\right)\left(h-v\right)+c+t}{\left(1-\mathrm{e}+\theta \mathrm{e}\right)\left(\omega+g_{\mathrm{S}}+h-v\right)}\right] . \tag{16}$$

If the parameter set that satisfies (16) is recorded as $\Re_2$, when all the parameters are in the collection, the retailer will take the initiative to select the λ_4^*'s commitment order quantity. Obviously $\Re_2 \supset \Re_1$. Also if

$$0 < \tau^* = \frac{p+g_{\mathrm{R}}-g_{\mathrm{S}}+\frac{c+t}{1-\mathrm{e}+\theta \mathrm{e}}+\frac{g_{\mathrm{S}}(p+g_{\mathrm{R}})}{p+g_{\mathrm{R}}+h-v}}{\omega(2+\frac{g_{\mathrm{S}}}{p+g_{\mathrm{R}}+h-v})} < 1$$

is established, the manufacturer's profit will yield higher than the price discount

mechanism is adopted, namely in the retailer commitment contract basis, the supplier adopts reasonable price discount mechanism can realize the downstream of the game on the side of the Pareto improvement [9].

4. Numerical analysis

The method of numerical simulation is used to analyze and explain the investment decision-making of RFID in the supply chain members. Assume that the market demand for the product obeys a uniform distribution $U(0, 2000)$ and the unit cost of wholesale products, value and price, respectively, are $\omega = 26$, $v = 9$ and $p = 35$. The supplier's unit production costs and inventory costs are $c = 16$ and $h = 4$. The two sides of the game out of stock losses are $g_{\mathrm{S}} = 6$ and $g_{\mathrm{R}} = 5$. The RFID tag cost $t = 1$. According to the calculation and analysis, we can see that the inventory loss rate of the supplier needs to meet the $0 < \mathrm{e} < 0.48$ and the cooperation intention of the VMI model can be reached. We can distinguish two cases

4.1. Situation 1

When the retailer uses a commitment contract, the effect of inventory loss rate and RFID technology identification success rate affects the revenue of supply chain members.

From the previous analysis, we can conclude that when retailer takes contractual commitments to encourage suppliers to implement RFID technology, if the commitment amount is small (such as Case 1) to downstream retailers, the choice of the optimal strategy is not the commitment contract [10]. Therefore, this situation cannot be verified numerically. If retailer's commitment ordering quantity satisfies the assumptions in Case 2, it can be shown by the numerical simulation that the retailer's optimal commitment order quantity $\lambda_2^* = 800$, this value is uncorrelated to the rate of supplier' shrinkage and the rate of RFID's identification. Figure 1 shows the impact of the vendor's inventory loss rate and the RFID technology on the identification success rate of the inventory inaccuracy problem on the optimal commitment amount of retailer selection Case 2. As shown in Fig. 1, when the rate of supplier' shrinkage and the rate of RFID's identification are below the curve, the commitment contract used by retailer will increase its maximal expected revenue, and make itself more positive to adopt incentive mechanism to promote the suppliers to implement RFID . Clearly, only the supplier's loss rate is at a high level ($\mathrm{e} > 0.345$), while the RFID technology is at a low level ($\theta < 0.3$) when the recognition success rate is at a low level. But in the actual situation, the recognition success rate of RFID technology is obviously not always at the level of 30 %.

On the other hand, in terms of different RFID identification success rate (order $\theta = 0.3$, $\theta = 0.6$ and $\theta = 0.9$, respectively), the impact of the retailer's commitment to its maximum expected return is shown in Fig. 2. Obviously, the retailer's commitment contract can significantly improve the supplier's maximum expected return, and the supplier's income difference increases with the increase of inventory loss rate, showing a trend of first increasing and then decreasing. With the increase

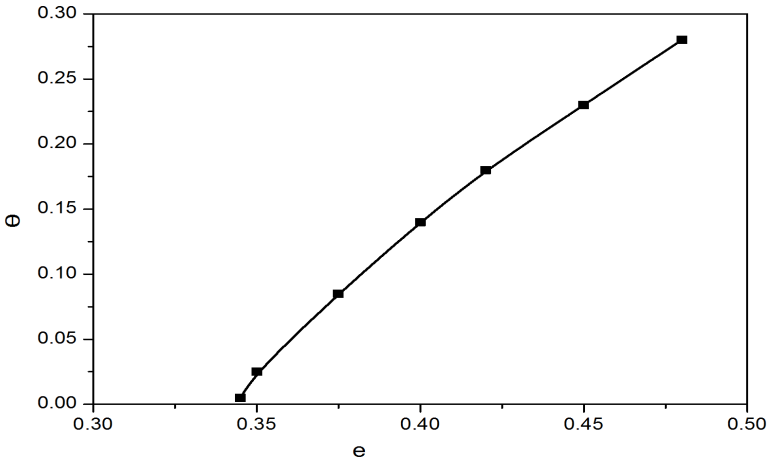

Fig. 1. Conditions for the retailer to select the optimal commitment order quantity in Case 2

of the recognition success rate of the RFID technology, the difference between the maximum expected return and the difference of the maximum expected return of the suppliers is basically a downward trend compared to the situation without using any incentive mechanism. This shows that the higher the recognition success rate of RFID technology is based on the above numerical assumptions, the more stable the retailer's incentive mechanism is to the supplier's effect.

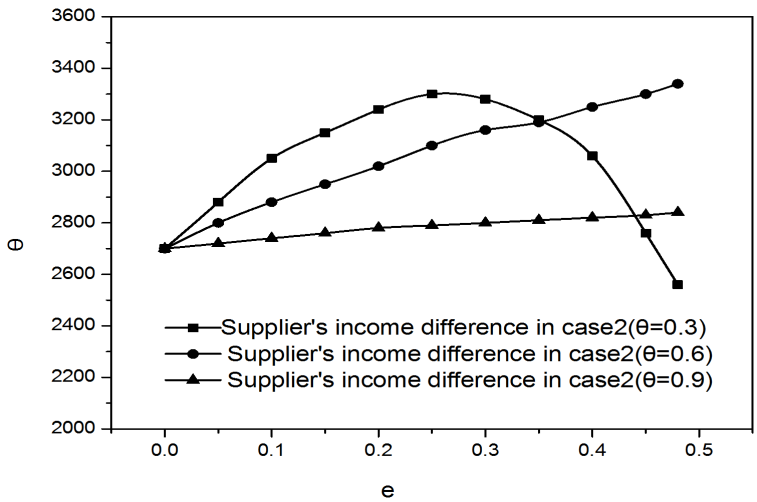

Fig. 2. Effect of different RFID recognition success rate on supplier's benefit in Case 2

It can be seen, based on the above numerical assumptions, that in case 2 retailers can benefit from the promise of the contract conditions being more limited, and in most cases, the supplier can get the expected revenue growth. Therefore, the promotion of RFID technology in the supply chain member enterprises only relies on the incentive contract of the retailer, cannot achieve the expected effect and the dominant side may be the first to abandon the implementation of RFID technology. This is why there is few successful cases that commitment contract or minimum order quantity contract is used by one side.

4.2. Situation 2

We will discuss the effect of the price discount coefficient on the income of the members of the supply chain when the supplier adopts the price discount mechanism in this part.

Case 3 shows that when the supplier uses the mechanism of price discounts, the more discount factor increases, the less retailer's profits increment is, and the less supplier's revenue decrement is. The trend of supplier's revenue decrement is more obvious. So this paper does not carry out numerical analysis for case 3. According to the above numerical simulation, in Case 4 it can be obtained that the supplier's optimal discount coefficient is $\tau* = 1.03 > 1$, which shows that the supplier's optimal decision is not to use the price discount mechanism.Figure 3 shows that the trend of supplier's expected revenue compared with the situation without contract in Case 4. Apparently, combination contract by both sides can increase the supplier's maximal expected revenue, but the changes of income is less than the numerical simulation results in Case 2. When the recognition success rate of RFID technology is maintained at a certain level, the amount of revenue increases with the increase of inventory losses, showing a trend of first increase and then decrease. At the same time, with RFID recognition success rate it continues to increase and income increased steadily, when $\theta = 0.9$ basically shows slowly increasing trend.

4.3. Situation 3

We will discuss the effect of the inventory loss rate and the RFID identification success rate on the retailer's parameter set when the supplier adopts the price discount mechanism.

Order $\tau* = 0.9$, the case 4 shows that, when the other factors are constant, the range of $\Re_1$ is larger than $\Re_2$ obviously, as shown in Fig. 4. It means that adopting price discount mechanism by suppliers can encourage the retailer to order in advance significantly. With the increasing of supplier price discounts, retailer are more willing to commitment contract and make more profit from the cooperation.

It can be seen that, based on the above numerical assumptions, that in Case 4 the value range that the retailer can benefit from the commitment contract significantly, increased. Moreover, in most cases, the supplier can get the growth of expected profit. Therefore, relying on the joint efforts of both sides he can significantly accelerate the implementation and promotion of the RFID technology in the supply

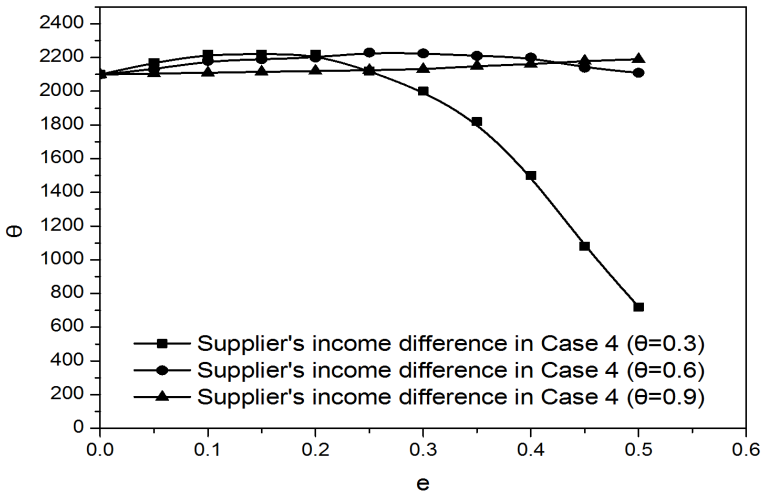

Fig. 3. Effect of different RFID recognition success rate on supplier's revenue in Case 4

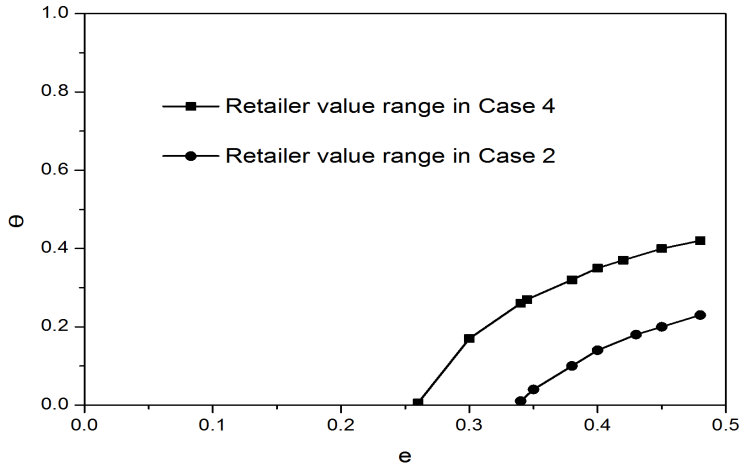

Fig. 4. The conditions for the retailer to adopt the promise contract in Case 4

chain system.

5. Conclusion

This paper studies the incentive mechanism of supply chain members to promote the RFID technology. First of all, retailers consider long-term cooperation to take

the commitment to promote the implementation RFID technology of the upstream supplier. Second, the supplier uses the incentive mechanism of price discounts in the amount of retail's commitment in advance. The results of theoretical research and numerical analysis show that when the retailer's commitment is smaller, the maximum expected return of both parties is consistent with that of no contract (case 1 and case 3). When the amount of retail's commitment is large, adopting $\lambda_2^* = D(p + g_{\mathrm{R}} - \omega)/(p + g_{\mathrm{R}} + h - v)$ as the retail's advanced order quantity, and meeting the requirement of $\Re_1$ for all the parameters can realize Pareto improvement for both sides. As well as using $\lambda_4^* = D(p + g_{\mathrm{R}} - \tau\omega)/(p + g_{\mathrm{R}} + h - v)$ as the retail's advanced order quantity, and $\tau*$ $(0 < \tau* < 1)$ as the supplier's optimal discount coefficient, and meeting the requirement of $\Re_2$ for all the parameters can realize Pareto improvement for both sides.

References

[1] T. Baldenius, J. Glover, H. Xue: *ARelational contracts with and between agents.* J Accounting and Economics *61* (2016), 369–390.

[2] O. E. Williamson: *The economic institutions of capitalism.* Free Press, New York, 1998.

[3] G. P. Cachon, M. A. Lariviere: *Contracting to assure supply: How to share demand forecasts in a supply chain.* Management Science *47* (2001), No. 5, 629–646.

[4] E. J. Durango-Cohen, C. A. Yano: *Supplier commitment and production decisions under a forecast-commitment contract.* Management Science *52* (2006), No. 1, 54–67.

[5] S. Y. Xie: *Economic game theory*, third edition. Fudan University Press, Shanghai, 2007.

[6] S. M. Disney, D. R. Towill: *The effect of vendor managed inventory (VMI) dynamics on the bullwhip effect in supply chains.* IJ Production Economics *85* (2003), No. 2, 199–215.

[7] J. H. Cai, W. L. Huang, G. G. Zhou: *Research on VMI model based on revenue sharing contract.* China Management Science *14* (2006), No. 4, 108–113.

[8] Y. Dong, K. F. Xu: *A supply chain model of vendor managed inventory. Transportation Research Part E 38* (2002), No. 2, 75–95.

[9] D. Liu, Z. X. Liu, C. Z. Zheng: *Replenishment policy for VMI and TPL supply chain with single supplier and two manufacturers.* Control and Decision *28* (2013), No. 1, 20–28.

[10] F. Bernstein, F. Chen, A. Federgruen: *Coordinating supply chains with simply pricing schemes: The role of vendor-managed inventories.* Management Science *52* (2006), No. 10, 1483–1492.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 113–122

SVM-based bridge health condition evaluation forecasting function method and evaluation criteria[1]

Jianying Ren[2], Mubiao Su[3]

Abstract. A new method for bridge structure health condition evaluation is proposed, i.e. the forecasting function method, and probes into the determination of evaluation criteria. First, Least Square Support Vector Regression (LS-SVR) is used to establish the forecasting function. Then, the confidence interval is used to determine the first evaluation criterion $\epsilon 1$ and the second evaluation criterion $\epsilon 2$ is determined for bridge structure health condition evaluation. Finally, the difference between measured node deflection and predicted node deflection is calculated and compared with both evaluation criteria for purpose of bridge structure health condition evaluation.

Key words. Deflection, forecasting function, least square support vector regression (LS-SVR), evaluation criterion, steel truss bridge.

1. Introduction

In view of the fact that frequent bridge accidents like sudden break-off etc. led to material casualties and property losses, people begin to pay close attention to bridge engineering safety [1]. Many factors may cause bridge collapse, but it is certain that the long-term effect of loads and the material fatigue, corrosion and ageing together with the lack of timely maintenance bring about internal damage accumulation and resistance deterioration of bridge structure, thereby resulting in accidents [2]. Effective means shall be timely employed to monitor and evaluate the health condition of a large number of bridge structures in service and other infrastructures so as to identify the structure damage in a timely manner and give

[1]This study is supported by National Natural Science Foundation of China (NSFC) (51278315, 11472180, 11602153), Natural Science Foundation of Hebei Province (E2015210020) and New Century Talent Foundation of Ministry of Education under Grant (NCET-13-0913).

[2]Dept. of Engineering Mechanics, Shijiazhuang Tiedao University, Shijiazhuang Hebei, 050043, China

[3]Structure Health Monitoring and Control Institute, Shijiazhuang Tiedao University, Shijiazhuang Hebei, 050043, China

early warning on possible disasters so as to avoid tragedies [3]–[5].

China has established health monitoring systems at scores of bridges like Jiangyin Bridge, Humen Bridge, Nanjing Yangtze River Bridge, Runyang Bridge, Sutong Bridge, Wuhu Yangtze River Bridge and Hangzhou Bay Sea-Crossing Bridge since the 1990s, having gathered abundant research findings [6]. Principal indicators for bridge structure health condition evaluation are structural dynamic characteristic indicators like frequency, vibration mode and modal damping [7] etc., which is vulnerable to external environmental noise, and their practical application effectiveness leaves much to be desired.

For the above reasons, this paper employs the latest data mining method, i.e. SVM based on bridge structure deflection to propose a new method for bridge structure health condition evaluation on basis of reference [8], and probes into the method for determining evaluation criteria.

2. Rationale of forecasting function method

2.1. Forecasting function model

Obtain substantive measured data samples of bridge structures in good condition using the bridge structural health monitoring system, and establish the forecasting function relationship between dependent variables and independent variables of bridge structure using such data mining methods as Support Vector Regression Machine

$$\{y\} = \{f(P, x, t,) \cdots\}\,, \tag{1}$$

where $\{y\}$ is a dependent variable, $f(\cdot)$ means a mapping function, P represents the load condition, x stands for the load position, T means the environmental factor, and so on.

2.2. Evaluation criteria $\epsilon 1i$ *and* $\epsilon 2i$

The determination of evaluation criteria $\epsilon 1i$ and $\epsilon 2i$, ($i = 1, 2, \cdots, n$ referring to the number of measuring point), as the key technique for forecasting function method, is currently considered one of the major and difficult issues that must be addressed during bridge structure health condition evaluation.

2.2.1. Determination of $\epsilon 1i$. The value of $\epsilon 1i$, $i = 1, 2, \cdots, n$ is equivalent to the "normal value" specified in the railway bridge assessment specification. Assign significance level α to the residual error between the predicted value and measured value that obey normal distribution, then $P\alpha = \alpha$ should be of low probability. Determine the value of m through table look-up based on significance level of α and the distribution regularity of residual error, and the confidence interval should be

$$[\hat{y}_i - mS_i\,, \hat{y}_i + mS_i]\,, \quad i = 1, 2, \cdots, L, \cdots n\,, \tag{2}$$

where $\hat{y}_i$ is the predicted value of variable dependent on intact bridge structure and S_i denotes the standard deviation whose value is given by

$$S_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(y_{ij} - \hat{y}_{ij})^2}\,. \tag{3}$$

wherey_{ij} means the jth measured value of the ith dependent variable, $\hat{y}_{ij}$ is the jth predicted value of the ith dependent variable and N stands for the number of samples. The measured value is considered normal when it falls in confidence in interval (2), otherwise, an anomaly is considered to exist. Thus, it can be seen that the evaluation criterion $\epsilon 1i$ could be determined using the following formula

$$\epsilon 1i = mS_i\,, \quad i = 1, 2, \cdots, L, \cdots n\,. \tag{4}$$

2.2.2. Determination of $\epsilon 2i$*.* According to the allowable values (e.g. allowable deflection and stress) of bridge structure response allowed by relevant bridge design codes and examination specifications, the structure is considered unsafe when the bridge structural health monitoring system detects that the measured value of a certain dependent variable exceeds the allowable value. However, the said allowable values are normally safety thresholds for structural response of designated section (mid-span section). Nevertheless, since the load applied is smaller than designed load during practical operation of bridge structure, bridge structure suffers from a certain degree of damage, and the response values of bridge structure detected by monitoring system normally fail to meet specification-defined safety threshold. As a result, it is impossible for monitoring system to release early warning information in a timely manner. To the end, this paper plans to employ the finite element model that is the closest to the practical structure after model modification, and to appropriately discount the integral stiffness of such finite element model (for example, 90% of stiffness in good condition after discounting indicates a damage degree by 10%).

2.3. Forecasting function method for bridge structure health condition evaluation

Figure 1 shows the process of forecasting function method for bridge structure health condition evaluation.

3. Support vector machine

Support Vector Machine [9] (SVM), as proposed by Vapnik et al. during period from 1992 to 1995, is the latest and most practical part of statistical learning theory and the youngest and fastest growing data mining method. SVM is divided into Support Vector Classification (SVC) and Support Vector Regression (SVR) by application. Since process for establishing forecasting function in this paper is a regression problem, the least squares-support vector regression (LS-SVR) [9] is used

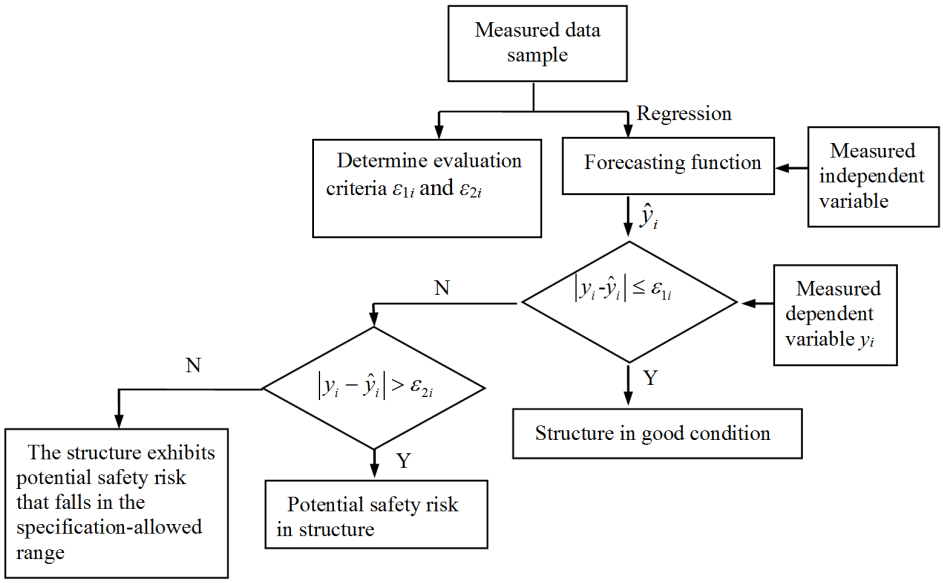

Fig. 1. Flow chart of forecasting function method

to build the forecasting function. As a variant of general support vector regression method, LS-SVR replaces inequality constraint with equality constraint and converts the solving problem of quadratic programming into solving problem of system of linear equations, thereby significantly simplifying the calculation and improving the training rate.

4. Calculation example

In view of the fact that the impact coefficient for long-span railway bridge structure is relatively small ($1 + \mu \le 1.05$ [7]), when a train passes through the bridge, the train load is taken as quasi-static load for field test to monitor the deflection of bridge structure (dependent variable) as well as the substantive data samples of train velocity, axle weight and ambient temperature etc. (independent variables), and to analyze the law that the deflection of bridge structure changes along with the change in such independent variables as train velocity, axle weight and ambient temperature by establishing the forecasting function relationship between dependent variable and independent variable using data mining method based on monitoring data.

A 64 m single line rail simply-supported steel truss bridge is taken as an example to verify the application of forecasting function method for bridge structure health condition evaluation. In order to describe the implementation procedure of this method, a main beam piece of the 64 m single line rail simply-supported steel truss bridge is used as research object (Fig. 2) to simplify the calculation. The dependent

variable is the deflection of lower chord nodes 2–8. The independent variables are load condition and load position in this paper. Since this is a simulation bridge structure, all data samples are obtained by adding computed result of finite element model to noise.

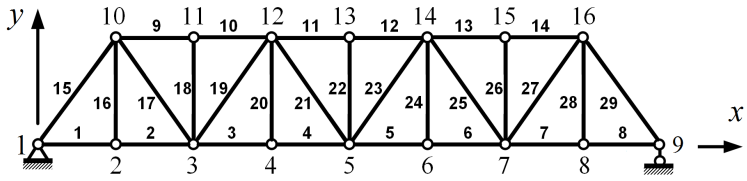

Fig. 2. Finite element model

4.1. Finite element model

The plane rod element is used to build the plane finite element calculation model for a piece of main truss. There are in total 16 nodes (Fig. 2, larger normal letters), and the rod piece between each two nodes is taken as a unit. A total of 29 rod units (Fig. 2, smaller boldface letters) are determined, the unit elasticity modulus $E = 210\,\text{GPa}$, the sectional dimension is as bridge drawing. Constrain the vertical and longitudinal line displacement simulation fixed hinge support along the bridge at node 1, and constrain the vertical line displacement simulation movable hinged support at node 9.

Load condition: single locomotive, dual locomotives, triple locomotives, single train (1 locomotive connected to 8 carriages, this arrangement makes it possible to cover the bridge and keep such a state for a certain period of time), dual trains (2 locomotives connected to 8 carriages). The locomotives are of Dongfeng-IV with an axle weight of 23 t as shown in Fig. 3, the carriage is of model C62 with an axle weight of 20.15 t as shown in Fig. 4 [10]. The effect of vehicle-bridge coupled vibration is not taken into account, the deflection at each lower chord node (nodes 2–8) is calculated by simplifying train load to a series of static loads moving on the bridge.

4.2. Establishment of deflection forecasting function for each lower chord node

Five load conditions are taken into consideration. The train moves from left to right with a loading step of 4 m, the calculation start when the first front wheel of the train gets in contact with the bridge (the position of the first wheel set is 0) until the last rear wheel set gets off the bridge (the position of the first wheel set: bridge length + train length). Independent variables of forecasting function: quantity of carriages x_1 (0 or 8), quantity of locomotives x_2 (1, 2 or 3), position of the first wheel set of train x_3 (value range from 0 to bridge length + train length), locomotive axle weight and carriage axle weight are not taken as independent variables since they are constants, the dependent variable is the deflection of seven lower chord nodes. Each

node is provided with 178 training data samples. In order to check the generalization of forecasting function, the load step for test sample calculation is 5 m (different from the step chosen for training sample, i.e. 4 m) so as to ensure most data in test sample are not included in training sample (142 in total). Use LS-SVR and training set for the regression of deflection forecasting function for each node (7 forecasting functions in total).

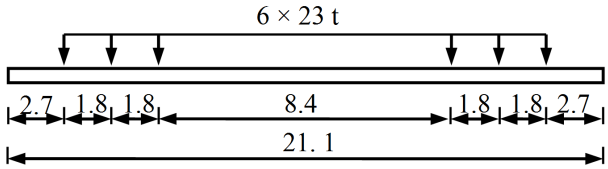

Fig. 3. Schematic diagram of axle weight and wheel base of Dongfeng-IV locomotive (dimensions in m)

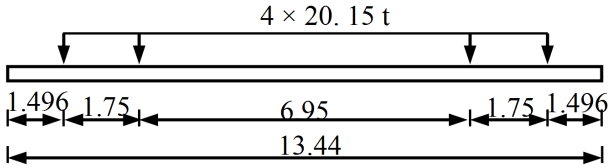

Fig. 4. Schematic diagram of axle weight and wheel base of carriage C62 (dimensions in m)

Substitute the independent variable $\{x_{i1}, x_{i2}, x_{i3}\}$, $(i = 1, 2, \cdots, 142)$ in test set sample into the deflection forecasting function for above 7 lower chord nodes, respectively, to work out the corresponding predicted value of deflection. Fig. 5 and Fig. 6 are the comparison charts between the predicted value of deflection of the five nodes and the deflection value calculated using finite element model. The condition is almost consistent at other nodes. Computational formula for residual error is determined from the formula

$$\text{error} = y_{\text{calculated}} - y_{\text{predicted}}\,, \tag{5}$$

where $y_{\text{calculated}}$ means the deflection calculated by finite element method and $y_{\text{predicted}}$ represents the deflection calculated from the forecasting function.

As shown in Figs. 5 and 6, the predicted deflection fits well with theoretical calculated deflection of the 5 lower chord nodes. The forecasting function for node deflection established with LS-SVR is reliable and well generalized, all residual errors being within the range of ± 0.2 mm and symmetrically distributed about zero line. The substantial consistency with normal distribution means that the predicted deflection has no abnormal value.

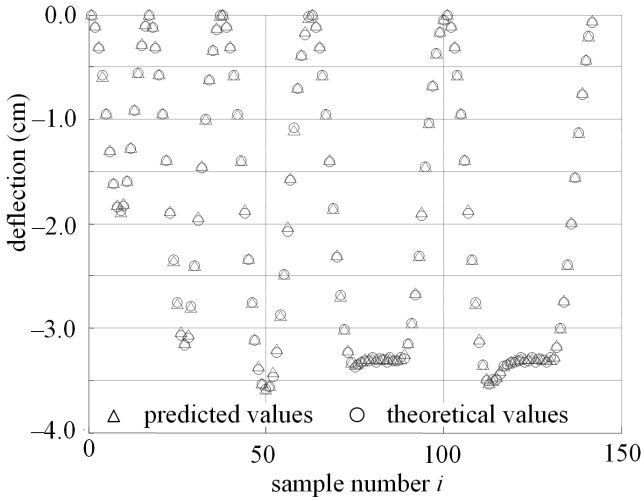

Fig. 5. Chart of comparison between predicted values and theoretically calculated values of lower chord nodes

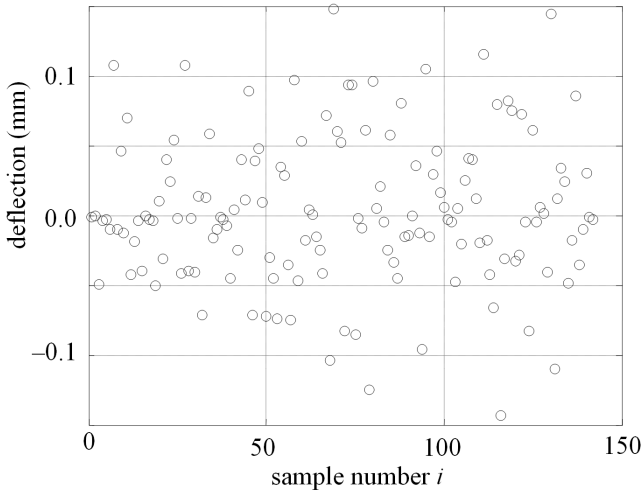

Fig. 6. Residual plot of predicted values and theoretically calculated values of lower chord nodes

4.3. Determination of evaluation criteria conclusions

The criteria ε_{1k} and ε_{2k} (k being the number of the node, $2 \leq k \leq 8$) for evaluation of the bridge with the forecasting function of each node are determined by the method described in Section 1.2.

4.3.1. Determination of the first evaluation criterion ε_{1k}. Since the data sample contains no measured data in this paper, the measured deflection is to be simulated by adding Gaussian random number to the deflection in calculated data sample

$$y_{i\text{measured}} = y_{i\text{calculated}} \cdot (1 + \varepsilon R_i)\,, \tag{6}$$

where i is the data sample number ($1 \leq i \leq 142$), $y_{i\text{measured}}$ denotes the ith simulated measured deflection, $y_{i\text{calculated}}$ represents the ith theoretically calculated deflection, R_i stands for the ith normal distribution random number with the mean equal to 0 and variance equal to 1 and ε is the level of noise.

The noise level ε in this paper is 3%, in other words $\varepsilon = 0.03$. Substitute now the simulated measured data sample with a noise level of 3% into formulae (3) and (4), and assume that the significance level $\alpha = 0.05$. Then $m = 1.96$ according to table look-up.

Table 1 shows the evaluation criterion ε_{1k} for each lower chord node.

Table 1. Evaluation criterion ε_{1k} in mm for each lower chord node

Node number k	2	3	4	5	6	7	8
ε_{1k}	0.8	1.2	1.4	1.6	1.4	1.2	0.8

According to the analysis of ε_{1k} of each node, the mid-span node 5 has the largest value of ε_{1k}, which decreases gradually towards both ends, while the end nodes 2 and 8 have the smallest values of ε_{1k}.

4.3.2. Determination of the first evaluation criterion ε_{2k}. According to the Fundamental Code for Design on Railway Bridge and Culvert [11], the allowable value of the mid-span deflection of simply-supported steel truss bridge is equal to 1/900 of the span. Since the span of steel truss bridge in this study is 64 m, the allowable mid-span deflection $f = 64/900 = 0.0711\,\text{m} = 71.1\,\text{mm}$.

The second evaluation criterion ε_{2qk} for each lower chord point under any condition is calculated, when elasticity modulus of all units in finite element model (Fig. 2) is reduced by 10%. Table 2 shows the values of ε_{2qk} for each lower chord point under various load conditions.

$$de(\%) = \frac{E \cdot A_{\text{Without damage}} - E \cdot A_{\text{With damage}}}{EA_{\text{With damage}}} \cdot 100\% =$$

$$= \frac{A_{\text{Without damage}} - A_{\text{With damage}}}{A_{\text{With damage}}} \cdot 100\%\,, \tag{7}$$

where E means the elasticity modulus of material, $A_{\text{Without damage}}$ is the sectional area determined when the rod piece is free of damage and $A_{With\ damage}$ represents the sectional area determined when the rod piece is damaged. Table 2 shows the evaluation criterion ε_{2qk} for each lower chord node under various load conditions.

Table 2. Evaluation criterion ε_{2qk} in mm
for each lower chord node under various load conditions

Load condition	node 2	node 3	node 4	node 5	node 6	node 7	node 8
Single locomotive	0.8	1.3	1.9	1.9	1.9	1.3	0.8
Dual locomotives	1.5	2.5	3.4	3.5	3.4	2.5	1.4
Triple locomotives	1.7	2.9	3.9	4.0	3.9	2.9	1.6
Single train	1.5	2.6	3.7	3.7	3.7	2.7	1.6
Dual trains	1.6	2.8	3.9	4.0	3.8	2.8	1.6

4.4. Health condition evaluation

Assuming there are two states of damage:

1. The unit 3 in finite element model as shown in Fig. 2 (indicated by smaller boldface letter) is damaged by 20%, while unit 5 is damaged by 30%.

2. The unit 5 is damaged by 50%, while the unit 11 is damaged by 40%.

Check the effectiveness of health condition evaluation performed by the above forecasting function method.

Simulated calculation of the difference between the maximum deflection of each node and maximum deflection in damage-free state is performed by the finite element method for the two above states of damage and under various load conditions as shown in Table 3.

Table 3. Difference between maximum deflection (in mm) in the two above states of damage and maximum deflection in damage-free state (DS means the damage state)

Load condition	DS	node 2	node 3	node 4	node 5	node 6	node 7	node 8
1 locomotive	1	0.2	0.4	0.7	0.8	0.9	0.5	0.2
	2	0.4	1.0	1.6	***2.4***	***2.5***	***1.5***	0.6
2 locomotives	1	0.5	1.0	**1.4**	1.5	**1.5**	1.1	0.5
	2	1.0	2.2	3.2	***4.4***	***4.3***	***3.0***	***1.4***
3 locomotives	1	0.5	1.1	**1.6**	**1.8**	**1.8**	**1.3**	0.6
	2	1.2	2.5	3.7	***4.9***	***4.9***	***3.3***	***1.7***
Single train	1	0.5	1.0	**1.6**	**1.6**	**1.7**	**1.2**	0.6
	2	1.1	2.3	3.5	***4.6***	***4.6***	***3.1***	***1.6***
Dual trains	1	0.5	1.1	**1.6**	**1.8**	**1.8**	**1.2**	0.5
	2	1.2	2.5	3.6	***4.9***	***4.9***	***3.2***	***1.5***

As indicated by boldface numbers in Table 3, the maximum deflection difference is greater than ε_{1k} at some nodes in damage state 1, when the program gives an alarm signal indicating structural safety risk in a specification-allowed normal service condition. As indicated by italic boldface numbers, the maximum deflection difference is greater than ε_{2k} at some nodes (which are normally near the units with significant damage) in damage state 2, when the program gives critical alarm indicating serious structural safety risk.

5. Conclusion

This paper proposes a deflection-based forecasting function method for bridge structure health condition evaluation. This method is composed of two major sections: First, establish correct forecasting function, second, determine two evaluation criteria. Demonstration is performed with a simply-supported steel truss bridge on railway as example. According to the result, forecasting function method could efficiently and accurately evaluate the health condition of simply-supported steel truss bridge structure on railway. In case of large bridge structures, it is advisable to establish as many forecasting functions for node or section as possible when building forecasting function for health condition evaluation, omission may happen if forecasting function fails to be established adequately.

References

[1] H. Li: *Structural health monitoring.* Dalian University of Technology Press, Dalian, 2005.

[2] H. Dai, A. Yuan: *Based on the sensitivity analysis of structure model updating.* Science Press, Beijing, 2011.

[3] H. X. He, W. M. Yan, H. Ma, Z. Wang: *Review and prospect of standardization of structural health monitoring system design.* J Earthquake Eng. Eng. Vibr. *28* (2008), No. 4, 154–160.

[4] B. Yu, H. Oiu, H. Wang, T. Guo: *Health monitoring system for Sutong Yangtze River Bridge.* J Earthquake Eng. Eng. Vibr. *29* (2009), No. 4, 170–177.

[5] M. B. Su, Y. L. Du, B. C. Sun, B. P. Chen, X. M. Wang: *Study on the long-term health monitoring and alarming system for the Wuhu Yangtze River Bridge.* J China Railway Society *29* (2007), No. 2, 71–76.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 123–130

A load forecasting method of hard-shaft coupling multi-motor diving system[1]

Zhou Yuanhua[2], Ma Hongwei[3], Yi Xianzhong[2]

Abstract. A Learning Machine Extreme (ELM)-based load forecasting method was proposed against the nonlinearity, strong coupling and other issues in the load measurement of hard-shaft coupling multi-motor diving system. Signal preprocessing was conducted by selecting motor current and speed as auxiliary variables and via wavelet transform, ELM theory was adopted to set up a load forecast model of hard-shaft coupling multi-motor diving system. Experimental tests were carried out in the cutting part of thin seam shearer, the results show that the method is of high accuracy and the forecast error is less than 5%, which meets the requirement of field control.

Key words. Hard-shaft coupling, multi-motor, load forecast..

1. Foreword

Multi-motor dragging system has three coupling patterns—motor independence, flexible coupling and hard-shaft coupling. Multi-motor that adopts multi-motor drives the same device, which cannot only save space and improve power, but is also conducive to reduce rotary inertia of devices, shorten the transition process, reduce energy consumption, improve the system's dynamic performance and flexibility .

In hard-shaft coupling multi-motor dragging system, motor speed is in forced synchronization and output torque are inter-coupled, due to the differences in motor parameters, torque output must be controlled in order to improve efficiency [1-4].

In order to protect system security, current is usually selected as s control parameter. Therefore, this paper selects motor current and motor speed as the measurement parameters to conduct load forecasting of hard-shaft coupling multi-motor and to improve the basis for power balance control.

[1]The support of Yangtze University Youth Foundation (2015CQN46) and Development of unconventional oil and gas drilling key instruments and automation tools (2016ZX05022-006-004) is gratefully acknowledged.

[2]School of Mechanical Engineering, Yangtze University, Jingzhou City, Hubei, 434023, China

[3]School of Mechanical Engineering, Xi'an University of Science and Technology, Xi'an City, 710054, China

2. Analysis of hard-shaft coupling multi-motor system

The hard-shaft coupling drive of the two three-phase induction motors in the cutting part of the thin seam shearer is taken as an example for analysis, and the hard-shaft coupling dual-motor system of the three-phase induction motors in the cutting part of the thin seam shearer are shown in Fig. 1 and are composed of two sets of three-phase induction motors, a transmission system, load, etc.

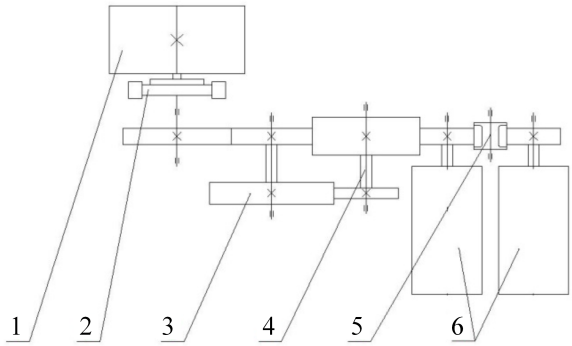

Fig. 1. Schematic diagram of hard-shaft coupling dual-motor transmission: 1–Load 2–Planetary gear train 3–Transmission gear 4–Elastic torque shaft 5–Idle gear 6–Motor

The dragging equation of the motor is:

$$T_{\mathrm{e}} - T_{\mathrm{f}} - T_{\mathrm{g}} = \frac{J_{\mathrm{p}}}{N_{\mathrm{p}}} \frac{\mathrm{d}\omega}{\mathrm{d}t}\,, \tag{1}$$

where T_{e} is the synthetic torque of motors M1 and M2, T_{f} is the resistance load torque, T_{g} is the drag load torque; J_{p} is the total rotary inertia of the system,, ω is the motor angular velocity, and N_{p} is the motor pole pairs.

According to the knowledge of motors, stator is static in $\alpha\beta$ coordinate system, and rotor rotates counterclockwisely in $\alpha'\beta'$ coordinate system with an angular velocity of ω. According to the mathematical model of the induction motor, the torque equation reads

$$T_{\mathrm{g}} = N_{\mathrm{p}}(i_{s\beta}\psi_{s\alpha} - i_{s\alpha}\psi_{s\beta}) - \frac{J_{\mathrm{p}}}{N_{\mathrm{p}}} \frac{\mathrm{d}\omega}{\mathrm{d}t} - T_{\mathrm{f}}\,. \tag{2}$$

Here, symbols $i_{s\alpha}$ and $i_{s\beta}$ are the components of the stator current in coordinates α and β, while symbols $\psi_{s\alpha}$ and $\psi_{s\beta}$ denote the components of the stator flux linkage in the same coordinate system. Formula (2) reflects the relationship between load as well as stator current and speed of the motor. This relationship is obviously nonlinear.

In this paper, the stator current and speed of the motor are selected as auxiliary variables to forecast the output torque of multi-motor driving system and provide a basis for the rational allocation of power.

3. Load forecast modeling based on ELM theory

3.1. ELM theory

Compared with traditional learning algorithms, Learning Machine Extreme (hereinafter referred to as ELM) has quick learning speed and good generalization performance, attracting more and more attention [5–10].

The feed forward neural network structure of a typical single hidden layer is shown in Fig. 2, the network has an input layer, a hidden layer and an output layer, wherein the input layer is fully connected to hidden layer and hidden layer is fully connected to output layer neurons.

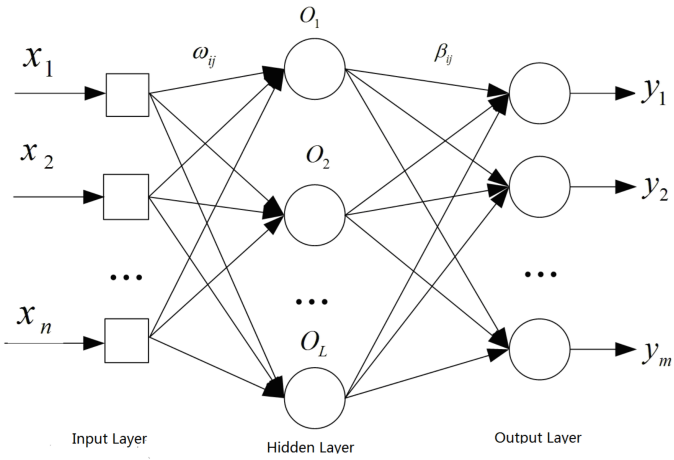

Fig. 2. Feed forward neural network structure of a typical single hidden layer

In this single hidden layer, the input layer of feedforward neural network has n neurons, corresponding to n input variables; hidden layer has l neurons, output layer has m neurons, corresponding to m output variables and the connection weights between input layer and hidden layer as well as between hidden layer and output layer are $\boldsymbol{\omega}$ and $\boldsymbol{\beta}$, respectively. The threshold value of hidden layer neuron is $\boldsymbol{b}$, the training set with Q samples has an input matrix X and an output matrix Y, the activation function of hidden layer neuron is $g(x)$, the output of single hidden layer's feedforward neural network is T. Then the output of neural network can be denoted by

$$H\beta = T^{'}, \tag{3}$$

where, $T^{'}$ is the transposition of matrix T and H is the output matrix of hidden layer in the neural network.

It can be proved on the basis of previous studies that when activation function $g(x)$ is infinitely differentiable, the parameters of neural network need no adjustment, $\boldsymbol{\omega}$ and $\boldsymbol{b}$ can be randomly selected before the training, and remain unchanged in the

training process, while the connection weight $\boldsymbol{\beta}$ between the hidden layer and output layer can be obtained by solving the set of equations

$$\min_{\beta} ||H\beta - T^{'}|| . \tag{4}$$

Its solution is

$$\hat{\beta} = H^{+}T^{'} , \tag{5}$$

where, $H^{'}$ is the Moore-Penrose generalized inverse of the output matrix H in the hidden layer.

The main steps of ELM learning algorithm are as follows:

Step 1: Determine the number of neurons in hidden layer, set the connection weight $\boldsymbol{\omega}$ between input layer and hidden layer as well as bias $\boldsymbol{b}$ of neurons in hidden layer randomly.

Step 2: Select an infinitely differentiable function as the activation function of hidden layer neuron, and then calculate the output matrix H of hidden layer.

Step 3: Calculate weight $\hat{\beta}$ of output layer: $\hat{\beta} = H^{+}T^{'}$.

It is, thus, clear that ELM is a very simple and quick learning algorithm. In order to ensure the convergence of algorithm, the number of hidden layer neurons in ELM network should be the same as the number of input learning samples.

3.2. Load forecast modeling

Signal preprocessing shall be conducted first in an ELM measurement modeling process, on account of the strong noise reduction ability of wavelet transform, wavelet transform is thereby selected for signal preprocessing. ELM neural network has fast learning speed and high learning accuracy, therefore, this paper adopts a soft measurement modeling method that integrates wavelet transform and ELM neural network in conducting soft measurement of the cutting load of shearer. First, conduct wavelet transform of signals and modeling of the transformed signals using multiple ELM networks [5–7], then get the forecast output signal by taking the output of multiple ELM models as weighted mean. The soft measurement model structure that integrates wavelet transform and ELM network is shown in Fig. 3.

4. Experimental study

4.1. Testing experiment

The principles of simulation experiment system of multi-motor driving system are shown in Fig. 3, which is mainly composed of the motor and its control circuit, loading device, connecting device, measuring device, etc.

During the experiment, continuous motor loading is achieved by adjusting the loading device. First, the operation data of single-motors are measured; then, the operation data of dual-motors are measured; finally, load measurement modeling is conducted using the experiment data. The experiment platform for design in this

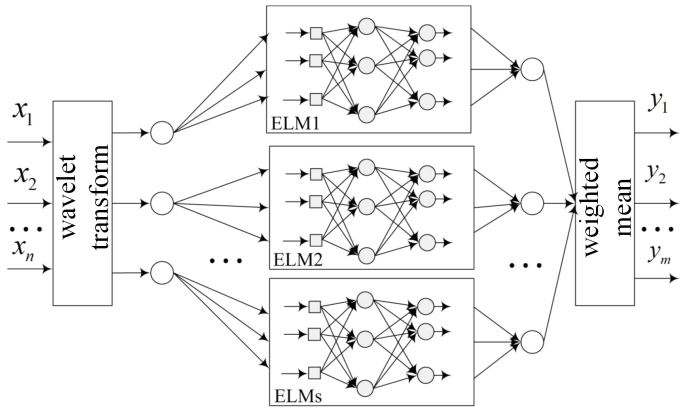

Fig. 3. Load measurement model integrating wavelet transform and ELM network

paper is shown in Fig. 4 and the experimental stand is depicted in Fig. 5.

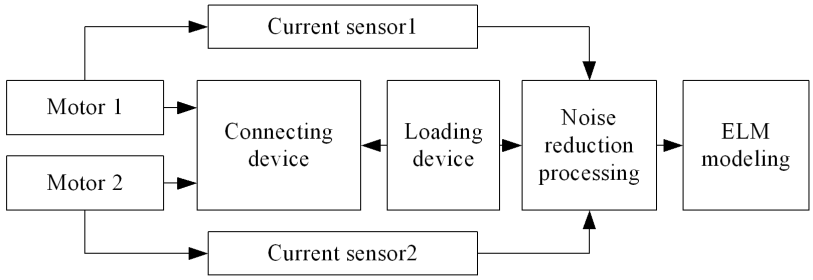

Fig. 4. Functional block diagram of multi-motor driving system and simulation experiment system

The experiment steps are as follows:

1. Keep motor speed unchanged, change motor load by adjusting the loading devices, and record the current value of motor stator.
2. Change motor speed and repeat step 1.
3. Repeat step 1 and 2 until getting enough data.
4. Divide the above data into test samples and training samples, train ELM network to get the load model of the motor.

4.2. Analysis of experiment results

The motor speed values in experiment process are, respectively, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0 times of the rated speed of the experiment, and the load is increased to 1.0 times from the rated 0.5 times at each speed.

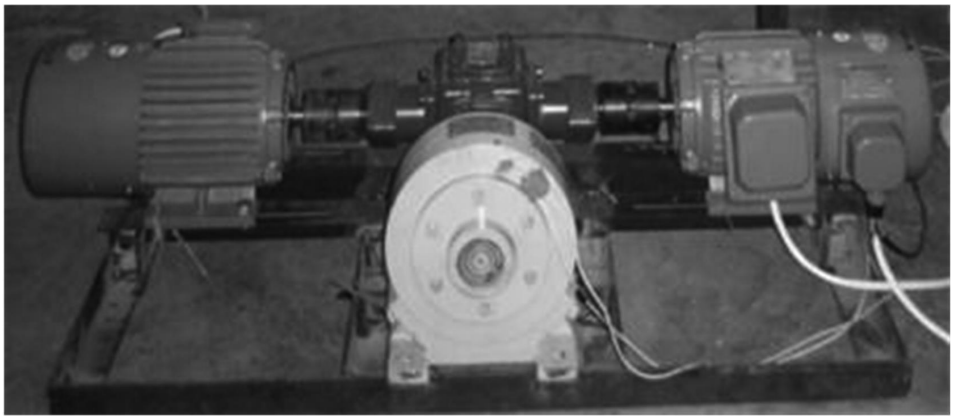

Fig. 5. Dual-motor driven simulation device

In the dual-motor driven experiment, 3 ELM models are used, modeling time is about 0.2 s. After several tests, one of the test results is shown in Fig. 6. We can see from the forecast results that the maximum error 0.02 A occurs at small load, the average value is adopted within the entire test range after 100 tests, and the average error is less than 5 %.

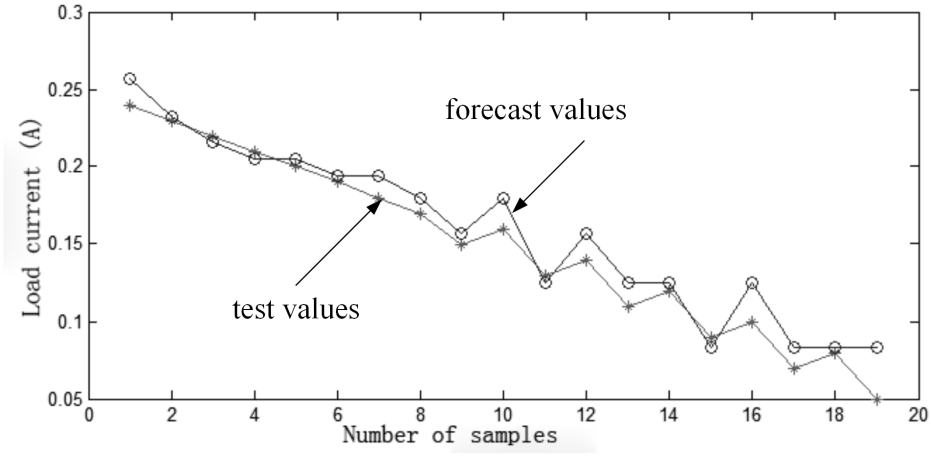

Fig. 6. The output result of dual-motor driven load forecast model

The above simulation results show that the forecast model set up by using ELM network has high measurement accuracy, the whole modeling process is 0.2 s, with high real-time capability, and the average forecast error is 5 %, which can be applied to engineering tests.

5. Conclusion

A load forecasting method that is based on wavelet transform and finite machine learning theory was proposed against the nonlinearity, strong coupling and other issues in the load measurement of current signal-based multi-motors, signal preprocessing was conducted by selecting motor current and speed as auxiliary variables and via wavelet transform, and a load measurement model was set up by using finite machine learning theory. Experiment data has proved the effectiveness of the method, modeling time is about 0.2 s and the load forecast error is less than 5 %, which can be achieved by online measurement.

References

[1] F. P. A. Vaccaro, M. Janusz, K. Kuhn: *Digital control of a Ward Leonard drive system.* Proc. AFRICON '92-3rd AFRICON Conference, 22–24 Sept. 1992, Ezulwini Valley, Swaziland, 123–127.

[2] M. Tomizuka, J. S. Hu, T. C. Chiu, T. Kamano: *Synchronisation of two motion control axes under adaptive feedforward control.* J Dyn. Sys. Meas, Control *114* 1992, No. 2, 196–203.

[3] Y. Koren: *Cross-coupled biaxial computer control for manufacturing Systems.* J Dyn. Sys. Meas, Control *102* (1980), No. 4, 265–272.

[4] S. A. Zaid, O. A. Mahgoub, K. El-Metwally: *Implementation of a new fast direct torque control algorithm for induction motor drives.* IET Electric Power Applications *4* (2010), No. 5, 305–313.

[5] P. S. Bhowmik, P. Purkait, K. Bhattacharya: *A novel wavelet transforms aided neural network based transmission line fault analysis method.* IJ Electrical Power and Energy Systems *31* (2009), No. 5, 213–219.

[6] S. Kelo, S. Dudul: *A wavelet Elman neural network for short-term electrical load prediction under the influence of temperature.* IJ Electrical Power and Energy Systems *43* (2012), No. 1, 213–219.

[7] N. A. Al-geelania, M. A. M. Piah, R. Q. Shaddad: *Characterization of acoustic signals due to surface discharges on H.V. glass insulators using wavelet radial basis function neural networks.* Applied Soft Computing *12* (2012), No. 4, 1239–1246.

[8] V. A. Akpan, G. D. Hassapis: *Nonlinear model identification and adaptive model predictive control using neural networks* . ISA Transactions *50* (2011), No. 2, 177–194.

[9] K. Salahshoor, M. H. Hajisalehi, M. H. Sefat: *Nonlinear model identification and adaptive control of CO_2 sequestration process in saline aquifers using artificial neural networks.* Applied Soft Computing *12* (2012), No. 11, 3379–3389.

[10] P. Kittisupakorn, P. Thitiyasook, M. A. Hussain, W. Daosud: *Neural network based model predictive control for a steel pickling process.* J Process Control *19*, (2009), No. 4, 579–590.

Received November 16, 2016

Acta Technica *62* (2017), 131–140

Parametric driven based generation and transformation of MBD mid-tolerance model[1]

Qiuzhong Zhou[2], Jutao Guo[3], Wanhong Xu[4], Tie Wang[2]

Abstract. A parametrically driven Model Based Definition (MBD) mid-tolerance model is proposed. First, the definition, vital significance and meaning of MBD mid-tolerance model are put forward. On the basis of comprehensively elaborating parameter design technology, one-to-one corresponding mathematical relation between feature constraint dimension and feature driving parameter is analyzed and mathematical relation between engineering dimension and feature driving parameter together with feature constraint dimension is further explained. Then mathematical expression between engineering dimension and parameter variable is established. Meanwhile, through analyzing the interrelation among dimensions by taking advantage of dimension chains, mathematical expressions between non-marking engineering dimension and driving parameter is established. Also the change law between marking engineering dimension and non-marking engineering dimension is specified and theoretical basis for generation and transformation of mid-tolerance model is provided. On this basis, a detailed process and specific algorithm of generation and transformation of mid-tolerance is proposed. Finally, a case used to verify the thought is feasible and relevant algorithm is proved to be accurate.

Key words. Parametrically driven MBD mid-tolerance model, feature parameter, engineering dimension, dimension chains..

1. Introduction

With the technological development of computer, software and digital manufacture, product definition technology has developed to fully 3D digital definition technology, called Model Based Definition (MBD), from two-dimensional CAD technology. The MBD technology is a technology to organize, express, show, operate

[1]Project supported by Program for Liaoning Excellent Talents in University (LJQ2015096) and Science and technology Project of Department of Education of Liaoning Province (L2015457).

[2]School of Automotive & Transportation, Shenyang Ligong University, Shenyang, 110159, China
[3]Shanghai Spaceflight Precision Machinery Institute, 201805, Shanghai, China
[4]Zhejiang Wanliyang Transmission Co., Ltd, Jinhua, 312073, China

and manage non-geometry manufacture information on product dimensions, tolerance, and manufacturing technology requirement on the basis of a 3D digital model, see [1–2]. It abandons two-dimensional engineering drawing, making the 3D digital model become the only vehicle of product information and making it the only evidence to convey design information in the manufacturing process. MBD technology changes the way of process design. Process design method and technology traditionally based on 2D engineering drawing or 3D model have not applied or satisfied requirement of this technology, so that computer-aided process design and planning technology based on fully 3D model is badly needed to realize fully 3D digital design and manufacturing, which transforms digital manufacturing technology [3–4].

Mid-tolerance is also named symmetry tolerance. Mid-tolerance model is a component geometry model built in the state of symmetry tolerance. MBD Mid-tolerance model is the component structure model built as nominal size in the state of symmetry tolerance and marks the engineering dimension with symmetry tolerance on the model by taking advantage of MBD technology. In the process of NC design, tool path of components has to be formed as mid-tolerance model or MBD mid-tolerance model, so as to improve working efficiency of NC program by technicians. When technicians design 3D process under current MBD technological condition, working procedure dimension is calculated under the condition of mid-tolerance [1–2]. After the 3D process design is completed, it is necessary to transfer the relevant working procedure dimensions to the maximum material dimension; finally the procedure MBD model is formed. Therefore, building of mid-tolerance model is meaningful and significant.

2. 3D parametric design based on feature

3D parametric design based on feature is (by establishing component parametric 3D model using a group of parameter to show or engage the feature dimensional relation) only adjusting or changing one or several parameter values in the group of parameters without entering all definite feature dimensional values to automatically change all relevant feature dimensions and revise and control all feature geometry forms of model, so as to realize accurate 3D modeling of components [5–6]. Parametric design technology may improve design efficiency and flexibility of components by taking advantage of powerful model modification method to avoid fussy repetitive working, becoming the efficient technological measure for original design, product modeling, deformation design of modification series, multi-plan comparison and dynamic and concurrent designs [7–8].

If $D_{\mathrm{F}} = \{D_{\mathrm{F}1}, D_{\mathrm{F}2}, ..., D_{\mathrm{F}n}\}$ is used to indicate n feature constraint dimensions of the component, every feature dimension corresponds to a parameter, and $X = \{X_1, X_2, ..., X_n\}$ is used to express n self-defined parameter variables of driving dimensional constraint components, then

$$D_{\mathrm{F}} = A \cdot X\,, \tag{1}$$

where A is an $n \times n$ unitary matrix.

Figure 1 shows a shaft sleeve component formed by rotation of sketch features or generated by single feature combination. The location relation of geometrical elements of the component on the axis is completely defined by three feature constraint dimensions (expressed by the dashed lines) as D(AB), D(BD), D(CD). If parameter variables AB, BD and CD drive such three feature constraint dimensions as D(AB), D(BD) and D(CD), the component location relation of geometrical element on axis will be determined by the AB, BD and CD parameter variables. Feature constraint dimension and parameter variable of the component have the following one-to-one corresponding relation.

$$\begin{bmatrix} \mathrm{D(AB)} \\ \mathrm{D(BD)} \\ \mathrm{D(CD)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{AB} \\ \mathrm{BD} \\ \mathrm{CD} \end{bmatrix} .$$

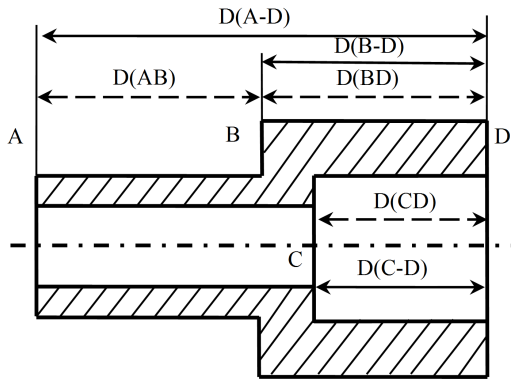

Fig. 1. Shaft sleeve component

3. Relation between design dimension and feature parameter

Engineering dimension can be expressed by feature constraint dimensions. If $D = \{D_1, D_2, ..., D_n\}$ expresses n engineering dimensions of the component P, then

$$D = B \cdot D_{\mathrm{F}} \,. \tag{2}$$

In the above formula

$$B = \begin{bmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{bmatrix} , \quad a_{ij} \in (-1, 0, 1) \,.$$

After substituting expression (1) into (2), then

$$D = B \cdot D_{\mathrm{F}} = BA \cdot X = B \cdot X \,, \tag{3}$$

in other words

$$\begin{bmatrix} D_1 \\ ... \\ D_n \end{bmatrix} = \begin{bmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{bmatrix} \begin{bmatrix} X_1 \\ ... \\ X_n \end{bmatrix}. \tag{4}$$

From (4) it can be seen that every engineering dimension can be expressed by algebraic relations of parameter variables. From (3) we easily obtain

$$X = B^{-1} \cdot D\,. \tag{5}$$

It is clear that from (5) every parameter variable can be expressed by algebraic relation of engineering dimensions. That is, the marked group of engineering dimensions can solely define a group of parameter variables.

As is shown in Fig. 1, shaft sleeve components have four geometrical elements on the axis. Therefore, the complete engineering diagram indicating components should be marked with three engineering dimensions. Symbols D(A-D), D(B-D), and D(C-D) in the figure represent one of annotation schemes of the axis engineering dimension. They have algebraic relation with dimensional parameters. Among which, engineering dimension D(B-D) is determined by parameter BD and engineering dimension D(C-D) is determined by parameter CD, while engineering D(A-D) is the sum of the parameters AB and BD. Thus, the engineering dimensions and parameter variables of the component have following corresponding relations

$$\begin{bmatrix} D(A-D) \\ D(B-D) \\ D(C-D) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} AB \\ BD \\ CD \end{bmatrix}.$$

4. Law of interaction between dimensions

4.1. Dimension relations

Interrelation of component engineering dimensions needs to be expressed and calculated by dimensional chains. Any dimensional chain is composed of component link and closing link and every dimensional chain has a single closed ring. Component link directly guarantees the dimension while closing link indirectly guarantees the dimension through other dimensions. According to marking requirements of component engineering dimension, dimensions directly marked on the component are component links, while the rest are closing links.

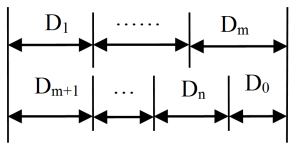

Fig. 2. Dimensional chains

As shown in Fig. 2 containing a linear dimensional chain composed by $n+1$ rings, the dimension D_0 is the closing ring and $D_1 - D_n$ are the component rings. The basic dimension of the closing ring equals to the sum of basic dimensions of increasing rings minus the sum of basic dimensions of decreasing rings, so that

$$D_0 = \sum_{z=1}^{m} D_z - \sum_{j=m+1}^{\mathrm{n}} D_j \,. \tag{6}$$

In this In this expression, D_0 is the basic dimension of the closing link, D_z denotes the basic dimension of the increasing link, D_j stands for the basic dimension of the decreasing link, m is the number of the increased links and n denotes the number of dimensional chains. Meanwhile, from the expression (6) we can deduct that

$$\Delta D_0 = \sum_{z=1}^{m} \Delta D_z - \sum_{j=m+1}^{\mathrm{n}} \Delta D_j \,. \tag{7}$$

From formula (7), we can see that when certain dimension in the component link is changed and other component links are not changed, the dimension of the closing links will change correspondingly. The details follow

1. When the component link is an increasing link, the closing link would change in the same direction. That is, the closing link increases with increasing of the component link and decreases with decreasing of the component link.

2. When the component link is a decreasing ring, the closing link would change in the reverse direction. That is, the closing link increases with decreasing of the component link and decreases with increasing of the component link.

4.2. Relation between closing link and driving parameter

From expression (6) we have

$$D_0 = \begin{bmatrix} 1 & \ldots & 1 & -1 & \ldots & -1 \end{bmatrix} \begin{bmatrix} D_1 \\ \ldots \\ D_m \\ D_{m+1} \\ \ldots \\ D_n \end{bmatrix} = \tag{8}$$

Introducing of expression (4) into (8), we can see that

$$D_0 = \begin{bmatrix} 1 & \ldots & 1 & -1 & \ldots & -1 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} \\ a_{(m+1)1} & \ldots & a_{(m+1)n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ \ldots \\ X_m \\ X_{m+1} \\ \ldots \\ X_n \end{bmatrix} =$$

$$= \begin{bmatrix} Q_1 & ... & Q_n \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ ... \\ X_m \\ X_{m+1} \\ ... \\ X_n \end{bmatrix} = Q_1 X_1 + ... + Q_n X_n, \tag{9}$$

where

$$Q_i = a_{1i} + a_{2i} + \cdots + a_{mi} - a_{(m+1)i} - \cdots - a_{ni}\,.$$

From expression (9) we can see that not only there is algebraic relation between marking engineering dimension and driving parameter of the component, but also non-marking engineering dimension can be expressed by a linear driving parameter. Therefore, under the condition of a known component of the driving parameter, all dimensions of the component are definite and can be obtained by calculating the driving parameter.

5. Results

An example was used to verify the thought that realizes the automatic generation and transformation of MBD mid-tolerance model through parameter driving. As shown in Fig. 3 the shaft sleeve is a component structure and axis is geometrical surface number. Mid-tolerance generation and transformation of radical dimension is the same as axis dimension. Radical dimension is ignored for simplifying. As shown in Fig. 4 the MBD model is created, when the component is marked with dimension and tolerance as the maximum material principle. According to its structural characteristics, the buildup feature constraint dimension and corresponding driving parameter are shown in Table 1. The relation between the design dimension and parameter of the component is shown in Table 2.

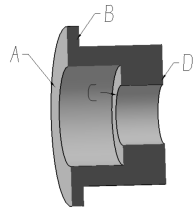

Fig. 3. Component structure and surface number

Table 1. Component feature parameters

constraint dimension	driving parameter
D(AB)	AB
D(AC)	AC
D(BD)	BD

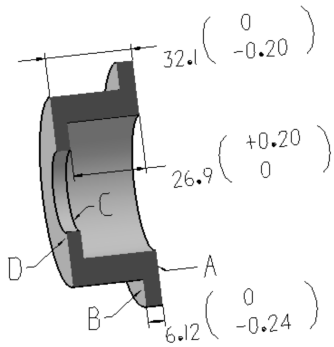

Fig. 4. Maximum material MBD model

Table 2. Relation between component design dimension and parameter

design dimension	parameter expression
D(A-B)	AB
D(A-C)	AC
D(B-D)	BD
D(A-D)	AB+BD
D(C-D)	AB+BD-AC

According to above information, the design dimension D, feature constraint dimension D_{F}, driving parameter X and its transition matrix B of the component are built as follows

$$D = \begin{bmatrix} D(A-B) \\ D(A-C) \\ D(A-D) \end{bmatrix} \quad D_F = \begin{bmatrix} D(AB) \\ D(AC) \\ D(BD) \end{bmatrix} \quad X = \begin{bmatrix} AB \\ AC \\ BD \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

The inverse matrix of B is

$$B^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

Therefore, according to (5)

$$X = \begin{bmatrix} AB \\ AC \\ BD \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} D(A-B) \\ D(A-C) \\ D(A-D) \end{bmatrix},$$

In addition, the component has two closing links of $D(B-D)$ and $D(C-D)$. The dimensional chains are shown as Fig. 5.

According to (8), the driving parameter expression of closing links is deduced as follows

$$D(B-D) = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} D(A-B) \\ D(A-C) \\ D(A-D) \end{bmatrix} =$$

$$= \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} AB \\ AC \\ BD \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} AB \\ AC \\ BD \end{bmatrix} = BD$$

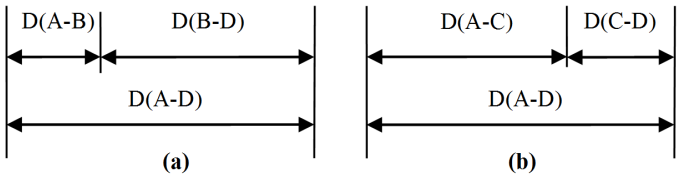

Fig. 5. Design of dimensional chains

and

$$D(C-D) = \begin{bmatrix} 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} D(A-B) \\ D(A-C) \\ D(A-D) \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} AB \\ AC \\ BD \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} AB \\ AC \\ BD \end{bmatrix} = AB - AC + BD\,.$$

Finally, when the maximum material model is transformed into mid-tolerance model (using the dimensions from Fig. 3), the design dimension

$$D = \begin{bmatrix} D(A-B) \\ D(A-C) \\ D(A-D) \end{bmatrix} \text{ becomes } \begin{bmatrix} 6 \\ 27 \\ 32 \end{bmatrix} \text{ from } \begin{bmatrix} 6.12 \\ 29.9 \\ 32.1 \end{bmatrix}.$$

Similarly, the driving parameter

$$X = \begin{bmatrix} AB \\ AC \\ BD \end{bmatrix} \text{ becomes } \begin{bmatrix} 6 \\ 27 \\ 26 \end{bmatrix} \text{ from } \begin{bmatrix} 6.2 \\ 26.9 \\ 25.98 \end{bmatrix}.$$

The closing link $D(B-D)$ becomes 26 from 25.98 and $D(C-D)$ becomes 5 from 5.2. The mid-tolerance MBD model is shown in Fig. 6.

The results are summarized in Table 3.

Table 3. Generation information table of component mid-tolerance

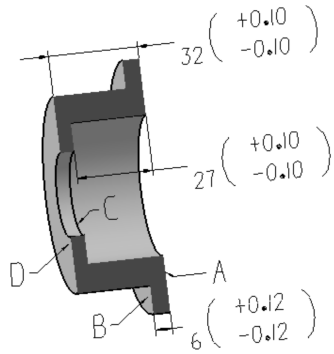

Fig. 6. MBD mid-tolerance model

MD	MMDV	MMTV	MTDV	STV	ID	DP	DV
D(A-B)	6.12	0/-0.24	6	±0.12	D(A-B)	AB	-0.12
					D(B-D)	BD	+0.12
D(A-C)	26.9	+0.2/0	27	±0.1	D(A-C)	AC	+0.1
					D(C-D)	—	-0.1
D(A-D)	32.1	0/-0.2	32	±0.1	D(A-D)	—	-0.1
					D(B-D)	BD	-0.1
					D(C-D)	—	-0.1

Note: MD–marking dimension, MMDV–maximum material dimension value, MMTV–maximum material tolerance value, MTDV–mid-tolerance dimension value, STV–symmetry tolerance value, ID–influenced dimension, DP–driving parameter, DV–dimension variation.

6. Conclusion

Under the MBD technology condition, generation and transformation of MBD mid-tolerance model is the major working content of 3D design process. Through the theoretical analysis, there is the solely definite mathematical relation between one group of engineering dimensions and one group of parameter variables, and also between marking engineering dimensions and non-dimensional engineering dimensions. Any change of engineering dimensions will lead to the change of relevant non-marking engineering dimensions. Therefore, parameter-driving technology may solve the generation and transformation problem of mid-tolerance model. The key part of generation and transformation technology of parameter driving mid-tolerance MBD model is to establish mapping and association relations between marking dimensions with feature driving parameters and feature constraint dimensions. Through developing 3D process design system which integrated 3D modeling system, marking dimension tolerance, driving parameter extraction, dimensional tolerance calculation, and marking dimension and driving parameter association setup are realized. Finally, there are realized automatic generation and transformation working of mid-

tolerance MBD model and improving working efficiency and accuracy of data, which provides technological support of integrated technology system of digital design and technique and production based on fully 3D model, which brings a reform to digital manufacturing technology.

References

[1] M. Alemanni, F. Desterfanis, E. Vezzetti: *Model-based definition design in the product lifecycle management scenario.* J Advanced Manufacturing Technology *52* (2009), No. 1, 1–14.

[2] V. Quintana, L. Rivest, R. Pellerin, F. Venne, F. Kheddouci: *Will model-based definition replace engineering drawings throughout the product lifecycle? A global perspective from aerospace industry.* Computers in Industry *61* (2010), No. 5, 497–508.

[3] *ASME Y 14.41-2003 [EB/OL].* http://en.wikipedia.org/wiki/ASME_Y14.41-2003, 2010.

[4] Q. Z. Zhou, Q. C. Fan: *MBD driven digital product collaborative definition technology.* Proc. 3rd IC Intelligent Networks and Intelligent Systems (ICINIS), 1-3 Nov. 2010, Shenyang, China, 661–664.

[5] L. Ding, D. Davies, C. A. McMahon: *The integration of lightweight representation and annotation for collaborative design representation.* Res. Eng. Design *19* (2009), No. 4, 223–238.

[6] R. Sacks, R. Barak: *Impact of three-dimensional parametric modeling of buildings on productivity in structural engineering practice.* Automation in Construction *17* (2008), No. 4, 439–449.

[7] H. A. van der Meiden, W. F. Bronsvoort: *Solving topological constraints for declarative families of objects.* Computer-Aided Design *39* (2007), No. 8, 652–662.

[8] Y. Jiang, R. Wang: *Parametric design of mannequins for virtual try-on system.* Proc. 7th IC Computer-Aided Industrial Design and Conceptual Design, 17–19 Nov. 2006, Hangzhou, China, 1–6.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 141–152

Applications of geophysical technology in natural gas hydrate prediction of LW21-1-1 well area

Yuanyuan Yi[1]

Abstract. A method of prediction of natural gas hydrate in sea deposits of oil and gas is proposed. It is found that three instantaneous processing and colored inversion can respectively mark out the natural gas hydrate characteristics of natural gas hydrate and amplitude of blanking zone, lithology, and provide a scientific basis for the prediction of natural gas hydrate. In this paper, natural gas hydrate of the Well LW21-1-1 zone is predicted by three instantaneous processing and colored inversion, which may represent a reference to drilling of other blocks.

Key words. Natural gas hydrate, deepwater drilling, three instantaneous profiles, colored inversion.

1. Introduction

Well LW21-1-1 is China National Offshore Oil Corporation in the South China Sea and Asia's first hit deep wells. The water depth is about 2460 m. According to the experience at home and abroad in deepwater drilling, drilling should be done before the well area geological disaster prediction, in order to prevent disasters and accidents. Geological disaster types usually appear in the sites where are the main submarine gas hydrate, shallow gas and shallow water flow.

Natural gas hydrate is a kind of solid crystals with domination of methane, which is composed from hydrocarbon gas molecules and water molecules. Its formation is controlled by temperature and pressure. Since the discovery that the natural gas hydrate can be connected with fuel energy, people have been paying more and more attention to it. But with the development of research, it was found that the possibility of gas hydrate formation to cause shallow geological disasters in the bottom of the sea is very large, and this kind of disaster environment for engineering activities at sea brought certain difficulties, especially for offshore drilling. Many domestic and foreign offshore drilling accidents display that it is easy to cause a major accident when natural gas hydrate is drilled.

[1]Key Laboratory of Oil-gas Resources and Exploration Technology of Ministry of Education ,Yangtze University, Hubei Wuhan, 430100, China

1.1. Lead to collapse of seabed

Shallow sediments of deep water basin [1] and continental slope usually contain natural gas hydrate [2]. Affected by change of pressure and temperature, the occurrence of natural gas hydrate can easily be subjected to change, the stability zone of natural gas hydrate can partially or even entirely disappear. When the sea level drops, the bottom of the stability zone of natural gas hydrate (GHSZ) in seabed sediments will become unstable and begin to decompose, which will make GHSZ change. As GHSZ band [3] starts to break down, water and gas begin to enrich. The resulting pressure will reach abnormal excess hydrostatic pressure, low-density gas and mud may penetrate through the natural gas hydrate layer, which can lead to geological disasters and even spray the surface threatening to drilling platforms. Meanwhile, the liquefied GHSZ can form a downward slide, along the sliding surface of bulk gas hydrate cementation wedge will be downward movement evolved into large-scale submarine landslide. Sea levels have changed so that these events repeated, eventually it will form a chaotic sediment landslide with a certain thickness in the lower slope. Therefore, in the environment of high pressure and low temperature, the natural gas hydrate is stable, change of pressure and temperature will lead to release amounts of gas, which may cause landslide and other major seabed disasters (see Fig. 1).

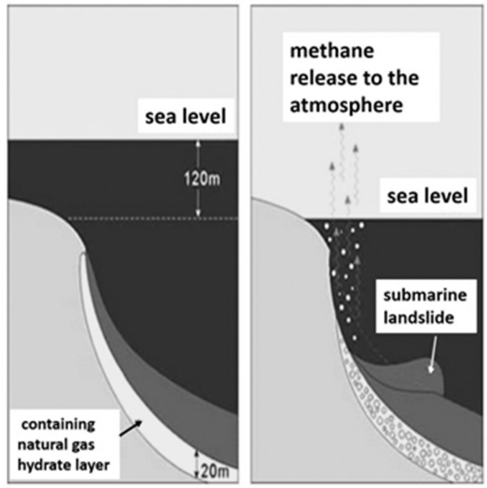

Fig. 1. Natural gas hydrate caused landslide

1.2. Threat of drilling security

Geological tectonism, sedimentation and maritime activities can also cause change of external conditions such as pressure and temperature in the continental slope and shallow sediments in deep water basin, and result in decomposition of natural gas hydrate. The movement of sediment and mudslide also can be triggered by outside

activities (such as drilling or seismic work in this horizon, etc.), then landslide and diapir phenomenon will occur, under the effect of high pressure for a long time, the high-pressure airbag will squeeze weaknesses of the upper layer, which will be a great threat to is its range operation drilling ships and oil platforms. At the same time, in the case of a large number of natural gas hydrate decomposing, amounts of methane will overflow the stratum and dissolve in seawater, so that the density of seawater will be reduced, and result in drilling ships and oil platforms sinking because of buoyancy sudden decrease.

2. Prediction of natural gas hydrate

Through these theories such as the previous analysis of the sensitivity of the characteristics of the natural gas hydrate development zone to various geophysical responses, study of seismic attribute analysis, AVO analysis, the establishment of the inversion velocity field in the formation of steep dip angle, geophysical prediction technique for pseudo well constrained wave impedance inversion. After analyzing the geological characteristics of the LW21-1-1 well, we think that this area is more sensitive to seismic instantaneous attributes, using the two geophysical methods to predict the natural gas hydrate by using three instantaneous profile and colored inversion.

2.1. Three instantaneous processing of seismic data

Three instantaneous profile is the instantaneous attributes extracted by Hilbert transform, namely the three instantaneous attributes, including instantaneous amplitude, instantaneous phase and instantaneous frequency and its corresponding properties. Instantaneous feature refer to concise information about horizon, using the instantaneous attributes need to be analyzed with the angle of the amplitude information, the decomposing process does not make any change for the basic information, but it gets different attribute profiles. Instantaneous profiles carrying a large number of information associated with oil and gas features, such as amplitude, lithology, frequency and phase, which can show geophysical phenomena which are different from conventional seismic profiles.

Three instantaneous attributes is established in the Hilbert transform and complex seismic trace analysis technology. When studying underground lithological change and stratigraphic structure, the Hilbert attribute extraction can be used. For a seismic signal, firstly calculate its Hilbert transform, the transformed result is deemed as the imaginary part, seismic signal is considered as the corresponding real part, they can compose complex seismic trace[4], then the instantaneous attributes of complex seismic trace are calculated. Three instantaneous attributes are on the base of Hilbert transform and complex seismic trace, so we can start from the collected seismic signals $s(t)$, its analytical signal as follows:

$$z(t) = s(t) + i\frac{1}{\pi t} * s(t)\,, \tag{1}$$

where $*$ denotes the convolution, and imaginary part of $z(t)$ in the form $\widetilde{s}(t) = i/(\pi t) * s(t)$ is the Hilbert transform of $s(t)$. Then the instantaneous amplitude, the instantaneous phase, the instantaneous frequency of the signal $s(t)$ can be denoted, respectively, as follows:

$$a(t) = \sqrt{s^2(t) + \widetilde{s}(t)}\,, \tag{2}$$

$$\theta(t) = \arctan \frac{\widetilde{s}(t)}{s(t)}\,, \tag{3}$$

$$\omega(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left[\arctan \frac{\widetilde{s}(t)}{s(t)}\right]. \tag{4}$$

Three instantaneous attributes can be used to locate underground lithologic change and fluid properties. The instantaneous amplitude is the intensity of earthquake wave size [5], instantaneous frequency is the mapping of a dielectric absorption properties, while the instantaneous phase can reflect the change of lithology stratigraphic sequence. In the gas hydrate bearing strata, as the earthquake wave speed increases which leads to increasing the reflection coefficient between the underlying strata, the corresponding strong reflection interface appears in the seismic profile. And if there is contained natural gas hydrate above the stratum, because sediment pores are filled and cemented by hydrate making the stratum become uniform, in the seismic reflection profile there usually appears a weak reflection or amplitude blanking zone. Natural gas hydrate in the seismic section forms usually a strong reflected wave roughly parallel with seabed, and the reflected layer like seabed is called BSR for short. According to seismic response of natural gas hydrate and corresponding characteristics of BSR, the instantaneous amplitude profile can clearly reflect characteristics of amplitude blank zone. The instantaneous phase profile can display the intersected state between the reflection of BSR and true reflection of the stratum better, while abnormal BSR area can be displayed on the instantaneous frequency profile [6]. Seismic profile combines with three instantaneous profiles and BSR reflection interface of natural gas hydrate can be roughly identified.

As shown in Fig. 2, the South China Sea of a crossline instantaneous amplitude profile, showing the suspected gas hydrate distribution area.

In the study, the profile of the LW21-1-1 well is processed. As shown in Fig. 3, there is no amplitude space, so the possibility of occurrence of gas hydrate is lower.

Instantaneous phase is a momentary phase at selected points within each seismic trace, which is a reflection of the phase of dominant frequency of seismic data. When the seismic wave is spread in different physical stratum, its phase can be changed. Instantaneous phase and reflection intensity are independent, not affected by the intensity of energy, and, therefore, they can be used to enhance the in-phase axis within the deep reservoir (seismic wave energy is weak in the depth). Instantaneous phase is a measure of the phase axis continuity in seismic profile. Instantaneous phase profile can more clearly represent the intersected phenomenon between BSR reflection and reflection of the real stratum. An instantaneous phase profile of a

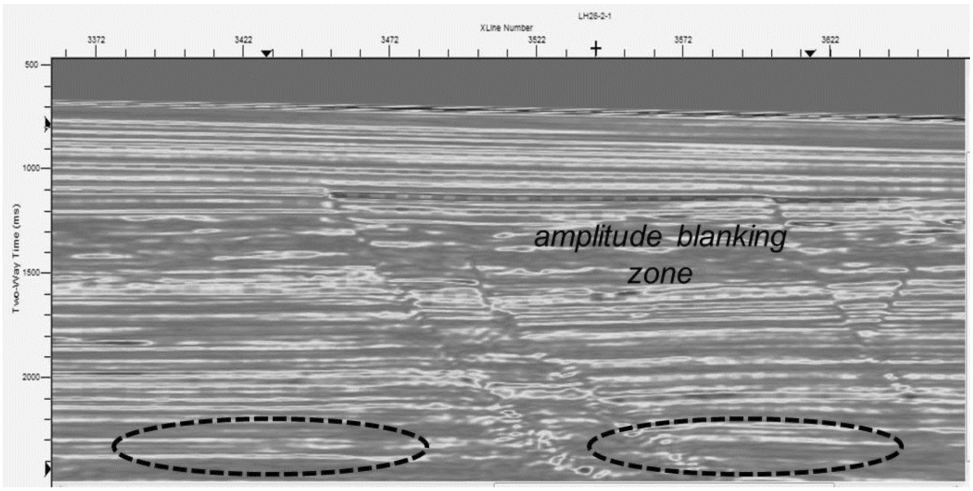

Fig. 2. Instantaneous amplitude profile of the South China Sea

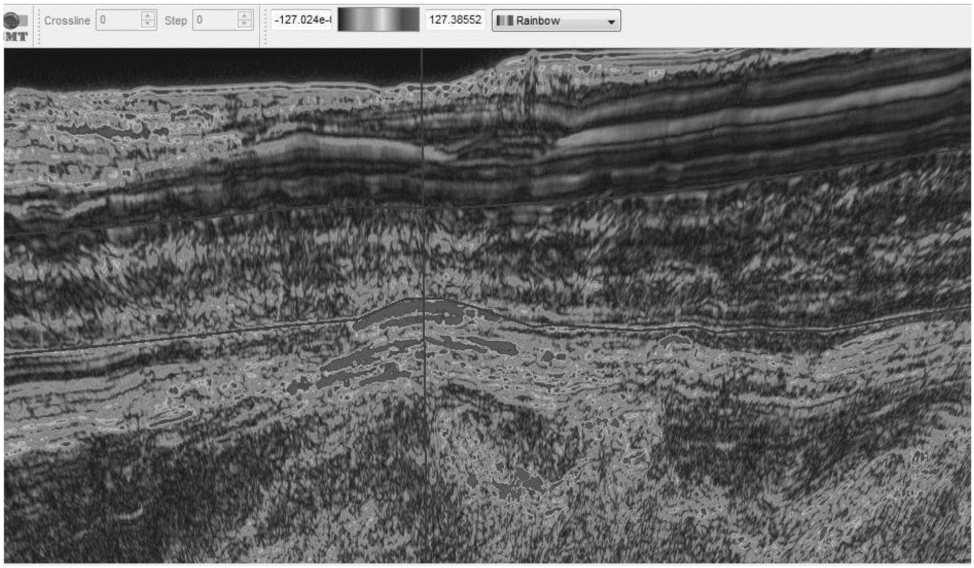

Fig. 3. Instantaneous amplitude profile of LW21-1-1 on crossline 2528

crossline in the South China Sea is depicted in Fig. 4. In Fig. 5, the study area on the phase anomaly is not obvious.

Instantaneous frequency is the time derivative of phase, and can reflect absorption and attenuation characteristics of the stratum. It can be advantageous for analyzing the lithologic changes. Instantaneous frequency is the time rate of change of phase and a reflection of dielectric absorption characteristics. The instantaneous frequency

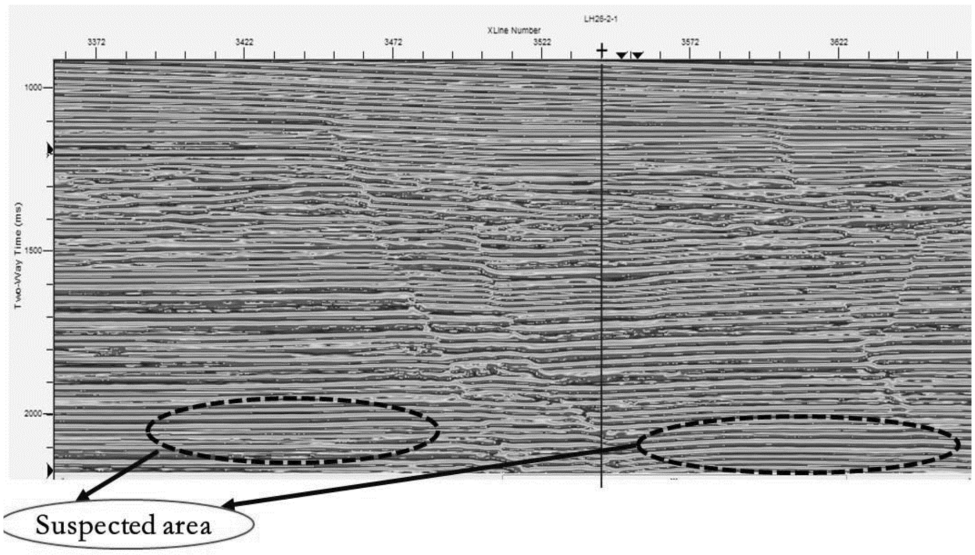

Fig. 4. Instantaneous phase profile of a crossline in the South China Sea

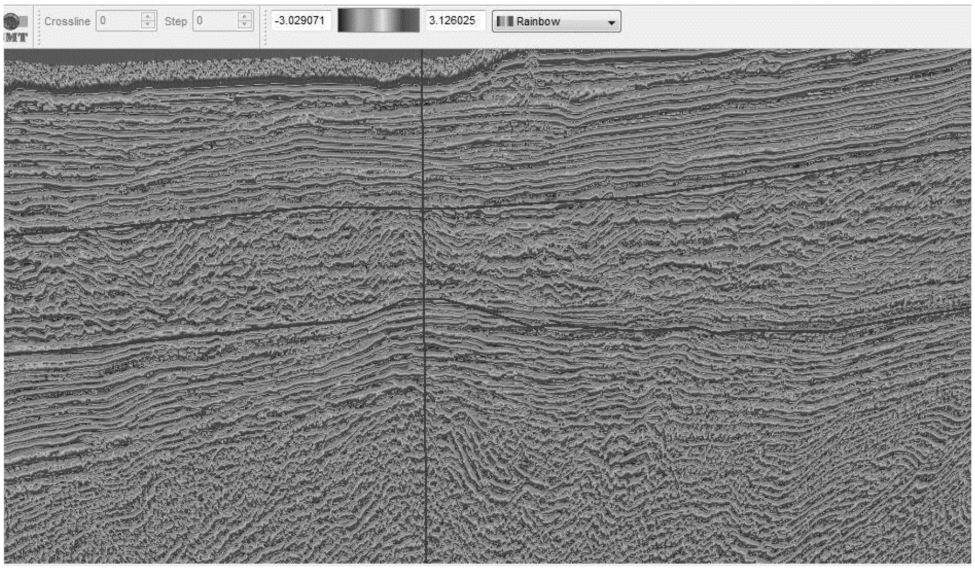

Fig. 5. Instantaneous phase profile of LW21-1-1 on line 2528

profile can reveal the BSR area, see Fig. 6.

In Fig. 7, the instantaneous frequency profile in the area does not show any abnormal BSR region. Instantaneous phase profile can more clearly represent the intersected phenomenon between BSR reflection and reflection of the real stratum,

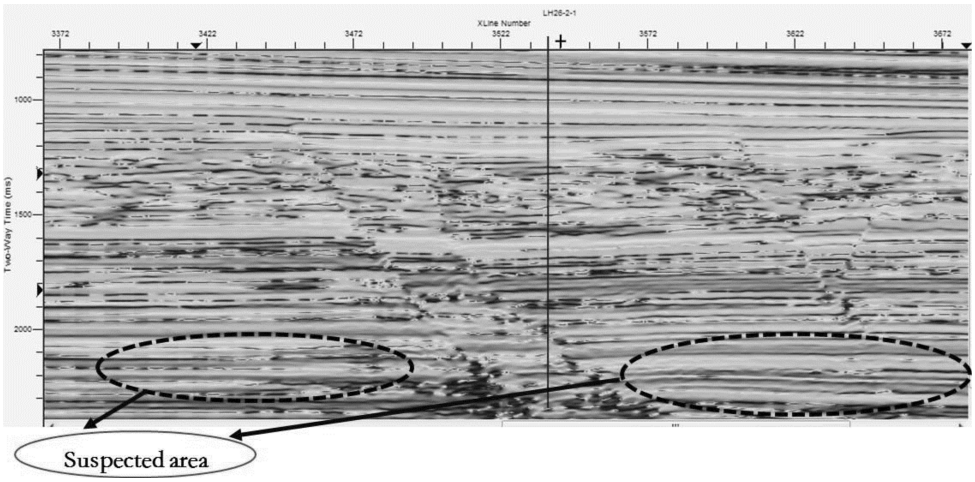

Fig. 6. Instantaneous frequency profile of a crossline in the South China Sea

and instantaneous amplitude profile can clearly reflect BSR interface and characteristics of amplitude blank zone. Instantaneous frequency profile can reveal abnormal BSR area. By the analysis, the study area shows no significant BSR interface and abnormal BSR region, based on three instantaneous profiles combined with seismic profile. We can roughly exclude BSR reflection interface of natural gas hydrate. Through analyzing three instantaneous profiles, it is concluded that there is little chance of natural gas water compounds in LW21-1-1 well area.

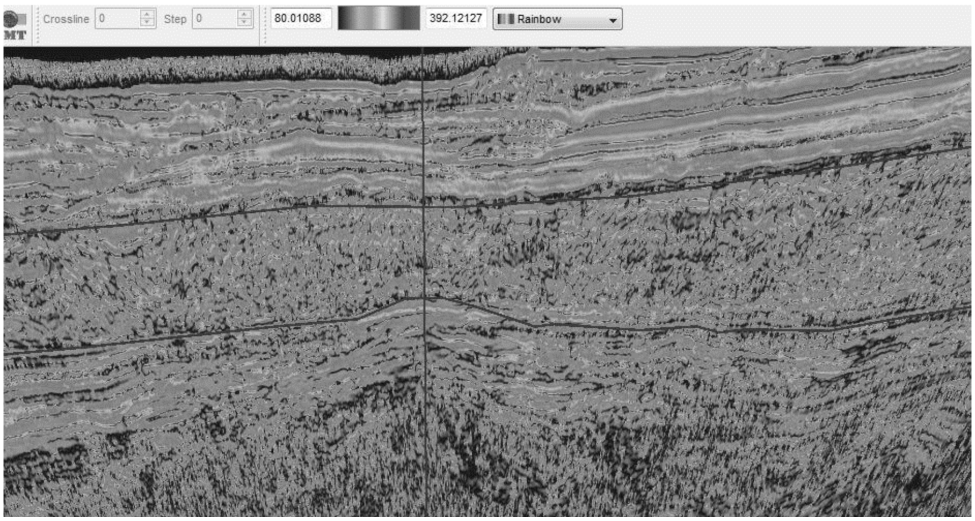

Fig. 7. Instantaneous frequency profile of LW21-1-1 on line 2528

2.2. Seismic colored inversion (SCI)

Inversion is based on a measured data to infer the original geological model. Color inversion is a kind of inversion [7] which uses well log data to constrain the wave impedance inversion in the frequency domain. The core is matching of wave impedance spectroscopy and seismic spectrum to complete inversion. This method does not have a wavelet extraction process, and also does not have to set the initial model to constraint. Its longitudinal resolution is higher than the sparse spike inversion, but lower than the model inversion. It is characterized by lower degree of dependence on the well, global optimization, better eliminating the interference of human analysis factors, so it can objectively reflect the geological phenomenon. The analysis formulae of the colored inversion follow.

The convolution model of seismic records is

$$s_i(t) = r_i(t) * \omega(t)\,, \tag{5}$$

where $s_i(t)$, $i = 1, 2, \ldots, n$ denotes the seismic record, $r_i(t)$, $i = 1, 2, \ldots, n$ denotes the reflection coefficient, n is the number of traces, and $\omega(t)$ denotes the seismic wavelet. The convolution model is transformed to the frequency domain, and then

$$S_i(\omega)\mathrm{e}^{-\mathrm{j}\phi_i(\omega)} = [R_i(\omega)\mathrm{e}^{\mathrm{j}\Psi_i(\omega)}][W_i(\omega)\mathrm{e}^{-\mathrm{j}\varphi_i(\omega)}]\,, \tag{6}$$

where $R_i(\omega)$ denotes the reflection coefficient amplitude spectrum, $W_i(\omega)$ denotes the wavelet amplitude spectrum, $\Psi_i(\omega)$ denotes the reflection coefficient phase spectrum and $\varphi_i(\omega)$ denotes wavelet phase spectrum. Then we can write

$$\left\{ \begin{array}{c} S_i(\omega) = R_i(\omega)W_i(\omega) \\ \phi_i(\omega) = \Psi_i(\omega) + \varphi_i(\omega) \end{array} \right\}. \tag{7}$$

and, therefore:

$$\ln S_i(\omega) = \ln R_i(\omega) \ln W_i(\omega)\,. \tag{8}$$

Assuming that impedance of each trace $Z_i(t)$, $i = 1, 2, \ldots, n$ is transformed to the frequency domain, we get

$$Z_i(\mathrm{j}\omega) = Z_i(\omega)\mathrm{e}^{-\mathrm{j}\Omega_i(\omega)}\,. \tag{9}$$

Among them $\Omega_i(\omega)$ denotes the wave impedance phase spectrum and$Z_i(\omega)$ denotes the wave impedance amplitude spectrum. According to the relation between the wave impedance and reflection coefficient we obtain

$$R_i = \frac{(Z_{i+1} - Z_i)}{(Z_{i+1} - Z_i)}\,. \tag{10}$$

Therefore, the reflection coefficient sequences can be obtained from the impedance sequences, in other words $R_i(\omega)$ can be obtained using $Z_i(\omega)$. The result is then substituted into (8) and in this way we can get the matched operator $W_i(\omega)$ in the frequency domain.

Notably, when the gap of the energy between the amplitude spectrum of seismic records and wave impedance spectrum of wells is large, before any inversion, the equalization process must be carried out for the wave impedance spectrum energy and amplitude spectrum energy of seismic records of each track on wells.

The colored inversion technique is an inversion method which is not directly dependent on the well data. When this method defines inversion operator in the frequency domain, well data only play the contrast effect for the desired results in the inversion operator estimation. The inversion is the deconvolution process, therefore, the inversion result is impedance, and the amplitude keeps good lateral changes.

The colored inversion process is schematically indicated in Fig. 8:

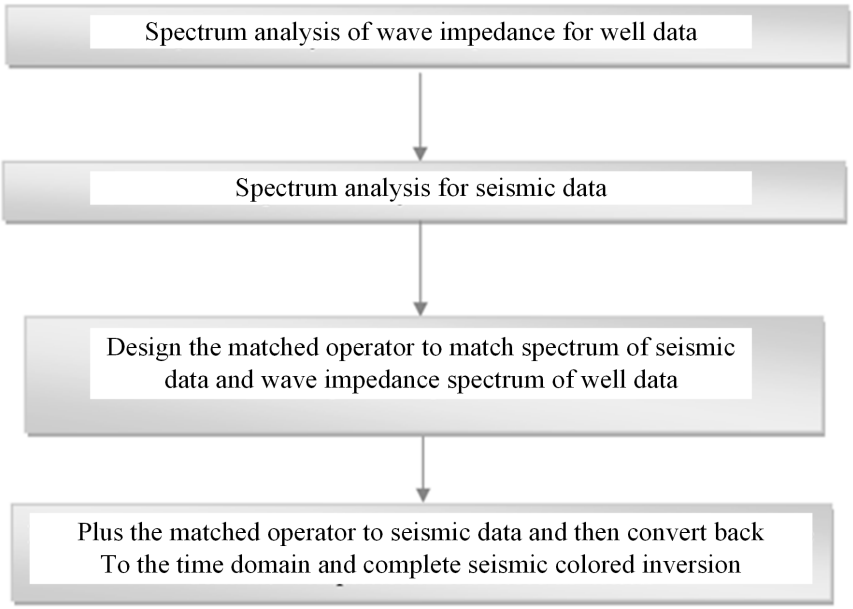

Fig. 8. Process of color inversion

This technology has the following characteristics: 1) no obvious wavelet extraction process; 2) no constraints of the initial model; 3) it can keep the lateral changes of amplitude and is conducive to the planar sedimentary imaging. Therefore, the results of colored inversion completely retain seismic data whose numerical magnitude is close to well data on the whole. This can be considered an attribute with inversion significance, that meets requirements of the seismic lithologic body [8–11].

Figure 9 depicts the impedance inversion section of the South China Sea.

Finally, Fig. 10 mainly shows inversion profiles of colored inversion about well LW21-1-1 and neighboring well LW3-1-1.

As can be seen from Fig. 10, as the phase axes of impedance of cross well profile are continuous and stable and do not exhibit any large impedance discontinuities,

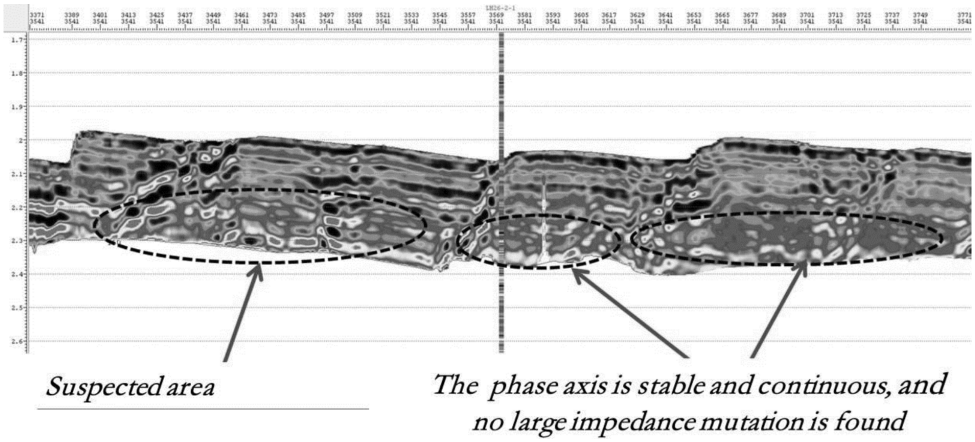

Fig. 9. Impedance inversion profile of the South China Sea

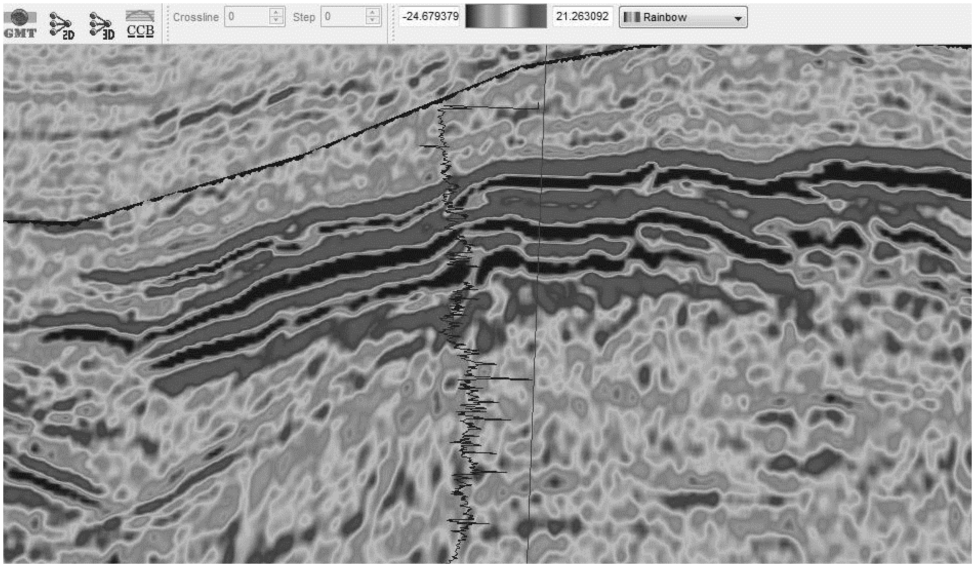

Fig. 10. 1569 line colored inversion profile of LW3-1-1

the figure suggests that the media is basically isotropic around the well and shows no evidence of gas hydrates occur.

The correct rate of pre-drilling shallow geologic disaster forecast depends largely on the selection of prediction methods, so that the correct method of disaster identification is very important. Choice of a correct approach must understand that parameters of shallow disasters we had studied have relationships and are most closely associated one to another. Researching natural gas hydrate can also use speed and other attribute parameters of seismic wave, and generally, there are many ways to

predict shallow disasters, improving the accuracy of recognition [12-14], which make the results more convincing.

3. Discussion

1. After a lot of research and field practice, we found that the research method also has some limitations. For some complex structure areas, the accuracy of a single technologies must still to be improved. To improve the accuracy of recognition we try to use a variety of methods, which will improve the accuracy of forecast and make the results more convincing.

2. Because of the lack of more detail of the deepwater seismic and logging data, it is difficult to get a better prediction results. But with the development of deepwater drilling technology, we will get more and more reliable data, so that the prediction accuracy will be better and better in the future.

3. The result in natural gas hydrate prediction of LW21-1-1 well area is correct. It means that our method is reliable.

4. Conclusion

The main results following from the research may be divided as follows:

- Three instantaneous profiles refer to instantaneous amplitude, instantaneous phase and instantaneous frequency. They respectively reflect the intensity of seismic wave, event connection and the time rate of change of phase. According to seismic response of natural gas hydrate layer and BSR characteristics, the process of three instantaneous attributes can reflect changes of reflected characteristics from another angle. BSR characteristic of amplitude blanking zone can be better reflected in the instantaneous amplitude profile. Intersected state between BSR reflection and reflection of the real stratum can be better reflected in the instantaneous phase profile. Instantaneous frequency is a reflection of dielectric absorption characteristic and abnormal BSR zone can be shown in the instantaneous frequency profile.

- Colored inversion technique of seismic data is an inversion method of direct conversion, fully established on the basis of seismic data. It retains basic characteristics of reflection on seismic data such as fault, occurrence, speed and so on. There also does not exist multiplicity based on the model inversion method. This method excludes interference of artificial analysis and can reflect the spatial change of the lithology obviously and objectively. It can also better reflect changes of physical properties of reservoirs based on stable condition of lithology. In addition, the inversion process does not require complex modeling, therefore it greatly improves the speed of inversion; on the other hand, it can remove inversion errors that impedance models are not allowed to bring.

- Analysis of the after-drilled data of well LW21-1-1 confirmed that the prediction was correct. Thus, it indicates that the application of technology of three instantaneous profile processes and colored inversion technique to predict natural gas hydrate is feasible.

References

[1] W. S. Borowski, C. K. Paull, W. Ussler: *Global and local variations of interstitial sulfate gradients in deep-water, continental margin sediments: Sensitivity to underlying methane and gas hydrates.* Marine Geology. *159* (1999), Nos. 1–4, 131–154.

[2] R. Boswell: *Is gas hydrate energy within reach?.* Science *325* (2009), No. 5943, 957 to 958.

[3] B. A. Hardage, H. H. Roberts: *Gas hydrate in the Gulf of Mexico: What and where is the seismic target?.* The Leading Edge *25* (2006), No. 5, 566–571.

[4] H. Hong, M. Pooladi-Darvish: *Simulation of depressurization for gas production from gas hydrate reservoirs.* J Canadian Petroleum Technology *44* (2005), No. 11, 39–46.

[5] I. R. MacDonald, N. L. Guinasso, Jr. R. Sassen, J. M. Brooks, L. Lee, K. T. Scott: *Gas hydrate that breaches the sea floor on the continental slope of the Gulf of Mexico.* Geology. *22* (1994), No. 8, 699–702.

[6] T. H. Shipley, M. H. Houston, R. T. Buffler, F. J. Shaub, K. J. McMillen, J. W. Ladd, J. L. Worzel: *Seismic evidence for widespread possible gas hydrate horizons on continental slopes and rises.* AAPG Bulletin *63* (1979) No. 12, 2204–2213.

[7] M. T. Taner, F. Koehler, R. E. Sheriff: *Complex seismic trace analysis.* Geophysics *44* (1979), No. 6, 1041–1063.

[8] H. W. Posamentier: *Seismic stratigraphy into the next millennium: a focus on 3D seismic data.* Proc. AAPG Annual Conference of Petroleum Geologies, 16—19 April 2000, New Orleans, LA, USA, 16–19.

[9] H. W. Posamentier, V. Kolla: *Seismic geomorphology and stratigraphy of depositional elements in deepwater settings.* J Sediment. Research *73* (2003), No. 3, 367–388.

[10] H. W. Posamentier: *Seismic geomorphology: imaging elements of depositional systems from shelf to deep basin using 3D seismic data: implications for exploration and development.* Geol. Soc. London Memoirs *29* (2004), No. 1, 11–24.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 153–164

Research and design of PFC control based on DSP

Ma Yuli[1], Ma Yushan[1]

Abstract. A realization scheme of single-phase power factor correction (PFC) circuit based on digital signal processor (DSP) is proposed and discussed. The parameters of the main circuit, signal conditioning circuit, current loop and voltage loop of PI control are designed. Finally, a prototype of the system is built and the actual measured results show that the power factor is above 0.99 at full load.

Key words. PFC, DSP, current loop, voltage loop, PI control.

1. Introduction

With the development of power electronic technology and wide application of power electronic devices in various areas, the current harmonic problem of the power grid is increasingly serious, and the harmonic pollution brings a series of harm not only to the system itself, but also to the surrounding electromagnetic environment [1]. The power factor correction (PFC) can effectively eliminate the harmonics of the rectifier, so it has a wide range of applications. Now, with the development of digital control technology, more and more control strategies is realized through digital signal processors (DSP) [2].

TMS320LF2406 is a kind of a low-price and high-performance chip introduced by TI company. Its processing ability is 30 MIPS, which is suitable for the control system of the motor and power supply transformation. Taking the widely used Boost-PFC circuit as an example, the current loop and voltage loop circuits are designed, using the TMS320LF2406 digital chip. Finally, a prototype is built, and the ideal dynamic and steady-state results are obtained in a practical application.

The basic parameters of the system are: output power $P_{\text{out}} = 800\,\text{W}$, input AC voltage range is $180 - 300\,\text{V}$, input voltage frequency is $45 - 65\,\text{Hz}$ and output DC voltage is $415\,\text{V}$.

The scheme of the system is depicted in Fig. 1.

[1]Qingdao Ocean Shipping Mariners College, 84 Jiangxi Rd, Shinan Qu, Qingdao Shi, Shandong Sheng, 266000, China

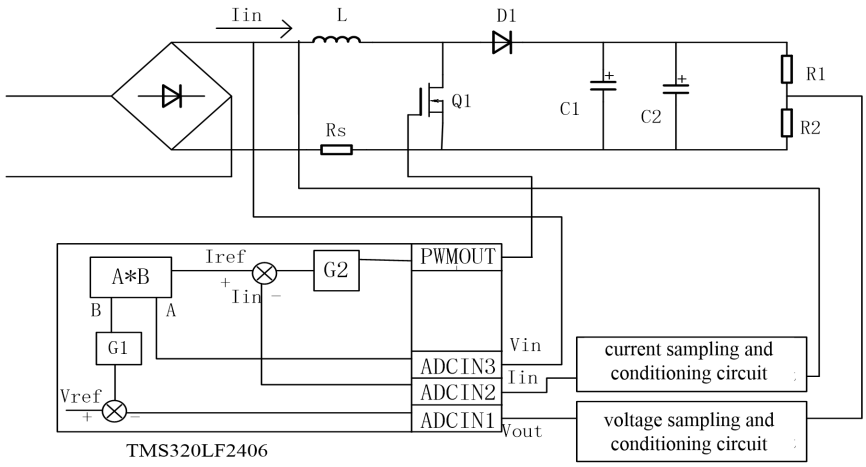

Fig. 1. Block diagram of a PFC converter based on DSP control

The rectified input voltage V_{in}, inductor current I_{in} and DC output voltage V_{out} are transformed, filtered, isolated by the corresponding sample and conditioning circuits, and then they are sent to the ADC module of DSP [3]. The output voltage V_{o} is sampled, then it is compared with the reference voltage V_{ref} and then the offset is sent into the voltage PI controller G1. The output of the voltage loop controller is defined as B, the input voltage is defined as A, and the two parameters are multiplied by the multiplier as the reference current I_{ref} of the current loop [4]. The inductor current I_{in} is compared with the reference signal I_{ref}, and the offset is taken as the input of PI controller of the current loop G2. The output value of current loop is compared with the value of the carrier signal in the DSP, the comparative value is used to change the PWM duty cycle in a switching period, and then through the isolation driving circuit makes the power switch turn-on and turn off [5]. In this way, the average value of the input current can track the input voltage waveform, so that the input current and input voltage are in the same phase, and the output voltage is stable [6].

The main circuit works in the CCM mode of the Boost converter, realizing voltage conversion and power factor correction. The digital control circuit is the core of the system, mainly converting analog signals to digital signals and achieving control algorithm of PI control of the voltage loop and current loop; the output signal is amplified by a switch tube driving circuit, then it is sent to the switch tube of the main circuit to control turn-on and turn-off of the switch tube [7].

2. The control circuit design

The objective of control circuit design is to determine the main circuit parameters [8], PI parameters of the voltage loop and current loop and the structure of signal sampling and conditioning circuits.

In the use of the Bode diagram method, the PI regulator is converted to the problem of zero point setting, only needing to set a reasonable zero, and then according to the ratio of the Bode diagram to determine the value of KP and KI [9].

2.1. The main circuit design

The main circuit inductance is operating in the current continuous mode (CCM). The continuation of the inductor current depends on the output power, switching frequency, inductance size and other parameters. According to the calculation, the value of the inductance is selected 78 μH. The output capacitor C depends on ripple size of the required output voltage, switching frequency, etc. According to the calculation, the design selects two 220 μF electrolytic capacitors. The current sampling resistor R_s is selected 30 MΩ. The PFC principle diagram of the main circuit is shown as Fig. 2.

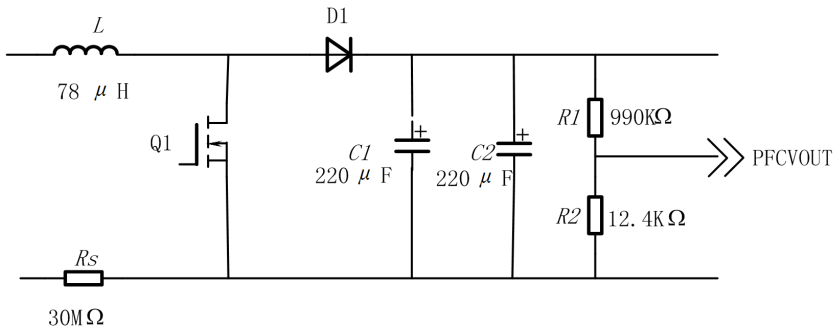

Fig. 2. PFC principle diagram of the main circuit

The switching frequency of the main circuit is selected 80 kHz. In order that the current loop cannot be disturbed by the noise of the switch, at the same time the current loop cannot also be affected by the low bandwidth, otherwise it will affect the effect of the current waveform to follow the voltage waveform [11]. In order to prevent from influencing the bus voltage ripple on the PFC controller, the crossing frequency of the voltage loop is controlled in less than half of the input electric frequency, in this paper, the calculate frequency of the voltage loop is selected 160 kHz/8=20 kHz.

2.2. PFC current loop design

When the PFC inductor current is CCM mode, the dynamic block diagram of the S-field of the PFC current loop is shown in Fig. 3.

Here, V_{Rs} is the voltage on the PFC input current sampling resistor, V_I is the PFC input voltage, V_o is the PFC output DC voltage, D is the duty cycle, L is PFC inductor, R_s is PFC input current sampling resistor. Considering 260 μH inductor of the analog LISN's low frequency impedance characteristic and the differential-mode component of the input filter, the circuit model of the PFC current loop is shown in

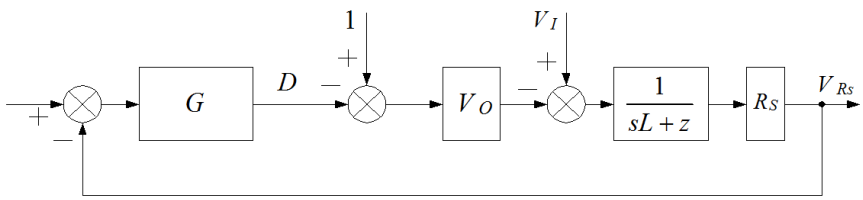

Fig. 3. Dynamic block diagram of S-field of PFC current

Fig. 4 with the source AC power supply.

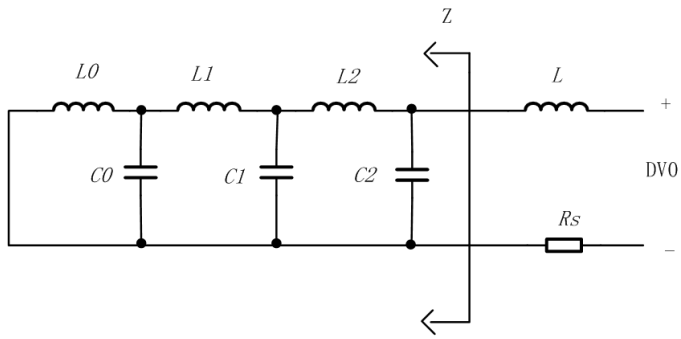

Fig. 4. Circuit model of the PFC current loop considering LISN'S inductor and filter

Here, L is 260 µH inductor of the analog LISN's low frequency impedance characteristic (actual loop test is 0). Symbol C_0 is 2.2 µF capacitor of the module input port, L_1 is 24 µH differential-mode component of input common-mode inductor, C_1 is 0.47 µF capacitor of between input common-mode inductors, L_2 is the sum of 13 µH differential-mode component of input common-mode inductor and 138 µH input differential-mode inductor. Finally, C_2 is 1 µF capacitor placed after the input rectifier.

Assuming that the impedance of the AC source is 0, the calculation formula of impedance Z is as follows:

$$Z = \frac{k_1 s^5 + k_2 s^3 + k_3 s}{k_4 s^6 + k_5 s^4 + k_6 s^2 + 1}\,, \tag{1}$$

where

$$k_1 = L_1 L_2 C_1 L_0 C_0\,, \tag{2}$$

$$k_2 = L_1 L_0 C_0 + L_2 L_0 C_0 + L_2 L_1 C_1 + L_2 L_0 C_1\,, \tag{3}$$

$$k_3 = L_0 + L_1 + L_2\,, \tag{4}$$

$$k_4 = L_1C_1L_2C_2L_0C_0\,, \tag{5}$$

$$k_5 = L_1C_2L_0C_0 + L_2C_2L_0C_0 + L_2C_2L_1C_1 + L_2C_2L_0C_1 + L_1C_1L_0C_0\,, \tag{6}$$

$$k_6 = L_0C_2 + L_1C_2 + L_2C_2 + L_0C_0 + L_1C_1 + L_0C_1\,. \tag{7}$$

Considering the influence of the input filter, the controlled object of the current loop is changed into

$$P_0(s) = \frac{V_{\mathrm{O}}R_{\mathrm{S}}}{sL + Z} = \frac{V_{\mathrm{O}}R_{\mathrm{S}}[k_4s^6 + k_5s^4 + k_6s^2 + 1]}{k_7s^7 + k_8s^5 + k_9s^3 + k_{10}s}\,, \tag{8}$$

where

$$k_7 = L_1C_1L_2C_2L_0C_0L\,, \tag{9}$$

$$\begin{aligned} k_8 = {} & L_1C_2L_0C_0L + L_2C_2L_0C_0L + L_2C_2L_1C_1L + L_2C_2L_0C_1L + \\ & + L_1C_1L_0C_0L + L_1C_1L_0C_0L_2\,, \end{aligned} \tag{10}$$

$$\begin{aligned} k_9 = {} & L_0C_2L + L_1C_2L + L_2C_2L + L_0C_0L + L_1C_1L + L_0C_1L + L_1L_0C_0 + \\ & + L_2L_0C_0 + L_2L_1C_1 + L_2L_0C_1\,, \end{aligned} \tag{11}$$

$$k_{10} = L + L_0 + L_1 + L_2\,. \tag{12}$$

This block diagram is effective in the frequency band between the resonant frequency (860 Hz) of the LC filter in the PFC circuit and switching frequency.

The PFC current signal sampling and conditioning circuit scheme is shown in Fig. 5. The circuit uses a differential amplifier to amplify the circuit, using in series the first-order RC filter with the cut-off frequency of 50 kHz and second-order RC filter with 47 kHz.

Through the analysis, we can derive that the transfer function of the PFC current signal sampling circuit is

$$H(\mathrm{S}) = \frac{8.12}{(3.198 * 10^{-6}\mathrm{s} + 1)(1.122 * 10^{-11}\mathrm{s}^2 + 1.73 * 10^{-5}\mathrm{s} + 1)}\,. \tag{13}$$

The 3V signal in PFC AD sampling port corresponds to the number 1023, while the output digital quantity 2032 of the PFC current loop controller corresponds to the PWM pulse width duty cycle 1, therefore, through the synthesis, the discrete dynamic block diagram of the PFC current loop is shown as Fig. 6, T being the calculation frequency of PFC current loop.

With the calculation frequency of 50 kHz, the Z transform is performed on the

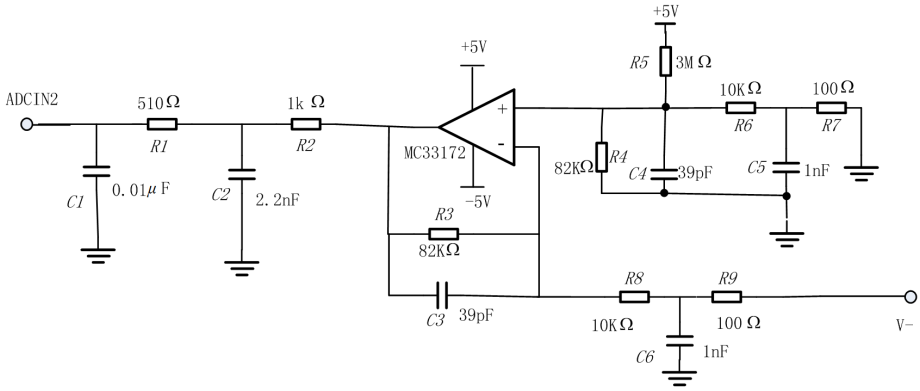

Fig. 5. PFC current signal sampling circuit

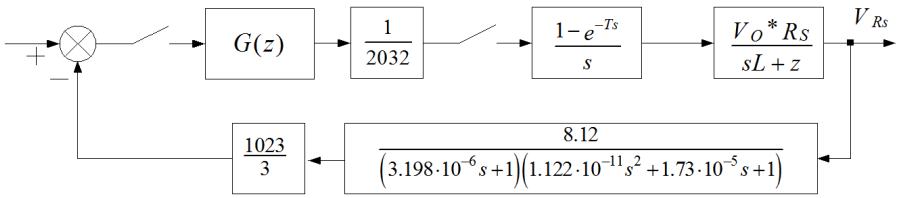

Fig. 6. Discrete dynamic block diagram of PFC current loop

part of the block diagram except the open transfer function. The resultant function is given as

$$P(z) = A(z)/B(z)\,,$$

where

$$\begin{aligned} A(z) &= 0.421z^{-1} + 3.72z^{-2} + 49.1z^{-3} + 42.7z^{-4} - \\ &- 25.4z^{-5} + 56.4z^{-6} + 36.2z^{-7} + 2.71z^{-8}\,, \\ B(z) &= 01 + 9.69z^{-1} + 109z^{-2} - 116z^{-3} + 114z^{-4} - \\ &- 159z^{-5} + 31.3z^{-6} + 9.71z^{-7} + z^{-8}\,. \end{aligned} \tag{14}$$

In order to use frequency analysis method to design, the Z domain can be changed to the W domain based on the bilinear transformation method. The Bode figure of $P(w)$ is shown in Fig. 7 (horizontal coordinate is for frequency, units are rad/s). In order to improve the PFC current inhibition ability on the third and fifth harmonics in the input current to obtain lower THD, PFC current loop controller $G(w)$ must contain an integral part, and the zero point cannot be too far below the cut-off frequency of the open loop transfer of the PFC current loop. Let us assume that after correction the open loop cutoff frequency of the PFC current loop is 5 kHz=$3.14\cdot 10^4$ rad/s. Due to the inherent existence of a calculating cycle delay $T = 12.5$ μs in the digital control, thus, at the 5 kHz frequency the phase lag of $\frac{5k}{80k}\cdot 180° = 11.25°$ is introduced. In order to get a sufficient stability, we can set two zero points for 3 kHz;

at the same time, in order to improve the anti-interference ability of the current loop, we set a pole point in 6 kHz. The structure of the current loop controller in W domain is as follows

$$G(w) = \frac{K * \left(\frac{w}{2\pi \cdot 3k} + 1\right)\left(\frac{w}{2\pi * 3k} + 1\right)}{w\left(\frac{w}{2\pi * 6k} + 1\right)}. \tag{15}$$

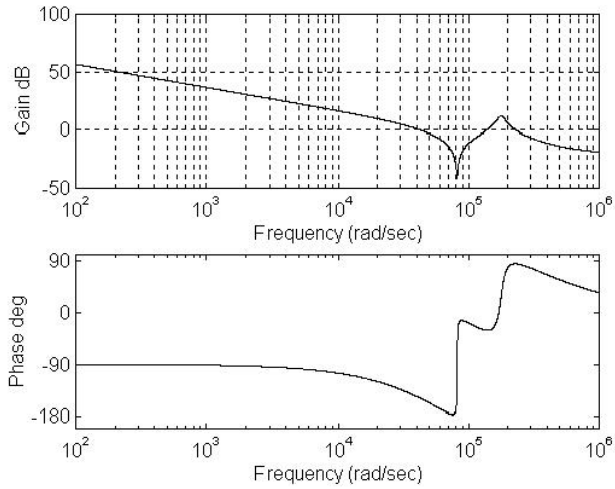

Fig. 7. W-domain Bode diagram of PFC current loop controlled object

In order that the cutoff frequency of the open loop of the corrected PFC current loop is 5 kHz=$3.14 \cdot 10^4$ rad/s, the open loop gain can be set $K = 6400$. Based on the bilinear transformation method $G(w)$ will be transformed from the W domain to the Z domain as

$$G(z) = \frac{0.6867 - 1.0839z^{-1} + 0.4277z^{-2}}{1 - 1.6186z^{-1} + 0.6186z^{-2}}. \tag{16}$$

After the correction of the Z domain for the PFC current ring open loop transfer function, the Bode diagram is shown in Fig. 8. From the graph we can see that the frequency of the first pass through the 0 dB line in the amplitude-frequency curve of the open loop transfer function is 4.89 kHz=$3.071 \cdot 10^4$ rad/s. The phase margin is 33.26°. In the simulation results, at the position of 16 kHz of the amplitude frequency curve, there are two crossing phenomena, the phase margin is negative in the position of two crossings, which will pose a threat to the stability of the module. In order to facilitate the preparation of the control program, $G(z)$ using a differential equation form is

$$I_{\mathrm{c}}(k) = 1.6186I_{\mathrm{c}}(k-1) - 0.6186I_{\mathrm{c}}(k-2) + 0.6867I_{\mathrm{e}}(k) -$$

$$- 1.0839I_{\mathrm{e}}(k-1) + 0.4277I_{\mathrm{e}}(k-2). \tag{17}$$

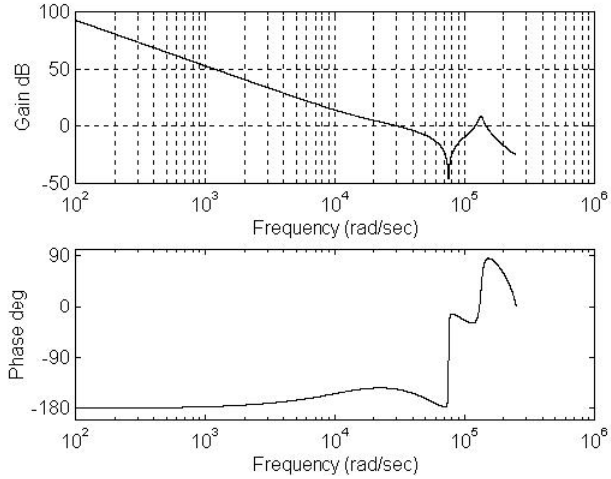

Fig. 8. Z-domain Bode diagram of PFC current loop open loop transfer function

Since the TMS320LF2406 is a 16bit DSP chip, in order to avoid overflow in the calculation, the coefficients in the differential equation are calibrated using Q10.

When the AC Source is 4 kVA, the PFC current loop is tested. The test conditions are as follows: the injection resistance is 100 Ω, sweep signal amplitude is 3 mV, the average number of times is 4, sweep speed is medium, and the frequency test range is 100 Hz–200 kHz.

The test results show that in the 220 V DC input, 15 A output, PFC inductor current is in the DCM state and the bandwidth of the PFC current loop is narrower than that of the CCM mode, only 763 Hz. The 180 V DC input, 15 A output, PFC inductor current is close to the CCM state and PFC current loop bandwidth will increase to 3.176 kHz.

2.3. PFC voltage circuit design

PFC current inner loop and power stage form a current source, while the controlled object of PFC voltage loop at low frequencies is equivalent to a current source for driving capacitor. The S-domain dynamic block diagram of the PFC voltage loop is shown in Fig. 9, in which P_{in} is the average input power of the PFC circuit, V_{O} is the output DC voltage, C is the output capacitor, and G is the voltage loop controller of PFC.

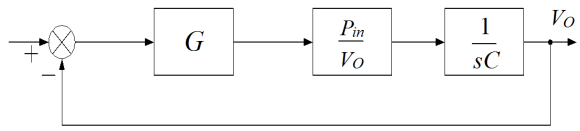

Fig. 9. S-domain dynamic block diagram of the PFC voltage loop

Under the premise of ensuring the stability of the PFC voltage loop, the bandwidth should be low enough to reduce the modulation effect of the 100 Hz voltage ripple on the PFC output capacitor to the PFC input current, otherwise the modulation will cause serious distortion of the input current.

The PFC voltage signal sampling and conditioning circuit scheme is shown in Fig. 10 (Fig. 2 comprehensively shows that the circuit uses a resistor divider to make sure that the output voltage is less than 3 V). A first-order RC filter with a cutoff frequency of 288 Hz is used to filter the high frequency noise. The voltage follower formed by LM2904 completes following and isolating, in order to enhance the output capacity of the circuit. Then in the AD port it adopts second-order RC filters with a cut-off frequency 33 kHz.

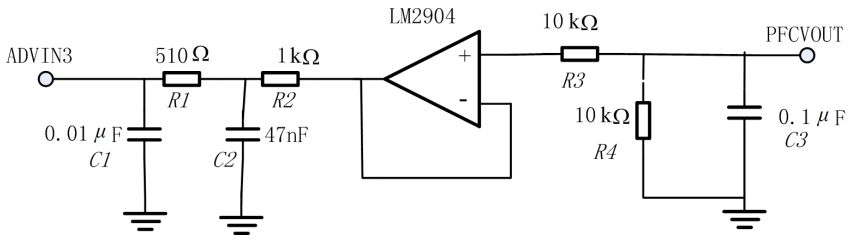

Fig. 10. Sampling circuit for PFC voltage signal

Through the analysis, (ignoring the second-order RC filter with a cut-off frequency 33 kHz in the AD port) we can see that the transfer function of the PFC voltage signal sampling circuit is

$$H(S) = \frac{0.00556}{0.5536 \cdot 10^{-3} S + 1}\,. \tag{18}$$

In the PFC voltage signal of AD sampling port, 3 V analog signal corresponds to the digital signal 1023, while in the PFC voltage loop controller output port the digital signal 2032 stands for "1". In order to further reduce the modulation effect of the 100 Hz voltage ripple on the PFC output capacitance to the PFC input current, the sampling value of the PFC output voltage is filtered by a first-order digital low-pass, and the cutoff frequency of the filter is 30 Hz. The filter in the W domain is expressed as

$$F(w) = \frac{1}{\left(\frac{w}{2\pi * 30} + 1\right)}\,. \tag{19}$$

At the calculation frequency of 20 kHz, $F(w)$ is transformed from the W domain to $F(z)$ in the Z domain, based on the bilinear transformation method.

$$F(z) = \frac{0.0047 + 0.0047 z^{-1}}{1 - 0.9906 z^{-1}}\,. \tag{20}$$

All the coefficients in the first-order digital low-pass filter are calibrated by the Q16 in the program.

After a comprehensive analysis, the discrete dynamic block diagram of the PFC voltage loop is shown in Fig. 11. Here T is the calculation frequency of the PFC voltage loop. At the frequency of 20 kHz, Z transform is used in the open loop transfer function of the block diagram in addition to the $G(z)$ and the digital low-pass filter, getting $P(z)$ as follows

$$P(z) = \frac{0.1054 \cdot 10^{-4} z^{-1} + 0.1023 \cdot 10^{-4} z^{-2}}{1 - 1.9136 z^{-1} + 0.9136 z^{-2}} \,. \tag{21}$$

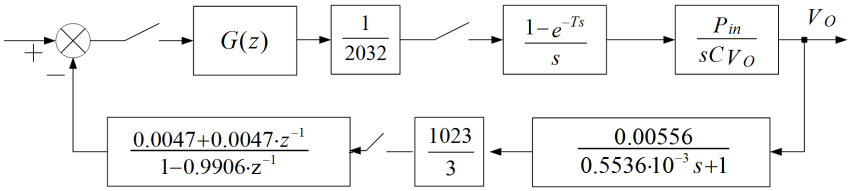

Fig. 11. Discrete dynamic block diagram of the PFC

Based on the bilinear transformation method, $P(z)$ is transformed from the Z domain to the W domain. The Bode diagrams of the W domain controlled object are shown in Figs. 12 and 13.

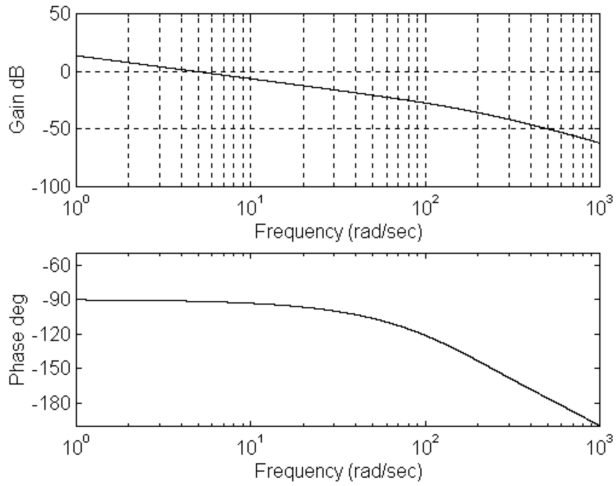

Fig. 12. Controlled object Bode diagram of W domain PFC voltage loop

The cutoff frequency of the open loop transfer function after adjustment of the PFC voltage is 12 Hz= 75.36 rad/s, and due to in digital control inherent delay there exists delay of about a calculation cycle time. The phase lag was introduced at frequency 12 Hz but as its value is small, it can be ignored. In order to get enough phase margin in the cutoff frequency of the PFC voltage loop, the controller of the PFC voltage loop adopts PI regulator, zero point is set to 2 Hz, and the controller

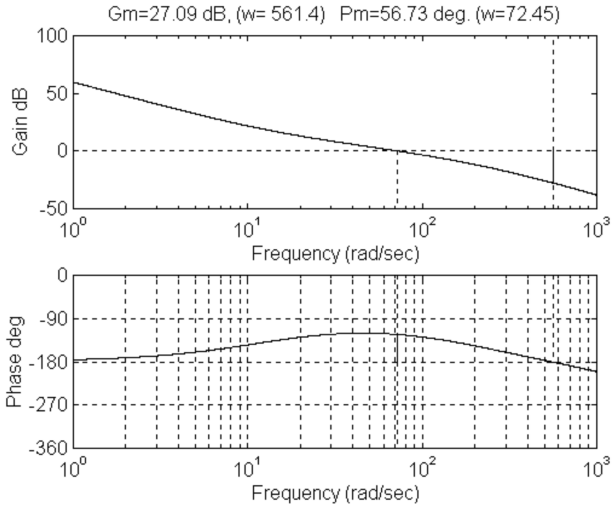

Fig. 13. Open loop transfer function Bode diagram of Z domain PFC voltage loop

structure of the W domain voltage loop is as follows

$$G(w) = \frac{K * \left(\frac{w}{2\pi * 2} + 1\right)}{w} . \tag{22}$$

In order to make the cutoff frequency of the open loop transfer function of the corrected PFC voltage loop 12 Hz= 75.36 rad/s, the open loop gain can be set $K = 200$. Based on the bilinear transformation method $G(w)$will be transformed from the W domain to the Z domain as

$$G(z) = \frac{15.9205 - 15.9105 z^{-1}}{1 - z^{-1}} . \tag{23}$$

Open loop transfer function Bode diagram of the corrected Z domain PFC voltage ring is shown in Fig. 14. The diagram shows that the amplitude cross-over frequency of the open loop transfer function is 11.5 Hz = 72.45 rad/s, the phase margin is 56.73°, and the gain margin is 27.09 dB.

In order to facilitate the preparation of control procedures and reduce the impact of truncation error, the PI controller uses the position algorithm, where the differential equations adopted are expressed in the form of

$$\mathrm{PI}(k) = \mathrm{PI}(k-1) + K_{\mathrm{I}} V_{\mathrm{e}}(k) , \tag{24}$$

$$V_{\mathrm{c}}(k) = K_{\mathrm{P}} V s_{\mathrm{e}}(k) + \mathrm{PI}(k) , \tag{25}$$

where

$$K_{\mathrm{P}} + K_{\mathrm{I}} = 15.9205 , \tag{26}$$

and, particularly

$$K_{\mathrm{P}} = 15.9105 , \quad K_{\mathrm{I}} = 0.01 , \tag{27}$$

All the coefficients in the first-order digital low-pass filter are calibrated by the Q16 in the program.

Adopting AC source 4 kVA for the DC power supply, the PFC voltage loop is tested. The test conditions as follows: the injection resistance is 1 k and the amplitude of sweep signal is 1 V, the average number of times is 16, the sweep speed is "Long", the frequency test range is 10 Hz–100 kHz. The experimental results show that the voltage has reached the design requirements.

3. Conclusion

In this research, the Boost PFC circuit based DSP control is studied, and the parameter selection of the main circuit is given. The design scheme and control flow of current loop and voltage loop are proposed for digital control, and an experimental prototype is designed. The actual measured results show that the input voltage is well tracked by the input current and the power factor is 0.99 at full load. Use of the DSP control parameter adjustment is more convenient, the system upgrade is easier and not easy to be influenced by aging and temperature drift. The number of devices is reduced and the anti-interference ability of the system is increased.

References

[1] C. K. Tse, M. H. L. Chow, M. K. H. Cheung: *A family of PFC voltage regulator configurations with reduced redundant power processing*. IEEE Trans. Power Electron. *16* (2001), No. 6, 794–802.

[2] A. M. Barry, D. Maksimović: *A simple digital power-factor correction rectifier controller*. IEEE Trans. Power Electron. *26* (2011), No. 1, 9–19.

[3] M. G. Umamaheswari, G. Uma, K. M. Vijayalakshmi: *Design and implementation of reduced-order sliding mode controller for higher-order power factor correction converters*. IET Power Electron. *4* (2011), No. 9, 984–992.

[4] K. T. Hassan, T. Kahawish: *A repetitive-PI current controller for boost single phase PFC converters*. Energy and Power Engineering. *3* (2011), No. 2, 69–78.

[5] G. Chu, C. K. Tse, S. C. Wong, S. C. Tan: *A unified approach for the derivation of robust control for boost PFC converters*. IEEE Trans. Power Electron. *24* (2009), No. 11, 2531–2544.

[6] C. Petrea: *Digital control of boost PFC converter working in discontinuous conduction mode*. Advances in Electrical and Computer Engineering. *7* (2007) No. 2, 16–22.

[7] M. Gopinath, S. Ramareddy: *Experimental results of bridgeless PFC boost converter*. Annals of Dunarea de Jos, University of Galati, Fascicle III *33* 2010, No. 1, 19–22.

[8] F. M. L. L. de Belie, D. M. van de Sype, K. De Gussemé, W. R. A. Ryckaert, J. A. A. Melkebeek: *Digitally controlled boost PFC converter with improved output voltage controller*. Electr. Eng. *89* (2007), No. 5, 363–370.

[9] D. M. Van de Sype, K. De Gussemé, A. P. Van den Bossche, J. A. A. Melkebeek: *A sampling algorithm for digitally controlled boost PFC converters*. IEEE Trans. Power Electron. *19* (2004), No. 3, 649–657.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 165–174

Single image super resolution algorithm based on local energy and anisotropic filter in NSCT domain[1]

Zhang Wei[2], Chen Wei[2]

Abstract. A novel super resolution algorithm is proposed for preserving the texture and edge information and improving the space resolution of single image (this technology is widely used, for example, in the field of military reconnaissance or space observation). The sub-band coefficients of the original image are obtained through Nonsubsampled Contourlet Transform (NSCT). Then, the noise of high sub-band is removed a new high frequency coefficient together with coefficient of variation weighted interpolation are obtained. The low sub-band coefficients are enhanced and estimated low frequency coefficients. Finally, the high resolution image is obtained by inverse NSCT. Experiments on both synthetic images and natural images demonstrate the effectiveness of the proposed method. Compared with traditional super resolution reconstruction algorithms, the proposed method can preserve the structure of image and resist Gaussian noise.

Key words. Nonsubsampled Contourlet Transform (NSCT), image interpolation, coefficient of variation, super resolution reconstruction..

1. Introduction

Super resolution is the process of combining a sequence of low-resolution (LR) noisy blurred images to produce a higher resolution image or sequence [1]. Super resolution from image sequences obtain better results than from single image since the former use these similar but not identical characters among the image sequences. Many effective methods have been proposed for super resolution based on multiple low-resolution images of the same scene such as reconstruction-based resolution [2], [3] or Learning-based resolution [4], [5]. In fact, it is very difficult to acquire sequence images for the same scene in the practical application such as military investigation and space observation. Therefore, the research of single image super resolution

[1]The research work is fully supported by the Science and Technology Program of Jilin (Grant No. 20150441003SC), the Science and Technology Program of Jilin (Grant No. 20140101197JC), and the Science and Technology Program of Jilin City (Grant No. 2015554003).

[2]Department of Computer Science and Technology, Beihua University, Jilin, 132021, China

possesses important significance in practical applications. Traditional single image super resolution methods always caused the image edge information blurred that cannot perform well for image characteristics since there is no redundant information in single image. The main purpose of this paper is to improve the space resolution of single image without losing the edge and texture features.

A single image super resolution algorithm is proposed in this paper. The basic idea is to decompose the LR image through NSCT and use coefficient of variation weighted interpolation along the edge direction. A nonlinear function is used to enhancement the dim-target. Then a super-resolution image can be obtained by inverse NSCT. Several experimental results using simulated images demonstrate that the new method is efficient for single image super resolution. Compared with traditional single super resolution algorithms, this algorithm can provide smooth edges and recovery the texture features, and also resists noise.

2. Proposed single image super resolution algorithm

2.1. Nonsubsampled Contourlet Transform

The contourlet transform is an efficient directional multi-resolution image representation. It is defined and implemented by the filter banks in the discrete domain. By a redundant pyramidal multi-resolution decomposition, the image was decomposed into a detail sub-band and approximate sub-band. Then a directional filter banks was applied on each of the detail sub-band to capture the singularity of direction. Different from the wavelet transform, the multi-directionality and the property that the support of basis functions has a variety of elongated shapes with different aspect ratios make it effectively capture the geometric structure features of the image information. More importantly, the contourlet transform can flexibly and effectively combine the multi-scale with the multi-directionality for image representation, therefore it can accurately and optimally describe an image [6]. The nonsubsampled contourlet transform(NSCT) is based on a nonsubsampled pyramid structure and nonsubsampled directional filter banks. It is a shift-invariant image decomposition that can be efficiently implemented and can well depress the Gibbs impact [7]. Compared to the contourlet transform, the NSCT has a better frequency selectivity and is sensitive to texture.

After NSCT, the energy of an image is distributed in various scales and directions. The coefficients have the feature of localization in space domain and exhibit dependencies across all of scales, space, and orientations. Figs. 1a–1e show the 'zoneplate' image, an exemplary frequency partition of the NSCT, low frequency coefficients and four directional sub-bands of the NSCT decomposition in two scales, respectively. It can be seen that the most energy concentrates in low frequency coefficients, so the low frequency coefficients reflect the main information of the image. The high frequency coefficients reflect the detail of image. Most coefficients are so small that they are invisible in high frequency sub-bands.

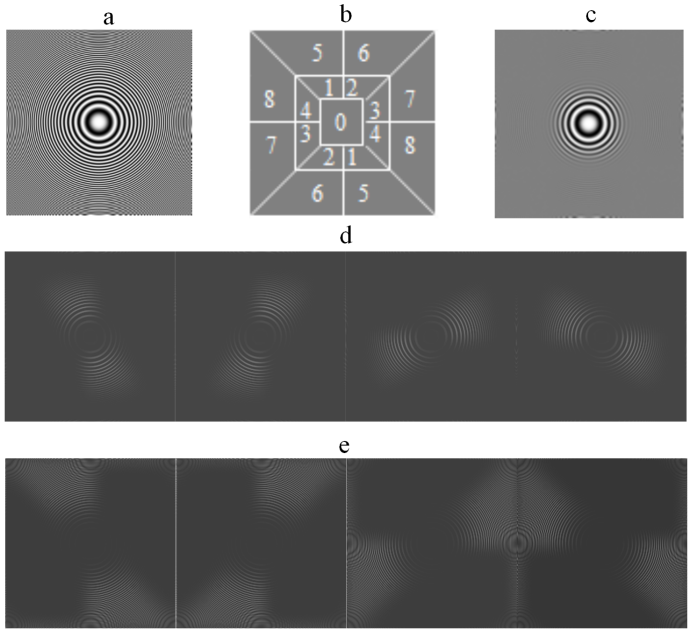

Fig. 1. Example of frequency partition and four directional sub-bands of NSCT coefficients: a–zoneplate image, b–frequency partition, c–low frequency coefficients, d–sub-bands 1–4 of NSCT coefficients, e–sub-bands 5–6 of NSCT coefficients

2.2. Coefficient of variation weighted interpolation in high frequency bands

To improve the resolution, the pixels in low resolution (LR) image should be mapped to the high resolution (HR) image. But there are many "holes" in the HR image to be filled. Interpolation is the simplest method to fill these holes. The traditional interpolation methods (bilinear interpolation, Newton interpolation) estimate interpolation point by neighborhood pixels, so blur is introduced in edge and texture, especially in weak edge and noise.

The NSCT coefficients of noise and weak edge are both small. Different from weak edge, most noises have not obvious geometrical structure, so they can be distinguished by the degree of correlation in different directions of the same scale. Reference [8] introduces a simple method that can remove the noise but preserves the edges as weak edges. First, it is necessary to calculate the high frequency bands threshold T_{B} by the Bayesian shrinkage and the coefficients that are lower than threshold are decided whether it is a noise by the correlation among the directional sub-bands in the same scale. The threshold is determined from the formulae

$$\overline{\delta}_{l,k} = \sqrt{\max\left(\frac{1}{m \cdot n}\sum_{i=1}^{m}\sum_{j=1}^{n} f_{l,k}^2(i,j) - \delta_{l,i}^2\right)}, \tag{1}$$

$$T_{\mathrm{B}} = \frac{\delta_{l,i}^2}{\overline{\overline{\delta}}_{l,i}}\,, \tag{2}$$

where $f_{l,k}(i,j)$ is the NSCT coefficient with coordinates i,j in the l-th scale and k-th direction. Symbol $\overline{\delta}_{l,k}$ is the standard deviation of the sub-band, $\delta_{l,i}$ is the standard deviation of noise and m,n are the sizes of sub-band image.

If the pixel value is lower than a threshold, its coefficient of variation c_{v} can be computed as

$$c_{\mathrm{v}} = \frac{\sqrt{\frac{1}{m\ldots n}\sum_{i=1}^{m}\sum_{j=1}^{n} f_{l,k}^2(i,j) - \delta_{l,i}^2}}{\frac{1}{m\cdot n}\sum_{i=1}^{m}\sum_{j}^{n} f_{l,k}^2(i,j)}\,. \tag{3}$$

The coefficient of noise is small and relatively stable in NSCT domain and the value of c_{v} is relatively small. The threshold $T_\gamma = 0.03\sqrt{m\cdot n}$ is set to differentiate between the noise and weak edges. The pixel will be labeled as noise if its value is lower than T_γ and be modified as the mean of neighborhood pixels.

For every pixel, interpolation should be performed along the edge directions in noise-free image to preserve the edges, so detect the information of edge direction is a necessary step. After the NSCT, contours and textures are mostly located in high frequency sub-bands and will be captured in responding directional sub-bands. Figure 1 shows the NSCT representation of the "Peppers" image. Taking a neighborhood with size of 3×3 for the pixel in the high frequency sub-band we can calculate the coefficient of variation. The maximum and second-maximum are $c_{\mathrm{v,m}}$ and $c_{\mathrm{v,s}}$, the responding directions are d_{m} and d_{s}, respectively. By the Newton interpolation along d_{m} and d_{s} we can obtain the f_{m} and f_{s}, respectively. The final value of the pixel is calculated as

$$f = \frac{cv_{\mathrm{s}}}{c_{\mathrm{v,s}} + c_{\mathrm{v,m}}} f_{\mathrm{s}} + \frac{cv_{\mathrm{m}}}{c_{\mathrm{v,s}} + c_{\mathrm{v,m}}} f_{\mathrm{m}}\,. \tag{4}$$

2.3. Dim target enhancement in low frequency sub-band

There are almost no noises in low frequency sub-band. But the reconstructed image is darker than LR image because of energy dispersion. To improve the visual effect, image enhancement is performed in low frequency sub-band. In addition, those dim-targets have smooth change of brightness from its neighborhood region, therefore they provide no clear edges. To enhance the dim-targets and make the background uniform, the smoothing process should be performed in the faultless areas and the sharpening process should be performed in the local dim areas. Therefore, a gray level transformation in the spatial domain is introduced to enhance the global contrast of reconstructed image effectively. Because human vision is sensitive to contrast, maximum contrast can be used to estimate the threshold between the background and dim-target. If this threshold is t, the absolute contrast $c(t)$ and best threshold T can be obtained as

$$c(t) = \min(f(u) - t, f(l) - t)\,, \tag{5}$$

$$T = \underset{t}{\operatorname{Arg\,max}}(c(t))\,, \tag{6}$$

where $f(u)$ is the mean of those pixels whose values are greater than t while $f(l)$ is the mean of those pixels whose values are less than t.

Contrast stretching uses a nonlinear function

$$f = \operatorname{sign}(f(i,j)) \cdot (|f(i,j)| - f_{\min}) \cdot \left(\sin\left(\frac{\pi}{2} \cdot \frac{f(i,j) - f_{\min}}{T - f_{\min}}\right)\right) + f_{\min}$$

for $f_{\min} < |f(i,j)| < T$ and

$$f = \operatorname{sign}(f(i,j)) \cdot (|f(i,j)| - f_{\max}) \cdot \left(\cos\left(\frac{\pi}{2} \cdot \frac{T - |f(i,j)|}{f_{\max} - T}\right)\right) + f_{\max} \tag{7}$$

for $T < |f(i,j)| < f_{\max}$.

2.4. The arithmetic flow

The arithmetic flow consists of two following steps:

1. The LR image is transformed by NSCT. The LR image is decomposed with three stages by dmaxflat NSDFB and 9-7 NSP of the algorithm proposed in this paper, and then eight directions are decomposed in each high-frequency sub-band.

2. Remove the noise of high sub-band using equations (1)–(3) and get new high-frequency coefficient with coefficient of variation weighted interpolation described by equation (4).

3. Bilinear interpolation is used in low-frequency image. Then the coefficients are enhanced by equation (8) and estimated are low-frequency coefficients.

4. High resolution image is obtained by inverse NSCT with estimated coefficients that are obtained in steps 2 and 3.

3. Experimental results

The synthetic image and natural image are tested with Bicubic, algorithm proposed in [9], algorithm proposed in [10], and the above method. The processed images are in Figs. 2 and 3.

3.1. Experiment on synthetic image

A synthetic image is obtained by double down-sampling and adding the Gaussian noise. Figure 2 illustrates five images. Figure 2a is the part of "boat". Figs. 2b to 2d are the HR images reconstructed by the Bicubic, the algorithm proposed in [9], and

the algorithm proposed in [10],respectively. Figure 2e is the HR image reconstructed by the method proposed in this paper.

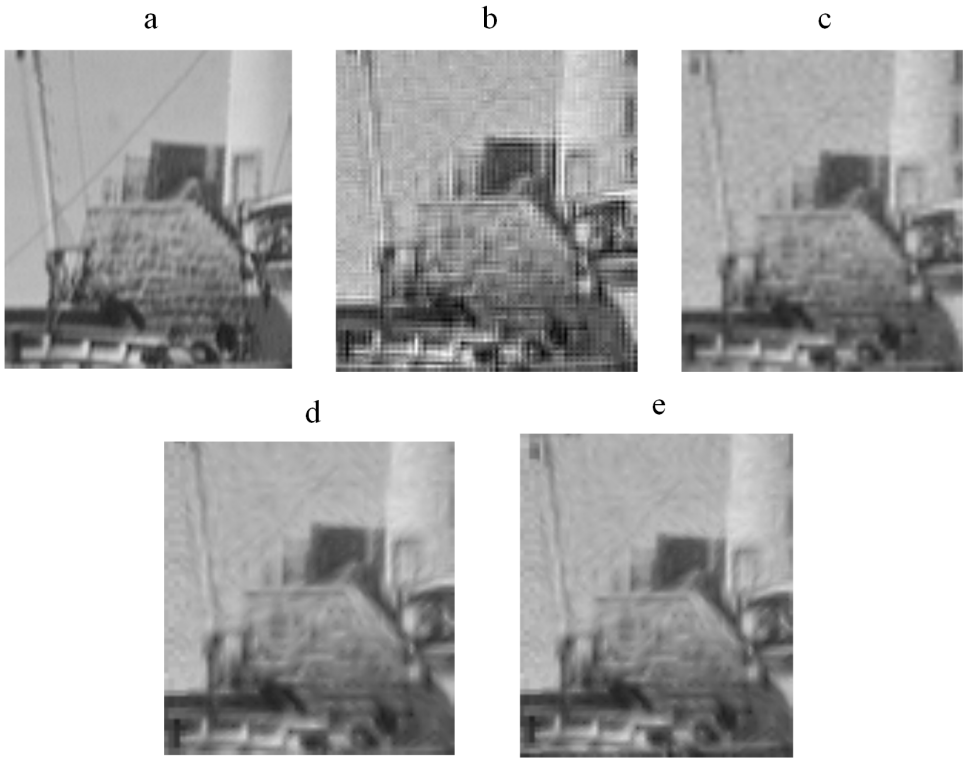

Fig. 2. Super-resolution results of part of boat: a–origin, b–Bicubic, c–processing by [9], d–processing by [10], e–processing by proposed algorithm

3.2. Experiment on natural image

We selected a natural image and tested it using various algorithm. There exist only global translation and rotation in the test image. Experimental results show the efficiency of the proposed algorithm in dim-target enhancement with low contrast.

4. Discussion

The simulation results show that the aliasing phenomenon is serious in image reconstructed by Bicubic. In Fig. 1a, the blocking artifacts are obvious on the mask and roof. SSIM, PSNR and entropy are tested in the LR image and reconstructed image and the results are showed in Tables 1–4. The proposed algorithm is more successful in noise suppression, edge enhancement and gets highest PSNR. The value of SSIM showed in Table 2 illustrates that the proposed algorithm can preserve the

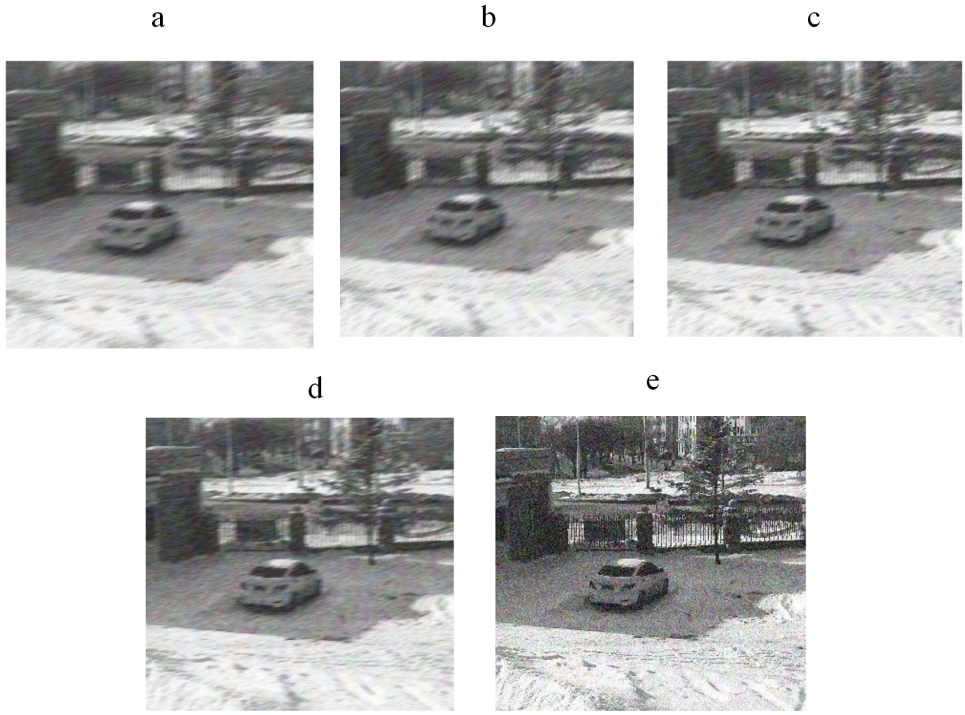

Fig. 3. Super-resolution results of natural image: a–origin, b–Bicubic, c–processing by [9], d–processing by [10], e–processing by proposed algorithm

structure of image effectively. Table 3 shows the corresponding result.

Compared with the bilinear interpolation, the PSNR of the proposed algorithm is improved by 2.4–16.7 %, the average PSNR increase is 5.8 %. Compared with the method of ref. [9], the PSNR of the proposed algorithm is improved by 0.1–3.8 %, and the average PSNR increase is 1.4 %. Compared with the method of ref. [10], the PSNR of the proposed algorithm is improved by 0.1–1.5 %, and the average PSNR increase is 0.35 %.

Table 1. Comparison of the PSNR obtained by different algorithms

image	Bicubic	ref. [9]	ref. [10]	proposed method
airport	29.257	30.784	30.785	30.808
baboon	26.868	28.140	28.424	28.613
barb	26.748	30.852	31.602	31.221
boat	28.676	29.644	29.987	29.988
couple	32.468	33.456	34.652	34.744
girl	31.458	31.782	32.413	32.440
moom surface	24.763	25.154	25.447	25.755
man	30.476	31.735	31.878	31.996
pixel ruler	29.367	30.452	30.553	30.576
testpat	31.782	32.418	32.568	32.669

Table 2. Comparison of the SSIM obtained by different algorithms

image	Bicubic	ref. [9]	ref. [10]	proposed method
airport	0.734	0.754	0.787	0.788
barb	0.759	0.812	0.816	0.828
boat	0.766	0.823	0.832	0.833
baboon	0.721	0.793	0.814	0.826
couple	0.785	0.812	0.825	0.842
girl	0.791	0.845	0.855	0.856
moom surface	0.719	0.763	0.784	0.797
man	0.782	0.803	0.826	0.832
pixel ruler	0.795	0.821	0.832	0.845
testpat	0.771	0.824	0.826	0.844

Compared with the bilinear interpolation, the SSIM of the proposed algorithm is improved by 6.2–14.1 %, the average SSIM increase is 8.8 %. Compared with the method [9], the SSIM of the proposed algorithm is improved by 1.2–4.5 %, the average SSIM increase is 3.0 %. Compared with the method [10], the SSIM of the proposed algorithm is improved by 0.1–2.2 %, the average SSIM increase is 2.1 %.

Table 3. Comparison of the entropy obtained by different algorithms

image	LR image	Bicubic	ref. [9]	ref. [10]	proposed method
airport	3.2569	3.2653	3.3219	3.3323	3.3732
baboon	2.3735	2.4159	2.5748	2.6137	2.6490
barb	2.4888	2.5132	2.6966	2.7345	2.7443
boat	2.6324	2.7032	2.7828	2.8344	2.8432
couple	2.9843	3.0352	3.1274	3.1435	3.2523
girl	2.6369	2.6697	2.7132	2.7312	2.7627
moom surface	3.2153	3.2408	3.3448	3.3562	3.3712
pixel ruler	2.9837	3.0221	3.0435	3.0521	3.0693
testpat	2.8732	2.9072	2.9546	2.9723	2.9952

Compared with the bilinear interpolation, the entropy of the proposed algorithm is improved by 1.6–9.5 %, the average entropy increase is 4.9 %. Compared with the method described in ref. [9], the entropy of the proposed algorithm is improved by 0.8–4.0 %, the average entropy increase is 1.9 %. Compared with the method described in [10], the entropy of the proposed algorithm is improved by 0.4–3.5, the average entropy increase is 1.0 %.

Table 4. Comparison of the MSE obtained by different algorithms

image	Bicubic	ref. [9]	ref. [10]	proposed method
airport	2.4654	1.9874	1.8456	1.8327
baboon	2.2853	1.8823	1.5436	1.4425
barb	2.1357	1.6349	1.4537	1.4218
boat	2.0458	1.4917	1.3627	1.2807
couple	1.9079	1.7543	1.6789	1.6325
girl	1.8894	1.7547	1.6756	1.6479
moom surface	2.5864	2.1464	1.9864	1.8768
man	2.0254	1.9854	1.8546	1.8103
pixel ruler	2.4712	2.0957	1.9954	1.8628
testpat	2.0569	1.9875	1.8539	1.8212

Compared with bilinear interpolation and method [9], the MSE of the algorithm is greatly improved. Compared with the method [10], the MSE of the proposed algorithm is improved by 0.7–6.6 %, the average MSE increase is 3.6 %.

5. Conclusion

The results show that the proposed algorithm provides clear edges, clear texture and overcomes the adverse effects caused by uneven illumination to a certain extent. Also, the effectiveness of the proposed method was confirmed in various simulation conditions.

References

[1] S. Farsiu, M. D. Robinson, M. Elad, P. Milanfar: *Fast and robust multiframe super resolution.* IEEE Trans. Image Processing *13* (2004), No. 10, 1327–1344.

[2] H. Su, Y. Wu, J. Zhou: *Super-resolution without dense flow.* IEEE Trans. Image Processing *21* (2012), No. 4, 1782–1795.

[3] A. K. Katsaggelos, R. Molina, J. Mateos: *Super resolution of images and video.* Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan and Claypool e-Books, 2007.

[4] B. Maihe, D. Barchiesi, M. D. Plumbley: *INK-SVD, learning incoherent dictionaries for sparse representations.* Proc. IEEE IC Acoustics, Speech and Signal Processing, 25–30 March 2012, Kyoto, Japan, 3573–3576.

[5] W. T. Freeman, T. R. Jones, E. C. Pasztor: *Example-based super-resolution.* IEEE Comp. Graph. Appl. *22* (2002), No. 2, 56–65.

[6] M. N. Do, M. Vetterli: *The contourlet transform: An efficient directional multiresolution image representation.* IEEE Trans. Image Processing *14* (2005), No. 12, 2091–2106.

[7] A. L. da Cunha, J. Zhou, M. N. Do: *The nonsubsampled contourlet transform: Theory, design, and applications.* IEEE Trans. Image Processing *15* (2006), No. 10, 3089 to 3101.

[8] Wei Zhang, Zhongcheng Fan: *A no-reference contourlet-decomposition-based image quality assessment method for super-resolution reconstruction.* Proc. SPIE 9273, Optoelectronic Imaging and Multimedia Technology III, 9–11 Oct. 2014, Beijing, China. Vol. 9273.

[9] Haiyuan Zhang, Nian Cai, Nan Zhang: *Using weighted parabolic interpolation to zoom images based on an error-amended sharp edge algorithm.* Computer Engineering and Applications. *7* (2011), No. 25, 194–197.

[10] Jiantao Wang, Yun Zhao, Qing Gao: *Multi-scale wavelet super-resolution image reconstruction based on nonuniform sampling multi-frame images.* Proc. 16th National Symposium on Remote Sensing, 24 Nov. 2008, Beijing, China, Vol. 7123.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 175–184

Study on soft sensor model of boiler thermal efficiency based on support vector regression and grid search algorithm[1]

Bian Heying[2], Zhang Xiaoli[3]

Abstract. A novel soft sensor model of boiler thermal efficiency is established by applying support vector regression(SVR) and grid search algorithm (GSA). This method divides data collected into training data and testing data by researching on the 1000 MW unit of Sanbaimen Power Plant in Datang Chaozhou, uses grid search algorithm to optimize the parameters C and g of the soft sensor model, and applies test data and random data to test the model accuracy and generalization capability. The simulation results show that the built soft sensor model has higher prediction accuracy where relative error is controlled within 1 % and better generalization ability, which provides a feasible scheme for online measurement of boiler thermal efficiency.

Key words. Support vector regression, grid search algorithm, boiler thermal efficiency, soft sensor, model.

1. Introduction

Boiler thermal efficiency is a key indicator for measuring the economic operation of the boiler. To improve its thermal efficiency, combustion should not only save energy, but also reduce pollution and greenhouse gas emissions [1]. Therefore, it is very important to measure accurately boiler thermal efficiency, which is helpful to adjust timely the boiler combustion optimization control strategy, improve the energy efficiency and reduce the cost of power generation. The boiler thermal efficiency model is a key factor to measure accurately the boiler thermal efficiency. There are

[1]The authors would like to thank for the support by Natural Science Foundation of the education department of Henan Province under the Grant 14B470009.

[2]College of Electrical and Mechanical Engineering, Pingdingshan University, Henan, Pingdingshan, 467000, China

[3]Food and Chemical Engineering Department, College of Henan Quality Project, Pingdingshan, 467000, China

two kinds of methods to establish the model at present: mechanism modeling and data modeling [2]. As the boiler combustion process has multivariable, nonlinear, strong coupling and large delay characteristics, it is very difficult to establish boiler thermal efficiency model by mechanism modeling. Data modeling method has attracted more attention In recent years [3], [4] the soft sensor model of boiler thermal efficiency was built according to the test data of boiler combustion by applying neural network and genetic algorithm; research results show that the model has better measurement accuracy and generalization ability. Literature [5] proposed the least squares support vector machines model between extracted feature and boiler efficiency by analyzing the boiler combustion historical data. Research results show that the prediction accuracy of the model was also obviously improved.

2. Support vector regression and modeling of grid search algorithm

2.1. Theory of modeling of support vector regression

The basic idea of modeling of support vector regression is that non-linear problems of a low-dimensional space are transformed into linear problems of a high-dimensional space by introducing the kernel function $K(x_i, x_j)$ [6].

If $\{(x_1, x_1), \ldots, (x_n, x_n)\}$ is the training sample set, then the regression function F is given by the formula

$$F = \left\{ f | f(x) = w^T \phi(x) + b, \quad w \in R^n \right\}, \tag{1}$$

where w stands for weight vector, T denotes the sample points sets, $\phi(x)$ is nonlinear mapping from input to output space and b is the threshold. Then

$$R_{\text{reg}}(w) = \frac{1}{2}\|w\|^2 + C \cdot R_{\text{emp}}[f(x)] = \frac{1}{2}\|w\|^2 + C \cdot \frac{1}{n}\sum_{i=1}^{n} |y - f(x)|, \tag{2}$$

where $\|w_i\|$ describes the model complexity, the role of the penalty coefficient C is a compromise between the experience risk and the complexity of the model, n stands for the number of training samples, R_{emp} is the experience risk function and R_{reg} is the structure risk function. The value of w can be determined from the formula

$$w = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \phi(x_i), \tag{3}$$

where α_i, α_i^* represents the solution of R_{reg} minimized, x_i is the support vector, n stands for the number of training samples, and $f(x)$ can be represented by the formula

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) [\phi(x_i) \cdot \phi(x)] + b. \tag{4}$$

Non-linear support vector regression function $f(x)$ can be defined according to the kernel function $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ using the formula

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x) + b = \sum_{i=1}^{n} w_i \cdot K(x_i, x) + b, \tag{5}$$

where w_i is the coefficient of support vector, $K(x_i, x)$ is the kernel function, and b is defined by the formula

$$b = y_i - w_i \cdot \phi(x_i) . \tag{6}$$

2.2. Parameter selection of SVR model

The radial basis kernel function $K(x_i, x_j) = \exp\left(-g \cdot |x_i - x_j|^2\right)$ is selected as kernel function of support vector regression model in this research. The accuracy of SVR model generally depends mainly on four parameters, particularly on the kernel function parameter g, penalty coefficient C, maximum allowable error e, and insensitive loss function ε. Parameters e and ε are, however, usually ignored because they are controlled by the human's behavior and exhibit only a little impact on the model of soft sensor ability [7]. On the other hand, parameters C and g affect directly the model soft sensor precision and generalization ability [8]. This research applies the grid search algorithm to optimize parameters C and g in order to obtain parameters that are as good as possible.

2.3. Method of GSA-SVR modeling

In this research, the method supporting vector machine combined with grid search algorithm is known as GSA-SVR modeling. The GSA-SVR training steps are as follows:

1. To determine training data and test data according to the data collected from the field.

2. To deal with data such as noise reduction and data normalization.

3. To optimize parameters C and g that are obtained by applying grid search algorithm.

4. To establish soft sensor SVR model of the boiler thermal efficiency by using the optimum parameters and training data.

5. To verify the accuracy and generalization ability of the model built by applying the test data.

6. To use the model to predict the data measured if the accuracy and generalization ability of the model meet the requirements; otherwise to build the model again.

3. Soft sensor model structure of the boiler thermal efficiency

3.1. Simplified model of boiler thermal efficiency calculation based on ASME

There are two calculation standards of the boiler thermal efficiency at present, one of them being the national standard of power performance test procedures of people's republic of China [9]; the other being the standard of power performance test procedures of the American society of mechanical engineers, referred to as ASME. The standard model must be simplified in order to improve the calculation efficiency because some parameters are not able to be measured online under the boiler operation condition. In this research, the model simplification principle based on ASME is as follows.

1. Water content heat loss is given by the formula

$$l_{\mathrm{m}} = \frac{C_{\mathrm{pH_2O}}}{Q_{\mathrm{d}}^{\mathrm{y}}} \cdot (k_3 + 0.01\,(k_4 + k_2 \cdot \alpha_{\mathrm{py}})) \cdot (t_{\mathrm{py}} - t_{\mathrm{R}}) \times 100\%\,, \tag{7}$$

   where $C_{\mathrm{pH_2O}}$ is mean specific heat of water vapor, $Q_{\mathrm{d}}^{\mathrm{y}}$ is the low-heat value of boiler efficiency, $K_2,\ k_3,\ k_4$ are the coefficients of $Q_{\mathrm{d}}^{\mathrm{y}}$, α_{py} is the excessive air coefficient of exhaust, t_{py} is the exhaust flue gas temperature and t_{R} is the temperature of cool air.

2. Dry flue gas heat loss is given as

$$l_{\mathrm{G}} = \frac{C_{\mathrm{py}}}{Q_{\mathrm{d}}^{\mathrm{y}}} \cdot (k_1 + k_2 \cdot \alpha_{\mathrm{py}})\,(t_{\mathrm{py}} - t_{\mathrm{R}}) \times 100\%\,, \tag{8}$$

   where C_{py} is the mean specific heat of exhaust and k_1 is the coefficient.

3. The heat loss caused by incomplete combustion of carbon is given as

$$l_{\mathrm{uc}} = \frac{33730}{Q_{\mathrm{d}}^{\mathrm{y}}} \cdot A^{\mathrm{y}} \cdot \left(0.9 \cdot \frac{C_{\mathrm{hf}}}{100 - C_{\mathrm{hf}}} + 0.1 \frac{C_{\mathrm{hz}}}{100 - C_{\mathrm{hz}}}\right) \times 100\%\,, \tag{9}$$

   where A^{y} is the application base ash, C_{hf} is is the fuel percentage in fly ash and C_{hz} is the fuel percentage in ash slag.

4. Radiation and convection heat loss is $l_{\mathrm{un}} = 0.5\,\%$.

5. Other heat loss is $l_{\mathrm{o}} = 0.5\,\%$.

6. Then the boiler thermal efficiency is

$$\eta = 100 - l\,, \tag{10}$$

   where $l = l_{\mathrm{G}} + l_{\mathrm{uc}} + l_{\mathrm{un}} + l_{\mathrm{o}}\,.$

The thermal efficiency of the boiler can be calculated according to the simplified model of the thermal efficiency of the boiler, but the calculation speed is slow and

it is very difficult to realize on-line measurement of the boiler thermal efficiency. Therefore, the soft sensor model of the boiler thermal efficiency is built by applying SVR and GSA algorithm in this research. The key parameters influenced by the boiler thermal efficiency are selected as soft sensor model input parameters, the partial boiler thermal efficiency values calculated according to the formulae (6)-(9) are taken as model output parameters during establishing soft sensor model.

3.2. Soft sensor model structure of the boiler thermal efficiency

The boiler thermal efficiency is influenced by many factors [10]. The model input parameters are selected by mechanism analysis and correlation analysis as follows: total amount of air, ventilation rates of six coal pulverizes, three values of the contents of oxygen in the flue gas, fuel component, six opening degrees of secondary air dampers, five opening degrees of the coal feeders, a tilting angle of the burners and the opening degree of fired-off air damper.

The total amount of fuel and air describes the load effect on the boiler thermal efficiency; effect of oxygen content is described by several parameters such as the opening degree of the secondary air damper, ventilation rate of coal feeders and oxygen content in the flue gas. The opening degree of the coal feeder describes the effect of the pulverized coal, the influence of other factors are described by opening degree of the fired-off air damper and tilting angle of the burners. The soft sensor model structure of the boiler thermal efficiency is built and shown in Fig. 1.

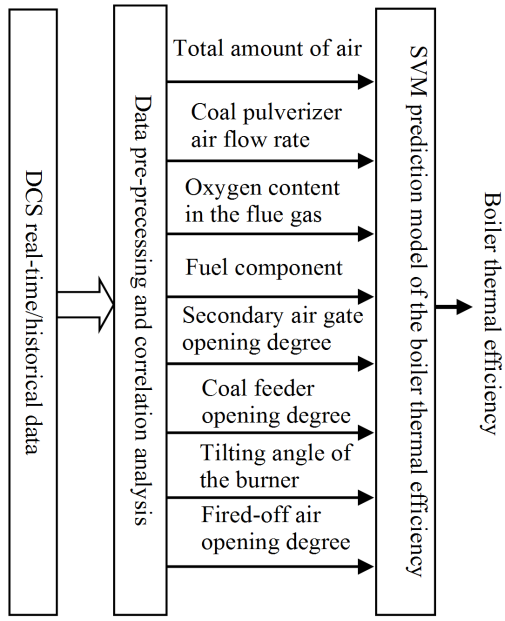

Fig. 1. Prediction model structure of boiler thermal efficiency

4. Soft sensor modeling of the boiler thermal efficiency based on GSA-SVR

The task was to establish a soft sensor model according to the method of GSA-SVR modeling introduced in the previous section. The simulation data come from the historical data of a 1000 MW ultra-supercritical unit of Datang Chaozhou power station.

Now how to determine the training sample data and testing sample data. First, the data coming from the distributed control system were preprocessed and selected randomly 200 groups data among them. 40 groups data were obtained by choosing a group data among every 5 groups, 40 thermal efficiency values were calculated according to the ASME model by applying the 40 groups, which served as the training sample output. Second, the data groups of 8 multiples were selected among the 200 groups data, 25 groups data obtained served as the test sample input, the thermal efficiency values match the 25 groups data served as the test sample output

To determine the variation range of the penalty parameter C and radial basis core parameter g according to the relevant theories, the variation range of the former parameter is [-8, +8], while the variation range of the latter parameter is [0, 300]. To optimize the parameters C and g by using the grid search algorithm, the fitness curve of the training sample set acquired by applying GSA-SVR algorithm is shown in Fig. 2.

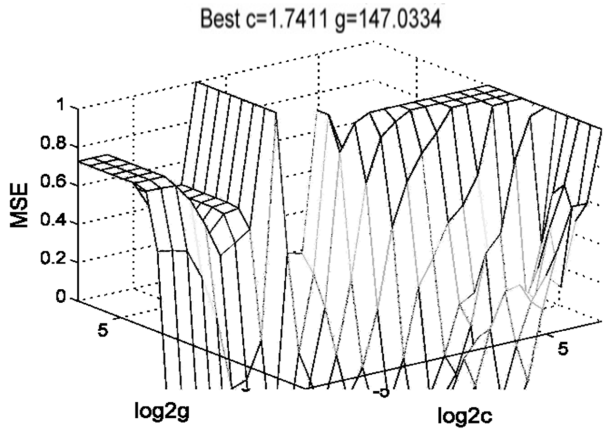

Fig. 2. Fitness curve of SVR optimized by GSA

Fig. 2 shows that the optimal penalty factor C equals to 1.7411, the kernel function parameter g equals to 147.0334 by applying the GSA-SVR model, and that the soft sensor model built has 17 support vectors. The decision function of the SVR model obtained by Matlab is

$$f(x) = \sum_{i=1}^{n} w_i \exp\left(-g\left|x_i - x\right|^2\right) + b\,. \tag{11}$$

5. Verification of validity of soft sensor model of boiler thermal efficiency

The final prediction effect of the soft sensor model is evaluated by using mean square error (MSE) and squared correlation coefficient (r^2). With the square correlation coefficient approaching more close to 1, the result of regression fitting is better, the value of MSE is smaller and prediction precision of the model is also better.

To verify the validity of the built soft sensor model, the model is tested by applying test data. The test results show that $r^2 = 0.99365$, MSE is 0.031, the regression curve fitting degree and the accuracy of the model built can meet the practical requirements. The simulation output curve of the actual value and soft sensor value of the boiler thermal efficiency by using Matlab is shown in Fig. 3.

The solid line represents the actual values of the boiler thermal efficiency; the dotted line represents the prediction values of the boiler thermal efficiency by applying the soft sensor model built.

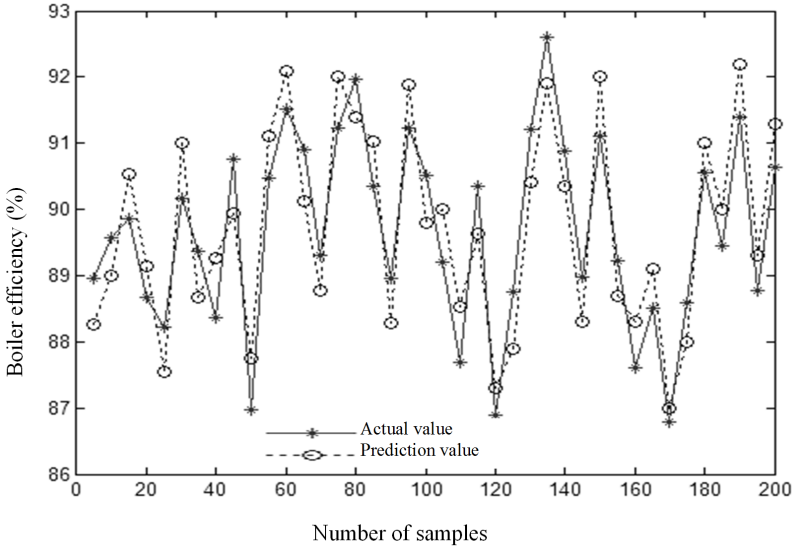

Fig. 3. Output curve of actual value and soft sensor value of boiler thermal efficiency

Fig. 4 shows the same trend of the rise and fall both of the soft sensor values and actual values, only the amplitude of variation is slightly different in the local range, which confirms that the soft sensor model built of the boiler thermal efficiency has higher prediction accuracy and better generalization ability. Fig. 4 contains the curve of the relative error and corresponds to Fig. 3.

The relative error is controlled basically within 1% and variance is 0.0043, as is shown in Fig. 4.

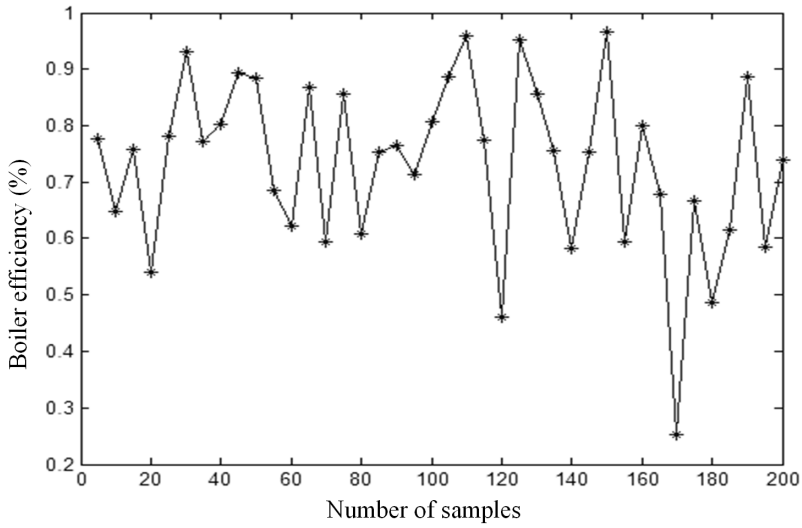

Fig. 4. Relative error curve of the boiler thermal efficiency

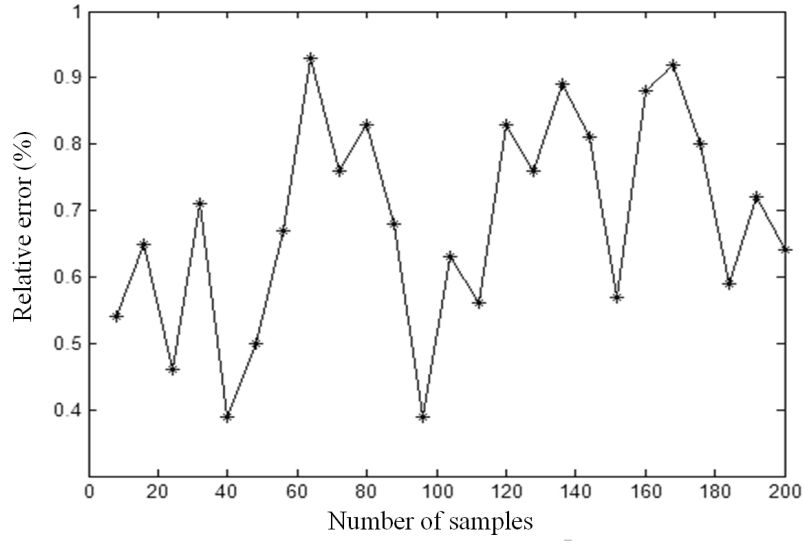

Fig. 5. Relative error of the boiler thermal efficiency

25 groups of data are selected randomly in order to further verify the generalization of the model established. The soft sensor results of the model show that the relative error can be limited in the range of accuracy and the soft sensor model has better generalization ability, as is shown in Fig. 5.

6. Conclusion

A novel soft sensor model of the boiler thermal efficiency is established by applying support vector regression and grid search algorithm about the problem that the boiler thermal efficiency is difficult to measure accurately. To optimize the parameters C and g of the prediction model are found using adaptive grid search algorithm. Soft sensor accuracy and generalization capability of the model are improved by preprocessing of acquisition data and verified by using test data. The simulation results show that the relative error of the soft-sensing model established is controlled within 1%, the variance is only 0.0043. In summary, soft sensor accuracy and generalization capability of the soft sensor model established can meet the practical requirements and provide an effective way of measuring the boiler thermal efficiency in power plant.

References

[1] Y. Torii, S. Abe: *Decomposition techniques for training linear programming support vector machines.* Neurocomputing *72* (2009), Nos. 4–6, 937–984.

[2] A. Borin, M. F. Ferrao, C. Mello, D. A. Maretto, R. J. Poppi: *Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk.* Analytica Chimica Acta *579* (2006), No. 1, 25–32.

[3] P. Han, F. Gao, Y. J. Zhai, Y. Lu : *Combustion modeling for utility boiler based on multi-output least squares support vector machine.* IJ Advancements Comput. Technol. *4* (2012), No. 14, 60–68.

[4] Y. He: *Combustion system model based prediction of boiler efficiency and NOx emission.* Thermal Power Generation *42* (2013), No. 4, 56–61.

[5] Y. Gu, W. Zhao, Z. Wu: *Online adaptive least squares support vector machine and its application in utility boiler combustion optimization systems.* J Process Control *21* (2011), No. 7, 1040–1048.

[6] L. X. Yin, M. C. Wang, Q. Shang: *The combustion optimization of a coal-fired boiler based on support vector machine and particle swarm algorithm.* Boiler Technol. *45* (2014) No. 4, 123–128.

[7] B. Ren, Z. P. Feng, X. Jiaotong: *Improved genetic algorithm and particle swarm optimization as well as comparison between them.* J Nanjing Normal University (Engineering and Technology Edition) (2002), No. 2, 14–20.

[8] M. A. Mohandes, T. O. Halawani, S. Rehman, A. A. Hussain: *Support vector machine for wind speed prediction.* Renewable Energy *29* (2004), No. 6, 939–947.

[9] F. Si, C. E. Romero, Z. Yao, E. Schuster, Z. Xu, R. L. Morey, B. N. Liebowitz: *Optimization of coal-fired boiler SCRs based on modified support vector machine models and genetic algorithms.* Fuel *88* (2009), No. 5, 806–816.

[10] H. Li, M. D. Li, H. J. Chen: *Soft sensors for thermal efficiency in industry oil boiler based on BPNN.* Energy Research & Utilization *3*, (2010), No. 4, 135–139.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 185–196

Research on big data processing and analysis architecture based on MongoDB[1]

Longge Wang[2,3], Tao Song[4], Junyang Yu[2,3]

Abstract. A novel big data processing and analysis architecture is proposed based on MongoDB. First of all, a data processing framework is present to guide the application of big data. Secondly, a new big data processing architecture consisting of non-relational database MongoDB and distributed framework MapReduce is put forward to satisfy the requirements of big data storage and parallel processing. Finally, the travel data of Beijing during the Spring Festival in 2014 is tested to verify the efficiency of the proposed big data processing architecture.

Key words. MongoDB, NoSQL, data processing, data analysis, big data.

1. Introduction

Over the past ten years, data has increased in a large scale in various fields. The application of internet technology, especially the rapidly development of the mobile internet technology, produces large amounts of data in the day-to-day use. And the novel data acquisition setups, such as GPS, new-type sensors, automatic tracking and monitoring system, also produce a large amount of data in everyday practice. Under the explosive increase of global data, the term of big data was used to describe enormous datasets. Big data is an evolving term that describes any voluminous amount of structured, semi-structured and unstructured data that has the potential to be mined for information.

[1]This work was supported by Open Foundation of State key Laboratory of Networking and Switching Technology (Beijing University of Posts and Telecommunications), project SKLNST-2016-2-23. Also, the authors greatly appreciate the reviewers' valuable comments on this paper.

[2]Software School, Henan University, Kaifeng, 475001, China

[3]State key Laboratory of Networking and Switching Technology (Beijing University of Posts and Telecommunications), Beijing, 100876, China

[4]Zhengzhou University, Zhengzhou, 450000, Henan, China

Big data brings the new opportunities and challenges for the data storage, data analysis and data processing. Big data is often characterized by 5 Vs [1], [2].

1. Volume: organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data.

2. Velocity: Data streams in at an unprecedented speed and must be dealt with in a timely manner; for example RFID tags, sensors and smart metering are driving the need to deal with torrents of data in near-real time.

3. Variety: data comes in all types of formats – from structured, numeric data in traditional databases to unstructured text documents, email, video, audio, stock ticker data and financial transactions.

4. Value: The critical problem in big data is how to discover values from datasets with an enormous scale, various types, and rapid generation.

5. Veracity. Big data veracity has an impact on the confidence reposed by the marketer to their database.

In a volatile big data environment, accuracy becomes an issue among digital marketers regarding the collected data for their business.

The application of big data has achieved very good results with the placement and use of a large number of sensors [2–4], but there are still some problems to be solved. Firstly, the traditional database is difficult to meet the needs of the data storage and management with the rapid growth of the sensor data. The sensor data has the high concurrency and large data volume characteristics. Secondly, the captured data has a certain degree of redundancy as the sensors usually collect data at intervals of several seconds. Therefore, finding a sensor data storage method with high data storage efficiency becomes an important research area for the application of big data.

In this paper, a novel big data processing and analysis architecture is proposed. Firstly, a data processing framework is presented to process and analysis the big data. Secondly, a MongoDB-based big data processing architecture is proposed to improve the efficient of storage and parallel processing, which composed of non-relational database (MongoDB) and distributed framework (MapReduce). Finally, the travel data of Beijing during the Spring Festival in 2014 is tested to verify the efficiency of the proposed big data processing architecture.

2. Processing of big data

Big data is basically consisted of the acquired data and exchanged data. The knowledge which was mined through the big data processing system can be used to support the upper decision or application. Five main phases [4] such as data preparation, storage management, calculation processing, data analysis and knowledge display are requisites. The processing of big data is shown in Fig. 1.

2.1. Data preparation phase

Big data requires some preprocessing as cleaning and sorting before storage and calculation. It is similar to ETL (Extracting, Transforming and Loading) in traditional data processing system. Compared with the traditional data analysis, the differences are not only in quantity and format but in data quality. In view of these, the format standardization and noise elimination are necessary steps in data preparation phase.

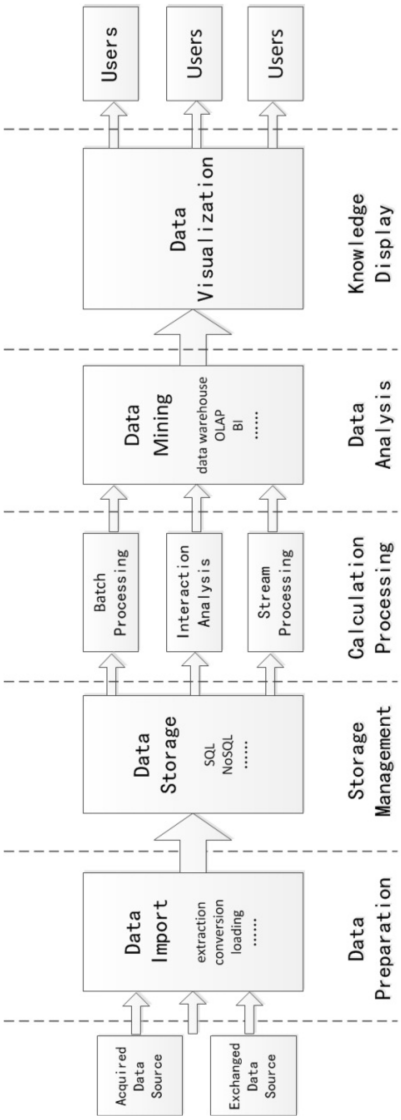

Fig. 1. Processing of big data

2.2. Storage management phase

The storage technology is facing great challenges in terms of performance and cost since the rapid growth rate of big data. Storage system of big data requires high adaptability of unstructured data management, high extensibility of various data formats and low storage cost of massive data.

2.3. Calculation processing phase

According to the data type and analysis target, it is necessary to introduce appropriate algorithm models to implement the fast data processing. Significant amounts of calculating resources are required in big data processing. Therefore distributed computation has become the mainstream architecture for big data processing.

2.4. Knowledge display phase

Presenting the visualized results to the user is an important process of big data analysis when the big data services engaged in supporting the decision application. In some closed loop big data services, visualized results generally are directly applied by the machine according to the algorithm without manual intervention.

3. Comparative analysis on Non-relational database

Storage technology is facing great challenges cause of the characteristics of big data such as heterogeneous, massive, real-time processing. Traditional relational database has been difficult to meet the performance requirements such as high efficiency accessing, high concurrent reading and writing, high availability and high scalability. Distributed storage technology of non-relational database [5], [6] has a greater advantage for massive data storing in each service node through the large-scale distributed structure correspondingly. Accordingly, the distributed storage based on the non-relational database provides an effective solution for the storage and management of big data. The comparative analysis of existing three Non-relational databases is shown in Table 1.

From three mainstream NoSQL databases characteristics analysis [5–7], MongoDB could better support the massive data sharding and quering compared to Hbase and Dynamo. Especially, the index function of MongoDB improves the query speed of acquired data.

4. Big data processing architecture based on MongoDB

According to the previous analysis, designing and constructing big data processing architecture which composed of non-relational database (MongoDB) and distributed framework (MapReduce) will effectively satisfy the core requirements of efficient storage and parallel processing.

The basic operation principle of MongoDB clusters system as follows [10], [11]. MongoDB clusters determine whether the data on a slice is more than a predetermined value when importing data occurred by users. Sharding mechanism will be activated if the value exceeded the limit of storage. Data set is divided into chunks and assigned to different shards. In this process, MetaData information in term of shard and chunk will be stored in Config Server. The architecture adopt with MongoDB clusters could be forming multi-partitions MongoDB servers. In the process of data storage, query and analysis, each partition node could be carried out concurrent processing due to distributed storage. With the increasing of the amount of data, data set can be added to the new shards. Furthermore, each data node is configured with the corresponding backup node. Consequently, the performance requirements of system such as availability, efficiency and scalability could be satisfied. Logical relationship of MongoDB clusters deployment is shown in Fig. 2.

Table 1. Comparative analysis of Non-relational databases

Database	HBase	Dynamo	MongoDB
Data models	Columnar	Key	Document
Single query only	Single query only	Key query only	Support most query
R/W performance	Write complex	Always write, read complex	Write complex
Usability of interface	Support most programming languages	Support simple interface only, RESTFUL	Support most programming languages
Expansibility	Add table service	Add node, table migration	Add shard, chunk migration
Data version	Time stamp	Vector clock	Real time
Data version	Not support	Not support	Support

Big data processing architecture mainly involves control and interaction among the MongoDB distributed cluster, Hadoop cluster, and master server [8–10]. The data processing procedure is as follows.

4.1. Control of Hadoop cluster

The master server is responsible for the coordination control of task assignment. according to the request of query or update on the data shard node, the task and information returned are distributed in a unified way to MongoDB configuration and routing node.

4.2. Control of mongo routers

The mongo routers are responsible for operations of routing and coordinating, thereby control clusters as a whole system. The mongos inquire about information of

MongoDB configuration and routing firstly, and then find out forward the request to the shard node for storing corresponding data. Furthermore, the processing results are returned to mongos separately when sharding node completed the operations. Mongo routers summarize all the results and return to master server at last.

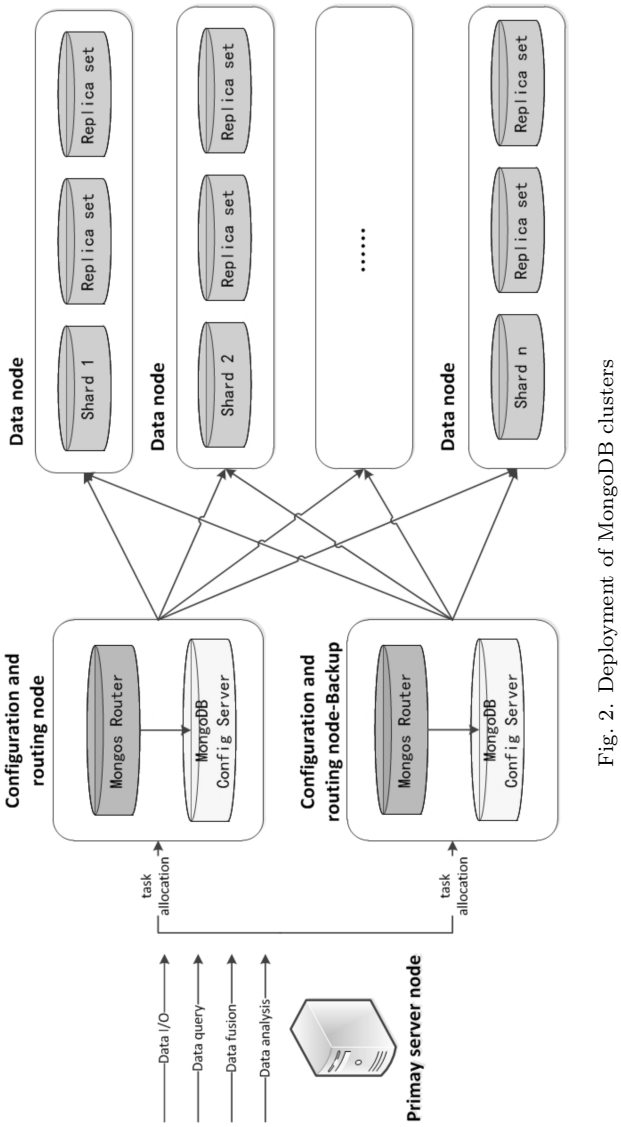

Fig. 2. Deployment of MongoDB clusters

4.3. Data source sharding of MongoDB

Many Chunks is creating as input data for parallel processing from MongoDB. If data is not sharding by MongoDB, then Hadoop is reading the data through the

route. On the contrary, Hadoop is reading the information of data Chunks from MongoDB router and configures server firstly, and then it reads and inputs the storing data on the shards.

4.4. MapReduce procedure

The system node calls MapReduce application to implement the Map and Reduce procedure through the above 3 processes. The final query results will be presented to the user or parallel write to the data shard on the MongoDB.

5. Experiment and discussion

In this section, the travel data of Beijing during the Spring Festival in 2014 are tested to verify the efficiency of the big data processing framework proposed by this paper. The obtained traffic data during the Spring Festival period are more than 2 TB, only part of the most important data are selected to test the proposed big data processing framework in this example.

5.1. Data preparation

In the road traffic data collection, the minimum time granularity, 5 minutes, traffic data are captured by using the various sensors located on the road, such as the traffic flow detector, induction coil, microwave sensors, video surveillance systems, and other equipment. Then the basic parameters of road network traffic flow can be obtained by using these sensor data in combination with the toll road data. Up to now, 693 sets of traffic volume collection equipment has been built in Beijing to cover 593 ordinary road sections in order to obtain road traffic data, where 110 sets in national road, 342 sets in provincial road, and 241 sets in country road. In addition, 18 highways have realized the collection of video, traffic flow and statistical data of import and export toll stations for key sections in Beijing.

In the inter-city passenger flow data collection, intercity passenger flow of real-time and expected travel data can be obtained through the railway ticketing system, inter-provincial passenger ticketing system, as well as civil aviation passenger flight dynamics information.

In this example, 17 days travel data between January 16, 2014 (the first day of the Spring Festival travel) and February 1 (the 17th day of the Spring Festival travel) are selected for experimental validation. Eight access to Beijing's national roads traffic data are selected as the ordinary road data, which captured by the cross-section traffic detection equipment. 11 Beijing's external (excluding urban) highway toll data are selected as highway data. And railway, civil aviation and inter-provincial passenger transport data are derived from the passenger data collected by their respective information systems.

5.2. Data calculation processing

The Map Reduce distributed batch computing method is adopted in order to meet the requirements of high efficient storage and parallel processing of large-scale traffic data in this example. MongoDB server cluster contains MongoDB server with multi-partition node, where each partition holds a portion of the data. Each node can be processed in parallel to ensure the efficiency of the system in the process of saving, querying and analyzing data.

Master acts as the master server to control the task distribution and process coordination of the entire system. Mongos is responsible for routing and coordinating operations, which achieves the overall control of the cluster. And MongoDB cluster creates many chunks from the source data as input data for parallel processing. The system node calls the MapReduce application to implement the Map and Reduce process of the data, and presents the final query result to the user, and writes the data to the MongoDB fragments.

5.3. Data analysis and mining

The proposed data processing framework is used to analyze the traffic data during the Spring Festival in this paper. First of all, the neural network algorithm [11] is applied for feature extraction of traffic data during the Spring Festival in Beijing, and then the key features of the key indicators are got. Secondly, the association rule algorithm [12] is used to analyze the relationship between different features, and the more valuable traffic features and models are obtained. Finally, the obtained models and information are used to guide the urban transport services, planning and management.

5.4. Knowledge display

The use of knowledge display to show the boring traffic data becomes intuitive and visible. The results of the feature analysis in this example are shown below.

- The effect of departure from Beijing is increasing sharply, and the urban passenger transport is the first way for residents to leave Beijing.

  The main feature of urban traffic is a large number of people leaving the city which led to a sharp drop in demand for travel within the city during the Spring Festival, as shown in Fig. 3.

  As can be seen from Fig. 3, the volume of intercity traffic and road traffic began to leave Beijing more than entering Beijing since the launch of the Spring Festival travel. The number of people leaving Beijing increased day by day with the cumulative effect of leaving Beijing, and reaching the maximum on February 1 (the second day of the Spring Festival). Intercity passenger transport (railway, civil aviation and highway inter-provincial passenger transport) is the most important way for residents to leave Beijing during the Spring Festival. The number of people leaving Beijing gradually increased with the Spring Festival approaching. The number of single-day leaving Beijing in January 29 reaches

Fig. 3. Visualization of changes in number of people leaving Beijing during Spring Festival

the peak, 800.000 people. This is because January 29 is the day before Chinese New Year's Eve, and the tradition of home New Year contributed to the increase in the number of people leaving Beijing.

- The traffic pressure distribution in the city is balanced, and the travel time constraint of citizens is obviously weakened.

  Compared to the average hourly traffic index during the Spring Festival and other normal periods, the result is shown in Fig. 4.

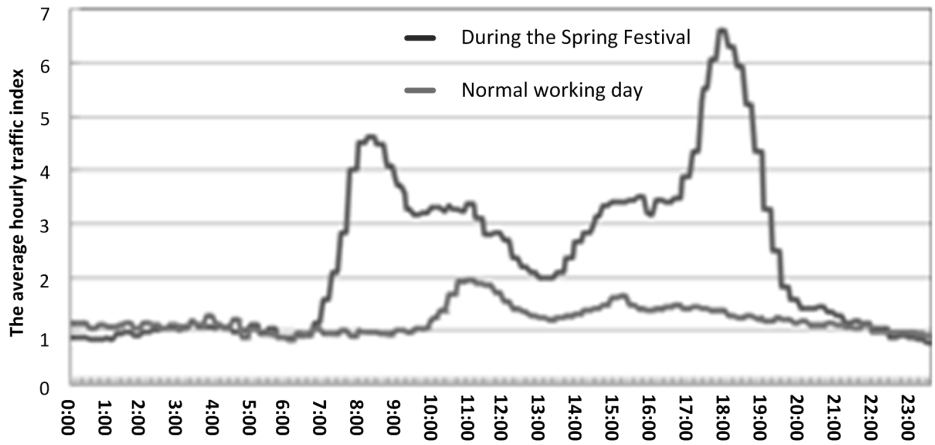

Fig. 4. Comparison of average hourly traffic index between Spring Festival and ordinary working day

It can be seen from Fig. 4 that the travel time constraints were significantly reduced during the Spring Festival, and the elastic travel time effectively balances the traffic pressure over time. There is no obvious morning and evening traffic peak during the Spring Festival in compare with the normal working days. The traffic pressure was balanced throughout the day, and the traffic index at all times was 2.0 or less during the Spring Festival. In other words, city traffic is at the unimpeded level all day. Therefore, the urban road network operation was good because of that a large number of people from Beijing to avoid the daily concentration of morning and evening peak pressure.

- The population is positively correlated with the total demand in the city, and is positively related to the operation of urban road network. igure 5 shows that the sharp drop in the total traffic demand with the return passenger flow, student traffic, and family visits and other traffic generated a large number of people and vehicles from Beijing. The average daily traffic index and population in Beijing are generally declining except on January 18 and 25 (two rest days before the holiday), and both the trends of them are basically consistent. The overall average daily traffic index decreased and the residents travel smooth traffic with the reduction of population in Beijing. All in all, the urban population has a significant impact on the operation of urban road network, and there was a strong positive correlation between them. Therefore, in the normalized congestion control work, we not only through policy-oriented to reduce the frequency of motor vehicle use, through the price mechanism to guide the peak-shifting time-sharing travel, but also should actively explore new management methods, reasonable planning of urban layout and function, strict control of city size and population. The proposed MongoDB-based big data processing architecture is verified through the realization and analysis of this example.

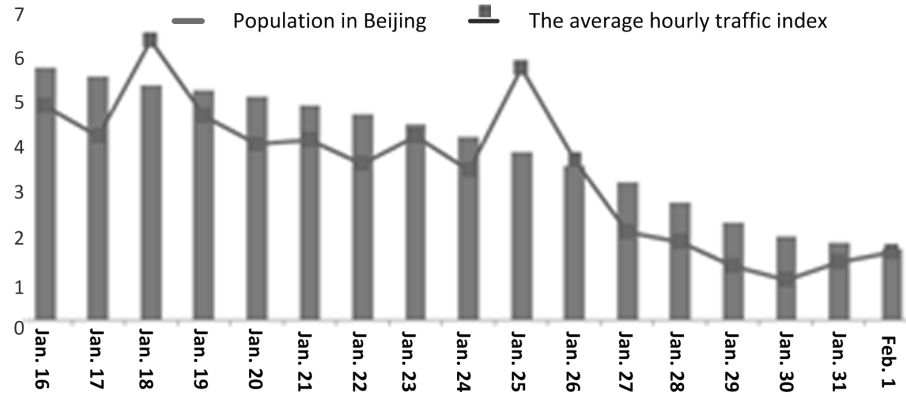

Fig. 5. Comparison between average daily traffic index and population changes in Beijing during Spring Festival

6. Conclusion

In this paper, a novel big data processing and analysis architecture was proposed to satisfy the core requirements of big data storage and parallel processing. Firstly, a data processing framework was used to guide the application of big data. Secondly, a variety of non-relational databases were compared to select the appropriate database. Thirdly, a big data processing architecture was proposed in order to improve storage efficiency, which composed of non-relational database MongoDB and distributed framework MapReduce. Finally, the travel data of Beijing during the Spring Festival in 2014 was tested in this paper. Experimental results illustrate that the proposed big data processing architecture is effective.

References

[1] S. Sagiroglu, D. Sinanc: *Big data: A review*. Proc. IC Collaboration Technologies and Systems (CTS), 20–24 May 2013, San Diego, CA, USA, 42–47.

[2] A. Katal, M. Wazid, R. H. Goudar: *Big data: Issues, challenges, tools and good practices*. Proc. IEEE Sixth IC Contemporary Computing (IC3), 8–10 Aug. 2013, Noida, India, 404–409.

[3] A. Zaslavsky, C. Perera, D. Georgakopoulos: *Sensing as a service and big data*. Proc. IC Advances in Cloud Computing (ACC), 26–28 July, Bangalore, India, 2012, CD-ROM.

[4] C. H. Lee, D. Birch, C. Wu, D. Silva, O. Tsinalis, Y. Li, S. Yan, M. Ghanem, Y. Guo: *Building a generic platform for big sensor data application*. Proc. IEEE IC Big Data, 6–9 Oct. 2013, Santa Clara, CA ,USA, 94–102.

[5] J. Han, E. Haihong, G Le, J. Du: *Survey on NoSQL database*. Proc. IEEE 6th IC Pervasive Computing and Applications (ICPCA), 26–28 October 2011, Port Elizabeth, South Africa, 2853–2857.

[6] B. G. Tudorica, C. Bucur: *A comparison between several NoSQL databases with comments and notes*. Proc. IEEE 10th Roedunet International Conference, 23–25 June 2011, Iasi, Romania, 1–5.

[7] V. Abramova, J. Bernardino: *NoSQL databases: MongoDB vs Cassandra*. Proc. IC Computer Science and Software Engineering, 10–12 Juli 2013, Porto, Portugal, 14–22.

[8] M. A. Khan, Z. A. Memon, S. Khan: *Highly available Hadoop namenode architecture*. Proc. IEEE IC Advanced Computer Science Applications and Technologies (AC SAT), 26–28 Nov. 2012, Kuala Lumpur, Malaysia, 167–172.

[9] E. Dede, M. Govindaraju, D. Gunter, R. S. Canon, L. Ramakrishnan: *Performance evaluation of a MongoDB and Hadoop platform for scientific data analysis*. Proc. 4th Workshop Scientific Cloud Computing, 17–21 June 2013, New York City, NY, USA, 13–20.

[10] L. Vasyliuk, V. Teslyuk: *Software model to MEMS data access based upon MongoDB*. Proc. IEEE XIXth International Seminar/Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory (DIPED), 22–25 Sept. 2014, Tbilisi, Georgia, 184–186.

[11] H. V. Jagadish, J. Gehrke, A. Labrinidis, Y. Papakonstantinou, J. M. Patel, R. Ramakrishnan, C. Shahabi: *Big data and its technical challenges*. Commun. ACM *57* (2014), No. 7, 86–94.

[12] G. H. Kim, S. Trimi, J. H. Chung: *Big-data applications in the government sector*. Commun. ACM *57* (2014), No. 3, 78–85.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 197–206

Access control model based on role and attribute and formal verification[1]

Hui Qi[2], Hongxin Ma[3], Xiaoqiang Di[2], Jinqing Li[2]

Abstract. The defects of the current access control models based on role and attribute (RABAC) and their causes are analyzed and more fine-grained, flexible and efficient RABAC model are proposed. The evaluation indicators of access control model are extended to describe the access control granularity, flexibility and decision performance of model. The model described in this paper and other models are evaluated theoretically in these three aspects.

Key words. RBAC, ABAC, Attribute-based access control, access control, RABAC.

1. Introduction

Since the late 1960s when the access control matrix was proposed, access control technology has received much attention and gained considerable development. At present, lots of access control models have been proposed and applied. Among these access control models, Role-Based Access Control (RBAC) model has made great success [1]. By introducing a mid-layer (role) between users and permissions, the RBAC model can maintain the flexibility and security of the access control system facing large number of users, large amount of data and large scale of business. With the further development of information technology, the emergence of new computing models (such as ubiquitous computing, mobile computing and cloud computing) and the expansion of network environment (from Internet to mobile Internet, Internet of things, space-ground integration network) make access control requirements become more complex. Access control decisions depend more on the context in which the

[1]This work is supported in part by the National High Technology Research and Development Program (863 Program) of China under Grant No.2015AA015701, the Science and Technology Planning Project of Jilin Province under Grant No. 20150204081GX and Grant No. 20140312038ZX, the 13th Five-Year Science and Technology Research Project of the Education Department of Jilin Province under Grant No. 2016-378 and the Industrial Technology Research and Development Project of Jilin Province Development and Reform Commission under Grant No. 2014Y113.

[2]School of Computer Science and Technology, Changchun University of Science and Technology, Changchun, 130022, China

[3]Training Department, Aviation University of Air Force, 130022, Changchun, China

access control requests reside and the security attributes of subjects and objects. ABAC model was born in this environment, which solves the problem that the RBAC model cannot well support a large number of context attributes (the problem of role explosion) and achieves the dynamic and fine-grained access control [2]. However, the access control decision of ABAC model is more complex, hence the safety analysis of access control rules is more difficult.

Both RBAC and ABAC have their particular advantages and disadvantages, and their advantages are complementary. Therefore, some scholars put forward the RBAC/ABAC hybrid approach (namely RABAC model) to hold the simplicity and security of RBAC, as well as the flexibility of ABAC. RABAC model is based on RBAC, using RBAC to manage static relationship between users and permissions to ensure the security of this relationship, while using ABAC to manage dynamic relationship between users and permissions, that is, dynamically applying attribute-based access control rules to user-role mappings, role-permission mappings and user-permission mappings. Nevertheless, there are some deficiencies in current RABAC models with respect to access control granularity, flexibility and decision performance. This paper deeply studies these problems, proposes the improved RABAC model and verifies the new model theoretically.

The rest of this paper is organized as follows. Section 2 gives a brief introduction to the research of access control model. Section 3 presents the framework of the proposed access control model. In Section 4 three important properties of the model are verified in theory. Section 5 concludes this paper.

2. Related work

RBAC model maps the users to the roles, and then maps the roles to the permissions. It makes the management of access control be divided into two parts: the user-role mapping and the role-permission mapping, which simplifies the management of access control. After the RBAC96 model, in order to further improve the security of access control management, several models such as ARBAC97 model (Administrative RBAC97), ARBAC99 model and ARBAC02 model appeared successively [1]. The emergence of these models makes RBAC model more mature, greatly improving the security and the ease of use.

But RBAC model is not suitable for the environment where it is necessary to dynamically and finely control the user-permission mapping [3, 4]. In this environment, the relationship between users and roles and that between roles and permissions frequently changes. More roles are needed to support different access control requirements, which makes the RBAC model more complex and difficult to manage. Therefore, people put forward the ABAC model, using access control rules based on attribute to directly determine the relationship between users and permissions to achieve the dynamic and flexible access control. But ABAC also has disadvantages. Unlike RBAC, it does not have a set of strict rules to ensure the security of user-permission mapping. In the face of a large number of access control rules, the security analysis becomes very difficult. For the problems existing in ABAC and RBAC, many scholars have put forward the new access control model (RABAC) that

combines ABAC and RBAC. Literature [5] proposed 3 ways to introduce ABAC into RBAC: Dynamic Roles, Attribute Centric and Role Centric, the third of which is the better RBAC/ABAC hybrid approach and is widely accepted, but Literature [5] does not make a detailed description of the scheme.

Literature [6] proposed a fine-grained access control model based on role and attribute for web applications and designed the method for verifying the model and that for automatically generating code. The model is based on RBAC and uses attribute-based policies to finely control permissions. Literature [7] proposed role-centric attribute-based access control model of which the access control decision-making process is divided into two stages: first use RBAC to determine all the permissions available to the user in the current session, and then use permission filtering policy (PFP) to extract the final available permissions. The PFP constrains the available set of permissions based on environment, subject and object attributes. The retrieval of PFP depends on the object attributes. Literature [8, 9] proposed a more flexible retrieval approach than Literature [7]. The retrieval approach not only supports object identifiers, but also supports query expressions based on object attributes. Literature [10] proposed a framework integrating attribute-based policies into RBAC. Different from the literature [7, 8], this framework does not apply the attribute-based policies after the establishment of all the permissions available to the user, but it applies the policies during the process of establishing available permissions, e.g. using attribute-based policies to filter user-role relationships and role-permission relationships independently just after the establishment of these relationships.

Literature [11] proposed a generic framework for access control model, under which it researched the two important properties: monotonicity and completeness and then it applied the two properties to evaluate and design different ABAC models. Literature [12], [13] studied the expressive power of access control model qualitatively and quantitatively, which is similar to the completeness of the literature [11]. This paper defines new properties for RABAC model on the basis of the literature [11]: access control granularity, flexibility and decision performance.

3. Model overview

The proposed access control model in this paper is based on the RBAC / ABAC hybrid approaches proposed in the above literatures, and solves some problems in these approaches. This section will show a framework for the new model which makes some improvements to the framework in the literature [10] and describes the composition of the framework and the access control workflow.

3.1. The composition of framework

The framework of the access control model is shown in Fig. 1. The model is divided into upper and lower parts. The upper part uses the RBAC model to control the static part of access control that is the static relationship between users and permissions, while the lower part uses the ABAC model to control the dynamic

part of access control that is reducing the number of permissions associated with a user by attribute-based access control rules.

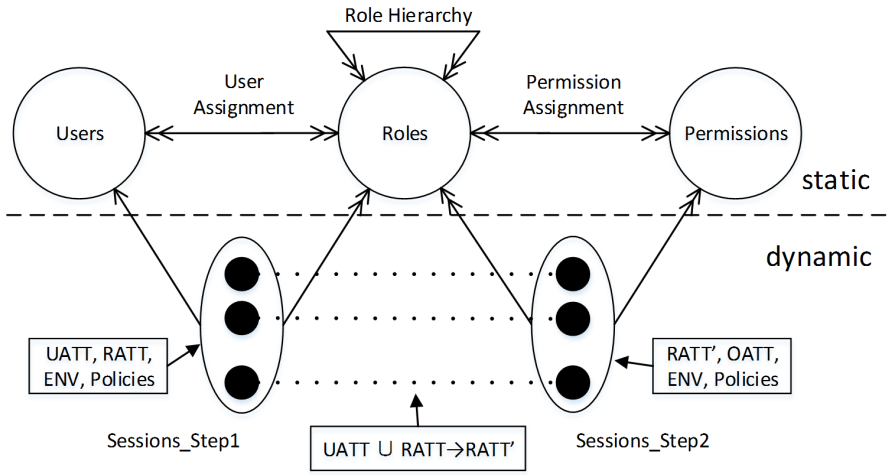

Fig. 1. Framework of the access control model

In the static part of the model, we retain most of the elements of the RBAC model, including users, roles, permissions (Permissions can be subdivided into operations (OPS) and objects (OBS), where Permissions $\subseteq$ OPS$\times$OBS). We regard objects as the attributes of operations so that Permissions=OPS), user-role relationships (UA), role-permission relationships (PA) and role inheritance relationships (RH). In addition, we also define the user attributes (UATT) for users, define the roles attributes (RATT) for roles and define the object attributes (OATT) for permissions. These attributes will be involved in the dynamic part of the model to filter permissions, namely to reduce the permissions of a user.

In comparison with the access control model proposed in the literature [10], the model of this paper is more fine-grained. The literature [10] also applies the attribute-based access control policies to the user-role relationships and the role-permission relationships respectively, but when applying the policies to role-permission relationships, it uses the simple role attributes (RATT), while this paper uses the new RATT' combined with UATT, that is introducing UATT into RATT to form the new RATT' during the transition from Sessions_Step1 to Sessions_Step2. Because RATT' contains the user attributes, the model of this paper can finely control the role-permission relationships. For example, using the model of this paper, we can define that the role of an organization (Organization is usually a user attribute.) is not able to activate a certain permission, which is not supported by the model of the literature [10].

In this paper, UATT is introduced into RATT, for in the RBAC model, the role can be seen as the agent of the user that represents the user to establish the association with the permission. So it is reasonable to add user attributes to role

attributes before performing Sessions_Step2.

3.2. Access control workflow

The above gives the overall structure of the model including the static composition and the dynamic composition, which is the static description of the model. This section will describe how the components of the model are involved in the process of access control decision, namely the dynamic description of the model. In this paper, the process is called the access control workflow, which is defined as follows:

1. Use the RBAC model to determine the user-role mappings, that is U→R.
2. Remove the mappings that violate the access control rules from U→R according to UATT, RATT, ENV and the corresponding policies so that the new user-role mappings (U→R', U→R'⊆U→R) are built.
3. Use the RBAC model to determine the role-permission mappings (R'→P), and combine UATT and RATT to build the new role attributes (RATT', RATT'=UATT∪RATT).
4. Remove the mappings that violate the rules from R'→P according to RATT', OATT, ENV and the corresponding policies so that the new role-permission mappings (R'→P', R'→P'⊆R'→P) are built.

After performing the above steps, we get the final user-permission mappings U→P', which is the subset of user-permission mappings determined by the RBAC model, that is U→P'⊆U→P. Since the security of U→P is guaranteed by the RBAC model and the access control model of this paper only removes some invalid mappings from U→P, the final user-permission mappings U→P' does not violate the constraints of the RBAC model. This is the strategy adopted by the current RABAC model. ABAC is used to filter the user-permission mappings U→P determined by RBAC. Therefor the security of the model is guaranteed and the dynamic and fine-grained access control can be achieved as well.

The access control workflow is actually the process of session establishment. The model of this paper extends the session in the RBAC model. In RBAC, the session only establishes the user-role mappings, while the session of this paper has to establish the role-permission mappings in addition to establishing the user-role mappings. When the establishment of session is completed, the process of access control decision will be over. Thus, it is unnecessary to apply the attribute-based policies after the session is established, which is different from the access control decision in the literature [7, 8]. The section 4 will compare the two decision-making methods.

4. Model verification

Verification of access control model, in particular, formal verification has been a difficult problem. Literature [7, 8, 10] has not verified the model, but these literature have pointed out that the formal analysis and verification of the model is

very important, and it is the direction of future efforts. Literature [11] defines two properties for access control model: completeness and monotonicity. Since these two properties refer to the calculation rules of access control decision and our model does not take into account the specific decision-making process, the two properties are not discussed in this paper.

This paper extends the properties of the literature [11], defines new properties suitable for RABAC model: access control granularity, flexibility and decision performance, and theoretically proves that the model in this paper is superior to the model in the literature [7, 8, 10] in these three aspects.

4.1. Access control granularity

Whether it is RBAC, or ABAC or hybrid model, the goal of the model is to determine the relationship between users and permissions, that is, the permissions a user can obtain in the current session. In order to achieve this goal, the RBAC model statically connects users with permissions through roles; ABAC uses attribute-based access control rules to dynamically determine the user-permission relationships; The hybrid model first uses RBAC to determine the static relationships between users and permissions, and then uses ABAC to dynamically reduce the user-permission relationships. Both the model in this paper and the model in the literature [7, 8, 10] use the approach of hybrid model, that is, first determine the static relationships, and then use attributes to dynamically reduce the relationships. The difference of these models is the way of reducing the user-permission relationships. In the process of reducing the relationships they require various attributes. Therefore, this paper uses the usage of attributes as the measure of access control granularity of the model.

The access control granularity of a RABAC model depends on the maximum number of attributes available in the process of reducing the user-permission relationships.

Based on the above measurement, we can compare the access control granularity of different RABAC models. First, we have to determine which attributes are used when reducing the user-permission relationships. Formal analysis shows that the model in the literature [7, 8] first uses RBAC to generate the ordered pairs consisting of user and permission (UP, UP$\subseteq$USERS$\times$PERMS, and then uses attribute-based policies to determine whether each pair <u, p> in UP is active or not, which in fact establishes the mapping of UP onto the set {T, F}. When building the mapping, the model needs to use the user attributes and the permission attributes. The mapping can be expressed as

$$\text{up\_tf} : \text{UP} \times 2^{\text{UATT}} \times 2^{\text{OATT}} \times 2^{\text{ENV}} \to \{\text{T}, \text{F}\}. \tag{1}$$

In the mapping up_tf, dom(up_tf) includes $2^{\text{UATT}} \times 2^{\text{OATT}} \times 2^{\text{ENV}}$, where UATT is the user attributes, OATT denotes the object attributes which can be seen as the permission attributes and ENV stands the environment attributes.

From the structure of dom(up_tf), it is concluded that the maximum number of attributes available in determining the relationship between UP and $\{T, F\}$ is: $|\text{UATT}| + |\text{OATT}| + |\text{ENV}|$. The key of the above analysis process is to make clear

the mapping structure of UP onto $\{T, F\}$, and then to calculate the access control granularity of the model according to the application of the attributes.

Applying this analysis method to the model in the literature [10], we can calculate the access control granularity of this model. The feature of the model is that the access control decision-making process is divided into two stages. The first stage uses RBAC to build the set of ordered pairs consisting of users and roles (UR, UR⊆USERS×ROLES), and then build the mapping of UR onto T, F. Because the role is the set of permissions, the mapping cannot determine whether each pair <u, p> is active or not. So the first stage does not imply the access control granularity of the model. In the second stage, the model first uses RBAC to establish the set of ordered pairs composed of roles and permissions (RP, RP⊆ROLES×PERMS). Because the role can be regarded as a user agent, the establishment of the set RP implies that the set UP is set up, and the subsequent attribute-based policies also act on the set UP. According to the description of the literature [10], the second stage of the establishment of the mapping of UP onto T, F can be expressed as

$$\text{up\_tf} : \text{UP} \times 2^{\text{RATT}} \times 2^{\text{OATT}} \times 2^{\text{ENV}} \to \{\text{T}, \text{F}\}. \tag{2}$$

From the structure of equation (2), it is known that the access control granularity of the model is |RATT|+|OATT|+|ENV|. In the equation (1), UATT contains RATT, namely RATT⊆UATT, we have |RATT|<|UATT|. Therefore, the model in the literature [7, 8] is more fine-grained than the literature [10].

The model of this paper is also divided into two stages, the first stage is the same as the literature [10]. The difference is the second stage. This paper modifies the structure of equation (2), which is changed to

$$\text{up\_tf} : \text{UP} \times 2^{\text{RATT}'} \times 2^{\text{OATT}} \times 2^{\text{ENV}} \to \{\text{T}, \text{F}\}. \tag{3}$$

In equation (3), RATT'=UATT∪RATT. So, equation (3) is equivalent to

$$\text{up\_tf} : \text{UP} \times 2^{\text{UATT}} \times 2^{\text{OATT}} \times 2^{\text{RATT}} \times 2^{\text{ENV}} \to \{\text{T}, \text{F}\}. \tag{4}$$

The structure of equation (4) shows that the access control granularity of the model in this paper is the same as the literature [7, 8], but it is better than that in literature [10].

4.2. Flexibility

We can find that the model of this paper not only has the mapping like equation (1), but also has the mapping of UR onto {T,F}

$$\text{ur\_tf} : \text{UR} \times 2^{\text{UATT}} \times 2^{\text{RATT}} \times 2^{\text{ENV}} \to \{\text{T}, \text{F}\}. \tag{5}$$

Without consideration of the impact of the attributes and the environment, the model of this paper contains two mappings: UR→{T,F}and UP→{T,F}. These two mappings can be rewritten as: U×R→T,F and U×P→{T,F}, where U represents USERS, R represents ROLES and P represents PERMS. Since R is a collection of

P, these two mappings can be integrated into the mapping

$$\text{upp\_tf} : \mathrm{U} \times 2^{\mathrm{P}} \to \{\mathrm{T}, \mathrm{F}\} . \tag{6}$$

The mapping of UP onto {T, F} in the literature [7, 8] can be written in this general form like the equation (6) without considering the attributes and the environment. So we can give the measure of flexibility.

The measure of flexibility is the maximum capacity of the set dom(upp_tf), namely the value of $|dom(upp\_tf)|$.

According to the above measure, we can easily draw that the model of this paper is more flexible than the model in references [7, 8] because of $|\mathrm{U}\times 2^{\mathrm{P}}|>|\mathrm{U}\times\mathrm{P}|$. The model proposed in this paper has the same flexibility as the model described in [10] because both of them have the same structure of dom(upp_tf), which is $|\mathrm{U} \times 2^{\mathrm{P}}|$.

4.3. Decision performance

In this paper, the decision performance of the model reflects the execution speed of the model. For the RABAC model, the decision performance is mainly determined by the speed of executing the access control policies. For in the access control decision-making process, the process of using RBAC to determine the user-role relationships and the role-permission relationships is the same to each RABAC model. The difference among these models is the attribute-based policies. The policies in this paper and the policies in the literature [10] refer to the mapping of the equation (1) and the mapping of the equation (5). The policies of the literature [7, 8] only refer to the mapping of the equation (1). Without considering the attributes and the environment, the size of the policy set of the two models is different, so the performance of the policy retrieval is also different. As a result, the size of the policy set is the measure of performance.

For the model of this paper and the model of the literature [10], the policy set can be simplified as: U×R→{T,F} and U×P→{T,F}. Since the latter is determined by the role-permission relationships, it can be rewritten as R×P→{T,F}. In this way, the maximum size of the policy set of the model is

$$|\mathrm{U} \times \mathrm{R}| + |\mathrm{R} \times \mathrm{P}| = |\mathrm{U}| \times |\mathrm{R}| + |\mathrm{R}| \times |\mathrm{P}| = |\mathrm{R}|(|\mathrm{U}| + |\mathrm{P}|) . \tag{7}$$

For the model in literature [10], the policy set can be simplified as U×P→T,F. The maximum size of the policy set of the model is

$$|\mathrm{U} \times \mathrm{P}| = |\mathrm{U}| \times |\mathrm{P}| . \tag{8}$$

Comparing with equation (8), the value of equation (7) is smaller because the number of roles is usually smaller than the number of users and the number of permissions. Assuming that the number of users is the maximum value, and the number of roles is the minimum value, in the equation (7), the value $|\mathrm{R}|\times|\mathrm{U}|$ plays a decisive role, which must be smaller than the value $|\mathrm{U}|\times|\mathrm{P}|$. Theoretically speaking, the model of this paper is superior to the model of the literature [7, 8] in decision

performance.

5. Conclusion

In this paper, we deeply analyse the problems existing in the RABAC model and propose an improved model. We also study the formal verification of the RABAC model, analyzing the access control granularity, flexibility and decision performance of the model, proposing the formal method for analyzing these 3 characteristics, and using this method to compare the differences between the new model and other RABAC models. We theoretically prove that the new model is superior to other RABAC models in these 3 characteristics. The formal verification method in this paper provides the evaluation method and the improvement direction for the RABAC model. In future research, we will carry out quantitative analysis of other properties of the RABAC model, will continue to improve the access control model using the quantitative measures, and will try to extend this formal verification method to other access control models.

References

[1] E. Sahafizadeh, S. Parsa: *Survey on access control models.* Proc. IEEE 2nd IC Future Computer and Communication (ICFCC), 21–24 May 2010, Wuhan, China, paper V1 1–3.

[2] V. C. Hu, D. Ferraiolo, R. Kuhn, A. Schnitzer, K. Sandlin, R. Miler, K. Scarfone: *Guide to attribute based access control (ABAC) definition and considerations.* NIST Special Publication 800–162. 2014

[3] L. Cirio, I. F. Cruz, R. Tamassia: *A role and attribute based access control system using semantic web technologies.* Proc. OTM Confederated IC On the move to meaningful internet systems, Vilamoura, Portugal, 25–30 Nov. 2007, Volume Part II, 1256–1266.

[4] V. Suhendra: *A survey on access control deployment.* Series "Communications in Computer and Information Science (CCIS)" *259*, 11–20.

[5] D. R. Kuhn, E. J. Coyne, T. R. Weil: *Adding attributes to role-based access control.* Computer *43* (2010), No. 6, 79–81.

[6] S. H. Ghotbi, B. Fischer: *Fine-grained role- and attribute-based access control for web applications.* Commun. Comp. Info. Sci. (CCIS) *411* (2013), 171–187.

[7] X. Jin, R. Sandhu, R. Krishnan: *RABAC: Role-centric attribute-based access control.* Proc. 6th IC on Mathematical Methods, Models and Architectures for Computer Network Security, 17–19 Oct. 2012, St. Petersburg, Russia, 84–96.

[8] Q. M. Rajpoot, C. D. Jensen, R. Krishnan: *Attributes enhanced role-based access control model.* Lecture Notes in Computer Science *9264*, Springer Verlag, 3—17.

[9] Q. M. Rajpoot, C. D. Jensen, R. Krishnan: *Integrating attributes into role-based access control.* Lecture Notes in Computer Science *9149*, Springer Verlag, 242—249.

[10] J. Huang, D. M. Nicol, R. Bobba, J. H. Huh: *A framework integrating attribute-based policies into role-based access control.* Proc. 17th Symposium on Access Control Models and Technologies, 20–22 June 2012, Newark, New Jersey, USA, 187–196.

[11] J. Crampton, C. Morisset: *Monotonicity and completeness in attribute-based access control.* Proc. 10th International Workshop Security and Trust Management, 10–11, Sept. 2014, Wroclaw, Poland, 33–48.

[12] W. C. Garrison III, A. J. Lee, T. L. Hinrichs: *An actor-based, application-aware access control evaluation framework*. Proc. 19th ACM Symposium on Access Control Models and Technologies, 25–27 June 2014, Ontario, Canada, 199–210.

[13] W. C. Garrison III, A. J. Lee: *Decomposing, comparing, and synthesizing access control expressiveness simulations*. Proc. IEEE 28th IEEE Computer Security Foundations Symposium and Affliated Workshops, 13–17 July, Verona, Italy, 2015, CD-ROM.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 207–218

Research on electro-hydraulic force servo system and its control strategy considering transmission clearance and friction[1]

Li Jianying[2], Zhao Zhongqiang[2], Wang Yanwei[3], Wang Xiaojing[2], Han Guihua[2], Li Shiming[2], Dong Fatang[2]

Abstract. The nonlinear factors such as transmission clearance and friction force in the electro-hydraulic force servo system are analyzed in detail; the corresponding mathematical model is established. In this paper the authors mainly aim at the clearance between the moving parts of the motion pair in the power mechanism, and the clearance is easy to cause large amplitude vibration of the whole system, and even instability and other adverse consequences. At the same time, it is easy to eliminate the influence of the friction in the mechanism, such as the tracking error of the system, and the appropriate friction force can improve the stability margin of the system. An evolutionary algorithm based on the ant colony algorithm and its control strategy for the above mentioned problems are studied. The ant colony algorithm is a heuristic search algorithm based on population optimization, and ant colony algorithm is to find the optimal path through positive feedback and distributed collaboration. The path optimization model can combine the rapidity of the problem solving, the global optimization and the rationality of the limited time answer. The effectiveness of the new simulated evolutionary optimization method, which is based on the ant colony algorithm, is determined and experimentally studied, and the corresponding experimental results are obtained. The nonlinear factors are controlled to improve the effect of comprehensive control performance index of the electro-hydraulic force servo system, such as the fast, dynamic and static accuracy and the stability of the system.

Key words. Electro-hydraulic force servo system, transmission clearance, friction, nonlinear factor, power mechanism, ant colony algorithm, control strategy.

[1]This work is supported by National Natural Science Foundation of China (51405113).

[2]Harbin University of Science and Technology, School of Mechanical and Power Engineering, Harbin, Heilongjiang, 150040, China

[3]Harbin University of Commerce, School of finance, Harbin, Heilongjiang, 150028, China

1. Introduction

Hydraulic power mechanism is the core component of electro-hydraulic servo-system, and is also an important part of electro-hydraulic force servo system [1]. The performance indexes of the hydraulic power mechanism mainly are the quality of the load (inertia), hydraulic natural frequency, hydraulic spring stiffness, hydraulic damping ratio, output force accuracy (output value) etc. The hydraulic power mechanism is usually composed of electro-hydraulic servo valve, servo cylinder, load and other parts [2]. The working principle of hydraulic power mechanism is realized by changing the voltage value of the input signal of the hydraulic power mechanism, and by transforming the input signal from voltage into current. By control of current of the valve, the direction of its opening can be changed and controlled effectively, and, consequently, the direction and amount of the fluid that is used to drive the load movement.

The electric hydraulic force servo system is an important part of electro-hydraulic servo control technology. Its role is based on the signal generator output size and its polarity. According to the strength and direction of the signal, the electro-hydraulic force servo system can assign the driving force the exact value and direction in order to control and drive the inertia load, elastic load and viscous load. Its schematic diagram is shown as Fig. 1. The electric hydraulic force servo system with excellent comprehensive performance should have the advantages of large power volume ratio, fast output response, and high anti-disturbance stiffness [3]. The electric hydraulic force servo system in the process of work is required to have good dynamic characteristics. For their determining, we need a detailed analysis of transmission clearance, friction, stiffness and other nonlinear factors, and need to find out their specific impact on performance, so as to provide a practical basis for good control of these factors in the design of the system, to ensure that the electric hydraulic force servo system can play its superior comprehensive performance [4].

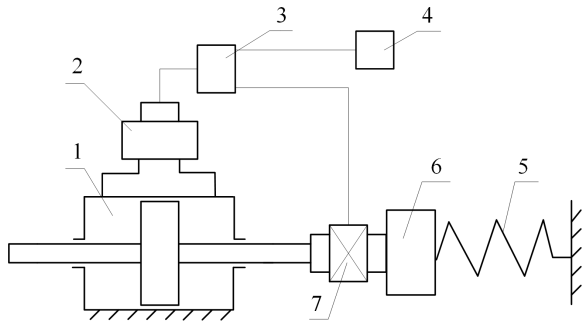

Fig. 1. Schematic diagram of electro-hydraulic servo system: 1–cylinder, 2–electro hydraulic servo valve, 3–controller, 4–signal generator, 5–spring load, 6–mass load, 7–force sensor

2. Electro-hydraulic force servo system

Electro-hydraulic force servo control system researched in this paper, which is the active force control system and its power mechanism is the structure form of the servo valve controlling hydraulic cylinder with its load. We can set the fuel supply pressure as a fixed value, and this pressure returns oil directly to the fuel tank. The slide valve flow equation, hydraulic cylinder flow continuity equation and hydraulic cylinder with the load force balance equation of electro-hydraulic servo control system [5] represent three classical equations, that must be supplemented with the expressions of the related electrical components [6]. We can conclude, that the mathematical analytic expression of the electric hydraulic force servo control system researched in this paper can provide mathematical fundamentals for analysis of the system and also strategy and development of a new controller [7].

$$U_{\mathrm{e}} = U_{\mathrm{r}} - U_{\mathrm{f}}\,, \tag{1}$$

$$U_{\mathrm{f}} = K_{\mathrm{f}} F_{\mathrm{g}}\,, \tag{2}$$

$$\Delta I = K_{\mathrm{a}} U_{\mathrm{e}}\,, \tag{3}$$

$$G_{\mathrm{sv}}(s) = \frac{\mathrm{X_v}}{I} = \frac{K_{\mathrm{sv}}}{\frac{s^2}{\omega_{sv}^2} + \frac{2\xi_{\mathrm{sv}} s}{\omega_{sv}} + 1}\,. \tag{4}$$

Here, U_{r}, U_{f} denote the command voltage signal and feedback voltage signal, respectively. Symbols K_{f} and F_{g} stand for the force sensor gain and hydraulic cylinder output force. Symbol K_{a} represents the servo amplifier gain, K_{sv}, ω_{sv}, and ξ_{sv} are the electro-hydraulic servo valve flow gain (m^3, natural frequency, and dimensionless damping ratio, respectively.

Secondly, we consider the conditions [8] satisfied, namely:

1. The flow of each throttle orifice in the electro-hydraulic servo valve is turbulent, and the influence of the compressibility of the liquid in the valve can be neglected.
2. The response capability of electro-hydraulic servo valve is perfect.
3. The hydraulic cylinder is an ideal double rod symmetrical hydraulic cylinder.
4. The oil supply pressure of the system is constant, and the oil return pressure is zero.
5. The pressure in each working chamber of the hydraulic cylinder is the same.
6. The internal and external leakage of the hydraulic cylinder is laminar flow.

Then the dynamic situation of the mechanism of the system can be described by the slide valve flow continuity equation, flow continuity equation of hydraulic cylinder, and force balance equation of hydraulic cylinder in the form

$$Q_{\mathrm{L}} = K_{\mathrm{q}} X_{\mathrm{v}} - K_{\mathrm{c}} P_{\mathrm{L}}\,, \tag{5}$$

$$Q_{\mathrm{L}} = AsY + \left(C_{\mathrm{tc}} + \frac{V_{\mathrm{t}} s}{4\beta_{\mathrm{e}}}\right) P_{\mathrm{L}}\,, \tag{6}$$

$$AP_{\mathrm{L}} = ms^2 Y + B_{\mathrm{c}} sY + KY + F\,. \tag{7}$$

Here, Q_{L}, K_{q}, X_{v}, K_{c} and P_{L} are, in turn, the load flow of electro-hydraulic servo valve, flow gain, opening of valve core, flow pressure coefficient, and load pressure of hydraulic cylinder. Symbols A, Y, C_{tc}, V_{t}, and β_{e} are, in turn, effective area of the piston of the hydraulic cylinder, piston displacement, total leakage coefficient of hydraulic cylinder, total volume of the two chambers of the hydraulic cylinder, and equivalent volume elastic modulus of hydraulic oil. Finally, m and B_{c} are, in turn, the comprehensive quality of the hydraulic cylinder piston assembly and viscous damping coefficient of the mechanical structure of the piston of the hydraulic cylinder, while K is elastic load stiffness and F is the external disturbance force.

In the simultaneous formulae (5)–(7), the displacement of the valve core can be input as X_{v} and hydraulic cylinder output force as F_{g}, so that the transfer function can be obtained in the form $K_{\mathrm{ce}} = K_{\mathrm{c}} + C_{\mathrm{tp}}$, Due to the usual conditions, conditional $(K_{\mathrm{c}} + C_{\mathrm{tc}})/A^2 << 1$ was established. So we immediately get

$$\frac{P_{\mathrm{L}}}{X_{\mathrm{v}}} = \frac{K_{\mathrm{q}}}{C\,K_{\mathrm{ce}}}\left(\frac{s^2}{\omega_{\mathrm{m}}^2} + \frac{2\xi_{\mathrm{m}} s}{\omega_{\mathrm{m}}} + 1\right), \tag{8}$$

where

$$C = (s/\omega_{\mathrm{r}} + 1)(s^2/\omega_0^2 + 2\xi_0 s/\omega_0 + 1)\,, \quad K_{\mathrm{ce}} = K_{\mathrm{c}} + C_{\mathrm{tc}}\,,$$
$$\omega_{\mathrm{h}} = (4\beta_{\mathrm{e}} A^2/V_{\mathrm{t}} m)^{1/2}\,, \quad \omega_{\mathrm{m}} = (K/m)^{1/2}\,, \quad \omega_0 = (\omega_{\mathrm{h}}^2 + \omega_{\mathrm{m}}^2)^{1/2}\,,$$
$$\xi_{\mathrm{m}} = B_{\mathrm{c}}/(2(mK)^{1/2})\,, \quad \omega_{\mathrm{r}} = 1/(1/\omega_1 + 1/\omega_2)\,, \quad \omega_1 = 4\beta_{\mathrm{e}} K_{\mathrm{ce}}/V_{\mathrm{t}}\,,$$
$$\omega_2 = K_{\mathrm{ce}}/A^2\,, \quad \xi_0 = (1/2\omega_0)(4\beta_{\mathrm{e}} K_{\mathrm{ce}}/(V_{\mathrm{t}}(1 + K/K_{\mathrm{h}})) + B_{\mathrm{c}}/m)\,.$$

3. Ant colony algorithm and intelligent control strategy

Ant colony algorithm is a kind of essential parallel algorithm [9–10]; it mainly solves the problem of path optimization and similar problems. This algorithm has a strong positive feedback structure in the design, and has strong robustness, but also has the characteristics of easy combining with other algorithms; it has been widely used in the relevant field of studying algorithms and control strategy.

3.1. Basic principle of ant colony algorithm

The basic principle of ant colony algorithm is described in [11]. If there exists a graph $G = (V, E)$, where V is a set of nodes in the graph and E is a set of two nodes connected to any nodes in the graph, the line is assigned a certain weight. The path planning problem of the ant colony algorithm is to find a path between these two nodes in the graph $G = (V, E)$ with the lowest weight. At the same time, it can also be understood as the control strategy in the structure of the graph theory, and as the theoretical and applied research on the parameter optimization of controller; in these areas the application of the ant colony algorithm may be considerably wide. The ant colony algorithm process is shown in Fig. 2. The specific implementation of ant colony algorithm can be described by the following steps.

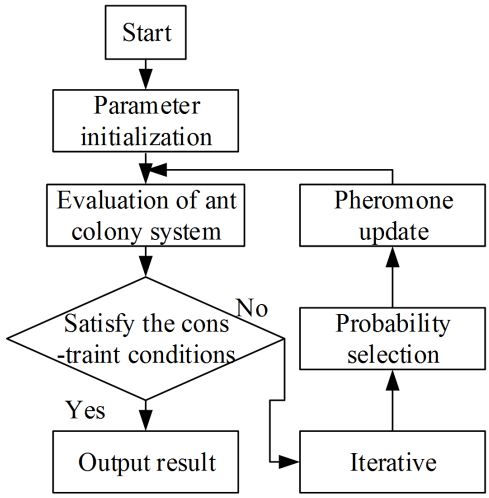

Fig. 2. Flow chart of ant colony algorithm

1. Initialization of parameters: The initial value of pheromone concentration is $\tau_{ij}(0) = \text{const.}$; its increment is also a constant.

2. According to the objective function to determine the merits of the ant colony system, to determine whether the iterative calculation.

3. In the process of searching, the ants always want to go from one node v_i to the other node v_j. Their motivation p_{ij} to use this path is given by the formula

$$\begin{cases} p_{ij}(t) = \dfrac{\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(t)}{\sum_{s\in M_k}\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(t)}\,, \ v_j \in M_k\,, \\ p_{ij}(t) = 0 \text{ otherwise}\,, \end{cases} \tag{9}$$

   where $\tau_{ij}^{\alpha}(t)$ is the pheromone concentration, $\eta_{ij}^{\beta}(t)$ denotes the visibility of side ($\eta_{ij}^{\beta}(t) = 1/d_{ij}$, where d_{ij} is the distance between the nodes v_i and v_j, $v_i, v_j \in V$). The visibility factor reflects the inspiration degree for the choice of side $e(v_i, v_j)$, and it remains a constant in the whole process. Finally, β denotes a desired heuristic factor, which indicates the relative importance of visibility and α is an information elicitation factor that represents the relative importance of the pheromone trail.

4. Pheromone update: once the cycle is completed, the pheromone on each path is updated according to the formulae

$$\tau_{ij}(t+1) = \rho\tau_{ij}(t) + \Delta\tau_{ij}(t, t+1)\,, \tag{10}$$

where

$$\Delta\tau_{ij}(t,t+1)=\sum_{k=1}^{n}\Delta\tau_{ij}^{k}(t,t+1)\,. \tag{11}$$

Here, symbol τ_{ij} is the intensity of the pheromone on the edge of $e(v_i,v_j)$, ρ is a residual volatile coefficient, which is used to show the persistence of pheromone less than 1, and $\Delta\tau_{ij}(t,t+1)$ stands for the increment of pheromone in the cycle.

In the research and practical application, the advantage of the ant colony algorithm is very obvious. At the same time, in the early stage of the search algorithm due to the lack of a certain quantity and effectiveness of strong pheromone, it will make the algorithm in the initial phase of the search speed is very slow, especially the impact of transmission clearance and friction factors existed in the system case. The two factors affect the initial threshold and the hysteresis, and the adverse effects on the system performance are difficult to be solved by ant colony algorithm. But after a gradual search, with the amount of information to a certain number and intensity, it will quickly improve the search for the best solution. The main idea of the study of the algorithm is, by using colony, global and fast random search, to generate an initial solution, and its applied directly to the initial information required for the distribution of ant algorithm, then ant algorithm is the use of feedback, tend to search the advantage of target value of the gradient to obtain optimal solution [12].

3.2. Research on control strategy and parameter optimization of control strategy based on ant colony algorithm

From the perspective of search optimization to understand, the tuning of the controller is to search for an optimal set of parameters from n-dimensional search space composed of the appropriate n parameters, so that the controller can achieve the best effect [13]. Its structure is shown in Fig. 3.

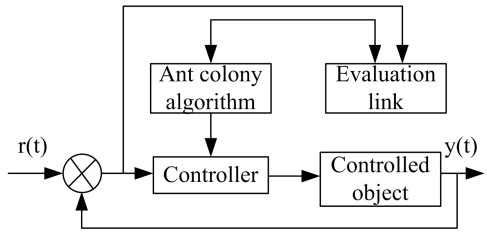

Fig. 3. Structure of controller satisfying ant colony algorithm

Taking into account the influence of transmission clearance and friction factor, the transfer function of the controller can be given by the second-order model with delay in the form

$$G(s)=\frac{K}{As^2+Bs+C}\mathrm{e}^{-\tau s}\,, \tag{12}$$

where A, B, C and K are appropriate constants and τ means the delay. The op-

timization of controller parameters is based on finding the minimum value of the objective function. In order to obtain satisfactory static and dynamic characteristics, reach a fast performance of the system, stability, overshoot and small control performance requirements, the objective function is selected as

$$F = \int_0^\infty t\,|e(t)|\,\mathrm{d}t\,, \tag{13}$$

where $e(t)$ is the time-dependent error.

In accordance with the following type of access to initialize the pheromone:

$$T_0(i) = \exp(-F(x))\,, \tag{14}$$

where $F(x)$ is the value of the evaluation function. The evaluation function value is negatively correlated with the concentration of pheromone. Then, we calculate the objective function of each ant, record the current initial optimal value, and obtain the global transition probability

$$p = \frac{\exp(T_0(\mathrm{BestIndex}) - T_0(j_\mathrm{g})}{\exp(T_0(\mathrm{BestIndex}))}\,, \tag{15}$$

where T_0 is pheromone update function to be optimized and j_g is the random function of independent variable value of the point g.

Further it is necessary to update pheromone. The pheromone update formula is a function that is related to the function to be optimized:

$$T_0(t_t) = (1 - \mathrm{Rou})T_0(t\_t) + (\exp(-F(x)))\,, \tag{16}$$

where t_t is an independent factor parameter of optimization process and Rou is the ratio of the lengths of two adjacent paths with different pheromone in each calculation. In this step, the optimization process is completed, and after the complete optimization is done, the optimal set of data is found, as well as the optimal result of the output controller.

At the same time, in the process of calculating and solving the equation, the following parameters need to be initialized: ant size, number of iterations, search step length and length of each sub region. The search step size can be expressed as follows:

$$step = 0.1rand\,, \tag{17}$$

where $rand$ is the random search step. The length of the sub range is set according to the specific problem, and can be used as the following equation

$$Len_j = (end_j - start_j)/Ant\,, \tag{18}$$

where the value of j depends on the dimension of the equations, end_j is the value at the end of the jth position search, $start_j$ is the value at the start of the jth position search and Ant is the ant search path. We also need to initialize the search location.

The initialization position can be expressed as

$$X(i,j) = (start_j + (end_j - start_j)rand(j)), \tag{19}$$

where X represents the parameters to be solved.

In order to shorten the time of solving nonlinear equation group with transmission clearance and friction force factors, we transform the solution of nonlinear equations into a function optimization problem, namely, nonlinear equation in the nonlinear equation group

$$G(x) = [g_1(x),\, g_2(x), \ldots,\, g_p(x)]^{\mathrm{T}}, \tag{20}$$

$$g_i(x) = 0. \tag{21}$$

The roots of this equation belong to the interval $(c_i,\, d_i)$. There is only one real root of the equation that lies in $[a_i,\, b_i]$, and $[a_i,\, b_i] \in (c_i,\, d_i)$. We can easily prove that the root $x*$ of the equation is equivalent to the function V_i in $[a_i,\, b_i]$ such that

$$V_i(x) = (g_i(x))^2. \tag{22}$$

It is a very small point in $[a_i,\, b_i]$.

At the same time, the solution optimum solution of nonlinear equations can be converted to

$$M(x) = \sum_{i=1}^{p} c_i (g_i(x))^2, \tag{23}$$

where $a_i < x < b_i$, $p \in N$, $i \in N$, a_i and b_i are the upper and lower bounds of the variable vector x_i, and $c_i,\ i = 1, 2, \cdots, p$ is the weight coefficient.

4. Experimental study of electro-hydraulic force servo control system

The experimental platform of experimental research includes industrial control computer, electrical part of signal transmission and processing, mechanical hydraulic parts of the test bench (including the experimental platform, servo valve control servo hydraulic cylinder, force sensor, parallel double cylinder support sliding track, split mass block, stiffness adjustable spring plate) and several more important parts. The hydraulic pipe and oil source and other auxiliary components are also necessary conditions for the experiment.

The realization of the control software of the experiment platform is based on the hardware in the loop simulation system which is generally used in the research with high confidence level. Fast control prototype (FCP) system based on virtual controller and real object is used in this experiment. It is a kind of hardware in the loop simulation experiment. This method can be used to study and experimental verify the new control strategy in the paper, which is efficient, fast and credible in real time hardware. We can on-line modify (through the actual test) the prototype structure or parameters, and then proceed repeatedly to the next round of testing, until we get the prototype of high precision and high reliability control of dynamic

and static index requirements [14–15].

In order to verify the accuracy of the mathematical model of electro-hydraulic force servo system considering the nonlinear factors of transmission clearance and friction, and the control strategy based on ant colony algorithm and the effectiveness of the controller design and parameter optimization on the system control, the experiment research contents of the electrical hydraulic force servo control system are studied in this paper. Fig. 4 shows the system response experiment curve of the system without using ant colony algorithm in the case of intelligent control strategy. Fig. 5 shows the response curve of the intelligent control strategy of ant colony algorithm to control the system, at the same time the accurate mathematical model of electro-hydraulic force servo control system based on the problem of transmission clearance and friction is established. Curve 1 and curve 2 in the figure, respectively, is the input and the output curve of the system.

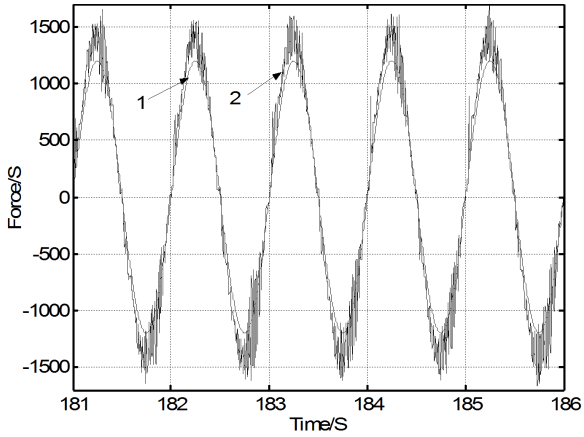

Fig. 4. Response curve of ant colony algorithm is not used

We can see from the diagram in Fig. 4 under the condition of the operating frequency is 1 Hz, output overshoot and peak area are affected by transmission clearance and friction force, and dynamic characteristics is bad. The curves in Fig. 5 are the dynamic characteristics of accuracy and rapidity in the condition of 3 Hz; especially the dynamic characteristics of the transmission clearance and friction in the peak area have obvious improvement. This above all shows the accuracy of the model and the effectiveness of the control strategy based on ant colony algorithm.

5. Conclusion

In this paper, the nonlinear factors of the transmission clearance and friction force in the electro-hydraulic force servo control system are considered, and a more accurate system mathematical model is established. At the same time, the control strategy based on ant colony algorithm to design the controller and optimize its parameters is researched. The research shows that it is effective for the control of the

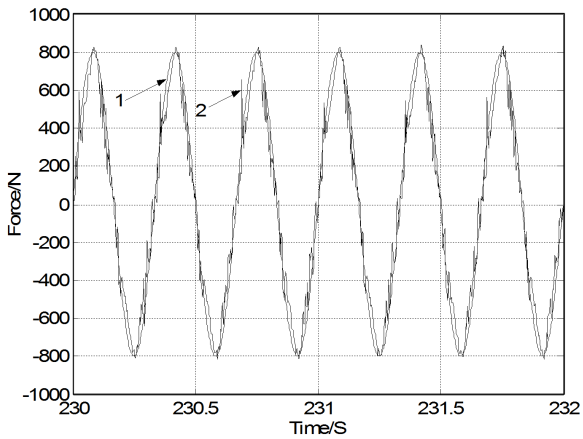

Fig. 5. System response curve of gap and friction and ant colony algorithm

system, and then the experimental study is carried out. In this paper, based on the algorithm of ant colony algorithm for the control of the controller parameter interval is a continuous optimization problem, the parameter interval of the controlled object is discretized, then the ant colony algorithm with global searching speed is used to find the optimal solution. We have completed the optimization process by means of research and implementation of initialization parameters, initialization optimization position, calculation of the optimization of the objective function, with the new pheromone, and repeated calculation and judgment. Finally, by using the ant colony algorithm to consider the nonlinear factors of the transmission gap and friction of the two steps in the process of adding hysteresis to do the experimental study. The result shows that the optimization effect of ant colony algorithm is quite good, and the control effect of the whole electro-hydraulic force servo control system is very obvious. The comprehensive dynamic characteristic of the system is improved obviously.

References

[1] H. E. Merritt: *Hydraulic control systems.* John Wiley & Sons, USA, 1967

[2] C. R. Burrows,: *Fluid power servomechanisms.* Van Nostrand Reinhold Co., 1972.

[3] D. McCloy, H. R. Martin: *The control of fluid power.* Longman group limited, London, England, 1973.

[4] G. Grossschmidt, M. Harf: *Multi-pole modeling and simulation of an electro-hydraulic servo-system in an intelligent programming environment.* IJ Fluid Power *17* (2016), No. 1, 1–13.

[5] M. H. Chiang: *The velocity control of an electro-hydraulic displacement-controlled system using adaptive fuzzy controller with self-tuning fuzzy sliding mode compensation.* Asian Journal of Control ACA, *13* (2011), No. 4, 492–504.

[6] C. Kaddissi, J. P. Kenne, M. Saad: *Indirect adaptive control of an electro-hydraulic*

servo system based on nonlinear backstepping. IEEE/ASME Trans. Mechatron. *16* (2012), No. 6, 1171–1177.

[7] J. E. Johnson: *Electrohydraulic servo system.* Industrial Publishing Company, Michigan, USA, 1973.

[8] J. Y. Li, Y. W. Wang, X. J. Wang, Z. Y. Mao: *Research on electro-hydraulic force servo system based on neural network and fuzzy intelligent control strategy.* J. Comput. Theor. Nanosci. *11* (2014), No. 4, 1205–1210.

[9] N. H. Viet, N. A. Vien, S. G. Lee, T. C. Chung: *Obstacle avoidance path planning for mobile robot based on multi colony ant algorithm.* Proc. IEEE First International Conference on Advances in Computer-Human Interaction (ACHI), 10–15 Feb. 2008, Sainte Luce, Martinique, 285–289.

[10] A. Karmel, C. Jayakumar: *Recurrent ant colony optimization for optimal path convergence in mobile ad hoc networks.* Proc. KSII Trans. Internet Info. Systems *9* (2015), No. 9, 3496–3514.

[11] M. C. Tsou, H. C. Cheng: *An ant colony algorithm for efficient ship routing.* Polish Maritime Research *20* (2013), No. 3, 28–38.

[12] M. Schyns: *An ant colony system for responsive dynamic vehicle routing.* European J. Operational Research *245* (2015), No. 3, 704–718.

[13] J. Dreo, P. Siarry: *Continuous interacting ant colony algorithm based on dense heterarchy.* Proc. Future Generation Computer Systems *20* (2004), No. 5, 841–856.

[14] K. H. Low, H. Wang, M. Y. Wang: *On the development of a real time control system by using xPC Target: solution to robotic system control.* Proc. IEEE IC Automation Science and Engineering, 1-2 Aug. 2005, Edmonton, Alberta, Canada, 345–350.

[15] J. E. Takosoglu, R. F. Dindorf, P. A. Laski: *Rapid prototyping of fuzzy controller pneumatic servo-system.* The IJ Advanced Manuf. Techno. *40* (2009), No. 3, 349–361.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 219–230

Temperature analysis on electromechanical composite brake of automobile based on numerical simulation[1]

J. H. Tang[2,3], K. Y. Wang[2,3], L. C. Zhang[2], G. Q. Li[2]

Abstract. The structure and principle of a new type of electromechanical composite brake for automobile are introduced. The mathematical models of electromagnetic field and temperature field of the electromechanical composite brake are established. The boundary conditions of electromagnetic field and temperature field of electromechanical composite brake are analyzed. The coupling analysis method of temperature field of brake disc is put forward. The temperature distribution of brake disc is calculated when the brake strength is small and the electromagnetic brake works alone, and the temperature distribution of brake disc is also calculated when the brake strength is large and the electromagnetic brake and the friction brake work at the same time. The results show that the temperature rise will not affect the magnetic performance of electromagnetic brake, and the temperature rise is reduced by 15.7 % compared to the traditional friction brake, which effectively enhances the performance of resisting heat recession of automotive brake. This provides a theoretical basis for the design of electromechanical composite brake.

Key words. Automotive engineering, electromechanical composite brake, temperature coupling analysis, numerical simulation..

1. Introduction

The electromechanical composite brake is a new type of brake for automobile, in which an electromagnetic retarder is equipped based on the traditional disc brake

[1]This research is supported by the University Science Research Project of Jiangsu Province (13KJB580005), the Jiangsu Province Ordinary University Graduate Research and Innovation Project (CXXZ13_0659), and the Basic and Applied Basic Research Fund Project of Jiangsu University of Technology (KYY14009).

[2]School of Automotive and Traffic Engineering, Jiangsu University of Technology, Changzhou City, 213001, China

[3]School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang, 212013, China

[1]. The electromagnetic retarder may change the braking energy of automobile in to heat energy and send off, so as to reduce the burden of traditional friction brake [2].It can effectively prolong the service life of friction brake and resist the heat recession of automotive brake, the braking performance of automobile will be improved at the same time. Therefore, the electromechanical composite brake is one of the development trends of automotive brake [3]. When the electromagnetic brake and the friction brake work at the same time, the brake disc not only has an external heat source, but also has an inner heat source. The inner heat source is mainly due to the eddy current loss in the brake disc caused by the brake disc's motion of cutting magnetic lines in the magnetic field. With the work of friction brake, the temperature of brake disc increases, which will weaken the magnetic conductivity between the coils and iron core of electromagnetic brake and affect the braking performance of automobile. In addition, there is a weak coupling relationship between the temperature field and the magnetic field of brake disc. There is a big difference between the coupling factors are considered and the coupling factors are not considered [4], [5]. Therefore, there is great practical significance to study the distribution of temperature field and electromagnetic field to the design of electromechanical composite brake. In this paper, the electromechanical composite brake is taken as research object, the sequential coupling method is adopted, and the coupling of electromagnetic field and temperature field of electromechanical composite brake are simulated and analyzed when the electromagnetic brake and the friction brake work at the same time.

2. Structure and principle of composite brake

The electromechanical composite brake consists of friction brake and electromagnetic brake. The friction brake is mainly composed of a brake disc and a brake caliper body, and the electromagnetic brake is mainly composed of iron core and coils. Its structure is shown in Fig. 1.

The brake piston and the friction block are pushed to the brake disc by the high pressure oil in the brake cylinder of friction brake. The friction between the friction block and the brake disc transforms the kinetic energy of automobile into the heat energy of brake disc, which makes the automobile slow down or brake. When the electromagnetic brake works, the electric current passes through the coil, the brake disc turns for cutting the magnetic line in the magnetic field. At this time, the brake disc will produce eddy current and the eddy current will produce the braking torque on the brake disc which hinders the rotation of brake disc. In the process of braking, the control unit can dynamically distribute the braking torque to the friction brake and the electromagnetic brake according to the braking conditions of automobile.

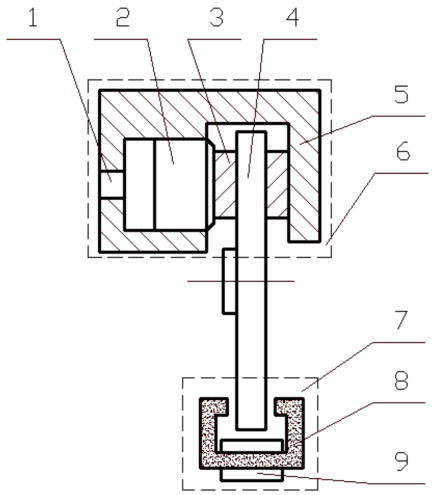

Fig. 1. Schematic diagram of electromechanical structure of composite brake: 1–oil hole, 2–brake piston, 3–friction block, 4–brake disc, 5–brake caliper body, 6–friction brake, 7–electromagnetic brake, 8–iron core, 9–coil

3. Establishment of mathematical models

3.1. Mathematical models of electromagnetic field

The formation mechanism of electromagnetic field of electromechanical composite brake is shown in Fig. 2. After the coil is electrified, the two terminals of iron core are changed into two magnetic poles. The brake disc turns relative to the iron core for cutting the magnetic line. The counter clockwise and clockwise vortices are generated separately in the front and rear sides of brake disc. According to the law of left hand, the ampere force generated by the vortices will generate braking torque on the brake disc, which will prevent the rotation of brake disc.

It can be seen from Fig. 2, the closer to the center of brake disc, the smaller the induction current generated in the magnetic field. Therefore, the whole brake disc is divided into a vortex area and a non-vortex area in the calculation. The magnetic field and electric field are required to be calculated in the vortex region, and only the magnetic field is calculated in the non-vortex region. According to the Maxwell equation group and the law of conservation of charge, the differential equations of the vortex region and the non-vortex area can be expressed as

$$\operatorname{curl} \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}\,, \quad \operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}\,, \quad \operatorname{div} \boldsymbol{D} = \rho\,, \quad \operatorname{div} \boldsymbol{B} = 0\,, \tag{1}$$

where $\boldsymbol{H}$ is the magnetic field strength vector, $\boldsymbol{J}$ is the eddy current density vector, $\boldsymbol{B}$ is the magnetic flux density vector, $\boldsymbol{E}$ is the electric field strength vector, $\boldsymbol{D}$ stands for the electric displacement vector and ρ is the electric charge volume density.

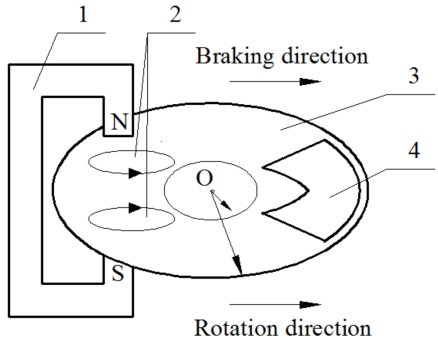

Fig. 2. Formation mechanism of electromagnetic field of composite brake 1–iron core, 2–eddy current, 3–brake disc, 4–friction block

This system must be supplemented with the constitutive equations in the form

$$\boldsymbol{D} = \varepsilon \boldsymbol{E}\,, \quad \boldsymbol{B} = \mu \boldsymbol{H}\,, \quad \boldsymbol{J} = \gamma \boldsymbol{E}\,, \tag{2}$$

where ε is the dielectric permittivity, μ is the magnetic permeability and σ is the conductivity.

The vectors used in the Maxwell equations represent the relationship of electromagnetic field are difficult to calculate. Therefore, in the analysis of electromagnetic field problems, certain potential functions are often introduced as the auxiliary quantity. In the vortex area, the scalar electric potential φ is used to describe the electric field and scalar magnetic potential ψ for magnetic field, while the magnetic vector potential $\boldsymbol{A}$ is used to express the magnetic field in the non-vortex area.

The potential function of the electric scalar potential in the vortex area in plane x, y is defined by

$$\boldsymbol{E} = -\text{grad}\ \varphi = -\mathbf{i}\frac{\partial \varphi}{\partial x} - \mathbf{j}\frac{\partial \varphi}{\partial y}\,. \tag{3}$$

Similarly, the potential function of the magnetic scalar potential in the vortex area in plane x, y is defined by

$$\boldsymbol{H} = -\text{grad}\ \psi = -\mathbf{i}\frac{\partial \psi}{\partial x} - \mathbf{j}\frac{\partial \psi}{\partial y}\,. \tag{4}$$

Finally, in plane arrangements the vector potential has only one component A_z in the z-direction and may be defined by the relation

$$\boldsymbol{B} = \text{curl}\ \boldsymbol{A} = \mathbf{i}\frac{\partial A_z}{\partial y} - \mathbf{j}\frac{\partial A_z}{\partial x}\,. \tag{5}$$

Substitution of equations (3) and (5) into (1) provides two equations of electric

and magnetic field distributions in the form

$$\nabla^2\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} = -\frac{\rho}{\varepsilon}\,,$$

$$\nabla^2 A_z = \frac{\partial^2 A_z}{\partial x^2} + \frac{\partial^2 A_z}{\partial y^2} = -\mu J_z\,. \tag{6}$$

In non-vortex area without currents there holds

$$\nabla^2 A_z = \frac{\partial^2 A_z}{\partial x^2} + \frac{\partial^2 A_z}{\partial y^2} = 0\,. \tag{7}$$

3.2. Mathematical models of temperature field

A small portion of the heat generated by the friction brake and electromagnetic brake is emitted into air, and the most of the heat is absorbed by the brake disc, which leads to the increase of the brake disc temperature. With the increase of the brake disc temperature, the heat conduction, heat convection and heat radiation will be generated in the interior and on the surface of brake disc. According to the principle of heat transfer, the 3D transient heat conduction equation of brake disc and friction block can be obtained [6].

$$k(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}) + q_{\mathrm{V}} = \rho c\frac{\partial T}{\partial t}\,, \tag{8}$$

where k is the thermal conductivity, T is the temperature, q_{V} represents the volumetric heat losses, ρ is the density of material, c is the specific heat of material at a constant pressure, and t is the time.

During the braking process in the brake disc, the eddy currents in it are the internal heat source, and its strength is [7]

$$q_{\mathrm{V}} = \gamma\omega^2\,|A|^2\ , \tag{9}$$

where γ is the electric conductivity of material and ω is the current angular frequency.

The following boundary condition of the temperature field is applied [8]

$$k\frac{\partial T}{\partial n} = -h(T - T_a) - C\sigma(T^4 - T_a^4)\,, \tag{10}$$

where T_a is the environmental temperature, h is the coefficient of convective heat transfer, C is the radiation coefficient, σ is the Stefan–Boltzmann constant, and n denotes the direction of the outward normal.

Another boundary condition should be satisfied on the contact area between the brake disc and the friction block:

$$q = F_{\mathrm{w}} \cdot f \cdot P(t) \cdot \omega(t) \cdot r\,, \tag{11}$$

where F_w is the weight of friction heat flux density input to the brake disc, f is the friction coefficient between the brake disc and friction plate, $P(t)$ is the transient variation of brake pressure, $\omega(t)$ is the transient variation of the angular velocity.

The third kind of boundary condition should be satisfied in the non-contact area between the brake disc and friction block. This condition has the form

$$h = 0.664\,\mathrm{Pr}^{1/3}\mathrm{Re}^{1/2}\,, \tag{12}$$

where Pr is the Prandtl number and Re is the Reynolds number.

4. Coupling calculation and analysis

The sequential coupling method is adopted for coupling between the electromagnetic field and the temperature field of electromechanical composite brake. The ANSYS software is used for simulation and calculation. In this paper, the Jetta car with 1.6MT CTX engine is taken as the research object. The total weight of the car is 1650 kg, the tire model is 245/45R18 and its structure parameters are 85 mm and 115 mm. The number of turns of a single coil is 820, the air gap is 2 mm, the current of exciting coil is 25 A, and the ambient temperature is 293 K. The material of brake disc is copper, its thermal conductivity is 383 W/(m °C), specific heat is 390 J/(kg °C), and its density is 8889 kg/m^3. Next step, the temperature rise of brake disc is calculated respectively in the two working conditions of small strength and high strength based on the ANSYS software.

4.1. Continuous braking with small strength

The first hypothesis is that the speed of automobile is 30 km/h, the slope of road is 10 degrees, and the observation time is 30 seconds. Because the strength of braking is small at this time, only the electromagnetic brake works in the electromechanical composite brake. Therefore, it is only necessary to consider the influence of eddy current and electromagnetic coil on the temperature rise of brake disc.

According to the characteristic of left and right symmetry of the electromagnetic brake with respect to the brake disc, the two dimensional computational domain model of the electromagnetic brake and the brake disc is established. In this time, the electric current of electromagnetic brake is 15 A and the speed of brake disc is 980 rpm.

Through analysis, it can be seen that the maximum magnetic flux density of brake disc is 1.454 T, the maximum eddy current density is $8.32 \times 10^4\,\mathrm{A\,m^{-2}}$, the maximum eddy current loss density is $9.028 \times 10^7\,\mathrm{W\,m^{-3}}$. The direction of eddy current density and the direction of magnetic induction intensity are perpendicular to each other. The position of eddy current loss is on the outer layers of two sides of the brake disc, which corresponds with the actual analysis. Therefore, the heat rate of brake disc can be calculated through the analysis of electromagnetic field. The heat rate of brake disc is the eddy current loss per unit volume of brake disc under the role of electromagnetic brake, which is used as the boundary condition

to analyze the temperature field of brake disc. Under the continuous braking with small brake strength, the temperature field of brake disc can be calculated, as shown in Fig. 5, and the temperature curves of different tori in the same radial direction is shown in Fig. 6.

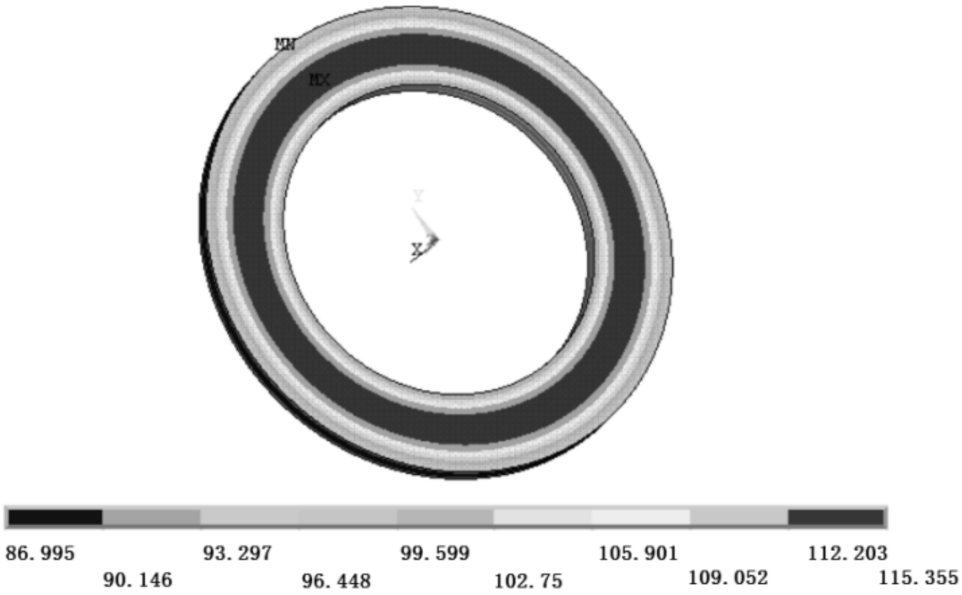

Fig. 3. Contour map of temperature field (MN denotes mimimum temperature, MX denotes maximum temperature)

The region with high temperature is mainly concentrated in the middle of brake disc, the temperatures of the inner and outer surfaces of the brake disc are lower than that of the middle torus. The maximum temperature is about 115 °C, the minimum temperature is about 87 °C and the temperature change trend of brake disc is same in the same radial direction. The main reason is that the vehicle speed is constant, and the inner heat source of brake disc is constant during the continuous braking, so that the heat generation rate caused by eddy current loss is constant. The internal heat source absorbed by the middle torus is greater than its heat dissipation, and the heat of the inside and outside tori which are far away from the action position of magnetic pole and brake disc is passed through the heat transfer. The distance from the middle torus is farther away, the lower the temperature.

4.2. Braking condition with high brake strength

Assuming that the automobile slows down from 120 km/h to 30 km/h, then runs at the speed of 30 km/h for 30 minutes. At this time, the electromagnetic brake and the friction brake play a role at the same time, and the braking torque of the two brakes is needed to be dynamically allocated. The total braking torque of

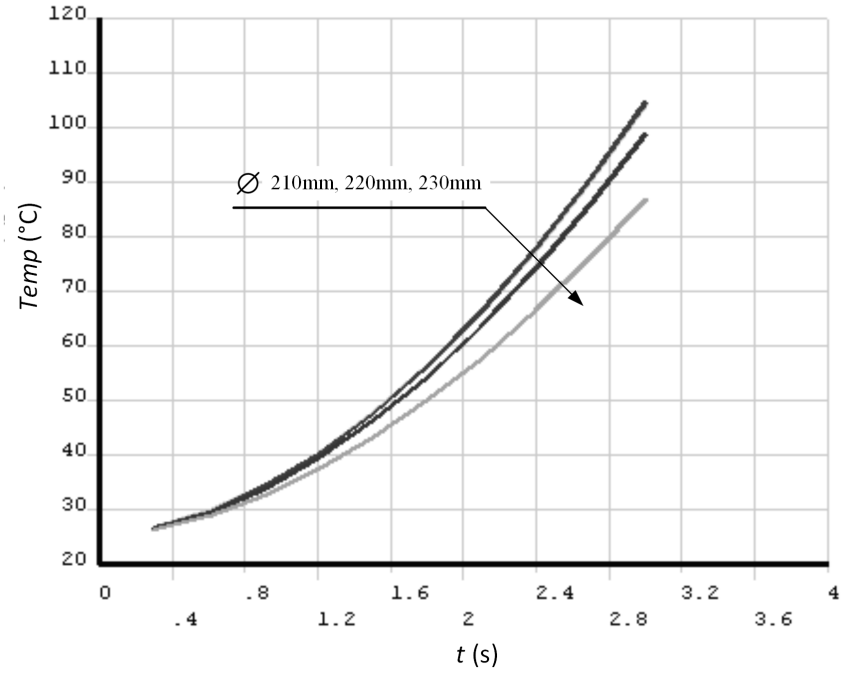

Fig. 4. Temperature curves of different tori

electromechanical composite brake is as follows.

$$T = T_{\mathrm{m}} + T_{\mathrm{h}} = \frac{1}{3} f \pi d^2 p \frac{\left(R_2^3 - R_1^3\right)}{\left(R_2^2 - R_1^2\right)} + \frac{16\sqrt{2}\pi\rho\Delta_{\mathrm{h}}\omega a b^2 L(\mu_0 N I)^2}{\left(16\pi\rho l_{\mathrm{g}} + \sqrt{2} a b \Delta_{\mathrm{h}} \mu 0 k_{\mathrm{e}} \omega\right)^2}, \tag{13}$$

where T_{m} is the brake torque of friction brake T_{h} is the brake torque of electromagnetic brake, d is the piston diameter, p is the brake fluid pressure, R_1 is the inner diameter of friction block, R_2 is the external diameter of friction block, ρ is the resistivity of brake disc, Δ_{h} is the skin depth of eddy current on the brake disc, ω is the angular velocity of the brake disc, L is the distance from the center of brake disc to the center of pole, I is the electric current of electromagnetic brake, μ_0 is the magnetic permeability of vacuum, l_{g} is the air gap between the brake disc and the magnetic pole, and k_{e} is the conversion coefficient.

In the above formula, there exist the relationship

$$\Delta_{\mathrm{h}} = \sqrt{\frac{2\rho}{\omega\mu_0\mu_r}}, \tag{14}$$

where μ_r denotes the relative permeability.

The integrated control strategy is adopted to dynamically allocate the braking torque between the friction brake and the electromagnetic brake. The allocation

results are shown in the Fig. 7.

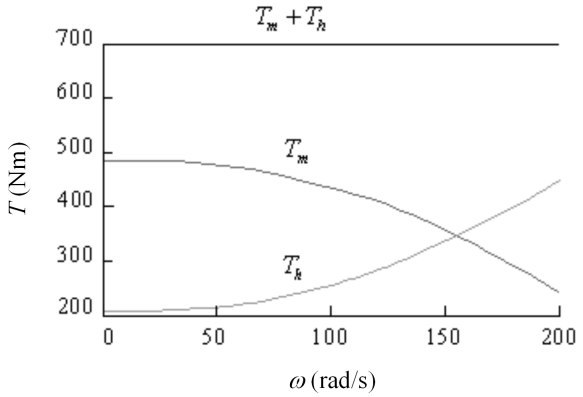

Fig. 5. Distribution curves of brake torque

The distribution results of brake torque in Fig. 7 are used to determining the above thermal boundary conditions of the electromagnetic brake and the friction brake. In the coupled analysis, the thermal boundary conditions comprise the heat generation rate of electromagnetic field, heat of electromagnetic coil, and heat flux density, thermal radiation, heat convection between the friction block and brake disc. Under the braking condition with high brake strength, the temperature field of brake disc can be calculated, as shown in Fig. 8, and the temperature curves of different tori in the same radial direction is shown in Fig. 9.

In the process of high-strength braking, the temperature of brake disc shows a gradient distribution from the middle of brake disc to the inner and outer sides. The high temperature zone is the contact area between the brake disc and friction block. The temperature of brake disc rises rapidly, and the temperature reaches the highest value which is about 365 °C at 1.8 s, and then begins to slow down. The reason is that the friction heat and current loss, which are just absorbed, do not have time to spread to the periphery. It results in a sharp increase of temperature on the inner surface of brake disc. With the extension of braking time, the speed of brake disc decreases continuously, and the heat flux density absorbed by it is also decreasing. However, the heat flux transmitted through heat conduction mode exhibits a little change in its interior. At this time, the heat absorbed by the inner surface of brake disc is lower than the heat loss, which causes the temperature of brake disc to drop, but the decline is slow.

Through the dynamic distribution of the brake torque and the temperature field analysis, the relationship between the temperature of brake disc and the brake torque can be simulated as shown in Fig. 10.

In the picture above, the curve of 1 is the temperature curve of friction braking torque, the curve of 2 is the temperature curve of integrated braking torque, and the curve of 3 is the temperature curve of electromagnetic braking torque. The trends of the three curves are basically the same. With the increase of the braking torque,

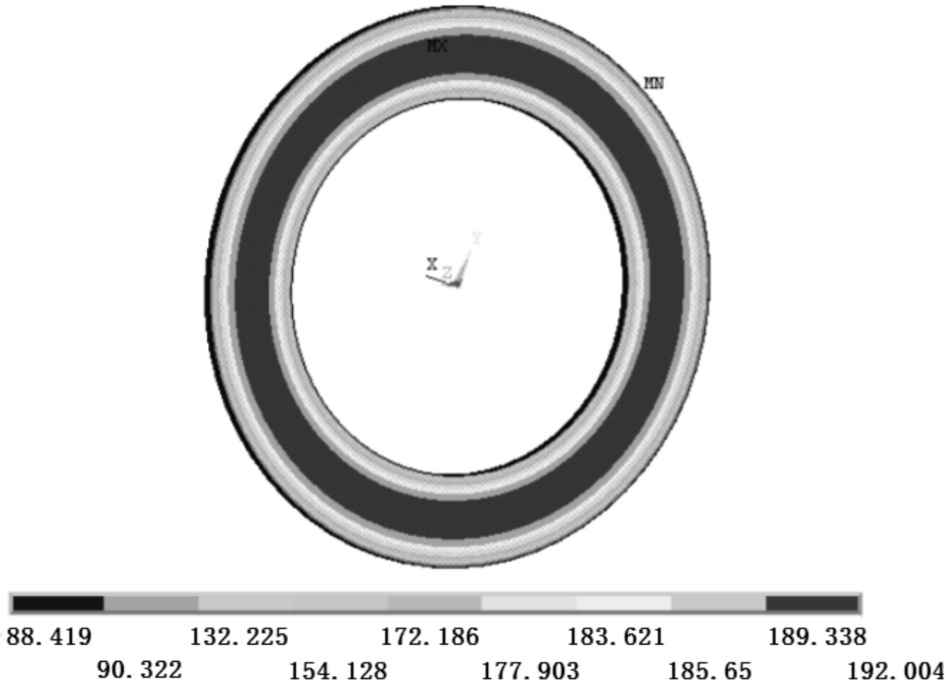

Fig. 6. Contour map of temperature field (MN denotes mimimum temperature, MX denotes maximum temperature)

the temperature of brake disc is increasing. Among them, the temperature rise of electromagnetic brake is slow, and the temperature rise of friction brake is the most obvious. When the braking torque reaches 700 N m, the temperature of friction brake is about 545 °C, and the temperature of integrated brake is about 440 °C, which represents a decrease by about 15.7 % compared with the traditional friction brake. And the maximum temperature of the integrated brake is much smaller than the Curie point of the iron material, so that it will not affect the magnetic conductivity of the integrated electromagnetic brake.

5. Conclusion

The thermal analysis of electromechanical composite brake is carried out, particularly, temperature fields under two kinds of typical braking conditions are obtained. The simulation results show that the temperature rise of electromechanical composite brake will not affect its electromagnetic force. The temperature rise of electromechanical composite brake has a clearer decline than that of the traditional friction brake. It can be seen that the models and calculation method presented in this paper are accurate and practical, which can meet the needs of engineering practice. This provides the theoretical basis for the design of electromechanical

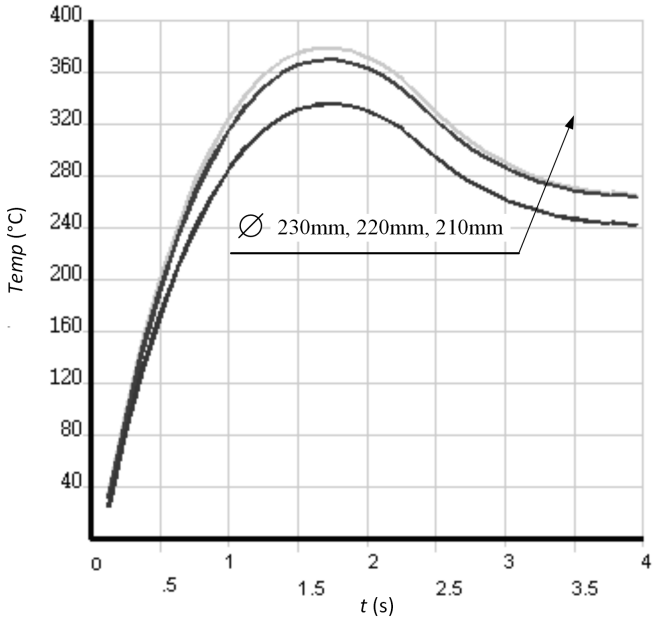

Fig. 7. Temperature curves of different tori

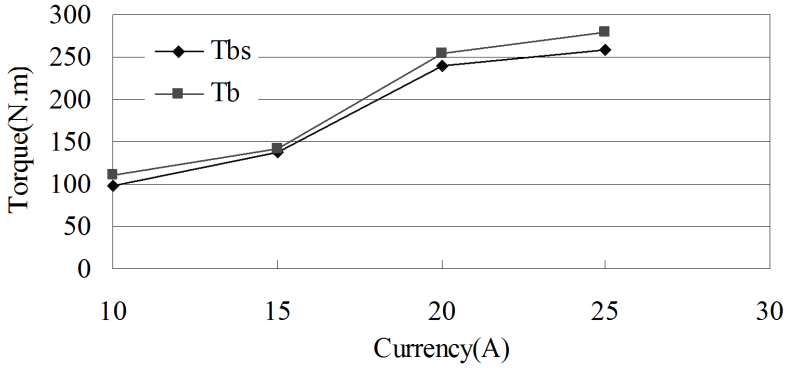

Fig. 8. Relation between temperature and brake torque

composite brake.

References

[1] J. H. Song: *Performance evaluation of a hybrid electric brake system with a sliding mode controller*. Mechatronics *15* (2005), No. 3, 339–358.

[2] S. Anwar, B. Zheng: *An antilock-braking algorithm for an eddy-current-based brake-*

by-wire system. IEEE Trans. Vehicular Technol. *56* (2007), No. 3, 1100–1107.

[3] S. E. Gay: *Contactless magnetic brake for automotive applications.* Doctoral dissertation, Texas A&M University, 2005.

[4] C. Meinecke, A. M. Miri: *Coupled electromagnetic and thermal analysis of magnets wound from cable-in-conduit conductors.* Physica C: Superconductivity *372–376* (2002), No. 3, 1427–1433.

[5] M. Montazeri-Gh, M. Mahmoodi-k: *Development a new power management strategy for power split hybrid electric vehicles.* Transportation Research Part D: Transport and Environment *37* (2015),79–96.

[6] T. J. Hughes: *The finite element method: linear static and dynamic finite element analysis.* Dover Civil and Mechanical Engineering, (2012).

[7] R. F. Harrington: *Field computation by moment methods.* Wiley-IEEE Press, USA, 1993.

[8] R. Shiraishi, K. Demachi, M. Uesaka: *Numerical simulation of coupled problem of electromagnetic field and heat conduction in superconducting magnetic bearing.* Physica C: Superconductivity *392–396* (2003), No. 1, 1–4, 734–738.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 231–238

Wireless thermal charging system based on GPRS technology[1]

Yu Yao[2], Xiao Zhigang[2], Kevin Gao[3]

Abstract. The thermal station dispatch department needs real-time monitoring of users in different geographical locations, the purpose of real-time monitoring is to record the heat transfer station temperature, pressure and other parameters, and according to the various heat transfer stations to monitor the parameters of the data, the operation of the heat transfer station to adjust to ensure the normal operation of winter heating. In this paper, General Packet Radio Service (GPRS) technology made a brief introduction, according to the characteristics of thermal enterprise monitoring system, a solution based on GPRS wireless thermal monitoring system is put forward. Aiming at the running effect and practical application of the system, the test method to meet the requirements is put forward, and the test results are analyzed. The function and design of the module can effectively meet the needs of thermal remote meter reading business.

Key words. GPRS, wireless network, TCP/IP protocol, database.

1. Introduction

For all users of different geographical locations in the city, thermal dispatching departments need real-time monitoring and meter reading of thermal data [1–3]. The monitored object is represented by the basic parameters in the thermal network: temperature, pressure, flow, average flow and instantaneous value. At present, the heat industry meter reading information management is based on the traditional manual meter reading, as there is practically no application of intelligent automatic meter reading system. If all the work has to be done by manual meter reading, it is a time consuming and quite inefficient way [4–6]. Collecting or reproducing heat meters for real-time data requires manual operation, seriously affecting collection

[1]This work is supported by the 2014 Baoding science and technology research and development plan (14ZG008), Research project of higher school science and technology in Hebei Province in 2014 (Z2014164), Agricultural University of Hebei science and Technology Fund(LG20150203), and 2015 Baoding science and technology research and development plan(15ZG015).

[2]Mechanical and Electrical College, Agriculture University of Hebei, 289 Lingyusi St, Nanshi Qu, Baoding Shi, Hebei Sheng, 071000, China

[3]Kana Racotech Geophysical Inc., Houston, TX, USA

efficiency [7–10]. Although many thermal power companies have implemented a variety of meter reading management mechanisms, the meter reading effect is still unable to be fundamentally improved and improved, increasing the workload of thermal enterprises and reducing the efficiency of instrument reading [11–12].

Based on the above reasons, a GPRS-based wireless network with remote calorimeter reading information system is established, which can realize the collection of field parameters of various users and the real-time communication of data commands.

2. System working principle

Remote wireless heat meter reading systems include data collectors, data concentrators, data servers, wireless networks, and Internet. The data collector receives the hot data, analyzes them and sends them to the concentrator over the wireless network. The data concentrator sends the data to the data server over the GPRS network, which stores the hot data in the database. When the administrator needs to monitor the thermometer in real time, the data server sends the data acquisition command to the concentrator through the GPRS network, and the data concentrator performs the operation. An example of the network structure of 4G in remote meter reading system is depicted in Fig. 1.

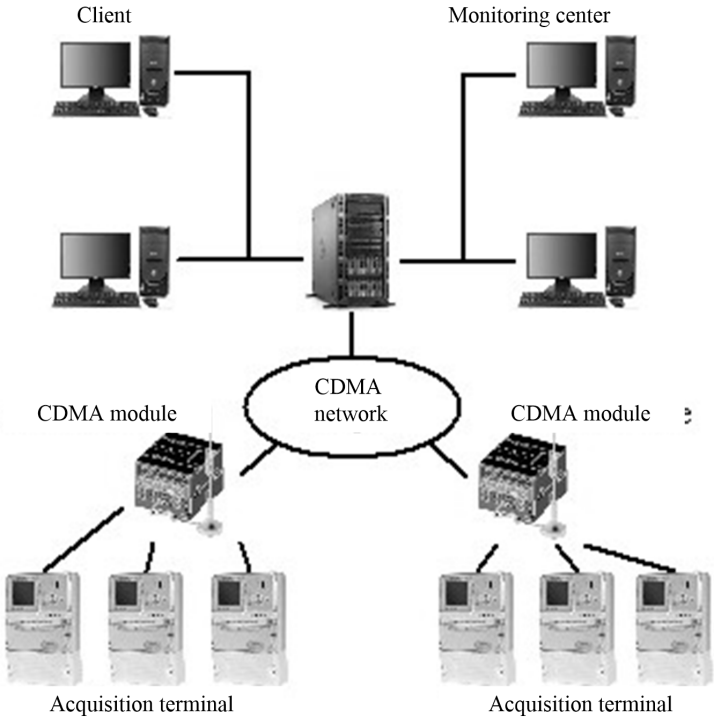

Fig. 1. Network structure of 4G in remote meter reading system

3. System outline design

3.1. Overall functional framework

Remote meter reading information system functions are: user management (import or export), heat table management (replacement table records), instrument readings display, system maintenance (rights management, system log, parameter maintenance), query statistics (user defined).

The system can automatically read the operating parameters of the heat meter, remote monitoring heat meter, read real-time heat values and real-time display hot-cost information.

Based on the demand analysis of the whole system summarized in the last section, the overall functional framework of the remote meter reading system is designed. The overall function frame diagram of the system is shown in Fig. 2.

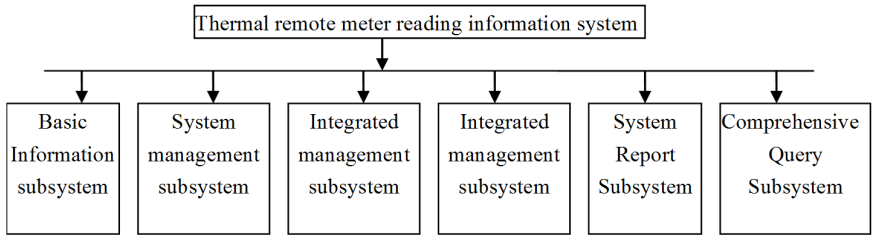

Fig. 2. System overall function

The software uses B/S structure and does not need to install the client software. It can be able to use the remote meter reading management function.

3.2. Database design

Information of the data involved in the system are mainly district information, heat exchanger station information, community information, building information, user information, user housing information, housing information, meter reading information, concentrator information, and heat meter and concentrator log information. Effective management and maintenance of these data is the core and the basis of the system. The performance of the whole system is good or not. The database design is based on the characteristics of different data and application processes, design science and reasonable database deployment plan to meet the unified management and efficient requirements on application of thermal remote meter reading information system.

The system database is a large database Oracle, version 11g Oracle. According to the analysis of the demand of the user data, the logical model of the system is established, and then the physical structure of the database is designed based on the logical model. Finally, the main physical structure of the database is as follows:

Large area information table xt_dqxx; heat exchange station information table xt_hrzxx; cell information table xt_xqxx; building information table xt_lfxx; user information table xt_yhxx; user housing information table xt_yhfwxx; housing information table xt_fwxz; meter reading information cb_cbxx; concentrator information table jzq_jbxx; heat meter information rlb_jbxx; concentrator log table jzq_log.

The main tables of the database are listed below in Table 1 and Table 2.

Table 1. Large area information table xt_dqxx

Serial number	Field name	Field	Type	Length	Remarks
1	Increment	id	integer		
2	Large area encoding	dqbm	VARCHAR2	30	Encoding: 00, 01 ···
3	Large area name	dqmc	VARCHAR2	50	
4	Operator	czr	VARCHAR2	50	
5	Operation time	czsj	DATE		
6	Remarks	bz	VARCHAR2	200	

Table 2. Exchange station information table

Serial number	Field name	Field	Type	Length	Remarks
1	Increment	id	Integer		Increment
2	The heat exchange station encoding	hrzbm	VARCHAR2	30	Encoding: big area code + station number: 0001, 0102 ···
3	The heat exchange station	hrzmc	VARCHAR2	50	
4	Large area ID	dqid	Integer		
5	Operator	czr	VARCHAR2	50	
6	Operation time	czsj	DATE		
7	Remarks	bz	VARCHAR2	200	

4. System detailed design and implementation

4.1. Basic information management subsystem

The basic information management subsystem includes regional information maintenance, heat exchanger station information maintenance, community information maintenance, storied building information maintenance, maintenance of the nature of housing, concentrator information maintenance, and heat meter information maintenance function.

Background class design: Create the Base Info Manage class and in Base Info Manage, the do Get do Post method of the Servlet class. Processing of the page region information maintenance, heat exchanger station information maintenance, community information maintenance, storied building information maintenance, maintenance of the nature of housing, concentrator information maintenance and maintenance of the heat meter.

Classes and methods used are: Check, Base Info Add Dao. The Base Info Manage and Base Info Change Dao classes are mainly designed in the basic information management module. Mainly for the user to submit a form of inspection, basic data add and modify, the use of the class and method presented is in Table 3.

Table 3. Classes and methods used in the design of basic information management

Classes and methods	Function
Check (form)	Check the integrity of the information
Do Get\ do Post method in Base Info Manage class	Processing information to add, modify, delete the form submitted by the page
Base Info Add method for Dao package Base Info Add Dao	Add data to the database table
Dao Base Info Change Dao class base Change	Modify, delete data in the database

4.2. Implementation effect of basic information management module

Figure 3 represents the basic information management module in the regional information maintenance list interface diagram.

5. System debugging and testing

5.1. Communication protocol testing

The information system of the thermal remote meter reading system based on GPRS is developed, which is based on the system software and communication protocol, and the function of the software is suitable, as well as its reliability and operability. Another important part of this system is data communication, which

relates to the transmission and two levels of communication protocol.

After the data concentrator starts and the data server establishes the connection, the login request information is sent first, the message type is 4 (04H) and the protocol test case format is as follows:

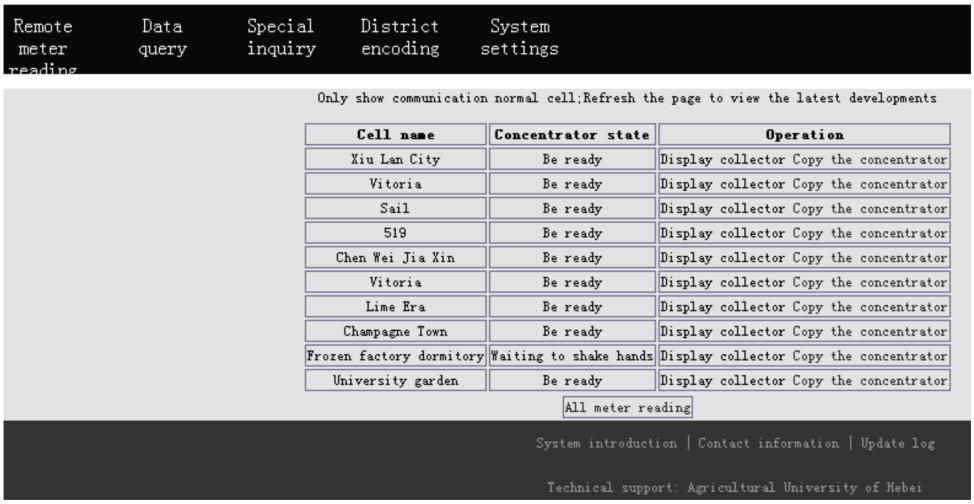

Fig. 3. Basic information management module

Table 4. Exchange station information table

Version number	Serial number	Message body length	Message type	Current time (1)	Test position
02H	00H	00H1AH	04H	13H06H16H15H00 H00H	92H
Equipment number		SIM card number (2)		Login password (3)	
30H30H30H31H		31H35H31H32H39H30 H38H32H39H, 35H39H		31H35H31H32H39H30H38H32 H39H35, H39H	

The data server receives a connection request for the data concentrator and analyzes the connection protocol: (1) the current time: 09 15 years 16 15 points 0 minutes 0 seconds (2) SIM card number: 18330226210 (3) login password: 18330226210 Analytical results: correct. The data concentrator connects the data server successfully and the data concentrator sends the hot table address for 0000 1011H and sends a request message to the current time: in September 15 16 15 0 0 seconds. Test for a communication request message is shown in Tables 5 and 6.

The data server receives the request message of the data concentrator, parses the message type, returns the data address of the heat meter to the data concentrator and returns the information format shown in Table 5-3.

Table 5. Test case for information of the heat meter

Version number	Serial number	Message body length:null	Message type	Current time (1)
02H	01H	00H00H	0BH	13H06H16H15H00H00H
Test position				
A4H				

Table 6. Test case returns results

Version number	Serial number	Message body length	Message type	Current time (1)	Test position
2H	01H	00H18H	0CH	13H06H16H15H00H02H	8BH
Total instrument number	Serial number	The number of the transmitting instrument	Instrument 1 address: 7Byt (E2)		Instrument 2 address:
03H	01H	03H	01H 00H 00H 00H 00H 11H 11H		02H 00H 00H 00H 00H 11H 11H
Instrument 3 address:					
03H 00H 00H 00H 00H 11H 11H					

The responses are

1. The current time: 09 15 years 16 15 points 0 minutes 2 seconds. The main content of the message returned by the server.
2. The address of the instrument: the low level in the front, the high position in the post, BCD encoding were 1000000001111, 2000000001111, 3000000001111.
3. Analytical results: correct.

After the test of the system, the test results show that the function of the system has been basically realized, the operation interface is simple and convenient, and the system function is perfect.

6. Conclusion

In this paper, we analyze and investigate the main business of remote meter reading system of thermal system, and analyze the information system of thermal

remote meter reading based on GPRS. We introduce the technology of GPRS, JS architecture, MVC mode, and develop the system according to the theory of software engineering, realize the centralized charging, centralized management, accounting, management, and also the function of the management system.

At present, the GPRS based remote meter reading information system is in the primary stage of application. There still exist many problems whose solutions need to be improved:

References

[1] N. Wu, Y. Guo, Y. Wei, A. Wei: *Design of the remote wireless meter reading system based on GPRS.* TELKOMNIKA Indonesian Journal of Electrical Engineering *11* (2013), No. 11, 6358–6366.

[2] Y. Wu, H. Wang: *Application of GPRS and GIS in boiler remote monitoring system.* TELKOMNIKA Indonesian Journal of Electrical Engineering *10* (2012), No. 8, 2159 to 2168.

[3] M. Chang, Q. Wang: *Application of wireless sensor network and GPRS technology in development of remote monitoring system.* TELKOMNIKA Indonesian Journal of Electrical Engineering *13* (2015), No. 1, 151–158.

[4] A. B. Ibrahim, M. N. Husain, A. R. Othman, M. S. Johal: *Low noise amplifier at 58 GHz with cascode and cascaded techniques using T-matching network for wireless applications.* IJ Electrical and Computer Engineering *1* (2011), No. 1, 1–8.

[5] J. Rezazadeh, M. Moradi, A. S. Ismail: *Fundamental metrics for wireless sensor networks localization.* IJ Electrical and Computer Engineering *2* (2012), No. 4, 452 to 455.

[6] Md. M. Ahamed, K. Bhowmik, Md. Shahidulla, Md. S. Islam, Md. A. Rahman: *Rectangular microstrip patch antenna at 2 GHZ on different dielectric constant for pervasive wireless communication.* IJ Electrical and Computer Engineering *2* (2012) No. 3, 417–424.

[7] J. Gomes, B. K. Mishra: *Performance evaluation of UWB wireless link.* IJ Information and Network Security (2012), No. 3, 188–199.

[8] A. Mehadji, H. Badaoui, Z. Berber: *A bow-tie Bluetooth/Wimax antenna design for wireless networks applications.* IJ Information and Network Security *1* (2012), No. 3, 207–215.

[9] F. Zeng, L. Yao, H. Chen: *Impact of topology and traffic on interference and routing in IEEE 80211 wireless mesh network.* TELKOMNIKA Indonesian Journal of Electrical Engineering *10* (2012), No. 4, 798–805.

[10] T. Brooks, J. Robinson, L. McKnight: *Conceptualizing a secure wireless cloud.* IJ Cloud Computing and Services Science *1*, (2012), No. 3, 89–114.

[11] V. Godbole: *Performance analysis of clustering protocol using fuzzy logic for wireless sensor network.* IAES IJ Artificial Intelligence *1* (2012), No. 3, 103–111.

[12] T. Dubey, O. P. Sahu: *Directional antenna assisted scheme to reduce localization error in wireless sensor networks.* International Journal of Information and Network Security *2* (2013), No. 2, 183–189.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 239–248

Environmental monitoring and intelligent irrigation system research[1]

Xiaoling Ding[2,3], Zenghui Zhang[3], Lixin Zhao[3,4], Yibin Li[2], Runguo Zuo[3]

Abstract. A remote irrigation system based on GPRS and LabVIEW is designed, aimed at the requirements of modern precision agriculture and water-saving society. We use K60 as the lower monitor recording the temperature, humidity and other environmental parameters which sends the data by the transmission medium of GPRS. Through the internet network, the remote data are transmitted to the host computer. At the same time, an intelligent water-saving irrigation is realized by introducing the fuzzy control system. It plays an important role in studying wheat growth in further depth. The simulation experiment results show that the system is simple, stable and reliable. The interface is friendly. It meets the purpose and requirement of expected design.

Key words. GPRS, LabVIEW, data acquisition, water saving irrigation, wireless sensor, networks, fuzzy control.

1. Introduction

The shortage of water resource is the basic national condition in China. Although the total amount of water resource is rich, our country has access to less water per capita. As a large agricultural country, agricultural water accounts for a high proportion of the total water consumption. According to statistics, in 2007, agricultural water accounted for more than 60 % of total water consumption. But nearly 6.7 million hectares of farmland do not have water enough for irrigation. Our

[1]The research was supported by a project grant from the Twelve-Five National Science and Technology Plan (2011BAD32B02-05-2), National Public Welfare Industry (meteorological) Research Projects (GYHY201106024-2-2) and Agricultural Machinery Equipment Research and Development Innovation Project of Shandong Province (2015YZ103).

[2]School of Control Science and Engineering, Shandong University, Jinan, 250061, China

[3]Mechanical & Electronic Engineering College, Shandong Agricultural University, Shandong, China Provincial Key Laboratory of Horticultural Machineries and Equipments, Tai'an, 271018, China

[4]Corresponding Author: Lixin Zhao, associate professor and master tutor of Mechanical & Electronic Engineering College Shandong Agricultural University, xlding103@163.com

country is short of 40 billion cubic meters of water every year.

Thus, water-saving irrigation is the strategic measure to realize precision agriculture construction, which has a great significance to safeguard the economic, ecological and social sustainable development of China. Form the 1980s, Israel, United States and other developed countries have been successful in use of remote sensing technology and data communication technology in the forecast of soil moisture and environment monitoring for crop growth [1–2]. In our country, the precise irrigation technology started late and generally uses the method of timing for irrigation. If the soil needs water, it is irrigated in time, but utilization of water resources is very low and waste is serious. Although the extensive irrigation phenomenon has been basically controlled, application of control-irrigation devices is very rare, so that agricultural water-saving technology also needs to be further improved [3–4]. In order to achieve the data remote transmission, real-time acquisition, display and storage, the project introduces environment monitoring and irrigation control using wireless networks by combined application of sensor technology, computer technology and virtual instrument technology [5–6]. Based on the change of soil moisture and crop demand in different periods, the system adopts the manual control solenoid valve and automatic control of two modes of switching to achieve intelligent monitoring irrigation. Data curve satisfy the need of accurate control of a wide range of drip irrigation, which can easily evaluate the change of parameters and collect an amount of information [7–8]. To reduce waste effectively, enhance adaptability of the control system and realize a certain amount of intelligence level, fuzzy control was added to the system.

2. Analysis of the problem

As shown in Fig. 1, the structure of the system is divided into lower monitor data acquisition, wireless data transmission module, host computer monitoring software module and several other parts. Sensor is set up at the monitory point, measuring current traffic data in the field. Transmitting data to the monitoring center using GPRS DTU allows the remote host computer a better real time monitoring and storage managing in the form of database.

3. System principle

3.1. Hardware structure

The system hardware mainly includes the main controller in the field, environmental sensors, solenoid valve, power supply system, wireless transmission module and interior control center. A part of signal is converted to digital signal by the single chip microcomputer (SCM) through the modular conversion circuit and then it is treated and transmitted to the remote monitoring host computer. It associates database and carries out remote control of electromagnetic valve switch to realize intelligent irrigation.

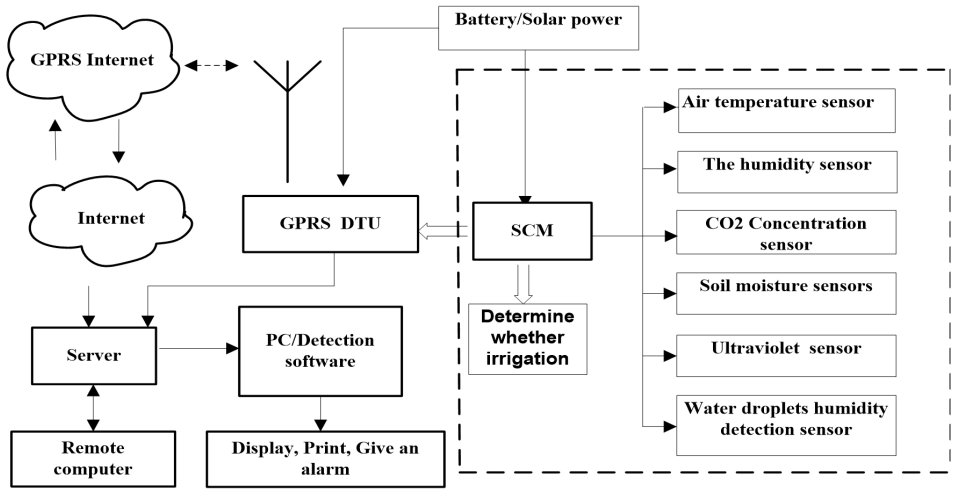

Fig. 1. Overall structure of the system

3.2. Sensor nodes

All of the environmental parameters, after being measured by the environmental sensors distributed in the field, are transmitted to the field main controller which transmit them to the control computing center through GPRS module [9–10]. Field data acquisition is a sensor network, which is made up of thousands of sensor nodes [11]. Each of the sensor nodes consists of various environmental elements of the sensor, node controller, STC15 SCM, and CC1100 wireless RF module. The data collected by CO_2 concentration sensor and soil moisture sensor is transmitted to the AD conversion circuit, then after passing through regulating circuit, filter circuit and amplifying circuit are transformed to the digital signal and transmitted to microprocessor. Other data (collected by the air temperature and humidity sensors and UV sensor) are transformed to digital signals directly, then the signal is stored in microprocessor internal storage and transmitted to gathering node through the CC1100 wireless transmission module.

The gathering node master controller of the system is the K60 SCM of Freescale-TM Semiconductor based on ARM Cortex-M4. It has a low power consumption and contains the mixed signal microcontroller. It is rich in serial interface and CAN bus interface and can collect 16 bit AD collection. The main controller deals with the data transmitted from nodes that is then stored. The processed data is sent to the GPRS module through the TTL turn RS232 module. The main controller judges whether it should open the solenoid valve according to the data that is sent by the GPRS module.

Its wireless transmission module uses the GTM900C wireless module Huawei (a three-frequency band GSM/GPRS wireless module, supporting the standard AT commands and enhancing them), and the highest rate of the GPRS data service can reach 85.6 Kbit. Both data acquisition module and GPRS DTU need the battery

to supply electricity. When battery power is less than 10 %, in order to charge up it in time, an intelligent control system send instructions to control monocrystalline silicon solar panels, and fully charged battery charge can supply the device for 3 months.

To increase the reliability of the system, μC/OS-II operating system is implemented in the SCM. Compared with the traditional SCM, it can also avoid the program run errors in the development work or falling into an infinite loop, makes the debugger easy and uses for the system extremely safe condition [12].

4. Software design

The server uses Windows 7 system to write static IP address and port into the initialization of the GTM900C. The host computer monitoring adopts TCP/IP network protocol in writing the program and performs on the server based on LabView2012.

4.1. System login

The login interface of the system is shown in Fig. 2, containing only input of the correct user name and password. The remote monitoring system mainly includes three parts: monitoring, data curve display and data processing.

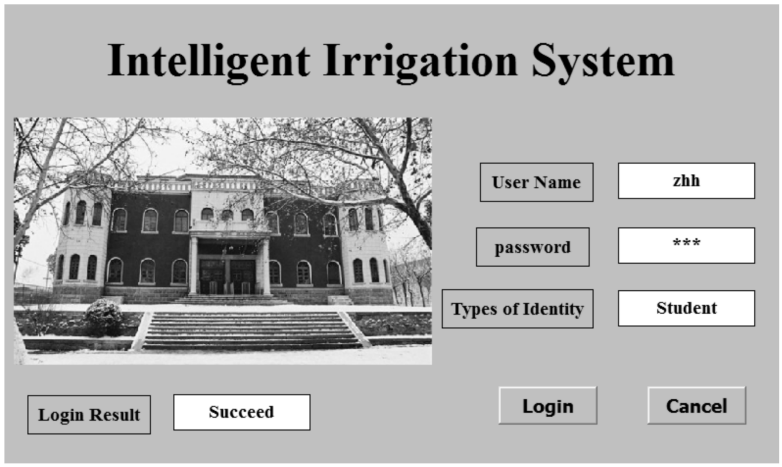

Fig. 2. Login screen

4.2. Design of the monitoring interface

Monitoring interface main sets TCP listening port, the server-side runs to monitor first of all, and then enters the same port with the DTU's configuration. Since DTU is within mobile intranet and has no fixed IP, DTU is assigned randomly to the remote address and remote port, then the address is connected to the IP but port assigning is not the same every time. The section of the alarm setting processes the

information collected and triggers alarm if the value is out of range. The part of solenoid valve control sets two control methods (hand control and self control) and it can be controlled manually or automatically, according to soil moisture changing. Intelligent control device automatically closes the solenoid valve when the ambient humidity is higher than the maximum humidity value that is set. Vice versa, when the humidity is below the minimum humidity value, the solenoid valve is opened automatically to achieve intelligent irrigation.

4.3. Choosing graph curves

Host computer improves the function of selecting curve, which allows viewing selectively a variety of data changes. Dynamic invoking sub VI in main VI can be loaded into the sub VI. The curve to be seen is selected on the left side of the check box in Fig. 3, displayed on the right side of this interface, and then printed or otherwise processed.

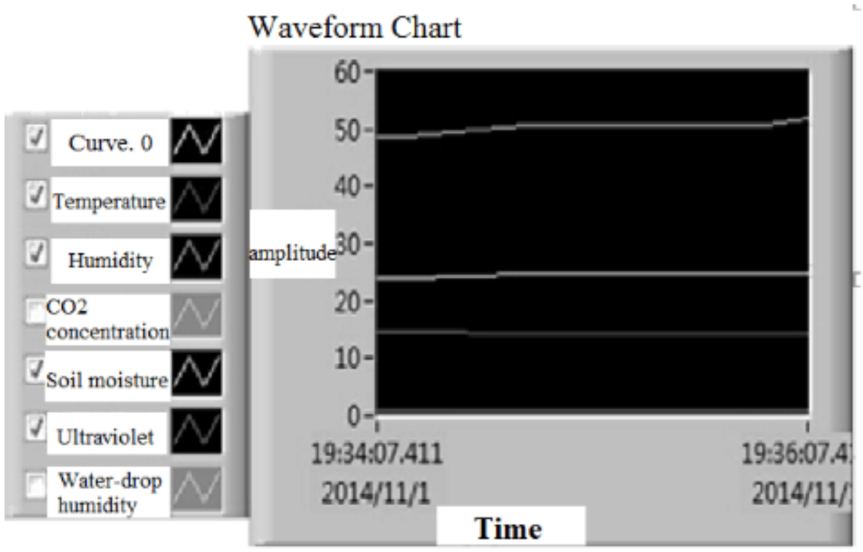

Fig. 3. Data processing interface

For example, checking four options in the check box of Fig. 3 (the curve 0 carried by the system is axis x), the curve can be shown on the side of the screen from top to bottom taking turns air humidity, soil moisture, temperature and ultraviolet (UV) curve (close to axis x), realizing the function curve of choice.

4.4. Remote monitoring realization

The research is carried out with the Widows 7 system as a server connected to VI panel through web browser [13]. VI is open remotely by inputting URL address

at the time of configuration in the address bar. If the remote-client wants to operate the host computer, the permission by server must be obtained to gain the control, in order to avoid operation without permission. Therefore, the system can realize local or remote client computer controlling the system environment.

5. System control strategy

5.1. Strategy selection

It is difficult to describe the automatic control process of irrigation system with precise mathematical expressions due to its time delaying and nonlinear process. To solve the complex and difficult problem accurately, the intelligence control irrigation system joins the fuzzy control and emulates the control system by MATLAB. Using the look-up table method, the system adopts two-dimensional fuzzy control structure of double input (soil moisture deviation E, EC deviation rate) and single output (irrigation time U). The variety of the choice of parameters of input and output variables is shown in Table 1 (NB: Negative Big; NM: Negative Medium; NS: Negative Small; ZO: Zero; PS: Positive Small; PM: Positive Medium; PB: Positive Big). The basic structure of fuzzy controller is shown in Fig. 4.

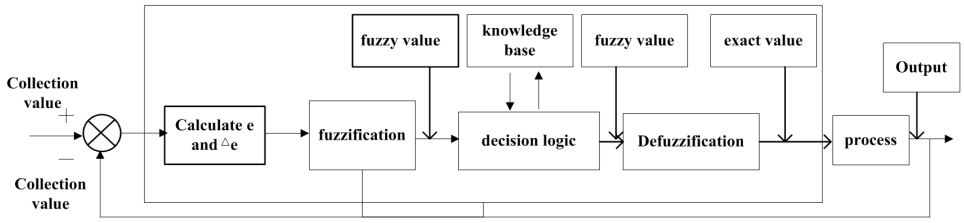

Fig. 4. Fuzzy control chart

5.2. Fuzzy rules

The fuzzy decision adopts Mamdani decision-making model and then uses maximum membership degree method to solve the fuzzy. Both EC and E choose 7 language values, so that there are 49 fuzzy inference rules in the system. According to the summary of agricultural practice in the long term for water-saving irrigation and experimental result of the research about drip irrigation, we can get the fuzzy control rule which is showed in Table 2.

5.3. Simulation experiment and analysis

In order to verify rationality and validity of the fuzzy control scheme, this design uses the simulation software Simulink (which belongs to the MATLAB) for the greenhouse water-saving irrigation system to make a simulated test. As shown in

Fig. 5, the system simulation model is composed of the step input signal, proportion, theory of domain range, fuzzy controller, zero step, function, delay module and display module.

Table 1. Various parameters of input and output variables

Input and output variables	The basic theory of domain	Language input and output variables	Fuzzy set theory domain	Membership function
E	–10 %–10 %	NB, NM, NS, ZO, PS, PM, PB	–3–3	Triangle Gaussian
EC	–2 %–2 %	NB, NM, NS, ZO, PS, PM, PB	–3–3	Gaussian
U	0–80 min	ZO, PS, PM, PB	0–3	Triangle

Table 2. Fuzzy control rule

U EC E	NB	NM	NS	Z	PS	PM	PB
NB	PB	PB	PB	PM	PM	PS	ZO
NM	PB	PB	PM	PM	PS	ZO	ZO
NS	PM	PM	PS	PS	ZO	ZO	ZO
Z	PM	PM	PS	ZO	ZO	ZO	ZO
PS	PS	PS	ZO	ZO	ZO	ZO	ZO
PM	ZO	ZO	ZO	ZO	ZO	ZO	ZO
PB	ZO	ZO	ZO	ZO	ZO	ZO	ZO

According to agricultural experts' experience and actual system, the relation function of soil water variation and irrigation time are showed in the following formula

$$\Delta y = 100 \sin\left(\frac{\pi \cdot t}{900}\right) . \tag{1}$$

Analyzing the output of fuzzy controller, we can set different values of E and EC to prove the validity of the system. When E = 3, EC = 3, and U = 2.68, the soil is very dry and dries quickly, so the irrigation time should be long, and this is consistent with output results of rules. As shown in Fig. 8, bottom part, when E = 3, EC = 3 and U = 2, the soil is a little dry, dries quickly, and the irrigation time is shorter than above, which also agrees to the output of rules.

The simulation experiment results analysis: when setting goal soil humidity at 25 %, the phase signal inputted to simulation model should be 25. It is the curve

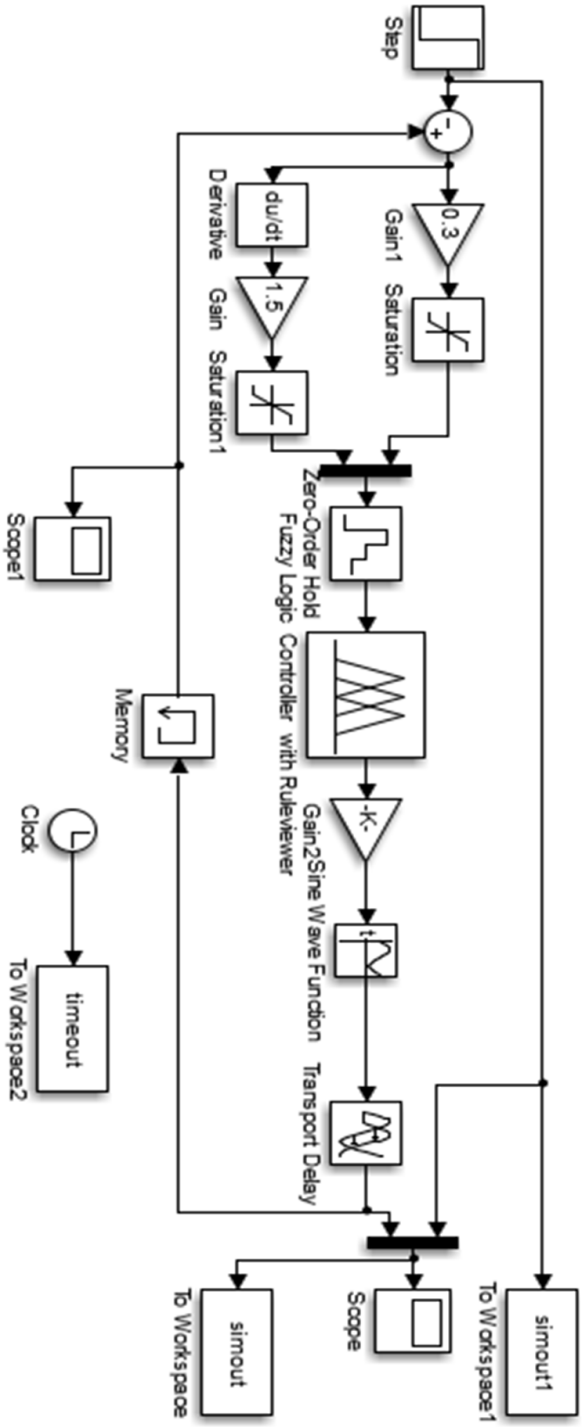

Fig. 5. Simulink model of the irrigation system

obtained by simulation. Figure 6 shows soil humidity simulation graph, where axis x is time in s and axis y is soil humidity in %. From the simulation graph, we can see the error range of simulated soil humidity is stabilized between the 2 % up and down.

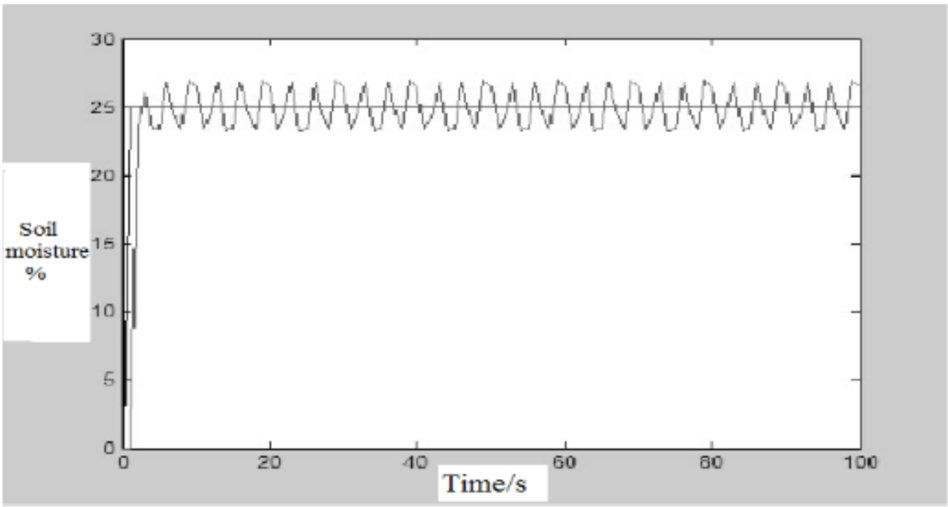

Fig. 6. Output of the rules

By comparing the actual soil humidity and simulation results, we can see that the error is almost the same. The simulated humidity is less than 2 %, and that can achieve the requirement of irrigation system. Obviously, the systematic accuracy can completely satisfy the need of field crops irrigation, and the stability of the system control is good.

6. Conclusion

The research was aimed at designing data management and remote intelligent control irrigation system based on GPRS, LabVIEW and fuzzy control. The system has the advantages of low cost, long transmission distance, fast speed and strong reliability. It is going to play a promoting role for the construction of modern precision agriculture and water-saving society. The results of simulation experiments show that the system can control electromagnetic opening and closing valves automatically, according to preset humidity range, meets the demand of field irrigation, and achieves goals and requirements of the desired design.

1. Completes the data acquisition and processing, and increases two option modes of manual and automatic control, that can intelligently irrigate according to actual humidity,

2. Increases the data curve comprehensive print function; adds the fuzzy control and is verified by simulation of MATLAB, and

3. Can realize monitoring for remote environments in the network, as well as intelligent control for remote water-saving.

References

[1] Y. C. Kuang, J. N. Duan, B. S. Yao: *Application of fuzzy control to automatic water-saving irrigation of rice.* Trans. J Agricultural Mechanization Research *27* (2011), No. 4, 18–21.

[2] J. Casademont, E. Lopez-Aguilera, J. Paradells, A. Rojas, A. Calveras, F. Barceló, J. Cotrina: *Wireless technology applied to GIS.* Computers & Geosciences *30* (2004), No. 6, 671–682.

[3] B. Q. Feng: *Analysis on variation characteristics of effective utilization coefficient of irrigation water and its affecting factors in different subarea of China.* Water-Saving Irrigation (2013), No. 6, 29–32.

[4] Y. X. Liu, T. S. Hong, X. J. Yue, X. Xu, Y. F. Wang: *Design and experiment of drip irrigation control device with low-power solar energy.* Trans. Chinese Society of Agricultural Engineering *28* (2012), No. 20, 206–26.

[5] G. L. Li, X. D. Li, Q. X. Zeng: *Development of automatic irrigation and soil moisture monitoring system based on solar energy in citrus orchard.* J Agricultural Mechanization Research (2009) *28* (2012), No. 12, 146–152.

[6] J. W. Ji, H. Ding, Z. M. Li, Y. Y. Zhao: *Monitoring system of irrigation for paddy fields based on wireless transmission.* J Agricultural Mechanization Research *29* (2013), Suppl. 1, 52–59.

[7] X. J. Fang, Y. M. Zhou, W. L. Cheng, X. H. Ding, X. L. Yang: *The design of wireless intelligent irrigation system based on ZigBee technology.* J Agricultural Mechanization Research (2009), No. 1, 114–118.

[8] S. P. Pei, B. R. Wu: *Soil moisture monitoring and irrigation systems based on Internet of things.* J Agricultural Mechanization Research (2013), No. 7, 106–109.

[9] J. Qiao, C. Wang, X. Wang, W. D. Zhuang, H. Yang: *Greenhouse environmental data acquisition system based on GSM wireless transmission.* J Agricultural Mechanization Research (2008), No. 9, 171–174.

[10] X. B. Xie, B. Wang, J. W. Li, X. Wang, W. J. Yang: *Field irrigation automatic control system design based on ZigBee.* J Agricultural Mechanization Research (2014), No. 9, 89–93.

[11] W. H. Ma, H. R. Wu, X. Sun, F. F. Li: *Wireless transmission based intelligent monitoring and alarm system for greenhouse environment.* J Agricultural Mechanization Research *30* (2014), No. 11, 188–194.

[12] J. J. Labrosse: *Embedded real-time operating system MC/OS-II.* Beijing University of Aeronautics and Astronautics Press, Beijing, 2007.

[13] Y. F. Liu, R. F. Cheng, Q. C. Yang: *Aeroponics control system for tomato cultivation in greenhouse based on LabVIEW.* J Agricultural Mechanization Research (2015), No. 1, 90–95.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 249–260

Study of the crack extension about splitting grouting based on fracture mechanics[1]

Teng Wang[2,3], Mingru Zhou[2,3], Yanmei Ding[4], Guowen Lu[2,3], Guangkang Zhou[2,3]

Abstract. The extended geometric shape of soil crack is an important factor in the strengthening design of split grouting, but the occurrence and development of soil split grouting is difficult to be visualized. This paper studies the cracks form of splitting grouting, establishment of the failure criterion and determination method of crack from the point of fracture mechanics. In order to establish the finite element model of crack extension we employed ABAQUS software together with engineering parameters of soil reinforcement from Lanzhou New Area fast road project K1 + 662.053. We also analyzed the crack length and effect of split grouting pressure and the impact of soil physical and mechanical properties. This research report demonstrates that when calculating split grouting pressure we should take the impact of crack extension length and soil strength parameters into major consideration. Crack length is affected dramatically by grouting pressure in which it shrinks non-linearly as the crack length increases.

Key words. Loess, split grouting, crack length, fracture mechanics, extension (propagation) behavior, effective paths.

1. Introduction

Grouting technology is a very professional academic branch of geotechnical engineering which has a good effect in improving the engineering geological conditions. Fracturing grouting is the exploitation of slurry pressure to promote the expansion of soil cracks, which changes the original structure of soil. And through the filling of the slurry it should also eliminate the soil of the overhead space, eliminate the collapsibility of soil, and, finally, to achieve the purpose of reinforcement [1]–[3]. Due to

[1]This work was carried out in the frame of Science and technology project of Gansu Province (JK2014-34).

[2]Western Engineering Research Center of Disaster Mitigation in Civil Engineering of Ministry of Education, Lanzhou University of Technology, Lanzhou, 730050, China

[3]Lanzhou University of Technology, Lanzhou 730050, China

[4]Architectural and Civil Engineering Institute, Weifang University of Science and Technology, Weifang 262700, China

the conspicuous soil thickness in the Loess Plateau region, utilizing grouting method to reinforce the building of settlements has become an effective mean of technology. However, most of the engineering design is blindness, and the selection of grouting parameters often depends on subjective experience. In addition, the research on the expansion behavior theory of the slurry in the soil body is very slow during the grouting.

It is very necessary to study the fracturing grouting crack propagation distance, which has a close connection with the effect of grouting reinforcement. The hydraulic fracturing process usually refers to the soil or rock mass that they produce crevice under high liquid pressure [4]. This theory is not only the theoretical basis of treatment engineering seepage control, but it is also the key method to solve the underground grouting engineering. The stress intensity factor is an important parameter to characterize the fracture of material, and it is one of the most important factors for solving various kinds of cracks of fracture mechanics [5]. In order to develop a calculation method for the fractured rock mass fracturing grouting pressure, Zou et al. adapted a strength criterion of nonlinear Hoek-Brown in fracture mechanics, and established the type II and mixed mode cracks of fracturing grouting pressure calculation method in the fractured rock mass, which introduces the hydraulic fracturing theory and fracture mechanics method into the theory of fracturing grouting [4]. Aiming at the plane strain of tunnel, plane model study of elastic stage of composite soil after split grouting was studied; on the basis of the study of the existing split diffusion model and area equivalent principle, a two dimensional simplified equivalent unit model of the composite soil after split grouting was put forward. Based on the theory of homogenization, the equivalent elastic parameters of the two dimensional simplified model were derived by the principle of deformation coordination [3]. Based on the Bingham constitutive model, a single plate crack slurry transport equation was established under the condition of constant pressure grouting [10]. However, most of studies were based on special rock conditions, and it was difficult to better explain the strengthening mechanism of split grouting in collapsible loess area of large thickness.

To research the criterion and determination method, a finite element model of crack extension was established by using ABAQUS software and combining Lanzhou New Area fast road project K1 + 662.053 engineering parameters of soil reinforcement. Some aspects of soil were also analyzed, such as the crack length, effect of split grouting pressure, and impact of physical and mechanical properties. It provides guidance and references for the engineering practice of split grouting in the large thickness of the west of China.

2. Fracturing mechanics analysis of fracturing pressure

2.1. Griffith crack assumption

The substance of fracturing grouting is to use high pressure grouting process, based on injecting cement or chemical agent. Then the grout puts additional compressive stress to the surrounding soil and the soil produces shear fracture and frac-

ture failure. According to the theory of Griffith crack [5], fracture is not the two part along the interface of crystal breaking, but the result of the soil mass crack extension under the pressure grout in the process of grouting soil.

In this paper, the substance of the loess in the process of fracturing grouting cracks under the uniform internal pressure was simplified for the elliptical hole process Griffith crack under the uniform internal pressure. As shown in Fig. 1, we assume that the soil in the distance is not subjected to force, and affects the distributed pressure p on the elliptical hole L evenly. Then, according to elastic theory, the boundary conditions are as follows [5]:

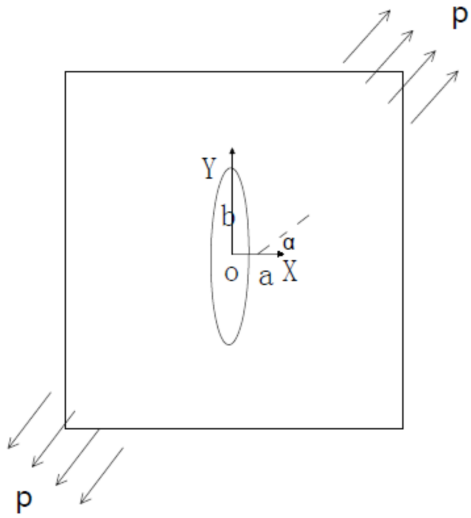

Fig. 1. Stress distribution in fracture model

$$\text{for } \sqrt{x^2+y^2} \to \infty : \ \sigma_{ij} = 0\,. \tag{1}$$

For $(x+y) \in L$ there holds

$$\left.\begin{aligned} \sigma_{xx}\cos(n,x) + \sigma_{xy}\cos(n,y) &= -p\cos(n,x)\,, \\ \sigma_{xy}\cos(n,x) + \sigma_{yy}\cos(n,y) &= -p\cos(n,y)\,. \end{aligned}\right\} \tag{2}$$

Here, $\cos(n,x)$ and $\cos(n,y)$ express any point outward normal direction of direction cosine. Symbol n expresses the the unit outward normal to the point.

According to fracture mechanics theory, the hole side stress can be obtained using the formula

$$\begin{aligned} (\sigma_{\varphi\varphi})_{\rho=1} &= p\tfrac{1-3m^2+2m\cos 2\varphi}{1+m^2-2m\cos 2\varphi}\,, \\ (\sigma_{\rho\rho})_{\rho=1} &= -p\,, \end{aligned} \tag{3}$$

where $m = (a-b)/(a+b)$.

The maximum normal stress is

$$\left.\sigma_{\max} = (\sigma_{\varphi\varphi})_{\substack{\rho=1\\ \varphi=0,\pi}} = (\sigma_{yy})_{\substack{x=\pm a\\ y=0}} = p\frac{1+3m}{1-m} = p(-1+2\frac{a}{b})\right\}. \tag{4}$$

Comparing the formula (3) and (4), the radius of curvature $\rho_0 \to 0$ at the top of the elliptic semi-major axis, internal pressure of elliptical hole and stretching of maximum normal stress of elliptic hole are almost the same. At the same time, the elliptic hole is transformed into a Griffith crack when $b = 0$.

2.2. Establishment of fracturing crack failure criteria

Stress distribution within the crack is composed of two parts when fracturing grouting, namely the original ground stress and pore fluid pressure. As stipulated in fracture mechanics, tensile stress is positive and pressure stress is negative, while the plus-minus sign is just contrary in soil mechanics [4]. To simplify the calculation, this paper considers the compressive stress positive. Then, the soil mass crack fracture mechanics analysis model diagram is shown in Fig. 2. Assuming that the soil is homogeneous and isotropic, according to the linear elastic theory of semi-infinite medium, the shear failure criterion satisfies Mohr–Coulomb criterion.

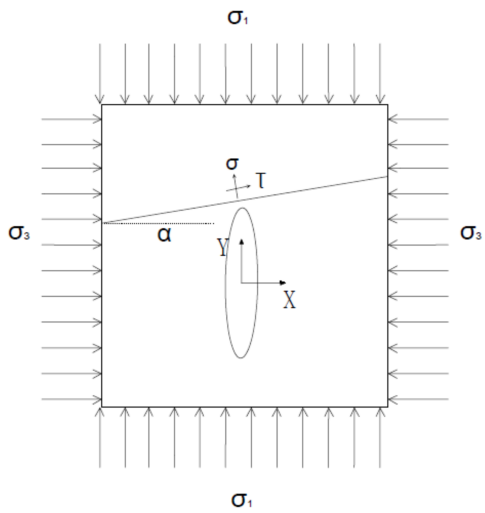

Fig. 2. Soil mass crack fracture mechanics analysis model diagram

The criterion expression is

$$\tau = C + \sigma_{\mathrm{n}} \tan\varphi\,, \tag{5}$$

$$\tau = \frac{\sigma_1 - \sigma_3}{2} \sin 2\alpha \,, \tag{6}$$

$$\sigma_{\mathrm{n}} = \frac{\sigma_1 + \sigma_3}{2} + \frac{\sigma_1 - \sigma_3}{2} \cos 2\alpha \,. \tag{7}$$

The symbols τ and σ_{n} represent the shear stress and normal stress on the surface of the shear failure, respectively; symbol α shows the shear failure plane method and the maximum principal stress direction angle, φ and C show the internal friction angle and cohesive force of soil, respectively.

According to the principle of Terzaghi effective stress, the pore pressure of soil shear failure occurs and can be expressed as

$$p_1 = \frac{\sigma_1 + \sigma_3}{2} + \frac{\sigma_1 - \sigma_3}{2} \cos 2\sigma - \frac{\left(\frac{\sigma 1 - \sigma_3}{2} \sin 2\alpha\right)}{\tan \varphi} \,. \tag{8}$$

In accordance with the above analysis, the serous fracturing surface usually occurs in the stress on the surface of least resistance. So we can see that once the normal stress on the surface of the shear fracture is equal to zero, cracks begin to expand, and the pore pressure is

$$p_2 = \frac{\sigma_1 + \sigma_3}{2} + \frac{\sigma_1 - \sigma_3}{2} \cos 2\sigma \,. \tag{9}$$

2.3. Fracturing grouting discriminant methods of crack extension

When fracturing grouting loess conforms to the rule of fracture mechanics of the propagation of the crack (namely during crack propagation), the end of each point stress intensity factor is equal to the critical stress intensity factor of the loess [5]. Fracture criterion of crack propagation is then as follows

$$K_{\mathrm{I}} = K_{\mathrm{IC}} \,. \tag{10}$$

The symbol K_{I} is the crack extension criterion which is the stress intensity factor of crack, while K_{IC} is the fracture toughness. Among them:

$$K_{\mathrm{I}} = \frac{p\sqrt{\pi}}{E(k)} \left(\frac{b}{a}\right)^{\frac{1}{2}} \left(a^2 \sin^2 \varphi + b^2 \cos^2 \varphi\right)^{\frac{1}{4}} \,. \tag{11}$$

The formula $E(k)$ is the second type of complete elliptic integrals, namely

$$E(k) = \int_0^{\frac{\pi}{2}} \left(1 - k^2 \sin^2 \alpha\right)^{\frac{1}{2}} \, \mathrm{d}\alpha \,, \tag{12}$$

where

$$k^2 = \frac{a^2 - b^2}{a^2} \,. \tag{13}$$

The length and width of the crack are equal when assuming that the crack front is infinitely small, namely $a = b$. The elliptical crack becomes a disc crack at this time and

$$E(k) = \frac{\pi}{2}\,. \tag{14}$$

Then a simplified formula can be obtained and at any point the stress intensity factor of homogeneous internal pressure along the elliptic crack front is

$$K_{\mathrm{I}} = \frac{2p\sqrt{a}}{\sqrt{\pi}}\,. \tag{15}$$

3. Model analysis

3.1. Project summary

The fast road project K1 + 662.053 is a section of engineering in Lanzhou new district, Formations for the top are loess, gravel layer and the bedrock. According to the field condition, the design of grouting depth is 5. m, the undisturbed loess under SEM microscopic structure is as shown in Fig. 3, undisturbed loess mechanics performances are shown in Table 1.

Table 1. Mechanical properties of intact loess

The moisture content ω (%)	Void ratio e	Modulus of elasticity E (GPa)	Poisson ratio μ	Cohesive force C (kPa)	Angle of internal friction φ (°)
20.3	1.10	20	0.3	32	26

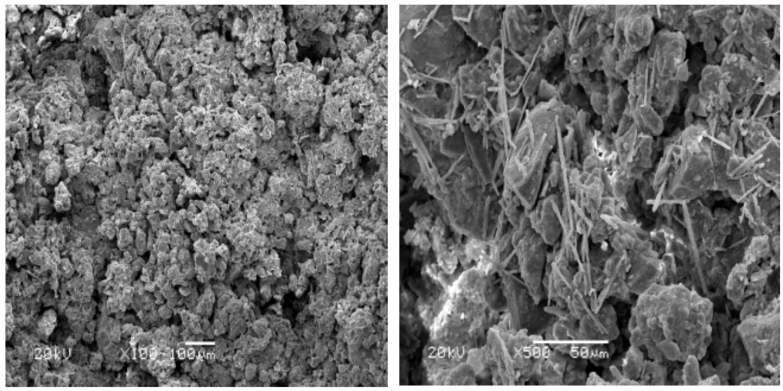

Fig. 3. Microstructure of undisturbed loess SEM: left–100 times magnified, right–500 times magnified

3.2. Model building

The grouting process of crack propagation was simulated by the finite element software ABAQUS. The stress intensity factor of the calculation unit can be calculated by the analysis of the stress and serous pore pressure values of the model, and whether the crack will extend can be evaluated by the criteria proposed in this paper. The model area is shown in Fig. 4.

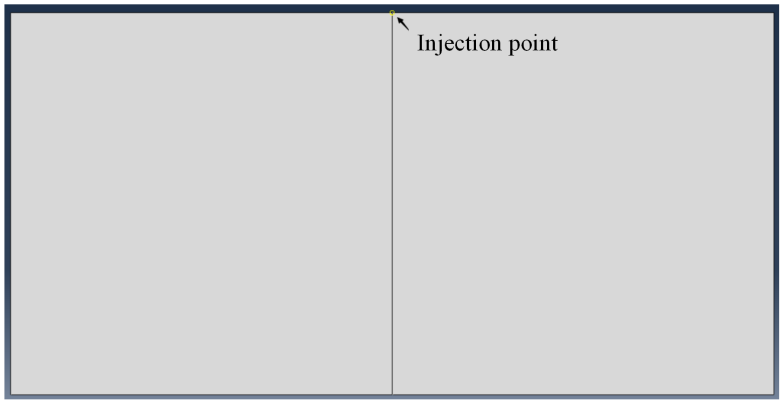

Fig. 4. Crack growth demonstration model

The parameters were selected in the following way: the fracture toughness of the loess is taken $K_{\mathrm{IC}} = 2.7\,\mathrm{MN/m^{3/2}}$ and the fracture energy in the shear mode is 54.7 J. Cohesive unit selection is COH2D4, other materials selection is CPS4R, the field test process can be determined using this method of the stability of the loess soil structure by analysis the calculation results. The analysis of fracture-initiation pressure and fracture morphology belongs one of the key technologies of fracture grouting. According to the changes of stress path in different times, the damage of the element can be analyzed, and the crack propagation in different periods can be obtained. The element boundary taken as a "Path" is shown in Fig. 5, the calculation path being conducted from the bottom up.

3.3. Data analysis

According to the changes of stress path in different times, the damage of the element can be analyzed, and the crack propagation in different periods can be obtained [11–12]. The calculation results of stress path are shown in Fig. 6.

In Fig. 6, we can see that the maximum value of S11 at the lowest end of the path is less than 2.7 MPa when time $t = 0.075\,\mathrm{s}$, and the interface did not appear failure behavior. The S11 at the lowest end of the path achieved the damage value of 2.7 MPa when time $t = 0.1143\,\mathrm{s}$,and the lower end of the unit began started to damage at this time. When time $t = 0.1720\,\mathrm{s}$, the path at the bottom of the interface is declining S11 value of damaging the original point. The value is not zero, which indicates that the damage value D does not reach 1 yet, although destruction was

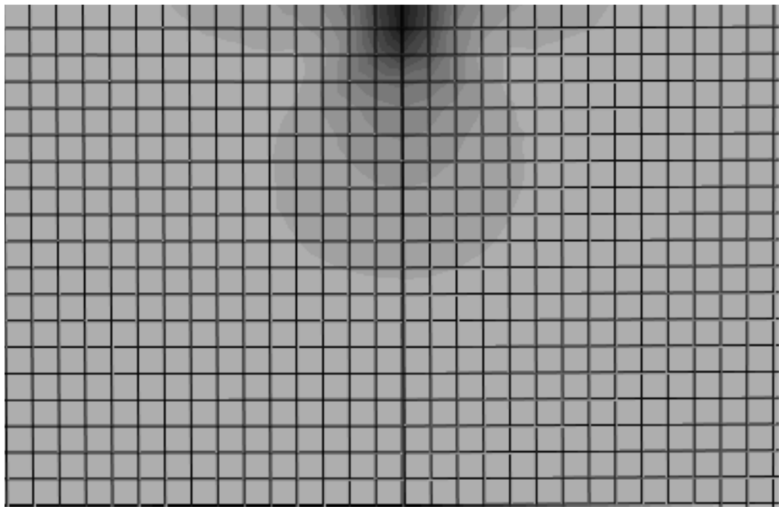

Fig. 5. Calculated interface path diagram

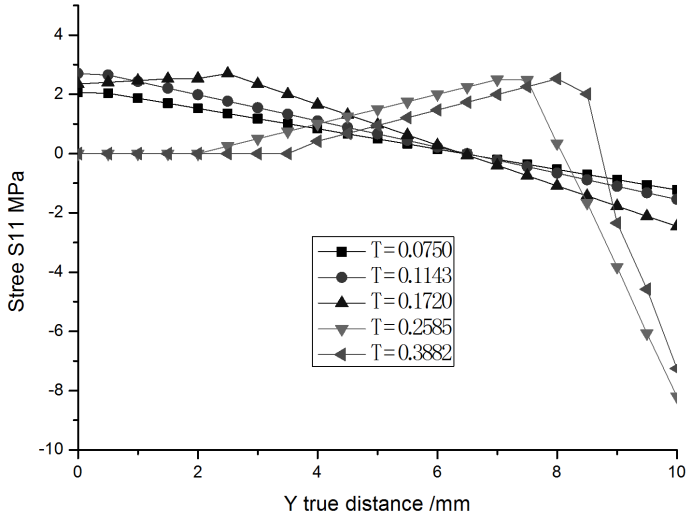

Fig. 6. Path diagram of calculation results

conducted here. The interface crack tip continues to spread along the path. When time $t = 0.2585$ s, one can see several S11 values on the top of the patch and bottom is zero. It is showed that the damage value of these units has reached 1, and it means a complete failure state. When time $t = 0.3882$ s, the failure units continues to increase, and crack continues to propagate along the path.

By analyzing the pressure of different periods, it is found that the damage of the unit is advancing slowly with time and the failure unit is increasing slowly. At the same time, the expansion of the crack also expands slowly with the increase of the failure unit, which shows that the stress intensity factor of the crack is greater than the stress intensity factor of the unit. With the expansion of the crack, the stress intensity of the crack gradually becomes smaller and smaller until it is less than the stress intensity factor of the element.

Cracks of soils are expanded through grout pressure in fracturing grouting to change the original structure of the soils. Overhead inter-spaces of the soils are removed through filling grouts, thus removing collapsibility of the soils and reaching consolidation [9]. The curves of crack length and grouting pressure are as shown in Fig. 7.

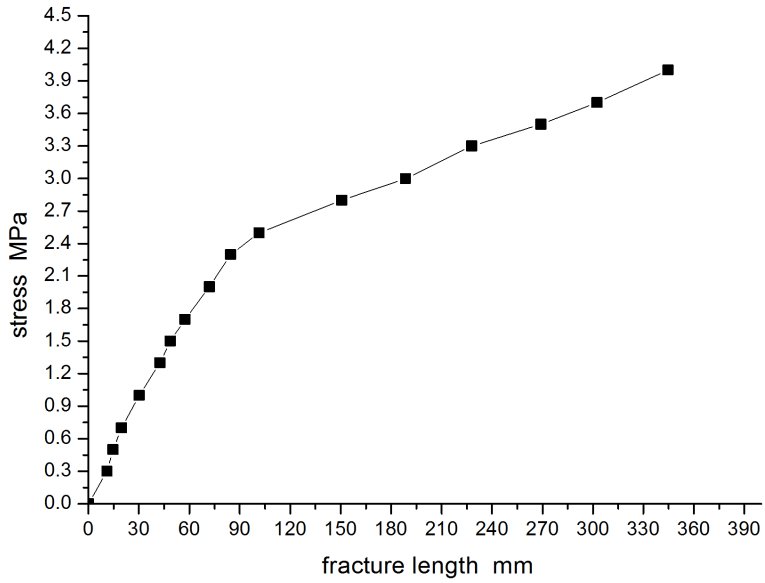

Fig. 7. Curves of crack length and grouting pressure

One can see in Fig. 7 that the crack length is affected by grouting pressure significantly; with the crack length increasing, the grouting pressure presents a nonlinear increase curve. At the same time, the nonlinear extension of the crack length indicates that the structural integrity of the local loess is better; the reason is that only the integral loess soil can form a long crack length. On the contrary, if the crack length changes not obviously with the increase of grouting pressure, it means that the structure integrity of the loess is not completed and may be damaged. In addition, the tension stress and shear stress that the slurry in the extension process in soil needs to overcome, increases with the increase of the crack length. Therefore, from Fig. 8 we can see that grouting pressure increases with the increase of the crack length.

Fracturing grouting cracks extension is not only related to the grouting construction technology, but also to the in situ stress and physical and mechanical properties of soil [6]. As one of two important parameters of the rock (soil) mass, internal friction angle is a shear strength index and an essential parameter of engineering design [7]. Internal friction angle of the soil reflects the frictional characteristics. Influence of internal friction angle of the soil on grouts flowing in the cracks is remarkable during fracture grouting.

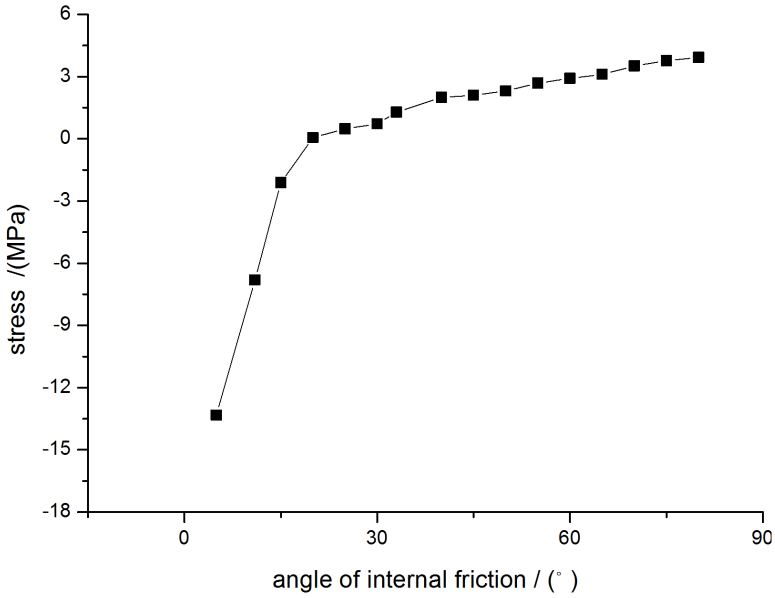

Fig. 8. Relationship between grouting pressure and internal friction angle

Based on literature [13], shear strengths of rock and soil material consist of two parts. However, the cohesion reaches its peak when in small deformation, while the friction can only give full play upon occurring rather large deformation. As a result, the cohesion displays fragility of the rock and soil, while the internal friction demonstrates plastic property. A curve of grouting pressure changing along with the variation of the internal friction angle is shown in Fig. 8. We can see from Fig. that gradually increasing grouting pressure is required with the increasing internal friction angle of the soil. As the destructive process of the rock and soil is just the process of the friction gradually play its role, the larger the internal friction angle of the soil; Grouting pressure shall overcome the strength of the soil, leading to the greater resistance of flowing grouts. Thus, the greater grouting pressure is required.

3.4. Test analysis

In the process of the actual test, fracturing grouting soil mass crack propagation can be analyzed by the change of displacement of soil. Because the test zone of

fracture form is more complex, it is difficult to evaluate the results by statistics tools. So we selected the test section which is more representative from a few viewpoints. The measured crack extension data are shown in Table 2.

Table 2. The test results of crack propagation

Number	Soil depth/m	Grouting pressure/MPa	Fracture length/m	Delay on highly/m	Decurrently height/m
Test 1#	4	0.9	0.31	0.031	0.029
Test 2#	4	0.9	0.33	0.027	0.025
Test 3#	4	0.9	0.32	0.030	0.026
Test 4#	4	0.7	0.25	0.024	0.021
Test 5#	4	0.7	0.24	0.022	0.025
Test 6#	4	0.7	0.25	0.021	0.020

According to the test of 1–6 reinforced soil monitoring data (Table 2), after the splitting, crack propagation in 0.24 m–0.32 m decreases with depth. For the soil under the different pressure of fracture, simulation results and experimental data are basically identical. Meanwhile, under low pressure, the diffusion distance is smaller, but near the injection pipe, the soil displacement is larger. Therefore, we can deduce that under low pressure, slurry long time in a bubble form, to compaction of soils, soil under the high pressure is given priority to with splitting, slurry flow far away. And fracture general trend of present along the split face next fracturing, rarely possible secondary splitting.

In Table 2, the soil mass crack propagation length is closely related to the ductility ratio and soil mass stress state, and the ratio measurement with the soil circle size a linear relationship between principal stress ratios. It also verifies the splitting cracks in the smallest little the conclusion of extending direction of principal stress plane

Microstructure analysis (see Fig. 3) found that soil structure differs from between space distributions. Lead to split face uncertainty guide, make each time grouting grout to different, cause of fracture of surrounding soils in the local discreteness. It also demonstrates the soil particle cementing with structural, and the expansion of crack is weak cementation with fracture.

In this paper, the studied object is loess medium. The grouting crack simulation results of tested soil layer under different pressure are basically consistent. The results show that the crack propagation of split grouting is consistent with the Griffith crack theory in fracture mechanics.

4. Conclusion

1. Based on the fracture mechanics, the crack forms, failure criterion and determination method of the crack expansion in splitting grouting process were studied and crack propagation model in Loess Area was established.

2. Through the model analysis, it is concluded that the crack growth is controlled by the fracture factor. The stress intensity factor of the computational element can be calculated according to the crustal stress and pore pressure, and through the decision criterion it is possible to evaluate whether the crack is extended.

3. The crack expansion of split grouting is not only related to the grouting construction technology, but also related to the crustal stress and the physical and mechanical properties of soil. The effect of crack length by grouting pressure is particularly significant, also the regularity of crack length illustrates the stability of the soil structure. Using this method can determine the stability of loess soil structure in the field tests, and provides a theoretical guidance for the follow-up field reinforcement design.

References

[1] X. Zhang: *Study on mechanism of slurry diffusion and sealing at the process of underground engineering moving water grouting and its application (Ph.D. Thesis)*. Shandong University, Jinan, China 2011.

[2] A. Bezuijen, R. T. Grotenhuis, A. F. Van Tol, J. W. Bosch, J. K. Haasnoot: *Analytical model for fracture grouting in sand*. J Geotechnical and Geoenvironmental Engineering *137* (2011), No. 6, 611–620.

[3] Y. W. Guo, S. H. He, X. M. Guan, X. B. Liu: *Theoretical study of plane equivalent elastic model of composite soils with fracturing grouting*. Rock and Soil Mechanics A *36* (2015), No. 8, 193–201.

[4] J. F. Zou, W. Tong, H. Luo, X. F. Wang: *Mechanism of fracture grouting for fractured rock based on Hoek-Brown failure criterion*. J Central South University (Science and Technology) *44* (2013), No. 7, 2889–2896.

[5] T. Y. Fan: *Theoretical basis of fracture*. Science Press, Beijing, China, 2001.

[6] Q. S. Zhang, L. Z. Zhang, R. T. Liu *Laboratory experimental study of cement-silicate slurry diffusion law of crack grouting with dynamic water*. Rock and Soil Mechanics (2015), No. 8, 2159–2168.

[7] P. Tu, X. H. Wang: *Study on experiment of subsea tunnel grouting materials*. J Railway Science and Engineering (2010), No. 6, 60–64.

[8] J. F. Kou, F. Xu, J. P. Guo, Q. Xu: *Damage laws of cohesive zone model and selection of the parameters*. J Mechanical Strength *33* (2011), No. 5, 714–718.

[9] Q. Wang, R. Zhang: *Experimental research on the trend of slurry about fracturing grouting and soil displacement in different pressure*. J China Railway Society *33* (2011), No. 12, 107–111.

[10] G. Gustafson, J. Claesson, Å. Fransson: *Steering parameters for rock grouting*. J Applied Mathematics *2013*, paper ID 269594.

[11] F. Sun, D. L. Zhang, T. L. Chen, X. P. Zhang: *Meso-mechanical simulation of fracture grouting in soil*. Chinese Journal of Geotechnical Engineering *32* (2010), No. 3, 474–480.

[12] J. Zhou, G. Zhang, G. Kong: *Meso-mechanics simulation of seepage with particle flow code*. J Hydraulic Engineering *37* (2006), No. 3, 28–32.

[13] Z. J. Shen: *Breakage mechanics for geological materials: an ideal brittle-elasto-plastic model*. Chinese Journal of Geotechnical Engineering *25* (2003), No. 3, 253–257.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 261–270

A loading balance model of virtual machine live migration in cloud computing environment[1]

Xin Sui[2,3], Li Li[2,4], Dan Liu[2], Hong Wei Yang[3], Xu Di[2]

Abstract. Virtual machine living migration technique provided a method for load balancing in cloud computing environments. In order to reduce the SLAV rate and improve server utilization rate, the virtual machines of overload server were migrated to other server and integrated low load server. The model proposed in this paper put the number of over-lowload servers, migration cost of virtual machine and power consumption of servers as the main target. In order to achieve these objectives, the paper defined the power consumption function of the servers, the migration cost function of virtual machines and the SLAV rate function, putting forward an ELBMLM model, and genetic algorithm was applied in the model. The experiment results showed that the proposed model was better than the others.

Key words. Cloud computing, virtual machine migration, load balancing, energy consumption.

1. Introduction

As a service available to the user, the stability of service resource and utilization rate is the key of influencing the intuitive feeling of users and economic benefits of service provider. So how to meet the needs of users and effectively manage the service resources becomes the key to the development of cloud computing [1].

[1]This work was supported in part by the project of "12th Five-Year" planning of the Education Science of Jilin Province (GH14049), the science and technology project of "13th Five-Year" planning of the Education Department of Jilin Province (2016), the key project of "13th Five-Year" planning of the Education Science of Jilin Province (ZD16024), and Key Science and Technology Project of Jilin Province (20160204019GX).

[2]Changchun University of Science and Technology, College of Computer Science and Technology, 130022, China

[3]Jilin Provincial Institute of Education, 130022, China

[4]Corresponding Author, ll@cust.edu.cn

Verma proposed an algorithm to solve the problems of virtual machine se lection (MADLVF), which would select the appropriate virtual machine from the overloaded host, migrate it to other server and put the load of the entire data center to tend into balance [2]. Jiao Zhang studied how much bandwidth was required to ensure the migration time and downtime during the migration of virtual machines [3]. In order to guarantee the service quality and reduce the cost, Mohamed Mohamed proposed an autonomic management model, which could effectively control the cloud resources, see [4].

2. Methodology

2.1. Problem formulation

2.1.1. Power consumption. From the document [5], the power consumption value of the physical server had a relationship with CPU usage rate at a certain time, which was presented as an approximate linear relationship. So the power consumption of the physical server could be calculated according to the CPU usage rate. The formula reads

$$P_i(u) = r_i * P_i^{\max} + (1 - r_i) * P_i^{\max} * u_i \,, \tag{1}$$

where $P_i(u)$ is the total power consumption of the physical server i, constant r_i stands for the power consumption ratio of the physical server i when it is idle and peak, $P_i^{\max}$ represents the maximum power consumption by fully utilized server, and u_i denotes the CPU usage rate of the physical server i.

The total power consumption of the physical server i between times t_1 and t_n is then given as

$$E_i = \sum_{t_1}^{t_n} P(u_i(t_j)) \,, \tag{2}$$

where $u_i(t_j)$ is the CPU usage rate of the server i at time t_j and $P(u_i(t_j))$ is the power consumption of the server i at time t_j.

The document research showed that the energy consumption of the network equipment was only related to the configuration of network devices. The calculation formula of network equipment energy consumption is shown as follows

$$P(C) = F(C) + A * X \,, \tag{3}$$

where $P(C)$ stands for the power consumption of the network equipment, C is the configuration parameters of the network equipment, $F(C)$ is the sum of standard and network line card energy consumption of network equipment. Symbol A denotes the power consumption of the network interface and X represents the number of the network interfaces.

The energy consumption of cloud data center is

$$E = \sum_{1}^{N} E_i + \sum_{t_1}^{t_n} P(C)\,, \tag{4}$$

where $N_(i)$ is the number of servers.

2.1.2. Live migration cost estimation. The living migration of virtual machine would lead to the performance degradation of application at the pre-copy stage and down time at the copy phase. The migration time of virtual machine is calculated by the formula

$$t_{V_i} = \frac{M_{V_i}}{B_{V_i}} \tag{5}$$

and the performance decline of the application is given by the formula

$$P_{V_i} = a \cdot \sum_{t_0}^{t_0+t_{V_i}} U_{V_i} + b \cdot M'_{V_i} + c \cdot B_{V_i}\,. \tag{6}$$

Here, t_{V_i}, represented the migration time, M_{V_i} denotes the memory size of the virtual machine, B_{V_i} stands for the network bandwidth, P_{V_i} shows the performance decline of the physical server, t_0 is the starting migration time, U_{V_i} denotes the CPU usage rate, a, b and c are the parameters of the performance degradation caused by the CPU utilization, memory occupancy and network band occupancy, and, finally, M'_{V_i} represents the memory occupied by the live migration.

2.1.3. Service-level agreement (SLA) violation. The violation of the SLA percent is given by the formula

$$\text{SLAVP} = \frac{N_{\text{SLAViolation}}}{N_{\text{total}}}\,, \tag{7}$$

where SLAVP denotes the percent of servers that were in violation of the SLA, $N_{\text{SLAViolation}}$ stands for the number of servers in violation of the service level agreement and N_{total} is the total number of servers.

2.2. Live migration algorithm

2.2.1. Load status of server. The status of the ith server may be described as follows

$$\left\{ \begin{array}{lcl} P_i^{\text{o}} = 1 & | & P_i^{\text{f}} = P_i^{l} = 0, U_i > a \\ P_i^{\text{f}} = 1 & | & P_i^{\text{o}} = P_i^{l} = 0, b \le U_i \le a \\ P_i^{l} = 1 & | & P_i^{\text{o}} = P_i^{\text{f}} = 0, U_i < b \end{array} \right\}. \tag{8}$$

In (8), P_i^{o}, P_i^{f}, P_i^{l} express that the ith server is on overload, full load, low load state or not, respectively, and meets the condition $P_i^{\text{o}} + P_i^{\text{f}} + P_i^{l} = 1$. Symbol U_i indicates CPU utilization, and a and b are the CPU utilization bounds of the full

load state.

2.2.2. Virtual machine migration model. The objective functions are

$$F_1 = \min \sum_{i=1}^{n} (P_i^{\mathrm{o}} + P_i^{l}) \,, \tag{9}$$

$$F_2 = \min \sum_{i=1}^{m} P_{V_i} \,, \tag{10}$$

$$F_3 = \min \sum_{i=1}^{n} E_i \,. \tag{11}$$

The constraint conditions of migrating virtual machines are

$$\left\{ \begin{array}{l} \sum_{j=1}^{m} V_j^{\mathrm{cpu}} \cdot P_{i,j} < C_i^{\mathrm{cpu}} \\ \sum_{j=1}^{m} V_j^{\mathrm{mem}} \cdot P_{i,j} < C_i^{\mathrm{mem}} \\ \sum_{j=1}^{m} V_j^{\mathrm{store}} \cdot P_{i,j} < C_i^{\mathrm{store}} \end{array} \right\}, \tag{12}$$

$$\sum_{i=1}^{n} P_{i,j} = 1, P_{i,j} \in \{0,1\} \,. \tag{13}$$

Formula (12) suggests that the sums of CPU, memory and hard disk for all virtual machines on one server, respectively, are smaller than the capacity of the server's CPU, memory and hard disk, where m is the number of virtual machine migration and n is the number of server integration. Symbols V_j^{cpu}, V_j^{mem} and V_j^{store} are, respectively, the need of CPU, memory and hard disk. Similarly, C_i^{cpu}, C_i^{mem} and C_i^{store} are the corresponding capacities. From formula (13), $P_{i,j}$ indicates whether the jth virtual machine is migrated to the ith server or not, and its value is 0 or 1, accordingly.

2.2.3. Virtual machine migration algorithm. In order to solve the virtual machine migration model, a global optimization genetic algorithm was designed and the algorithm included the encoding method, crossover operator, mutation operator, correction operator and local search operator.

- Encoding method: Assuming that the number of virtual machines needed to be migrated in the data center is m, the number of servers needed to be consolidated is n, the mapping of virtual machine and server can be expressed by an $n \times m$ matrix

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \dots & P_{1,m} \\ P_{2,1} & P_{2,2} & \dots & P_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n,1} & P_{n,2} & \dots & P_{n,m} \end{bmatrix} .$$

Initial matrix P is the all zero matrix. From the formula (13), there was only one number on one column of the matrix, which was equal to 1, and the other were all 0.

- Crossover operator: First, four integers a, b, c, d were generated randomly: $a, c \in [1, N]$ and $b, d \in [1, M]$. The cross method then was

$$\begin{aligned} P1 &= (P^1_{1,1}, ..., P^1_{a,b-1} | P^1_{a,b}, ..., P^1_{c,d} | P^1_{c,d+1}, ..., P^1_{n,m}), \\ P2 &= (P^2_{1,1}, ..., P^2_{a,b-1} | P^2_{a,b}, ..., P^2_{c,d} | P^2_{c,d+1}, ..., P^2_{n,m}), \\ P3 &= (P^1_{1,1}, ..., P^1_{a,b-1} | P^2_{a,b}, ..., P^2_{c,d} | P^1_{c,d+1}, ..., P^1_{n,m}), \\ P4 &= (P^2_{1,1}, ..., P^2_{a,b-1} | P^1_{a,b}, ..., P^1_{c,d} | P^2_{c,d+1}, ..., P^2_{n,m}). \end{aligned}$$

- Crossover operator: First, random generation of a real number $a \in [0, 1]$. If $a \leq P_m$, then jump to the following step. Choice of a column of the matrix, and random generation of integer $b \in [1, N]$, $P_{i,j} = 1, b \neq i$, make $i = b$, $P_{b,j} = 1, P_{i,j} = 0$. Repetition of the previous step from the first to the mth column.

- Correction operator: Let

$$\mathrm{CPU}^{\mathrm{e}} = \sum_{j=1}^{m} (V_j^{\mathrm{cpu}} \cdot P_{i,j}) - C_i^{\mathrm{cpu}},$$

$$\mathrm{MEM}^{\mathrm{e}} = \sum_{j=1}^{m} (V_j^{\mathrm{mem}} \cdot P_{i,j}) - C_i^{\mathrm{mem}},$$

$$\mathrm{Store}^{\mathrm{e}} = \sum_{j=1}^{m} (V_j^{\mathrm{store}} \cdot P_{i,j}) - C_i^{\mathrm{store}},$$

if $\mathrm{CPU}^{\mathrm{e}} \leq 0 \&\& \mathrm{Mem}^{\mathrm{e}} \leq 0 \&\& \mathrm{Store}^{\mathrm{e}} \leq 0$ return P'
else select a column in the matrix $P_{i,j} = 1$, randomly generate integer $a \in [1, N]$, $a \neq i$, and make $P_{a,j} = 1, P_{i,j} = 0$ satisfying the formula (12)
endif.
The last step is repeated from the first to the mth column.

- Local search operator: Find the population individual P, which has the largest F_1 value, from the population after correction. Then choose a population, where $P_{i,j} = 1$, randomly generate an integer $a \in [1, N]$, $a \neq i$, make $P_{a,j} = 1, P_{i,j} = 0$, and generate population individual P'.
if $F_1' < F_1$, go to the previous step; else return P', endif.

3. Results

3.1. Experiment parameter setting

The experiment simulated that the cloud data center consisted of 500 servers of different configurations and two types of virtual machines. Among them, 200 servers were HP Proliant ML110 G4 and 300 servers were HP Proliant ML110 G5. The energy consumption values of the two kinds of servers are shown in Table 1. Configuration parameter of the virtual machine are given in Table 2.

Table 1. Energy consumption of two types on different servers

Server	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
G4	86	89.4	92.6	96	99.5	102	106	108	112	114	117
G5	93.7	97	101	105	110	116	121	125	129	133	135

Table 2. Virtual machine parameters

Parameter	VM1	VM2
Number of CPU	1	1
Frequency of CPU	0.5GHz	1GHz
Memory	1GB	2GB
Disk	50GB	50GB

3.2. Experiment comparison results

3.2.1. Number of over-lowload servers. In the simulation process, when the number of requests for virtual machines increased, the number of servers in the over-lowload increased. The experimental result showed that the migration scheme proposed by this model could reduce the number of overloads and low loads to the minimum and the rate of growth to the slowest, which was compared with the other three models. The results are shown in Fig. 1.

3.2.2. Migration cost of virtual machines. Figure 2 shows the migration cost of virtual machines under four models with variable number of virtual machine requests. The experiment result indicates that the migration cost of ELBMLM model increased 6.6 times when the number of requests for virtual machines was increased 10 times. However, as the maximum migration cost model—the FRAware model—increased 7.34 times, and it showed that the proposed scheme could effectively reduce the live migration cost of virtual machines, which the model was compared with other three models, and it could reduce the total energy consumption of cloud data center.

3.2.3. Power consumption of cloud data center. The experiment result reflects the power consumption condition by increasing the number of requests for virtual machines. The total power consumption of the proposed scheme, which was com-

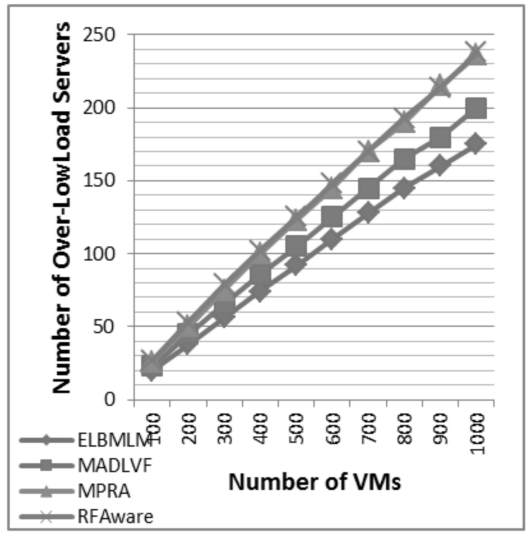

Fig. 1. Number of over-lowload servers

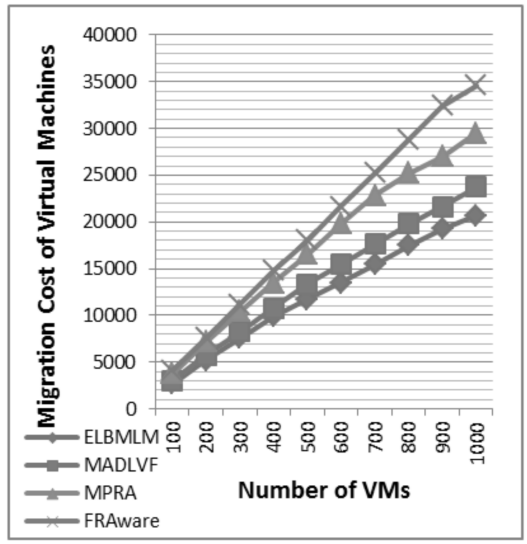

Fig. 2. Migration cost of virtual machines

pared with the other three models', was not the lowest but the second. In order to realize the minimum over-lowload servers' number and the cost of virtual machines' live migration, the power consumption of cloud data center must be influenced. The result is depicted as Fig. 3.

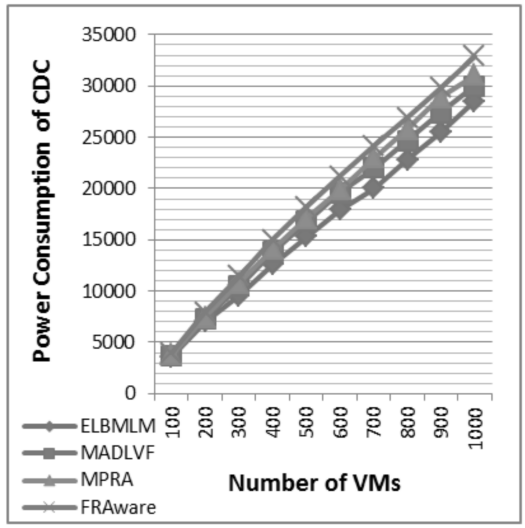

Fig. 3. Power consumption of CDC

3.2.4. Average violation rate of service level agreement. Figure 4 shows the average violation rate of service level agreement to the cloud data center under the variable number of virtual machine requests. The experiment result showed that four models' average violation rate of service level agreement with the increasing requests' number of virtual machine, and the violation rate of this model was among the four models', presenting a relatively stable state and effecting well.

4. Discussion

The proposed model compared with MADLVF, MPRA, and RFAware model in the simulation experiment not only took the cost of power consumption and live migration into consideration but also reduced the number of over-lowload servers and energy consumption of cloud data center to the minima.

It could be seen through the research, that the live migration of virtual machine could effectively regulate load balancing, prevent resource aggregation and finally reduce the energy consumption of data center. During the simulation experiment, this paper used genetic algorithm as optimization algorithm and this algorithm could fix ELBMLM model well.

5. Conclusion

The parameters of computation using the model proposed in the paper were compared with parameters of other three models on the number of over-lowload servers: migration cost of virtual machines, power consumption of servers and SLAV average

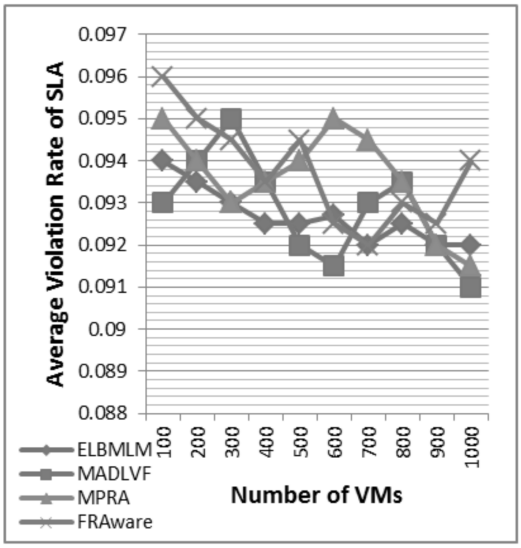

Fig. 4. Average violation rate of SLA

percent of four aspects. The results of experiment show that the proposed model is better than other three models both in the number of over-lowload servers and migration cost of virtual machines, which leads to savings in energy consumption.

References

[1] A. Beloglazov, J. Abawajy, R. Buyya: *Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing.* Future Generation Computer Systems *28* (2012), No. 5, 755–768.

[2] J. K. Verma, C. P. Katti, P. C. Saxena: *MADLVF: An energy efficient resource utilization approach for cloud computing.* IJ Information Technology and Computer Science *6* (2014), No. 7, 56–64.

[3] J. Zhang, F. G. Ren, C. Lin: *Delay guaranteed live migration of virtual machines.* Proc. IEEE Conference on Computer Communications, 27 April–2 May 2014, Toronto, Canada, 574–582.

[4] M. Mohamed, M. Amziani, D. Belaid, S. Tata, T. Melliti: *An autonomic approach to manage elasticity of business processes in the cloud.* Future Generation Computer Systems *50* (2015), 49–61.

[5] G. Chen, W. He, J. Liu, S. Nath, L. Rigas, L. Xiao, F. Zhao: *Energy-aware server provisioning and load dispatching for connection-intensive internet services.* Proc. Symposium on Networked Systems Design and Implementation (USENIX), 16–18 April 2008, San Francisco, California, USA 337–350.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 271–280

A study on the construction of spatial database and its application on regional agricultural planning[1]

Gan Yangying[2], Xiao Guangjiang[2], Liu Xu[2], Su Zhuhua[2], Zhang Jinge[2], Du Yuanyuan[2]

Abstract. The structure of spatial database is introduced. Through an actual case of agricultural planning, the spatial database is constructed by SuperMap. In the end, the application of spatial database on regional agricultural planning are systematically analyzed. Our objective is to provide technical support for multi-source data management and to make the work of regional agricultural planning more scientific, efficient and multi-functional.

Key words. Regional agricultural planning, SuperMap, spatial database, 3D modeling..

1. Introduction

Regional agricultural planning plays an important role in the development of regional economics and beneficial to the improvement of the living standard of local farmers (Fan & Zhu [1], Liu [2]). Traditional planning methods are mostly artificial, paper-based, unilateral and inefficient. Over the past years, a number of innovative approaches were applied in regional agricultural planning, such as the GIS techniques (Petrisor [3], Wu et al. [4], Liu et al. [5]), computer database (Chen et al. [6], Liu et al. [7]) and 3D modeling techniques (Tang et al. [8], Kumar & Kushwaha [9]). Spatial database, by the way of relational database management, can realize the seamless connection, storage, query, management and analysis application of spatial data and attribute data (Song et al. [10], Hu et al. [11]). Therefore, the construction of spatial database of regional agricultural planning can realize the

[1]This work was financially supported by the Guangdong Science and Technology Innovation and Entrepreneurial Talent Service Project (No. 2015A020224013), the Soft Science project of Guangdong Province (No. 2016A070705039), and the President fund of Guangdong Academy of Agricultural Sciences (No. 201533).

[2]Institute of Agricultural Economics and Rural Development, Guangdong Academy of Agricultural Sciences, Guangzhou, 510640, China

uniform management of data based on uniform projection coordinates on maps, and make the comprehensive analysis on agricultural situation, spatial feasibility, spatial layout and ecological environment much easier (Zhou et al. [12], Liu et al. [13]), and provide multi-disciplinary, multi-level, multi-target services and decision assistant for planning compilation, spatial layout and evaluation of implementation management (Guan et al. [14]). In this paper, we took an agricultural park planning as an example, elaborate in detail the process of spatial data acquisition, database construction and its application on the platform of SuperMap. Our research may provide technical support for multi-source data management and to make the work of regional agricultural planning time-saving, efficient and multi-functional.

2. Overview of spatial database

A spatial database refers to the collection of geospatial data in a certain area by the means of scanners, keyboards and other input devices and usually organized and stored according to GIS format (Wang & Gong [15]). It is the foundation and core of GIS spatial analysis, decision-making assisting and system development. The integrity and accuracy of database construction has direct impacts on the realization of GIS analyses and management functions (Peled & Gilichinsky [16]).

Spatial database of regional agricultural planning includes spatial data and attribute data, which usually requires regional topographic maps, remote sensing images, economic and social statistical data, soil data, climate data, and other multi-source data. By using the map editing and property management functions of GIS, multi-source data can be classified and processed to realize the seamless connection of maps and attribute data. Figure 1 shows the flow diagram of spatial database construction. In this paper, we took "Master Plan of Rongjiang County (Guizhou province) Modern Agricultural Demonstration Park (2010–2020)" as an example to elaborate the process of spatial database and attribute database construction and their application on regional agricultural planning.

3. Spatial database construction

3.1. Data acquisition and storage

3.1.1. Data acquisition. Data of spatial database usually includes spatial data and attribute data. While spatial data are often acquired by scanning vectorization, field measurements, aerial survey and remote sensing images etc., attribute data are obtained through field surveys, statistics and historical data etc. In this study, spatial and attribute data were acquired by site surveys, remote sensing images, statistical data and historical data of Rongjiang County (Guizhou Province) modern agricultural demonstration park.

1. Remote sensing images (RGB, resolution: 1 m), December 28, 2015, JPG.

2. Park topographic map (1:1000), March 2015, AutoCAD.

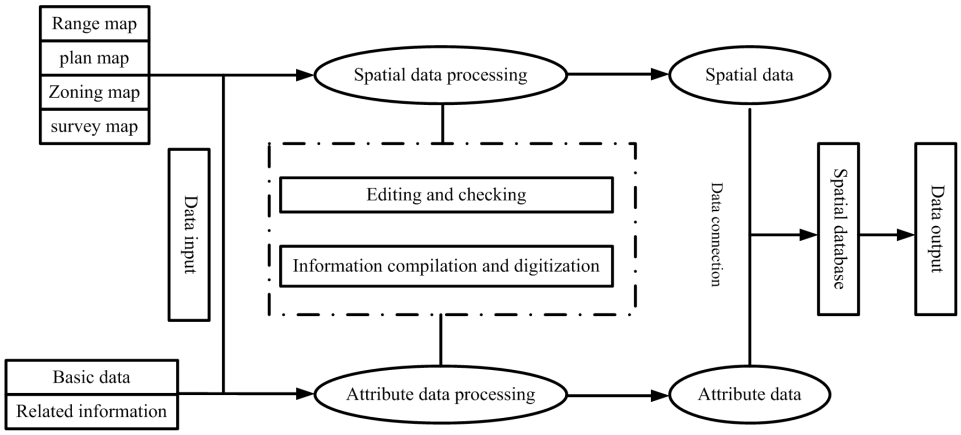

Fig. 1. Process of spatial database construction

3. Land use status of Rongjiang County, Guizhou province, Y2010, paper version.

4. Land use planning map of Rongjiang County, Y2010–2020, Guizhou province (1:10000), paper version.

5. Administrative map of Guizhou Province (1:760000), Y2013, paper version.

6. Statistical Yearbook of Rongjiang County (2010–2015).

7. Related materials on agricultural development situation of Rongjiang County, Y2010–2015.

3.1.2. Data storage. There are two methods for non-electronic spatial data storage: hand tracking digital technology (with digitizer) and scan technology (with scanner) (Zhu et al. [17]), the former is becoming more and more popular. Generally, maps are scanned with accuracy of 200–300 DPI. The obtained image primitives are then vectorized with GIS software. Electronic format spatial data are directly imported to GIS platform for processing and analysis.

In this paper, the GIS spatial databases were constructed by importing digitized maps to SuperMap Deskpro 6R platform. Remote sensing images and topographic map (CAD format) were directly imported while land use maps, land use planning maps, administrative maps were imported in JPG format by 300DPI precision scanning.

3.2. Data classification and coding

3.2.1. Principles of classification and coding. The principles of classification and coding need to be established so that the spatial data and attribute data can be recognized and fast processed by computer. Specific principles are (He et al. [18]) unicity, scalability, legibility, simplicity and integrity.

3.2.2. Data classification methods. There are three data classification methods: linear classification, faceted classification and mixed classification (Zhou et al. [19], Zhang et al. [20]). In this study, mixed classification was applied according to the characteristics of different maps.

3.2.3. Data classification and coding. Data classification and coding refers to establishing code symbol sets and setting the length and size of codes for spatial and attribute data. In present study, the vector and raster data were encoded, respectively. Vector data were classified and encoded based on point, polyline and polygon data, in which land data were classified and encoded in two levels based on "land use classification" (GBT21010–2007) (Table 1), other data were encoded according to the actual demand of the project based on the information nature of the maps. Raster data were represented by specific data that representing the actual value.

Table 1. Codes of land classes

Land class 1		Land class 2	
arable land	01	paddy field	011
		irrigated land	012
		dry land	013
garden land	02	orchard	021
wood land	03	forest land	031
		bush forest land	032
grazing land	04	natural grassland	041
residential land	07	rural homestead	072
traffic land	10	rural road	104
water and water facility land	11	river	111
		pond	114
		inland beach	116
		ditch	117
other land	12	agricultural facility land	122

4. Data processing and spatial database construction

4.1. Construction of attribute database structure

The attribute data were divided into current status data and planning data according to the planning phase of the park and stored in two packets. The current status data packet included elevation points, the park range, remote sensing images, DEM, land use status, elevation grade, slope classification, current roads etc. The planning data packet included industrial evaluation, field planning, pool planning, pumping stations planning, water monitoring points, water pipes, drains planning, road planning and pig farms planning etc.

For every map, an attributive structure table was established according to layer class and purpose. For every layer, a relational attribute table was established with the point, polyline and polygon data. Point layers, such as the elevation point layer, need to determine the elevation data, thus the "elevation points" field was added to default field in order to save the elevation data (Table 2). Polyline layers, such as the road layer was added to a "class" field in order to separate the roads of different class (Table 3). Polygon layers, such as the land use status, were added ("Code of land class I, Name of land class I, Code of land class II, Name of land class II, Area of land class II"), see Table 4.

Table 2. Attribute structure table of the elevation point layers

Field names	Alias	Field types	Field length
*SmID	SmID	32 bit integer	4
Elevation points	Elevation points	Single precision	4

Note:"*" means the default field of SuperMap, the same below.

Table 3. Attributive structure table of the road layers

Field names	Alias	Field types	Field length
*SmID	SmID	32 bit integer	4
*SmLength	SmLength	Double precision	8
*SmToPoError	SmToPoError	32 bit integer	4
Type	32 bit integer	Text type	12

Table 4. Attributive structure table of the land use status layers

Field names	Alias	Field types	Field length
*SmID	SmID	32 bit integer	4
*SmArea	SmArea	Double precision	8
*SmPerimeter	SmPerimeter	Double precision	8
Code of land class I	Code of land class I	Text type	3
Name of land class I	Name of land class I	Text type	50
Code of land class II	Code of land class II	16 bit integer	2
Name of land class II	Name of land class II	Text type	50
Area/Mu	Area	Single precision	4

4.2. Map editing and attribute data input

The map data were edited by the way of artificial description and automatic vectorization. Attribute data of each layer were inputted and edited according to its corresponding information. Take park status layer for example, the plotting CAD base map was used for automatically capture and editing. The attribute data

were inputted according to GBT21010-2007 with the aid of remote sensing images for status classification and identification. The area data were calculated through SuperMap default field "SmArea". For scanned layers like the land use status and land use planning map, their attribute data were digitally scanned and inputted artificially. Similarly, for other layers, their maps data were edited and attribute data were inputted according to data features and actual research demand.

4.3. Mapping and visualization

According to different thematic needs and corresponding drawing specification, map layers can be selected, combined and displayed through human-computer interaction. Various spatial analysis process and results, such as the land use planning map (Fig. 2) and the slope classification map etc. (Fig. 3) can be displayed intuitively in a graphical view.

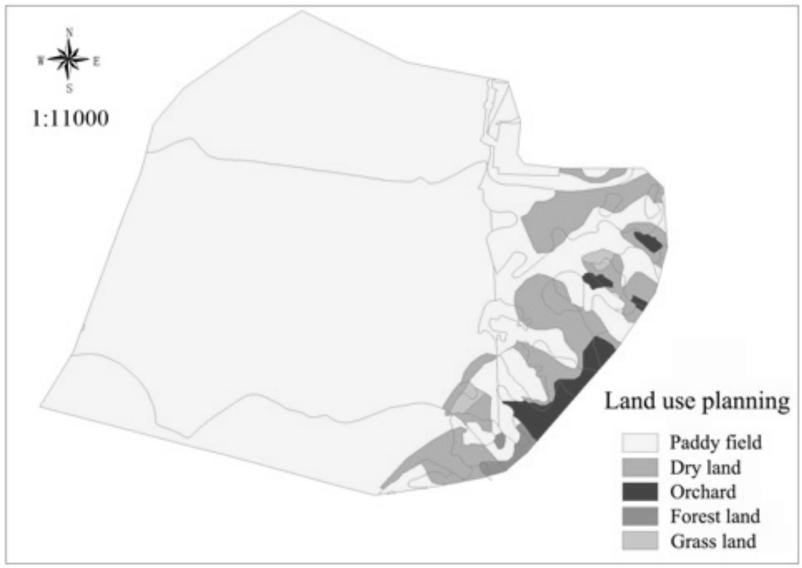

Fig. 2. Land use planning map of Rongjiang County

5. Application of spatial database

Through the combination of planning data, information and maps, the spatial database can provide scientific planning basis and management support for regional agricultural planning, strengthen the quantitative analysis of map data, further enhance the feasibility and implementation of regional agricultural planning to facilitate the analysis and planning of industrial layout, acquire more precise planning data and improve planning efficiency. The application of spatial database of regional

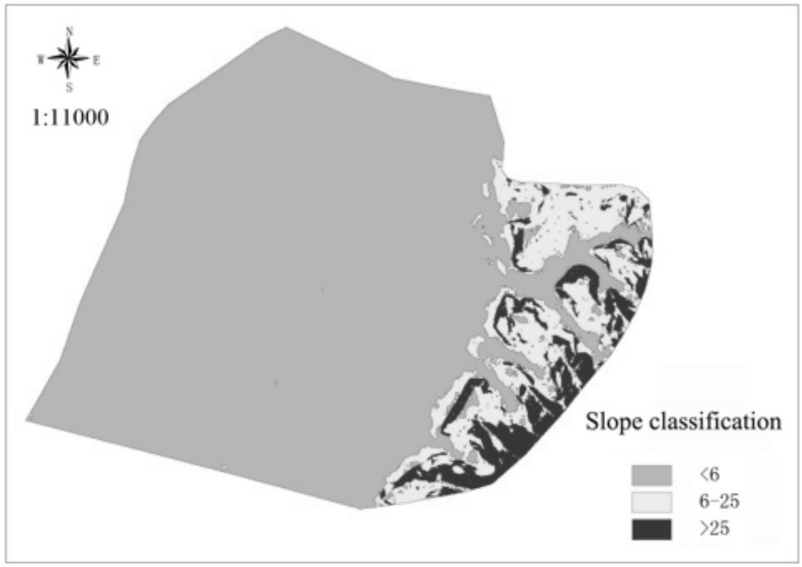

Fig. 3. Slope classification map of Rongjiang County

agricultural planning mainly includes the following four aspects:

Firstly, spatial database is the basis of comprehensive analysis of status data, objective and accurate evaluation of regional agricultural planning. With spatial database, the status data such as land, soil, hydrology, climate, transportation and location etc. are standardized and modeled, making it easy to analyze the advantages vs disadvantages of status production, processing and distribution of regional agricultural, and provide technical support for regional agricultural planning.

Secondly, with spatial database, it is convenient to analyze the favorable and unfavorable conditions of the upper planning and related planning in the project areas, thus to provide accurate information for spatial feasibility of planning. In addition, regional agricultural planning is often constrained and controlled by higher territorial planning, land use planning, town planning, transportation planning and other related planning. With the help of spatial database, it is possible to collect multisectoral spatial planning simultaneously and to provide some important information, such as the planning area, for regional agricultural planning.

Thirdly, based on spatial analysis of topography, water source, roads and farmland water conservancy facilities, we can accurately plan the layout of industrial land, set infrastructure layout to provide accurate data for the investment and help to guide the construction design (Dong, et al. [21]). Spatial database is a collection of multi-source data involved in agricultural planning, with it we can select agricultural industry through the establishment of different planning models, and provide support for objective analysis and evaluation of the implementation of agricultural planning.

Finally, based on GIS platform, some application systems can be developed and

make the planning process and results integrated, thus to facilitate the implementation of planning. With the addition of some other modular such as the planning implementation modular and planning management modular, it is possible to realize the information management of the whole process, from preparation to implementation of regional agricultural planning.

References

[1] Y. Fan, Y. P. Zhu: *A reasoning component's construction for planning regional agricultural advantageous industry development.* Computer and Computing Technologies in Agriculture II *2* (2008), 1291–1297.

[2] X. B. Liu: *Study on regional modern agricultural development planning.* Dissertation for doctorate, Shenyang Agricultural University, Shenyang, Liaoning, China (2011).

[3] A. I. Petrisor: *GIS-based methodology for the analysis of regional landscapes and their changes based on land cover and use: a planning perspective aimed at conserving the natural heritage.* Planning and Designing Sustainable and Resilient Landscapes, series Springer Geography (2014), 87–104.

[4] W. Wu, H. B. Liu, H. L. Dai, W. Li, P. S. Sun: *The management and planning of citrus orchards at a regional scale with GIS.* Precision Agriculture *12* (2011), No. 1, 44–54.

[5] Y. Liu, X. J. Lv, X. S. Qin, H. C. Guo, Y. J. Yu, J. F. Wan, G. Z. Mao: *An integrated GIS-based analysis system for land-use management of lake areas in urban fring.* Landscape and Urban Planning *82* (2007), No. 4, 233–246.

[6] X. Chen, Y. C. Wu, B. W. Zhao: *Study on agricultural park planning methods based on omni-directional information processing technology.* Computer and Computing Technologies in Agriculture VI *393* (2013), 221–228.

[7] X. Liu, G. J. Xiao, B. Z. Lei, J. G. Zhang: *Evaluation of modern agricultural park planning program based on fuzzy comprehensive evaluation..* Chinese Journal of Tropical Agriculture *1* (2016), 90–94.

[8] L. Y. Tang, C. C. Chen, H. Y. Huang, D. Lin: *An integrated system for 3D tree modeling and growth simulation.* Environmental Earth Sciences *74* (2015), No. 10, 7015–7028.

[9] S. Kumar, S. P. S. Kushwaha: *Modelling soil erosion risk based on RUSLE-3D using GIS in a Shivalik sub-watershed.* J Earth System Science *122* (2013), No. 2, 389–398.

[10] L. H. Song, M. X. Shen, R. Y. He, J. H. Cong: *Study on building forestry GIS spatial database based on ArcGIS.* Computer Engineering and Design *19* (2008), 5117 to 5122.

[11] G. F. Hu, Y. H. Qin, C. B. Yang, Y. Q. Wang: *Construction of spatial database of land consolidation planning based on GIS platform and its application.* J Anhui Agricultural Sciences *19* (2015) 350–354.

[12] C. F. Zhou, X. Liu, G. J. Xiao, Y. Zhang, J. Tan, F Li, J. W. Zeng, D. C. Cai: *Application of GIS technology in status analysis of regional agricultural plan.* Guangdong Agricultural Sciences *22* (2013), 193–195.

[13] X. Liu, G. J. Xiao, C. F. Zhou, J. Tan, Y. Zhang, F. Li, J. W. Zeng, D. C. Cai: *Application of GIS technology in ecological environment protection analysis in regional agricultural planning.* Guangdong Agricultural Sciences *9*, (2015), 160–162.

[14] K. Guan, G. Wang, T. Feng, X. F. Liu: *Agricultural area development planning assisted construction based on GIS.* Agriculture Network Information *3* (2013), 22–25.

[15] Y. D. Wang, J. Y. Gong: *A construction schema for provincial spatial database of China.* Geo-spatial Information Science *12* (2009), No. 1, 25–32.

[16] A. Peled, M. Gilichinsky: *GIS-driven classification of land use using IKONOS data and a core national spatial information database.* Applied Geomatics *5* (2013), No. 2, 109–117.

[17] E. L. Zhu, J. H. Li, L. Chen, eds.: *Geography information system basic and application course*. Machinery Industry Press, *52*, Beijing, China, 2004

[18] J. B. He, X. T. Li, J. T. Bi, Y. R. Cao: *Research on classification and coding of resources and environment information and its association with ontology*. Geomatics World *5* (2003), 1–6.

[19] J. X. Zhou, D. Zhang: *Discussion on classification and coding standard system of logistics information*. Logistic and Material Handling, (2000), No. 2, 37–39.

[20] M. Z. Zhang, T. Y. Song, X. M. Tang, S. S. Huang, P. J. Liu, Y. S. Xie: *Classification and coding of forest resources information*. J Fujian College of Forestry *2* (2015), 147–152.

[21] C. Dong, G. H. Huang, Q. Tan, Y. P. Cai: *Coupled planning of water resources and agricultural land use based on inexact-stochastic programming model*. Frontiers of Earth Science *8* (2014), No. 1, 70–80.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 281–290

Design and realization of pork anti-counterfeiting and traceability IoT system

Xiaojin Zou[1]

Abstract. In order to improve the correctness of pork traceability, the IoT is applied to construct the pork anti-counterfeiting and traceability system. Firstly, the basic theory of internet of things is analyzed. Secondly, the pork anti-counterfeiting and traceability system is constructed. Secondly, the main technologies of anti-counterfeiting and traceability system are discussed. The system tags are designed, the EPC coding is confirmed, and the sensor node location algorithm is designed.

Key words. RFID technology, EPC coding, anti-counterfeiting and traceability, pork.

1. Introduction

In recent years, the pork supplying market lacks the united management, and the health problem is every serious, and this phenomena prevent the healthy development of pork market badly. Pork tracking and identifying is carried out based on label, which is corresponding to the pork needed to be identified by a certain means, then the relating property of pork can be traced and managed timely. The pork anti-counterfeiting and traceability can offer an good condition for checking the healthy of pork and improve the breeding quality (Zhu et al. [1]).

In recent years, the pork safety events happen frequently, and the pork safety is paid attention by the whole society. The pork is necessary food for the dining table, and the safety of the pork is more and more concerned, and the pork anti-counterfeiting and traceability system is an critical technology for eating unworried pork by the consumer. In recent year the IoT technology has been developed quickly, the extensive consumers can be easy to find out the information of pork. The IoT can achieve the connection between the things and Internet based on a certain protocol based on radio frequency identification technology, positioning system, and infrared sensor. The functions of IoT conclude identification, location, tracking and

[1]Jiangxi College of Engineering, No.229, West Xianlai Avenue, Xinyu City, Jiangxi, 338029, China zouxiaojin1970@163.com

management. The IoT can achieve the information management, and it can obtain the information of pork safety correctly (Huang [2]).

Some researching achievements have been obtained currently, Alavi et al. [3] analyzed the privacy of RFID authentication protocols that proposed in recent years, the privacy of all the analyzed protocols has some problems and can not offer user privacy, and two improved versions of analyzed protocols were proposed, privacy analysis showed that the privacy of improved protocols were resistance against different traceability attacks. Pang et al. [4] proposed a value-centric business technology joint design framework, and the assessed the income-centric added values, and the developed the corresponding sensor portfolios, and the field trials and an implemented prototype system showed the effectiveness of the system). Ray et al. [5] proposed a secure object tracking protocol to ensure the visibility and traceability of an object along the travel path to support IoT, the protocol is achieved based on radio frequency identification system, and the lightweight cryptographic primitives and physically unclonable function were applied by the system, and results showed that the new system is more secure. Zhang et al. [6] proposed a real-time production performance analysis and exception diagnosis model, and the hierarchical-timed-colored Petri net with smart tokens was applied, and the decision tree was applied to diagnose exceptions from the critical production performance. And simulation results showed that the new system had good performance. Kang et al. [7] proposed a sensor-integrated radio frequency identification data repository-implementation model using MongoDB, the data repository schema is revised to integrate and store the heterogeneous IoT data sources, and an effective shard key to maximize query speed and uniform data distribution over data servers, and simulation results showed that the new system is ana efficient for IoT generated RFID/Sensor big data.

2. Basic theory of internet of things

Pork-anti-counterfeiting and traceability system is constructed based on TCP/IP protocol, and achieves the in-formation exchange between the pork safety monitoring platform and pork safety status, can offer the corresponding data for pork safety management platform, and can manage the communication interface with different protocol underground coal mine, and the IoT communication platform of pork anti-counterfeiting and traceability is shown in Fig. 1 (Li et al. [8]).

The hierarchy diagram of pork anti-counterfeiting and traceability system based on IoT is shown in Fig. 2.

3. Construction of pork anti-counterfeiting and traceability system

Information model in Anti-counterfeiting and traceability system is shown in Fig. 3. As can be seen in Fig. 1, through information collection layer, information preprocessing layer, information transmission layer, information intelligence processing and storage layer, the pork information in the Anti-counterfeiting and traceability

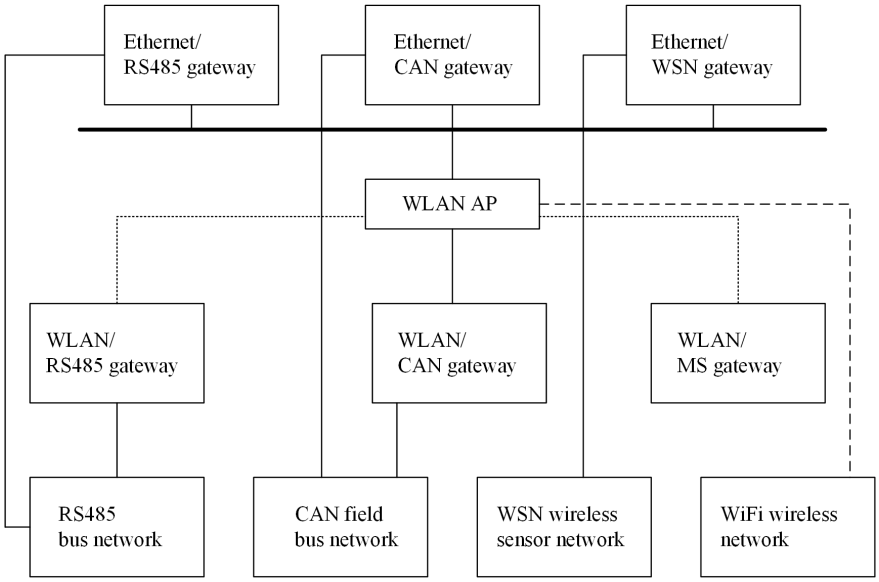

Fig. 1. Communication platform diagram of pork anti-counterfeiting and traceability

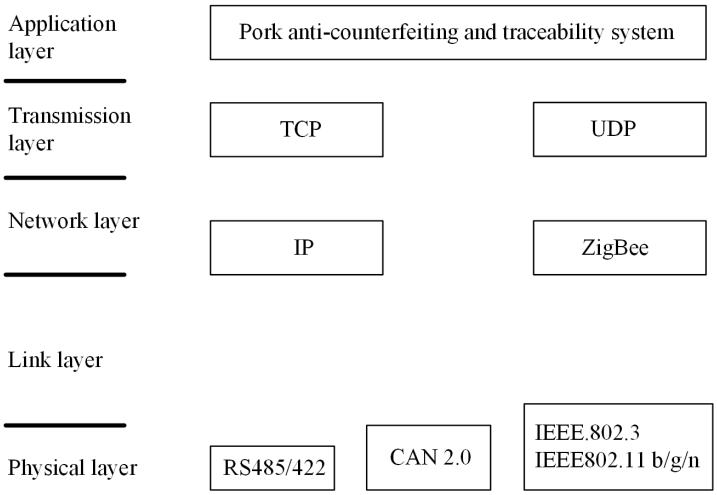

Fig. 2. Hierarchy diagram of pork anti-counterfeiting and traceability system based on IoT

system based on IoT technology is finally stored in the back-end database of WEB application system. The information collection layer mainly gathers information on pig farms and slaughter houses (birth information, feeding information, medical information, vaccination information), inspection and quarantine information,

pork production time, pork transportation information (number of goods, transport vehicles, transport time), and sales information (Papanagiotou et al. [9]).

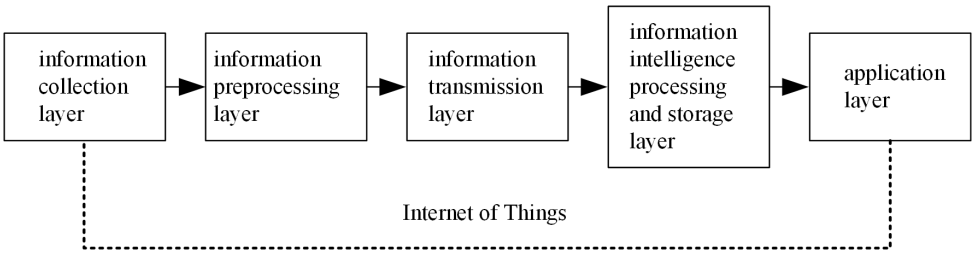

Fig. 3. Information model in the anti-counterfeiting and traceability system

Information preprocessing layer is mainly responsible for a series of processing work of necessary review, integration, conversion, and protocols of the original data before processing the data to complete the incomplete data, correct the data error, remove the redundant data and select the required data for data integration, convert the data format into the required format, remove redundant data attributes and hence to meet alienation of data type, consistency of data format, precision of data information standardization of data storage; Information transmission layer is responsible for transmitting the collected information and storing in the database; through regulation definition, Information Intelligent Processing processes repeating Read, Read failed, false readings, data validation and other issues on the same electronic tag, and makes high-speed processing for vast amounts of information through intelligent information processing technology in real-time, and intelligent mining, management, control and storage of data; and the Information storage layer will store the information collected from the physical aspects in the back-end database after preprocessing (Ren et al. [10]).

Given the respective advantages of NFC and RFID, the system uses a combination of NFC and RFID to achieve bidirectional traceability of pork. NFC tag is used to store pigs ID, each pig has a unique ID generated by the farm, subsequently attached a digital signature, and saved to NFC tag, and other feeding information, vaccination injection information, slaughter information, inspection and quarantine information, logistics distribution and other information stored in the RFID tag encoded by the EPC. NFC and RFID tag information will be written and uploaded to the back-end database system through middleware or hand-held reader. Through NFC smartphone or NFC reader terminal, consumers can inquiry information of pork. System architecture is shown in Fig. 4.

The traceability system mentioned in the literature is a one-way mode, while it is difficult to determine the direction of the condemned pork for the enterprise and total sales point once the pork has problem. The system designed in this paper can conduct upward anti-counterfeiting and traceability for customers to trace back to the logistics information, inspection and quarantine information, cultivation information and other information flows, in addition, downward decision-making of statistics and traceability to trace back to the distribution point information and

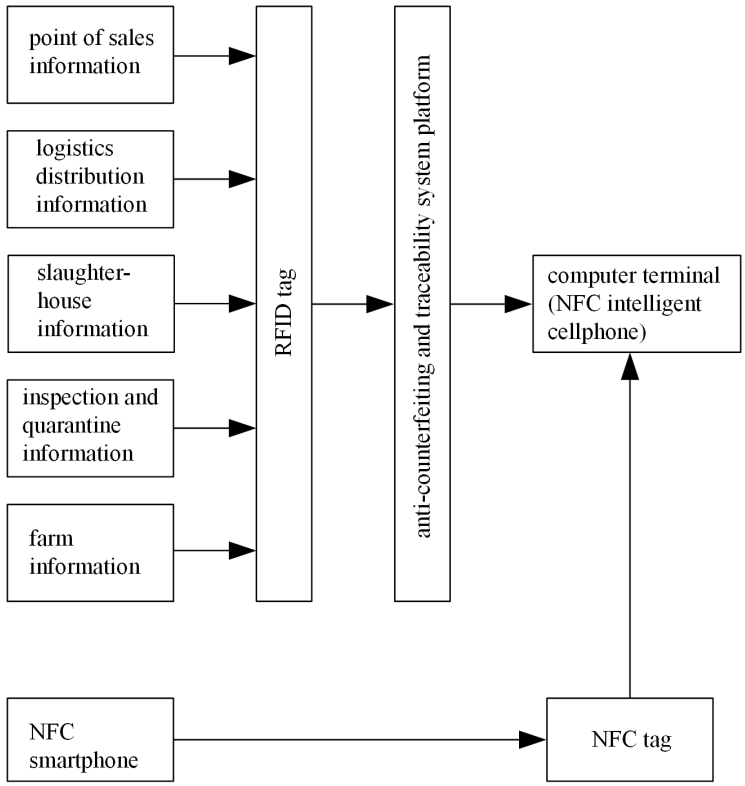

Fig. 4. Anti-counterfeiting and traceability system architecture diagram

consumer group for the enterprises and wholesale market. When problem was found in pork, we can help governments to deal with the problem, at the same time, to prevent the phenomenon of changing products in the supply chain. Bidirectional traceability pattern is shown in Fig. 5.

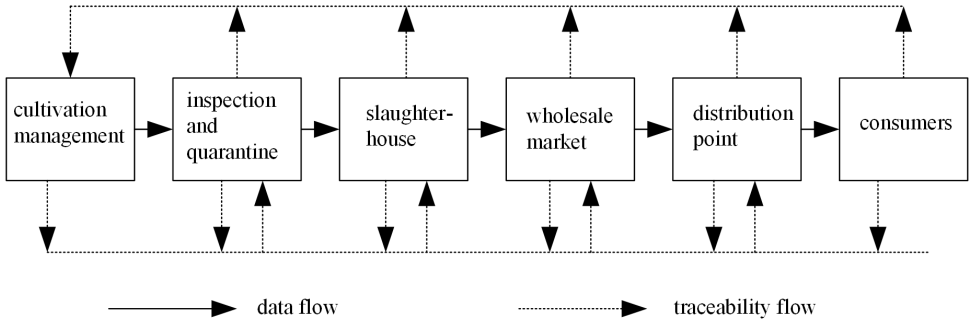

Fig. 5. Diagram of bidirectional traceability model

4. Key technology of anti-counterfeiting and traceability system

NFC tag has four main types (type 1 to type 4). Type 1 is based on ISO14443A, currently exclusively supplied by Innovision Research and Technology Company (TopazTM), which has 96 bytes of memory, with low cost and wide application. Type 2 is also based on ISO14443A, currently exclusively supplied by Philips (MIFARE UltraLight), whose memory is half of type 1 tag. type 3 is based on FeliCa, currently exclusively supplied by Sony, with larger memory (at present 2 kilobytes) and high transmission rate(212 kB/s), suitable for more complicated application. Type 4 is entirely compatible with ISO14443A/B, can be produced by most suppliers including Philips (typical products like MIFAREDESFire), with larger memory and reading rate between 106 kB to 424 kB per second. The anti-counterfeiting and traceability system designed in this paper mainly adopts NFC smart phone to read data, type 1 of NFC electronic tag, with 106KB/S transmission rate and whose recognition distance is about 10 cm (Xiao and Gao [11]).

According to operating frequency, RFID tag is mainly divided into low frequency RFID tag, medium-high frequency RFID tag, UHF and RFID microwave range tag. The typical frequency of low frequency tag is 125 kHz–134.2 kHz. Typical frequency of medium-high frequency tag is 13.56 MHz; UHF and RFID microwave range tag is generally short for "microwave tag", typical UHF working frequency is 860 MHz–928 MHz, and microwave range working frequency is 2.45 GHz–5.8 GHz. Microwave tag mainly has two categories: passive tag and active tag. The working frequency of the microwave passive tag ranges between 902 MHz–928 MHz. Finally, the working frequency of the microwave active tag is mainly in the 5.8 GHz–2.45 GHz. Microwave tag works in the far field region of the reader antenna radiation. In this paper, RFID in the anti-counterfeiting traceability system mainly used to store cultivation information, inspection and quarantine information, logistics information, and select the ultra-high frequency RFID tag for RFID electronic tag.

EPC (electronic product code, the electronic product code) technology is developed by the US Auto ID, through the Internet platform, using wireless data communication technology and RFID technology to construct a network platform which can achieve information sharing of global items, and can realize the tracking and back-tracking of goods information. At current, coding models applied in EPC system mainly have 64 bits, 96 bits and 256 bits. EPC coding consists of version number, serial number, product domain name management and product classification. This system adopts 96 bit EPC code, as shown in Fig. 6, 8 bits are used to store the EPC version number, 28 used to store pork brand and farm information, 24 bits are used to store the pork classification code, 36 bits used to store pork inspection and quarantine, logistics and distribution information, date of slaughter, wholesale market, distribution points, and other information, and fill 0 in the back when the actual code length is less than the length of the reserved (Gu et al. [12]).

The sensor node is defined by $S = \{s_1, s_2 \cdots, s_n\}$ and sensor range of different node is defined by A_i. The sensor node number of IoT at mth moment is defined by X_n, where $\{X_n, n = 1, 2, \cdots\}$ is a Markov chain. When the Markov chain enters

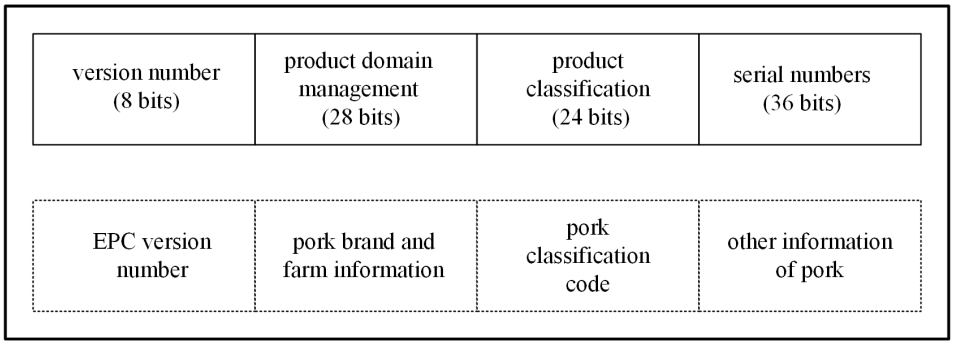

Fig. 6. EPC coding meaning diagram

the status j, the monitoring value of node s_j is changed based on probability $p(j|i)$.

If

$$s_j \in \mathrm{nbr}_{s_i}\,,\ \text{then}\ \sum_{s_j \in \mathrm{nbr}_{s_i}} p(j|i) = 1\,,$$

while if

$$s_j \ni \mathrm{nbr}_{s_i}\,,\ \text{then}\ \sum_{s_j \in \mathrm{nbr}_{s_i}} p(j|i) = 0\,,$$

symbol nbr denoting the neighborhood.

The monitoring value of node i at tth moment is defined as $Y_t^i \in \{0, 1\}$, where 1 denotes that the objective value is monitored, 0 shows that the objective value is not monitored. The whole monitoring vector of all sensor nodes in IoT before Tth moment os defined as $Y_{1,T} = \{Y_1, Y_2 \cdots, Y_T\}$ the status vector is defined as $X_{1,T} = \{X_1, X_2 \cdots, X_T\}$ and the combined distribution function is defined by Zhao et al. [13] as

$$P\left(X_{1,T}, Y_{1,T}\right) = P(X_1) \prod_{t=2}^{T} P\left(X_t|X_{t-1}\right) \prod_{t=2}^{T} P\left(Y_t|X_t\right)\,. \tag{1}$$

The maximum probability $\delta_t(i)$ at tth moment in i is expressed by the formula

$$\delta_t(i) = \left[\max_{j \neq nrb_i} P\left(z_t^i|z_{t-1^j}\right) \delta_{t-1}^j\right] \eta_i^{y_1^i} (1-\eta_i)^{1-y_1^i}\,. \tag{2}$$

Here, $P\left(z_t^i|z_{t-1^j}\right)$ is the probability that the object located at the jth node moves to the ith node and η_i is the monitoring probability of node S_i.

The weight value of $\delta_i(i)$ can be calculated using information of the ith node, which is expressed as

$$\delta_i(i) = \pi \eta_i^{y_1^i} (1-\eta_i)^{1-y_1^i}\,. \tag{3}$$

The weight value of different sensor node in IoT can be expressed as

$$\omega_i = \frac{\delta_t(i)}{\sum_{\delta_t(i)>0} \delta_t(i)} \,. \tag{4}$$

The coordinates x and y of the monitored object are given by the expressions

$$x = \sum_{i=1}^{n} \omega_i x_i \,, \quad y = \sum_{i=1}^{n} \omega_i y_i \,. \tag{5}$$

The sensor node location can be obtained using the above algorithm correctly.

5. Test of pork anti-counterfeiting and traceability IoT system

The system test can verify the performance of the pork anti-counterfeiting and traceability IoT system, the direct user of the system is the pork processing enterprises, and the indirect user of the system is consumer, the performance of the system can decide the satisfaction degree. Therefore it is important to carry out a test of the system. Every module of the system can cooperate, the test software and hardware environments are listed as follows:

Hardware environment: the processor of the application server is Intel Xeon quad core 2.83 GHz, the cache is 12 M, the internal storage is 8 GB, the hard disk capacity is 300 G. The processor of the database server is Intel 6500 quad core 2.66 GHz, the cache is 24 M, the internal storage is 4 GB, the hard disk capacity is 500 G. Software environment: the operation system is Linux Redhat 5.5, and the database is Oracle 11g, the system software is CIN-SCF platform. The logic modules of the pork anti-counterfeiting and traceability IoT system are tested, and the corresponding testing results are listed in Table 1.

Table 1. Testing results of logic modules of system

Logic module	Number of testing samples	Testing result
Cultivation management	103	100 % pass
Inspection and quarantine	46	100 % pass
Slaughterhouse	15	100 % pass
Wholesale market	75	100 % pass
Distribution point	55	100 % pass
Consumers	42	100 % pass

As can be seen from Table 1, all logic modules of the pork anti-counterfeiting and traceability IoT system passed the test, and the performance of the system can satisfy the business requirement.

6. Conclusion

Combining RFID technology and NFC technology, the Pork Anti-counterfeiting and Traceability System based on Internet of Things technology proposed in this paper realizes bidirectional traceability pattern of upward anti-counterfeiting and traceability and downward decision-making of traceability from the cultivation, inspection and quarantine, slaughter, wholesale, distribution and other supply chains, which have a certain value in implementing traceability of meat and vegetables.

References

[1] Y. N. Zhu, L. J. Long, F. Chen: *Construction of black tea anti-counterfeiting and traceability system based on bidirectional traceability model.* Hunan Agricultural University (Natural Science) *40* (2014), No. 5, 552—555.

[2] J. Z. Huang: *Design and research on tea anti-counterfeiting and traceability system based on NFC.* Rural Economy and Science *24* (2014), No. 3, 23–26.

[3] S. M. Alavi, K. Baghery, B. Abdolmaleki, M. R. Aref: *Traceability analysis of recent RFID authentication protocols.* Wireless Personal Communications *83* (2015), No. 3, 1663–1682.

[4] Z. B. Pang, Q. Chen, W. L. Han: *Value-centric design of the internet-of-things solution for food supply chain: Value creation, sensor portfolio and information fusion.* Information Systems Frontiers *17* (2015), No. 2, 289–319.

[5] B. R. Ray, M. U. Chowdhury, J. H. Abawajy: *Secure object tracking protocol for the internet of things.* IEEE Internet of Things Journal *2* (2016), No. 37, 544–553.

[6] Y. F. Zhang, W. B. Wang, N. Q. Wu, C. Qian: *Real-time production performance analysis and exception diagnosis model.* IEEE Trans. Automation Science and Engineering*13* (2016) No. 2, 1318–1332.

[7] Y. S. Kang, I. H. Park, J. Rhee, Y. H. Lee: *MongoDB-based repository design for IoT-generated RFID/sensor big data.* IEEE Sensors Journal *16* (2016), No. 2, 485–497.

[8] W. I. Li, C. H. Sun, X. Q. Liu: *Encryption algorithm design and realization of embedded agricultural product traceability code.* Trans. Chinese Society of Agricultural Engineering *280* (2012), No. 17, 253–258.

[9] P. Papanagiotou, I. Tzimitra-Kalogianni, K. Melfou: *Pork quality in the eye of the Greek consumer.* British Food Journal *114* (2012), No. 5, 647–660.

[10] H. T. Ren, L. Y. Cheng, X. Y. Song, L.z,Z. Yang: *Technology of quick-frozen Dongpo pork.* J Chinese Institute of Food Science and Technology *13*, (2013), No. 2, 27–34.

[11] K. Xiao, G. D. Gao: *Research on detection of pork freshness system based on mobile photograph image analysis.* ICIC Express Letters *7* (2013), No. 12, 3265–3269.

[12] H. N. Gu, Q. Li, W. Z. Li, Z. G. Zhang, S. Guo, C. Cheng: *Quality change and shelf-life prediction of chilled pork during storage.* Modern Food Science and Technology *29* (2013), No. 11, 2621–2626.

[13] F. Zhao, Q. H. Zhao, H. B. Chen: *Tracking and localization algorithm for mobile station based on extended Kalman filtering and data fusion.* Computer Engineering *38* (2012), No. 22, 244–247.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 291–300

The reliability model of cloud service system with recovery mechanism[1]

Dan Liu[2], Xin Sui[2,3], Huan Wang[2], Xu Di[2], Li Li[2,4]

Abstract. A variety of failure issues of CSS (cloud service system) is analyzed, and then proposed the FRRM reliability mode based on the failure nodes. In the model, the faults of CSS's soft/hardware and LAN communication are all discussed. Moreover, in order to reflect the influence of time on system reliability, the paper introduces time influence factor. For effective verification of FRRM reliability model, the paper makes a comparison experiment with the classical NHPP reliability model. The simulation results show that the FRRM reliability model is effective. It can make the reliability of cloud service system optimized.

Key words. Cloud service system, reliability model, modeling, faults.

1. Introduction

How to improve the reliability of cloud service system (CSS) had attracted the attention of many scholars both at home and abroad. Scholar He Li proposed a cloud service reliability measurement model which was based on checkpoint rollback strategy. The cloud service reliability model was established with the checkpoint rollback strategy over the changed system service reliability problems which were caused by failure nodes in CSS's virtual machine tasks. The theoretical deduction proved that the service reliability could be optimized when the test point setting time was increased by a certain index, but the failure of the calculation node was considered only in the study. Scholar Yadan Zhang proposed a hidden Markov

[1]This work was supported in part by the project of "12th Five-Year" planning of the Education Science of Jilin Province (GH14049), the science and technology project of "13th Five-Year" planning of the Education Department of Jilin Province (2016), the key project of "13th Five-Year" planning of the Education Science of Jilin Province (ZD16024), and Key Science and Technology Project of Jilin Province (20160204019GX).

[2]Changchun University of Science and Technology, College of Computer Science and Technology, 130022, China

[3]Jilin Provincial Institute of Education, 130022, China

[4]Corresponding author: Li Li, Changchun University of Science and Technology, College of Computer Science and Technology, 130022, China, ll@cust.edu.cn

model based on simulated annealing algorithm in the problem of faults detection. By solving the initial parameter of hidden Markov model, the detection accuracy was improved. Most scholars paid more attention to the faults of a single reason in the existing research. However, the existing reliability model could hardly solve the problem because of the failure of CSS's heterogeneity was diversified. From the viewpoint of system failures, the paper analyzes the failure types of CSS, and considers the influence of time on the system reliability. Finally, the reliability model called FRRM [1–4] was proposed.

2. Problem formulation

2.1. The architecture of CSS

CSS can provide services to the users because of the effective management, allocation and scheduling of virtual resources. The operation process is as follows: when cloud computing management platform (CCMP) receives users' service request, the task can be divided into multiple child-tasks, in order to improve the efficiency of the work. Then, according to the resource requirements of child-tasks, scheduling of some executive order or LAN-communication child-tasks which is based on each resource node can start. After that, task results commit into CCMP and are integrated, and the final result redrawn to the users. During the execution of child-tasks, CCMP nodes or LAN all may exhibit failure, which deteriorates the quality of the CCMP service [5–6]. So the paper established the reliability of the system services model based on the analysis of node failure.

2.2. The analysis of CSS node failure

In the actual cloud computing environment, due to the software, hardware and other reasons, the failure problem in the implementation process of the cloud service system is inevitable. Therefore, it is necessary to analyze the failure mechanism of the nodes according to the behavior characteristics of the nodes. The paper discusses the reliability of CSS from the perspective of resource allocation. Node failures may be caused by software or physical breakdown (including hardware or LAN fault), and they can be separated into recoverable failures and no-recoverable failures [7–9].

3. The proposed model

According to the definition of reliability in reliability engineering theory, cloud services platform can be defined as: CSS had the ability within the specified time to provide services to users in a given condition. CSS provides users with computing, storage, applications and other services in the form of service. All these services should be completed in the shortest possible time, and redraw the final result to the users. Based on the above situation, the paper makes the following assumptions: in the cloud service system, the virtual machine manager who is responsible for the

task distribution is reliable, and the resource nodes are independent of each other. The child-tasks are independent of each other. All kinds of faults are independent during each child-tasks' implementation processes. It can be considered as a discrete probability distribution during the failure and recovery processes of each resource nodes and LAN.

3.1. The reliability model of CSS

3.1.1. Brief introduction of modeling method. In essence, cloud computing is a fusion of distributed processing, parallel computing and virtualization technology. It is a new computing model of information system development and evolution, it has a significant scalability, flexible and easy to configure. In CSS, unavailable computing resources, software failure, link failure, etc. are likely to hinder the completion of the task and the reliability of cloud model. Therefore, as a focal point of node failure, the paper proposes an abstract description of the reliability of the service system and pushes out a reliability system model. In the process, the reliability of cloud computing service system can be expressed as

$$R(t) = 1 - F(t)\,, \tag{1}$$

where $R(t)$ stands for the CSS reliability function and $F(t)$ denotes its distribution function.

3.1.2. The establishment of reliability model. CSS divides the task Q submitted by user into a number of m child-tasks, and names the set of child-tasks as $Q = q_1,\, q_2,\, q_3, ..., q_m$. Then, m child-tasks are assigned to the l resource nodes; there is an inevitable redundancy allocation in the allocation process. And the handle time representing the allocation time of kth child-task to node i can be denoted as $\tau_{i,k}$.

$$\tau_{i,k} = \frac{W_i}{S_k}\,, \tag{2}$$

where W_i stands for ith child-task workload and S_k is the computing speed of the kth node.

Due to CSS usually adopts the fault-tolerant technique. Therefore, in addition to the execution time $TE_{i,k}$ and total failure recovery time $TR_{i,k}$ the total LAN communication time $TS_{i,k}$ may also be added, and this is the actual total execution time T of child-task. In interval $(0, T]$ it is assumed that the total number of recoveries of kth node N_k is a random variable, such that $N_k = n > 1$. In the course of ith child-task execution, two node failure cases happen: one is a recoverable failure, and the other is a non-recoverable failure. So, the the total time T of the ith child-task execution on node k can be written as

$$T = \{ \begin{array}{ll} \tau_{i,k} + TS_{i,k} & N_k = 0 \\ TE_{i,k} + TR_{i,k} + TS_{i,k} & N_k > 0 \end{array} . \tag{3}$$

From the hypothesis from the last section, the failures of soft/hardware and LAN

communication obey certain Poisson distribution. And during the execution of the process, the failure distribution parameters are constant. In many kinds of failures, only the hardware failure is recoverable. Therefore, in the time interval $(0-T]$ there holds

$$TE_{i,k} = \sum_{1}^{n+1} TE_{i,k}^{(j)} \,. \tag{4}$$

and

$$TR_{i,k} = \sum_{1}^{n+1} TR_{i,k}^{(j)} \,. \tag{5}$$

where $TE_{i,k}^{(j)}$ stands for the jth execution time of the ith child-task at the kth node, $j \in [1, n+1]$, and $TR_{i,k}^{(j)}$ denotes the jth failure recovery time.

It can be assumed that $TE_{i,k}^{(j)}$ is independent and identically distributed random variable, and the parameter of the exponential distribution is λ_k. The value of $TR_{i,k}^{(j)}$ is also independent and identically distributed random variable, and the parameter of the exponential distribution of μ_k, which is called recovery rate. Finally, $TS_{i,k}^{(j)}$ is also independent and identically distributed random variable, and the parameter of the corresponding exponential distribution is ν_k. In order to describe the failure recovery capability, the random variable $X_k^{(j)}$ is defined, which is used to express whether the node k is recoverable or not. If the kth node becomes non-recoverable, it is expressed by $X_k^{(j)}$. Then, the child-task distributed at the failure node is stopped. So, when $g > j$, then $TE_{i,k}^{(g)} = 0$ and $TR_{i,k}^{(g)} = 0$. When the node k becomes recoverable, it is expressed by $X_k^{(j)} = 0$.

In CSS, the failure recovery has a certain probability. So it should be divided into different situations to discuss. It should be also considered the influence of reliability time when discussing the system reliability model. Therefore, the paper introduces the time influence factor $\delta^{(t-\tau_{i,k})}$, where $\delta \in (0,1)$.

All the faults belong to non-recoverable failures except for hardware failures. In interval $(0-T]$, it is assumed that the ith child-task is executed at the kth node and follows the exponential distribution of the parameter λ_k. Then, the probability of software failure is

$$F_{\mathrm{s}} = \delta^{(t-\tau_{i,k})} \mathrm{e}^{-\lambda_k \tau_{i,k}} \,. \tag{6}$$

Likewise, the probability of LAN failures of system communication is

$$F_{\mathrm{c}} = \delta^{(t-\tau_{i,k})} \mathrm{e}^{-\lambda \nu_k \tau_{i,k}} \,. \tag{7}$$

Then, it is analyzed the probability of hardware failure denoted as F_{h}. Because of certain recoveries of the hardware failure, not only the probability of non-failure node should be considered but also the probability of failure at the recovered node. So, $F_{\mathrm{h}} = F_{\mathrm{tf}} + F_{\mathrm{fr}}$.

Here, F_{tf} represents the failure of node, and the exponential distribution of the parameter was ρ_k, so that $F_{\mathrm{tf}} = \mathrm{e}^{-\rho_k \tau_{i,k}}$. Symbol F_{fr} stands for the failure at the

recovery node, so that $F_{\mathrm{fr}} = F_{\mathrm{fr}}^{(1)} + F_{\mathrm{fr}}^{(2)} + F_{\mathrm{fr}}^{(3)} + \ldots + F_{\mathrm{fr}}^{(n)}$. Generally, $F_{\mathrm{fr}}^{(n)}$ represents the recoverable failure to ith execution of the child-task on kth node and the number of this failure is n. So, it can be written as

$$F_{\mathrm{fr}}^{(n)} = F_{\mathrm{fr}}\{\sum_{j=1}^{n} TE_{i,k}^{(j)} < \tau_{i,k}\,,\ \sum_{j=1}^{n} TE_{i,k}^{(j)} + TE_{i,k}^{(n+1)} > \tau_{i,k}\,,\ \sum_{j=1}^{n} X_k^{(j)} = 0\}\,. \tag{8}$$

As the times $TE_{i,k}^{(j)}$, $(j = 1, 2, \ldots n)$ are independent of each other, it can be taken as the Erlang distribution of random variables which follows parameters n, ρ_k. According to this, $\sum_{j=1}^{n} TE_{i,k}^{(j)}$ and $TE_{i,k}^{(n+1)}$ are independent of each other. So, the joint density can be written as

$$f_{i,k}^{(n)} = \begin{cases} \mathrm{e}^{-\rho_k x - \rho_k y} \frac{\rho_k^2 (\rho_k x)^{n-1}}{(n-1)!} & x \geq 0, y \geq 0 \\ 0 & \text{elsewhere} \end{cases}, \tag{9}$$

$$F_h = \delta^{(t-\tau_{i,k})} \sum_{n=1}^{\infty} F_{\mathrm{fr}}^{(n)} \cdot \mathrm{e}^{-\rho_k \tau_{i,k}} = \delta^{(t-\tau_{i,k})} \mathrm{e}^{(x_k - 1)\rho_k \tau_{i,k}}\,. \tag{10}$$

Therefore, in interval $(0, T]$, the probability F of total failure of the ith child-task at the kth node is:

$$F = F_{\mathrm{s}} F_{\mathrm{c}} F_{\mathrm{h}} = \delta^{3(t-\tau_{i,k})} \mathrm{e}^{-\lambda_k \tau_{i,k} - \lambda \nu_k \tau_{i,k} + (x_k - 1)\rho_k \tau_{i,k}}\,. \tag{11}$$

The above paragraph just discussed the failure probability of ith task at kth node. In the actual system, on must consider the running mode of parallel execution of multi-tasks and multi-nodes. In the process of distribution and execution, not only redundant allocation should be considered, but also the influence of time on the system reliability. Therefore, the reliability R of CSS should be written as

$$R = \prod_{i=1}^{m} [1 - \prod_{k=1}^{l} (1 - \delta^{3(t-\tau_{i,k})} \mathrm{e}^{-\lambda_k \tau_{i,k} - \lambda \nu_k \tau_{i,k} + (x_k - 1)\rho_k \tau_{i,k}})]\,. \tag{12}$$

4. Results

In order to verify the validity of the FRRM model, the paper designes the simulation experiment for it. The authors research the impact of node failure recovery mechanism on the reliability of CSS. In the same experimental conditions, FRRM is compared with the model of non-homogeneous Poisson process (NHPP).

The initial conditions: in the simulation experiment, the task number is set 0∼100, and the allocation of resource nodes for CSS is $0 \sim 200$. The value of $\lambda_k \in [0.001, 0.004]$, $\mu_k \in [0.001, 0.003]$, $\nu_k \in [0.002, 0.003]$. All the initial conditions of the experiment followed these conditions.

4.1. The influence of node failure recovery rate on reliability

In the simulation experiment, it can be obviously seen that the probability of successful execution of the FRRM model is higher than that of the NHPP model under the same failure recovery rate. When the failure recovery rate approaches 0, it indicates that the failure recovery mechanism has no effect, and the successful execution probability is low; when the failure recovery rate increases, the probabilities of successful executions of both models also increase. When the failure recovery rate approaches 1, both probabilities of successful executions of the two model reach the maximum values. The process shows that the failure recovery mechanism plays an important role as the probabilities of successful executions of both models do not reach 100 %. It describes that two failures of software and LAN belonging to no-recoverable faults play a role, as shown in Fig. 1.

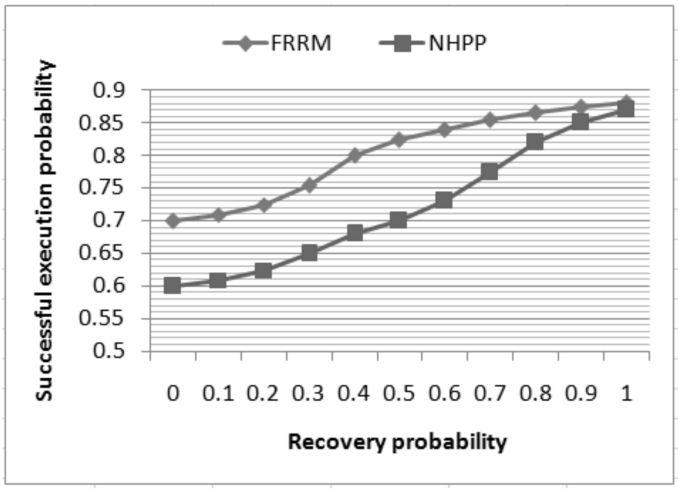

Fig. 1. Relationship between recovery probability and probability of successful execution

4.2. Impact of number of tasks requested by user to system reliability

Through simulation results, with the increase of the number of tasks requested by the user, the successful execution probability is reduced. In the process of experiment, when the number of requests increased from 10 to 50, the system reliability of both the FRRM model and NHPP model decreased, and the downward trend was relatively slow. But when the number of requests increased from 50 to 100, the system reliability of the two models significantly decreased. The successful execution probability of the FRRM model was reduced by nearly 8 % and that of the NHPP model was reduced by 25 %. In contrast, the probability of successful execution of the FRRM model is higher than that of the NHPP model, and this shows that the

FRRM model performs better, as is shown in Fig. 2.

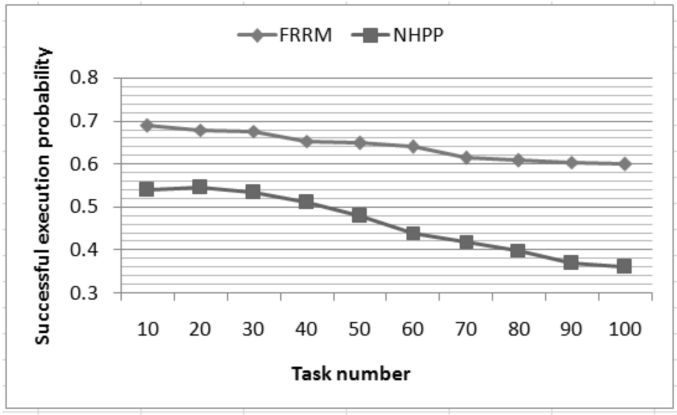

Fig. 2. Relationship between number of requested tasks and probability of successful execution

4.3. Impact of number of CSS resources on system reliability

Through simulation results, the probability of successful execution was also increased with the increase of the number of system resources. In the process of simulation experiment, the reliability of the cloud service system under the two models was both improved under the increasing number of available resource nodes. In contrast, the probability of successful execution of the FRRM model was higher than that of the NHPP model. Especially when the number of resources was between from 0 to 80, the probability of successful execution increased by nearly 21 %, as is shown in Fig. 3.

4.4. Verification of time influence factor on system reliability

In the FRRM model, the time influence factor $\delta^{(t-\tau_{i,k})}$ was added. The experiment compared the reliability model before and after adding the attenuation factor in the simulation experiment, as shown in Fig. 4.

5. Discussion

During the simulation experiment, NHPP was chosen to compare with the FRRM model. As the modeling object of NHPP was similar to the research object described in the, they were both targeted by modeling with "discrete event". The experiment results showed that FRRM model could effectively improve the software reliability of CSS and had an advantage over the others. But in the experiment, it was found

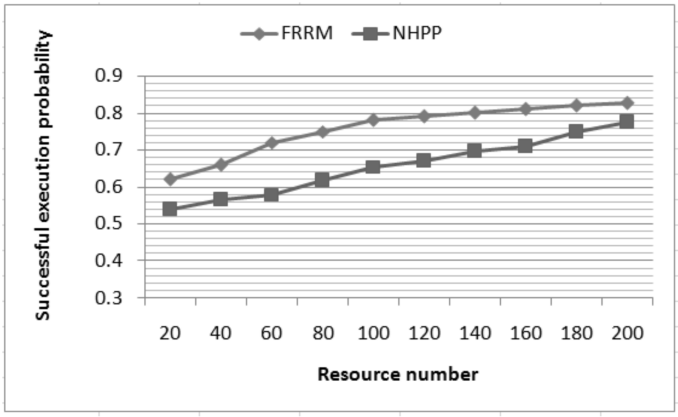

Fig. 3. Relationship between number of requested resources and probability of successful execution

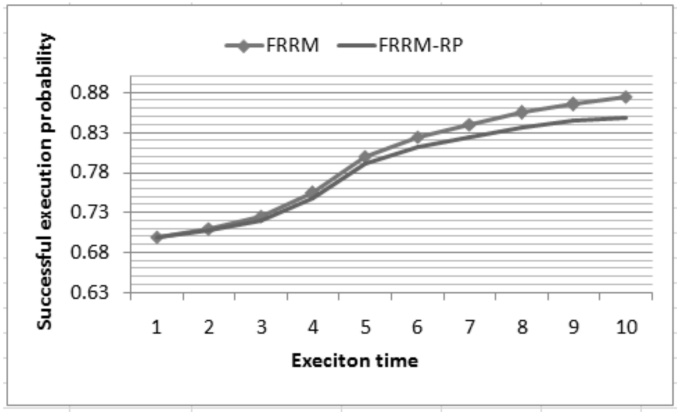

Fig. 4. Influence of time influence factor on system reliability

that the convergence of the algorithm was not very well. The optimization algorithm will be researched in further periods of work in the field.

6. Conclusion

The paper mainly researched the reliability of CSS and established the reliability model by analyzing the failure mechanism of CSS. In the process of model building, the relationship between the operation time and system reliability was also considered. In order to verify the validity of the model, a simulation experiment was carried out. In the experiment, the proposed model proved to improve the reliability of CSS.

References

[1] G. Levitin, Y.-S. Dai: *Service reliability and performance in grid system with star topology.* Reliability Engineering and System Safety *92* (2007), No. 1, 40–46.

[2] A. Dogan, F. Ozquner: *Matching and scheduling algorithms for minimizing execution time and failure probability of applications in heterogeneous computing.* IEEE Trans. Parallel and Distributed Systems *13* (2002), No. 3, 308–323.

[3] Y. S. Day, M. Xie, K. L. Poh: *Reliability of grid service systems.* Computers and industrial Engineering *50* (2006), Nos. 1–2, 130–147.

[4] N. Yadav, V. B. Singh, M. Kumari: *Generalized reliability model for cloud computing.* IJ Computer Applications *88* (2014), No. 14, 13–16.

[5] H. R. Faragardi, R. Shojaee, H. Tabani, A. Rajabi: *An analytical model to evaluate reliability of cloud computing systems in the presence of QoS requirements.* Proc. IEEE/ACIS IC Computer & Information Science, 16–20 June 2013, Niagata, Japan, *446*, 315–321.

[6] J. Zhao, L. Hu, Y. Ding, G. Xu, M. Hu: *A heuristic placement selection of live virtual machine migration for energy-saving in cloud computing environment.* Plos One *9* (2014), No. 9, paper e108275.

[7] W. Wang, G. Zeng, D. Tang, J. Yao: *Cloud-DLS dynamic trusted scheduling for cloud computing.* Expert Systems with Applications *39* (2012), No. 3, 2321–2329.

[8] X. Qiu, Y. Dai, Y. Xiang, L. Xing: *A hierarchical correlation model for evaluating reliability, performance, and power consumption of a cloud service.* IEEE Trans. Systems, Man & Cybernetics Systems *46* (2016), No. 3, 401–412.

[9] J. Rahme, H. Xu: *A software reliability model for cloud-based software rejuvenation using dynamic fault trees.* IJ SW Engrg. & Knowledge Engrg. *25* (2015), No. 10, 1491–1513.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 301–306

Corrosion inhibitors for steel oilfield equipment[1]

Dinar D. Fazullin[2], Gennady V. Mavrin[2]

Abstract. The rate of corrosion and protection of steel are determined as a test environment using the model which produces water with the addition of inhibitors or corrosion inhibitors ($CaCl_2 \cdot 6H_2O$, $MgCl_2 \cdot 6H_2O$, NaCl, $CaSO_4 \cdot 2H_2O$). The proposed steel protection "Steel 20" was 77.8 % in comparison to the same conditions conducted for tests of commercially available corrosion inhibitors. The protective properties of the proposed corrosion inhibitor are not inferior with respect to the currently used inhibitors.

Key words. Petroleum recovery, concentrate, corrosion inhibitor, “Steel 20”, corrosion rate, hydrogen sulfide.

1. Introduction

Most of the oil and gas fields are characterized by a high water content, which greatly complicates the process of production, collection and treatment processes associated with the formation of stable emulsions of oil, deposits of inorganic salts and corrosive destruction of equipment and pipelines [1]. Destruction of oilfield equipment is determined by the physicochemical properties of water and hydrocarbon components of the system, their composition, quantity ratio, presence of dissolved gasses (hydrogen sulfide, carbon dioxide, oxygen, etc.). At higher flow velocities, ensuring intensive mixing phase emulsion system, it is formed by an oil-in-water or water in oil while defending their separation into two immiscible phases. In all cases, the corrosive environment is water. The most common problem for the oil industry today are carbon dioxide corrosion, hydrogen sulfide corrosion, hydrogen embrittlement, etc. The experience of corrosion indicates that the safe operation of process equipment can be achieved through the use of inhibitors. At the same time, compared to other methods of corrosion protection technology for inhibiting aggressive media they are relatively simple and do not require the use of significant

[1]The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

[2]Kazan Federal University, Naberezhnye Chelny, Prospect Mira, 68/19, Russia, 420015

logistical costs [2]. Corrosion inhibitors are classified to organic (95 %) and inorganic (5 %). Organic corrosion inhibitors—surfactants—can be classified as artificial and natural. Artificial surfactants are corrosion inhibitors obtained in chemical and petrochemical plants by methods of oxidation, suffocation, nitration, alkylation, etc. Corrosion inhibition has complex mechanism and depends on the formation of mono- or multidimensional protective layers on the metal surface. The protective nature of the surface layer depends on many factors: interaction between inhibitors and substrate, incorporation of the inhibitor in the surface layer, chemical reactions, electrode potentials, concentration of the inhibitor, temperature and properties of the corresponding surface, etc. The corrosion inhibitors as surfactants can be divided into two groups: the water-oil-soluble (WOS) surfactants and oil-soluble (OS) surfactants. Also, according to the mechanism of action of WOS and OS in nonpolar (hydrocarbon media) they are divided into corrosion inhibitors chemisorptions type donors, electron acceptors, corrosion inhibitors adsorption and dewatering agents. To protect steel pipes from corrosion in oil production, various types of surfactants [3–10] currently used are such IR as "Affinor," "SNPCH-1004." These reagents have proved decisive; however, the problem is the reagents' high expense. That is why there raises the question of development of new chemicals that exhibit not inferior protective qualities of new IR corrosion inhibitors and their costs are lower. In this regard, the aim of this work is to reduce the human-made hazards in the oilfield equipment, through the use of waste coolant concentrate obtained by membrane separation as a corrosion inhibitor.

2. Methods

The concentration of oil, grease and nonionic surfactant was determined by corrosion inhibitor spectrometry using "Kontsentratomer KN-3". The water content and solids was determined gravimetrically. The PH value (pH) of concentrate was determined by the potentiometric method. The density of the inhibitor was found using a technique based on GOST 3900 "Oil and oil products methods for determination of density". The essence of the method consisting in immersing the hydrometer in the tested liquid is to consider the effects of the hydrometer scale on the temperature determination and based on the results of the density at 20 °C. The concentrate solubility was determined in the following manner. Tests to determine the solubility, dispensability are performed in a transparent glass container with a stopper chemical at room temperature. They are prepared in the test inhibitor mixture solvent of various concentrations (0 to 50 %). For this purpose, the mixture is shacked and stirred vigorously to be dissolved or formed as emulsion structure. Both indicators are evaluated visually against a sheet of white paper in transmitted light periodically for a time interval of 1 hour to 7 days. Determination of the rate of corrosion and protection of metal with the use of waste coolant concentrate were performed according to GOST 9.506-87 "Inhibitors of metal corrosion in aqueous environments of oil" by gravimetric method. The method consists in determining the weight loss of metallic specimens during their stay in the inhibited and uninhibited test environments, followed by evaluation of the protective ability of the inhibitor to change the

rate of corrosion. The aggressive environment conditions as produced in water-oil fields should be considered. As the metal samples we used metal plates of rectangular shape of dimensions of $70\times35\times0.5$ mm which are made of "Steel 20" used for pipelines to pump oil. To activate the surface of the samples before the test, the plates were immersed for 1 minute in a 15 % solution of hydrochloric acid (HCl), then washed thoroughly with running water and distilled water and dried with filter paper. Just before the samples are weighed on analytical scales with an accuracy test of 0.0001, the samples were hung on the suspension and placed in a glass beaker with the test environment. The solutions were stirred by magnetic stirrers to create dynamic conditions. As used test medium, inhibited and uninhibited water reservoir model prepared according to the composition described in GOST 9.506-97 was used of volume of 130 cm^3. Such an inhibitor of corrosion was added to the emulsion concentrate of the spent coolant, and the corrosion tests were carried out for 24 hours to determine the mass loss of samples purified from surface corrosion products as gasoline, alcohol, and soft elastic corrosion, washed thoroughly with tap water and distilled water, dried with filter paper and weighed on a mathematical scales with an accuracy of 0.0001 g.

3. Results and discussion

Used coolant constitutes 3–10 % soluble, that involves industrial oils, as an idol, sodium nitrite, ethylene glycol, and other materials. In the machine construction companies, they use a coolant with the short life service which can be employed for a several month. After cleaning by using membrane methods [11–14], the waste coolant are used as a corrosion inhibitor. Their composition is presented in Table 1 for the second usage.

Table 1. Component composition of the waste coolant concentrate

No.	Ingredients	Content (%) ($P = 0.95$, $n = 2$)
1	Humidity	4.56
2	The organic part (oil and non-ionic surfactant...)	94.7
3	Contamination	0.74
Total		100

Currently used corrosion inhibitors are composed mainly of nonionic surfactant-quaternary ammonium; phosphites and phosphates; alkyl- and acryl zinc and other metals, nitrate oil, oxidized petrolatum, complete and incomplete (acid) esters oxypropylation and oxy-methylation products. For comparison, the main properties of the spent coolant concentrate with commercially available corrosion inhibitors that are studied are the following physicochemical properties: pH, density, and solubility. The results of studies of physical and chemical properties compared with the most commonly used in the oil industry corrosion inhibitors are listed in Table 2.

Table 2. Physico-chemical properties of corrosion inhibitors

Corrosion inhibitor	Index		
	pH	Density (g/cm^3)	Solubility
Concentrate spent coolant	6.9±0.1	0.940±0.047	Emulsifies in water
"SHNPH"	6–9	0.880±0.044	Soluble in water, alcohols
"TNHC-7"	6–9	0.890±0.045	Soluble in water, alcohols
"NAPOR-1007"	6–9	0.890±0.045	Soluble in aromatic hydrocarbons, alcohols

Results showed that the investigated indicators concentrate waste emulsion coolant characteristics are close to the corrosion inhibitor properties of the brand "NAPOR-1007".

Consequently, because of the similarity of physical and chemical properties, it was used as infrared and conducted tests on corrosion protective properties. Table 3 presents the model of produced water with a corrosion inhibitor and without inhibitor ($CaCl_2 \cdot 6H_2O$; $MgCl_2 \cdot 6H_2O$; NaCl, $CaSO_4 \cdot 6H_2O$), prepared according to GOST 9.506-87, see Table 3.

Table 3. Properties of the model of water formation

pH	ρ (g/cm^3)	Specific conductivity (mS/cm)	I (mol/dm^3)	Content (mg/dm^3)		
				Fe_{total}	O_2	H_2S
6.0±0.1	1.12	194±19	3.51	0.7±0.1	6.2±0.9	305±61

Indicators for the model with the produced water show the weakly acidic medium, which may be the result of the high content of hydrogen sulfide and sulfides. These conditions are sufficient to identify the corrosion properties of the medium, which has been investigated in our tests. According to the literature [14], the effectiveness of corrosion inhibitors can increase the input of the latter hydrophobic reagent. This condition results in reducing the hydrophobic character in polypropylene with the content range of 1 to 30 %. For application according to [15] it can be used in oil pipes under the earth as inhibitor of polyethylene corrosion. The comprehensive study shows that using these material is useful for protection of pipes. In other words, by calculation of total strain energy as well as the thermoelastic energy of material before any fracture or corrosion grows in the advanced materials, these inhibitors of polyethylene corrosion can prevent it from damage [16–18]. The results indicate that the optimal condition for the coolant are 9 pH and 10 % polypropylene glycol (PPG). With these conditions the test results show that the protection level for the steel 20 is about 78 %. We conducted comparative laboratory corrosion tests

with corrosion inhibitor brands "SHNPH" and "TNHS-7" which are applied in oil gathering pipeline system, see Table 4.

Table 4. Results of laboratory corrosion tests of corrosion inhibitors

Corrosion inhibitor brand	Corrosion inhibitor dosage	The mean corrosion ratio (mm/year)	Protection degree (%)
"TNHS-7"	3.8	0.0383	72.0
"SHNPH"		0.0209	84.7
Concentrate spent coolant		0.0304	77.8

4. Conclusion

Due to hazards and environmental problems in the oilfield equipment, we developed an approach using the waste coolant which creates membrane separation as a corrosion inhibitor. The high degree of protection observed in the corrosion inhibitor brands "SHNPH" and modified concentrate coolant on the level of protection were only slightly inferior with respect to the present corrosion inhibitors. Results showed that the worst performance of corrosion and protection degree of the steel for the corrosion inhibitor "TNHS-7" brand occurs at a dose of $3.8\,\mathrm{g/dm^3}$. The results demonstrated the effectiveness of corrosion inhibitors for steel oilfield equipment.

References

[1] A. S. Yaro, K. R. Abdul-Khalik, A. A. Khadom: *Effect of CO2 corrosion behavior of mild steel in oilfield produced water*. J Loss Prevention in the Process Industries *38* (2015), 24–38.

[2] H. M. Jol, D. G. Smith: *Ground penetrating radar surveys of peatlands for oilfield pipelines in Canada*. J Applied Geophysics *34* (1995), No. 2, 109–123.

[3] J. Biomorgi, S. Hernández, J. Marín, E. Rodriguez, M. Lara, A. Viloria: *Internal corrosion studies in hydrocarbons production pipelines located at Venezuelan Northeastern*. Chemical Engineering Research and Design *90* (2012), No. 9, 1159–1167.

[4] M. J. A. Prince: *Experimental study on nonionic surfactants for minimizing surface adsorption as an improved oil recovery (IOR) process*. Indian Journal of Science and Technology *7* (2014), No. 6, 1–4.

[5] S. Paranthaman: *Overcoming the low efficiency of box type solar cookers*. Indian Journal of Science and Technology *8* (2015), No. 2, 82–84.

[6] M. M. Osman, R. A. El-Ghazawy, A. M. Al-Sabagh: *Corrosion inhibitor of some surfactants derived from maleic-oleic acid adduct on mild steel in 1 M H2SO4*. Materials Chemistry and Physics *80* (2003) No. 1, 55–62.

[7] Z. A. Hamid, T. Y. Soror, H. A. El-Dahan, A. M. A. Omar: *New cationic surfactant as corrosion inhibitor for mild steel in hydrochloric acid solutions*. Anti-Corrosion Methods and Materials *45* (1998), No. 5, 306–311.

[8] G. Latha, S. Rajeswari: *Evaluation of non-ionic surfactants as corrosion inhibitors in seawater*. Anti-Corrosion Methods and Materials *43* (1996), No. 5, 19–22.

[9] N. Pebere, M. Duprat, F. Dabosi, A. Lattes, A. De Savignac: *Corrosion inhibition study of a carbon steel in acidic media containing hydrogen sulphide and organic surfactants.* J Applied Electrochemistry *18* (1988), No. 2, 225–231.

[10] M. L. Free: *Understanding the effect of surfactant aggregation on corrosion inhibition of mild steel in acidic medium.* Corrosion Science *44*, (2002), No. 12, 2865–2870.

[11] D. D. Fazullin, G. V. Mavrin, R. G. Melkonyan: *Removal of oil products and phenol from waste water by composite sorbents under dynamic conditions.* Chemistry and Technology of Fuels and Oils *50* (2014), No. 1, 88–94.

[12] D. D. Fazullin, G. V. Mavrin, M. P. Sokolov: *Utilization of waste lubricating-cooling fluids by membrane methods.* Chemistry and Technology of Fuels and Oils *51* (2015), No. 1, 93–98.

[13] D. D. Fazullin, G. V. Mavrin, I. G. Shaikhiev: *Particle size and zeta potential changes in the disperse phase of water-emulsified waste waters in different treatment stages.* Chemistry and Technology of Fuels and Oils *51* (2015), No. 5, 501–505.

[14] K. C. Chang, M. H. Hsu, H. I. Lu, M. C. Lai, P. J. Liu, C. H. Hsu, W. F. Ji, T. L. Chuang, Y. Wei, J. M. Yeh, W. R. Liu: *Corrigendum to "Room-temperature cured hydrophobic epoxy/graphene composites as corrosion inhibitor for cold-rolled steel" [Carbon 66 (2014) 144—153].* Carbon *82* (2015), 611.

[15] C. H. Zhang, Y. C. Zhao, M. Björling, Y. Wang, J. B. Luo, B. Prakash: *EHL Properties of polyalkylene glycols and their aqueous solutions.* Tribology Letters *45* (2012), No. 3, 379–380.

[16] S. M. N. Mehrian, A. R. Daneshmehr, A. Hadi, A. Atf: *Presentation of new model to calculate strain energy release rate for a crack in composite structure.* Caspian Journal of Applied Sciences Research *2* (2013), No. 5, 26–34.

[17] S. M. N. Mehrian, M. H. Naei, S. Z. Mehrian: *Dynamic response for a functionally graded rectangular plate subjected to thermal shock based on LS theory.* Applied Mechanics and Materials *332* (2013), 381–395.

[18] M. R. Vaziri, S. M. N. Mehrian, M. H. Naei, Y. J. S. Ahmad : *Modification of shock resistance for cutting tools using functionally graded concept in multilayer coating.* J Thermal Science and Engineering Applications *7* (2015), No. 1, 011014-8.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 307–314

Sorption properties of carbon waste pyrolysis product for biological wastewater treatment

Ilnar A. Nasyrov[1], Stanislav V. Dvoryak[2], Ildar G. Shaikhiev[3]

Abstract. Influence of pH, temperature and time of contact for adsorption of heavy metals are studied. Experiments on model solutions of heavy metals ions at 20 °C against constant pH value are provided. From the obtained data, isotherms of sorption are constructed and sorption size is calculated. Dependence of sorption extent of heavy-metal ions on volume of the past model solution is studied. Also, the influence of temperature and time on the sorption of heavy-metal ions is investigated. It is found that the adsorption time is 15 minutes for copper ions and 10 minutes for the iron and chromium ions. When the temperature rises to 60 °C, the time of sorption of heavy-metal ions is reduced insignificantly (on 5–10 min). According to the obtained results, the maximum degree of adsorption on iron ions (III) was 95.9 %. After passing through the sorbent model solution obtained in 4 dm^3 volume, the degree of adsorption decreased to 50.9 %.

Key words. Silt rainfall, pyrolysis, sorbent, waste, heavy-metal ions.

1. Introduction

Urban agglomeration with domestic and industrial human activities generate liquid waste in the form of wastewater that is taken off into the sewer. Purification of domestic and industrial wastewater is an urgent problem for urban areas. Sewage treatment plants come on, performing cleaning stages. Steps of wastewater treatment are traditional and include the following main operations [1]: mechanical wastewater treatment in primary sedimentation tanks, biological treatment in aeration tanks, cleaning of suspended solids of activated sludge in secondary sedimentation tanks, advanced treatment of water, and disinfection. During the passage of wastewater through the purification steps in treatment plants, there is formed sludge sediment in the majority, not amenable to any processing except dewatering sludge

[1]Federal University, Naberezhnye Chelny, Prospect Mira 68/19, Russia, 423810
[2]Lomonosov Moscow State University, Moscow, Leninskiye Gory, GSP-1, Russia, 119899
[3]Kazan National Research Technological University, Kazan, Karl Marx street 68, Russia, 420015

on fields in vivo [2]. This process takes a long term and large areas under sludge. Also, storage of silt rainfall leads to the spread of negative gas background and does not preclude the pollution of soil, surface water and groundwater, vegetation by toxic components that are part of rainfall. Russian Federation annually produces more than 2 million tons of sewage sludge regarding dry matter [3]. Large amounts of sediment, multicomponent, and the presence of their composition in heavy metal compounds, along with other pollutants, as well as the lack of appropriate recycling technologies lead to their accumulation and, consequently, the rejection of land for storage. Existing methods for recovery and recycling of deposits of biological wastewater suggested in the past as a source for producing activated carbon 4–6, complex sorption materials intended to remove heavy-metal ions, bioremediation of oil-contaminated soil, and use of sewage sludge as fertilizer. Application of sewage sludge to produce activated carbon was advantageous due to the high carbon content in the dry matter of sludge and low mass loss during the carbonization. During pyrolysis, a solid residue - Pyro Carbon is produced or its mineral compositions. Research has shown that the pyrolysis product is an available sorption material, the cost of which is much lower than that of industrial sorbents. The most expedient use of the resulting material as a sorbent is at a local emergency spills of oil and oil products, as well as for deep purification of sewage, and biological treatment of the past. Organic and mineral composition can be used for remediation and detoxification of waste sludge and canned cards. Introduction of them to the sludge storage area will contribute to the structuring of processes and humification of the emerging soils. The pyrolysis of sewage sludge with precipitation fields at 775–825 °C allows obtaining activated carbon with a high specific surface area and microporosity. Note that during the carbonization of dry matter of sewage sludge, the maximum amount of volatile compounds is distilled at 265–330 °C, and the carbonized precipitate is formed, when temperature without air access reaches 420–655 °C. Recycling sludge in this way allows getting rid of the sludge pit and improving the quality of wastewater.

2. Experimental part

To determine the optimum pH range in a series of conical flasks, 1 g of sorbent, and 250 ml of model solution corresponding heavy metal ion at a concentration of 1–50 mg/dm^3 were taken, (the pH factor being adjusted from 2 to 12), capped and mixed in a shaker for 2 hours under thermostatic control (20±0.1 °C). Next, the filtrate was separated from the sorbent and we determined the initial and final concentration of heavy metals ions by atomic absorption spectroscopy with a spectrometer marks "QUANT-Z". Sorption time varied between 10 and 120 minutes. The concentration of heavy metal ions was 1–50 mg/dm^3. When studying the effect of temperature on the degree of extraction, it ranged from 20 to 60 °C. At the end of each experiment, the degree of sorption was calculated. The data obtained were used to construct graphs "the degree of sorption - time" that make possible to determine the optimal time needed to achieve the maximum degree of sorption. To obtain sorption isotherms in series conical flasks, we used 1 g of sludge pyrolysis

product, added 250 ml of a solution containing heavy metal ions with concentrations in the range 1 to 2000 mg/dm^3, capped and stirred in a shaker. It was found that at this ratio and weight linkage metal concentration in the solution, there was reached equilibrium 15 minutes after the start of the experiment [7]. The solution was then separated from the sorbent and its concentration was determined in the investigated metal.

3. Results and discussion

In this paper, we studied the particle size, structure and elemental composition of activated carbon sorbent, see Fig. 1. Then we investigate the sorption properties of carbon sorbent in static conditions. The study results indicate that the sorbent has a porous structure with a grain size of 200–500 microns, preferably rectangular. The porous structure allows predicting the sorption properties of carbon sorbent obtained by pyrolysis of the sludge. The results of the study of the elemental composition of the resulting carbon sorbent are shown in Table 1.

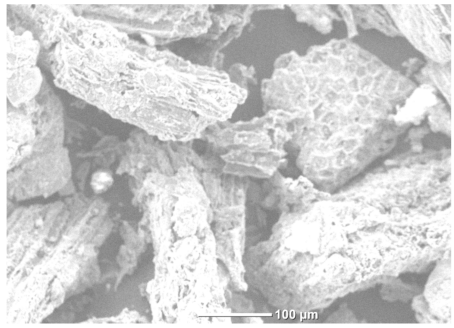

Fig. 1. Activated carbon sorbent received as a result pyrolysis of silt rainfall

The high content of organic compounds in the dry matter of the silt rainfall leads to the total content of carbon and oxygen higher than 86 % of the elemental composition of the analyzed sorption material. The mineral part of carbon sorbent accounts for less than 14 % of the elemental composition. A laser particle size analyzer brand "Microsizer 201C" defined particle size distribution. According to a granulometric analysis of sorbent it consists of 34.1 % of particles with sizes ranging from 200 to 300 microns, 59.8 % of particles with sizes from 300 to 600 microns and 3.1 % of particles with sizes from 10 to 200 microns. Later we determined sorption characteristics of the resulting product of the pyrolysis of silt rainfall. Figure 2 depicts the percentage of heavy metals including Fe (III), Cu (II) and Cr (VI) in various concentration.

Table 1. Elemental composition of carbon sorbent

Element	C	O	Ca	Fe	Si	P	Al	K	Na	Mg	S
Mass fraction (%)	59.18	26.87	4.02	2.83	2.44	1.78	1.27	0.48	0.18	0.53	0.42

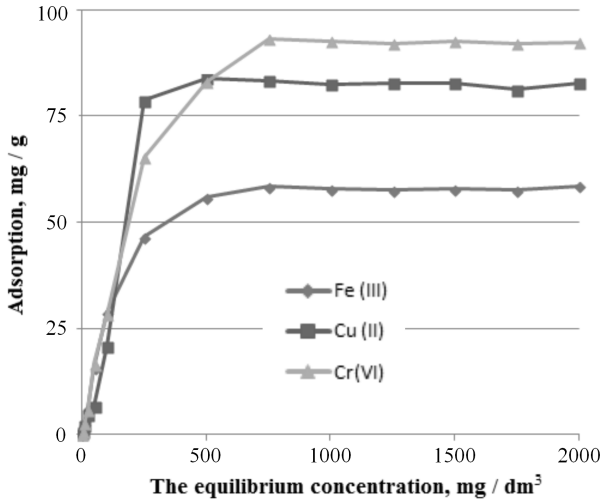

Fig. 2. Adsorption isotherms of iron ion, copper and chromium on obtained sorbent

Based on the values of calculated sorption value (a, (mg/g)) according to the formula [7]

$$a = \frac{(C_0 - C_1) \cdot V}{m_c}, \tag{1}$$

where V is the solution volume (cm^3 and m_c is the sorbent mass (g), we can draw the dependencies of the adsorption on the equilibrium concentration. Experiments were carried out on the model solutions at 20 °C with the constant value of pH. Data for the analysis of the isotherms in Fig. 2 give grounds to assert that the saturation point of the projection on the y-axis indicates the magnitude of the sorption capacity of carbon sorbent on heavy metal ions [8]. It was determined that the maximum adsorption capacity for Cu_2^+ ions is 84 mg/g, for Fe_3^+ ions it is 59 mg/g, and for Cr_6^+ ions it is 93 mg/g. For comparison, Table 2 shows the adsorption capacity of activated carbon (BAU-A) and activated carbon fiber before treatment (AYV) and after treatment with nitric and sulfuric acids (OAYV-sulfuric and OAYV-nitric) [9].

Table 2. Sorption characteristics of carbon adsorbents

Adsorbent	Conductivity adsorption (mg/g)		
	$Cr_2O_7^{2-}$	Fe^{3+}	Cu^{2+}
BAU-A	28.0		3.3
AYV	36.0		12.2
OAYV-sulfuric	37.2	13.5	12.7
OAYV-nitric	42.0	10.2	16.6

According to the classification [10], the obtained isotherm belong to class L (Langmuir isotherms), L-2 type. This type of isotherms (L) is characterized by the independence of the heat of adsorption on the surface coverage and the lack of competition from the solvent. With the increase in the employment share of adsorption sites, adsorptive molecules can with more difficulty find a vacant place, especially if they are prone to the formation of large associates to increase their concentration in the solution due to the intermolecular interaction. L-2 type isotherm characteristic at a certain concentration reaches the saturation adsorption of the adsorptive. Table 3 shows the Langmuir equation constants (K_{L}) for the adsorption of iron ions, copper, and chromium on pyrolysis of the silt rainfall product.

Table 3. Parameters of Langmuir equation for sorption of heavy metal ions on sorption material

heavy metal ions	a_m (mg/g)	$K_{\mathrm{L}} \cdot 10^{-3}$
Fe^{3+}	58.6	1.39
Cu^{2+}	83.8	1.90
Cr^{6+}	93.0	2.62

To select the optimal conditions for the sorption, we carried out studies of the sorption characteristics of carbon sorbent in the static mode, determined by the effect of pH, temperature, and time on the processes of sorption by ions heavy metal. Figure 3 illustrates the effects of pH percentage on absorption of heavy metals.

The degree of adsorption (R (%)) was calculated by the formula

$$R = \frac{C_0 - C_1}{C_0} \cdot 100\,, \tag{2}$$

where C_0 is the initial concentration, (mg/dm^3) and C_1 is the concentration of heavy metal ions (mg/dm^3) after sorption.

Based on the results of the experiment we plotted graphs $R = f(\mathrm{pH})$ and determined the region of optimal acidity values in them, in which the adsorption of heavy metal ions is maximized. The value R for every such point was taken as the average of three independent parallel experiments. The results showed the greatest degree of sorption of heavy metal ions observed for pH ranging from 6 to 12. Increasing the effectiveness of purifying at pH>8 indicates the precipitation of metal hydroxides

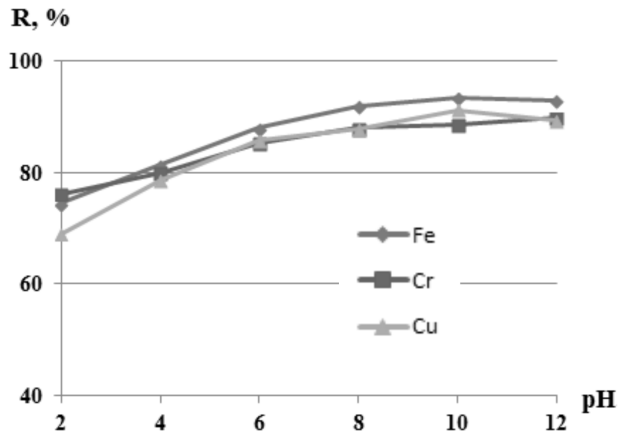

Fig. 3. Effect of pH on sorption of heavy-metal ions

and the flow of reagent purification. In the adsorption process it is influenced by factors such as time and temperature. The graphic dependence of the degree of sorption of heavy metal ions on the phase contact time at various temperatures in a neutral medium is shown in Fig. 4.

Studies of the kinetics of adsorption of heavy metal ions indicate that coal sorbent obtained by pyrolysis of the silt rainfall is characterized by a relatively high speed of the process. By evaluating the absorption time of ions by high ratios at initial times, it is found that the adsorption time is 15 minutes for copper ions and 10 minutes for the iron and chromium ions. When the temperature rises to 60 °C during the sorption of heavy metal ions, this time is reduced insignificantly (to 5–10 minutes). Increased temperature also results in a slight increase in the degree of adsorption (not more than 5 %). In an acid medium (pH < 3) we get similar results, but with a lower extent of sorption ($R(Cr^{6+}) = 82\,\%$; $R(Cu^{2+}) = 77\,\%$; $R(Fe^{3+}) = 86\,\%$). Subsequently, experiments were conducted to study the carbon sorbent sorption parameters on iron ions in dynamic conditions. For this purpose, to a series of glass columns of height 150 mm and diameter 10 mm there was placed 1.5 g of the reagent and model solution proceeded with a concentration of iron ion (III) 5 mg/dm^3.

4. Conclusion

According to the obtained results, the maximum degree of adsorption on iron ions (III) was 95.9 %. After passing through the sorbent model solution obtained in 4 dm^3 volume, the degree of adsorption decreased to 50.9 %. Therefore, it is shown that the method of recycling of deposits of biological sewage treatment by pyrolysis provides a comprehensive sorption material. Studies of the structure and sorption properties showed the possibility of using silt rainfall pyrolysis product as a sorption material for the removal of heavy metal ions from aqueous media.

*** March 8, 2017 ***

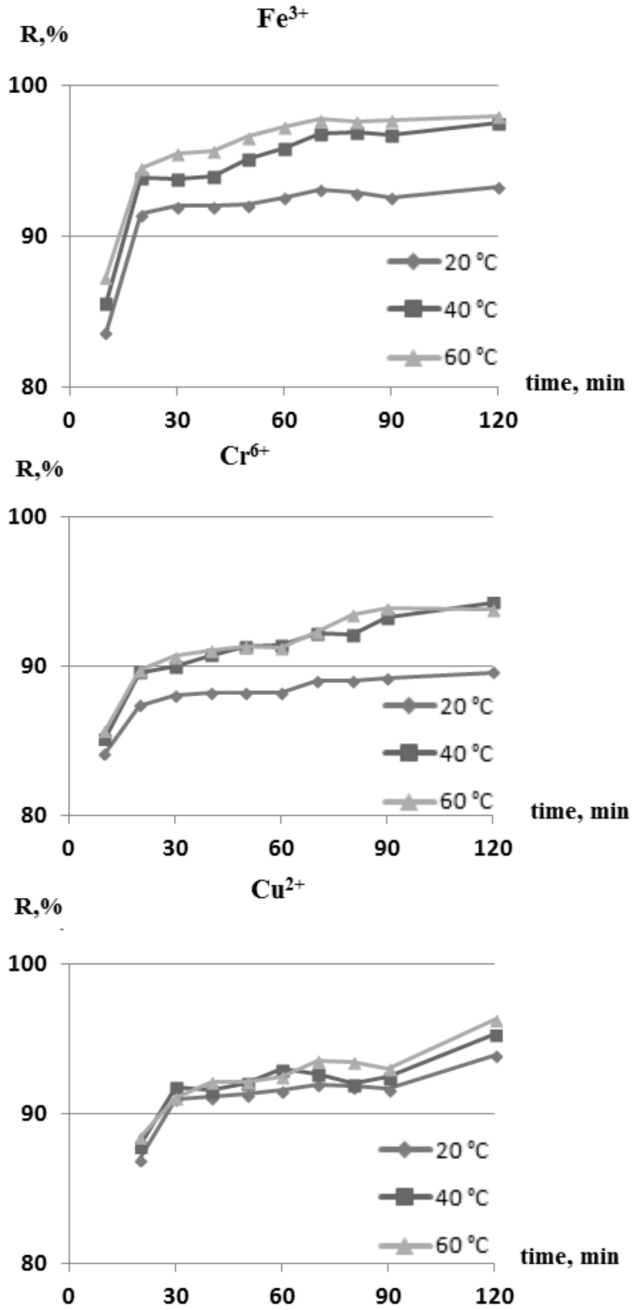

Fig. 4. Influence of temperature and time on sorption of heavy metal ions (pH = 7.0±0.1

References

[1] S. M. Seyedhosseini, M. J. Esfahani, M. Ghaffari: *A novel hybrid algorithm based on a harmony search and artificial bee colony for solving a portfolio optimization problem using a mean-semi variance approach.* J Central South University *23* (2016), No. 1, 181–188.

[2] Y. Pourasad, M. Mahmoodi-k, M. Oveisi: *Design of an optimal active stabilizer mechanism for enhancing vehicle rolling resistance.* J Central South University *23* (2016), No. 25, 1142–1151.

[3] C. Palanichamy: *A sustainable energy option to the expanding chennai metropolitan area.* Indian Journal of Science and Technology *8* (2015), No. 22.

[4] B. Mashadi, M. Mahmoodi, A. H. Kakaee, R. Hosseini: *Vehicle path following control in the presence of driver inputs.* Proc. Institution of Mechanical Engineers, Part K: J Multi-body Dynamics *227* (2013), No. 2, 115–132.

[5] R. R. Zinnatov, G. V. Mavrin, I. A. Nasyrov, M. P. Sokolov, D. A. Kharlyamov: *Sorption concentration of ions of copper (II) and lead (II) by magnetic sorbent.* Research Journal of Pharmaceutical, Biological and Chemical Sciences *6* (2015), No. 5, 1623–1628.

[6] G. H. Yunusa, A. Kassim, N. Gofar: *Effect of surface flux boundary conditions on transient suction distribution in homogeneous slope.* Indian Journal of Science and Technology *7* (2014) No. 12, 2064–2075.

[7] D. A. Kharlyamov, G. V. Mavrin, I. Y. Sippel: *About the possibility of sorption concentration of heavy metals using magnetite.* Life Science Journal *11* (2014), No. 8, 607–610.

[8] D. A. Kharlyamov, G. V. Mavrin, E. A. Danilova, R. R. Zinnatov, E. V. Prytkova, S. V. Dvoryak: *Application of a magnetic composite sorbent on the basis of woodworking waste for sewage treatment from heavy metals.* Research Journal of Pharmaceutical, Biological and Chemical Sciences *7* (2016), No. 3, 1667–1670.

[9] G. Singh, A. Perwez: *Estimation of assimilative capacity of the airshed in iron ore mining region of goa.* Indian Journal of Science and Technology *8* (2015), No. 23.

[10] G. Parfit, K. Rochester: *Adsorption from solutions on solid surfaces: Translation from English.* Publishing house Mir, Moscow, Russian. (1986).

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 315–326

Design and analysis of electromagnetic-hydraulic composite brake based on electromagnetic importing method[1]

Liu Xue Jun[2], Liu Cun Xiang[2], Lin Tu Gan[2]

Abstract. The interaction theory of the electromagnetic brake and hydraulic booster frictional brake is analyzed, and structure of vehicle brake mechanism which can be produced the braking torque by the electromagnetic brake and frictional brake, is illustrated. The electromagnetic brake and hydraulic brake can be controlled more flexibly. The problem that the electromagnetic braking torque is too small and difficult to control braking energy diversion can be solved effectively, which provides the conditions for commercial marketing of the electromagnetic-hydraulic composite brake. According to some references, braking torque of the electromagnetic-hydraulic composite brake is proposed. The simulation results show that the torque value is related to the material and current in the field coil, etc. The torque value obtained from the proposed equation approaches the value obtained from simulation.

Key words. Electromagnetic-hydraulic composite brake, electromagnetic importing, design.

1. Introduction

At present, the braking system of passenger car mainly depends on the friction braking powered by hydraulic system. However, with the continuous improvement of engine power and vehicle speed, many problems of hydraulic braking system appear that may be divided into several categories. For example, powder brought about by friction causes pollution of environment, increased wear of hydraulic system leads to higher maintenance costs and, heat fade and thermal failure of hydraulic braking system is an important reason to vehicle accident.

[1]This work is supported by the national natural science foundation of Guangxi China (2014GXNSFBA118262), Transportation science project of Guangxi Chinam(2013-100) and Transportation science project of Guangxi China(2015-261-8).

[2]Department of Automobile Engineering, Guangxi Communication Vocational & Technical Institute, Nanning, 530023, China

In order to overcome the above problems of the traditional hydraulic braking system, electromagnetic-hydraulic composite braking system becomes a new research direction. One of the issues to be solved is technical and is connected with reaching a sufficiently high electromagnetic braking torque: this requires a large coil and the space around the braked wheel is too narrow to install there such a coil.

Meanwhile, the front wheels and the electromagnetic-hydraulic composite brake mechanism must be able to move together with the wheel hub. Scholars in China and abroad just basically study the theory of electromagnetic-hydraulic composite brake [1–4], but some problem such as how to design the structure of the composite braking system and how to mount the composite brake on the vehicle are not solved yet. So, after the technology is developed and used in practice, the paper provides a new possibility how to bring magnetic field to vehicle wheels. The task is analyzed theoretically and the results may contribute to successful solution of connected practical problems.

2. Structure and function of electromagnetic-hydraulic composite brake system

In view of the difficulties with installing larger field coil into a narrow space on the wheel, this paper offers an electromagnetic-hydraulic composite brake system based on electromagnetic importing method. Due to different material requirements of electromagnetic braking system and hydraulic braking system, the electromagnetic-hydraulic composite braking system is designed as a dual disc brake. A catheter of high magnetic permeability with low residual magnetism and coercive force is applied to transfer magnetic flux of high density from the field coil to the braking disc.

The aim of technical scheme of the electromagnetic-hydraulic composite brake system is to improve the existing passenger car wheel hub and lengthen the brake disc, thus to ensure that it can be mounted without problems. Another aim is that the electromagnetic brake friction plate with bolt holes, wheel, brake disc and electromagnetic friction brake disc fastening must rotate together with the wheels, which must be realized by the steering wheel via relative rotation of the bearing section.

The basic structure of the electromagnetic - hydraulic composite brake is shown in Fig. 1.

3. Description and theory of electromagnetic-hydraulic composite brake

As shown in Fig. 1, when the vehicle brakes, according to the intensity of braking, the control unit of the braking system starts supplying appropriate current to the field coil 17. The magnetic flux generated by the coil passes through the core 18, pole 16 and catheter 15 into the the electromagnetic guide block 14. Each coil has two electromagnetic guide block, corresponding to the N and S magnetic poles (shown in

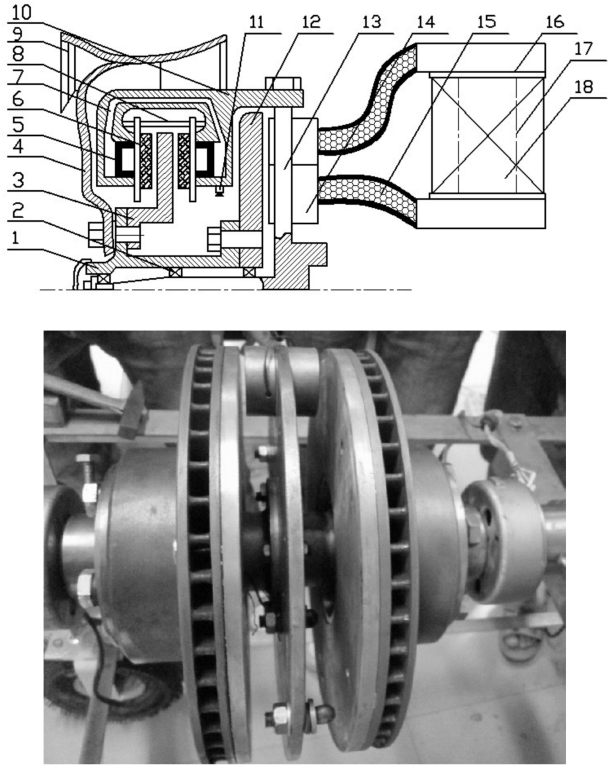

Fig. 1. Structure of electromagnetic-hydraulic composite brake: 1–hub, 2–wheel hub bearings, 3–friction brake disc, 4–wheel rim, 5–piston, 6–brake block assembly, 7–wheel cylinders, 8–guide support pin, 9–wheel plate, 10–clamp bracket, 11–inlet of wheel cylinder, 12–electromagnetic brake disc, 13–steering knuckle, 14–electromagnetic guide block, 15–electromagnetic catheter, 16–magnetic pole, 17–field coil, 18–iron core, 19–guide vane

Fig. 2). The block 14 and electromagnetic brake disc 12 is very close to one another (typically 1 mm). As the electromagnetic brake disc rotates, it will cut the force lines leading from the N-pole of the electromagnetic guide block to the S-pole, which leads to generation of eddy currents of rotational origin that produce the braking torque.

As the electromagnetic brake works, the oil from the brake master cylinder will pass through the oil inlet of the wheel cylinder 11 into the wheel cylinder 7. The cylinder pushes the piston 5 and the brake block assembly 6 starts moving towards the center of the brake pad assembly under the guide of the guide support pin 8. This leads to clamping friction brake disc 3 tightly and achieving the effect of the friction brake.

As shown in Fig. 3, the electromagnetic brake disc closed to the side of the friction braking mechanism is processed with guide vanes 19. The vanes can dissipate the heat generated by the electromagnetic brake. At the same time, the guide vane can also cool the friction brake mechanism. In order to improve vehicle's braking torque

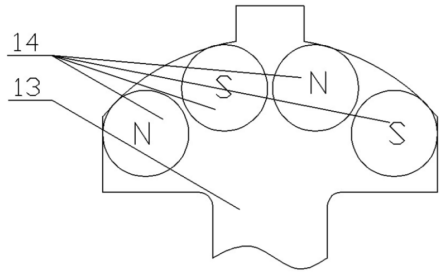

Fig. 2. Arrangement of steering knuckle electromagnetic guide block

at low speed, a copper plate of a certain thickness is installed at the electromagnetic brake disc side.

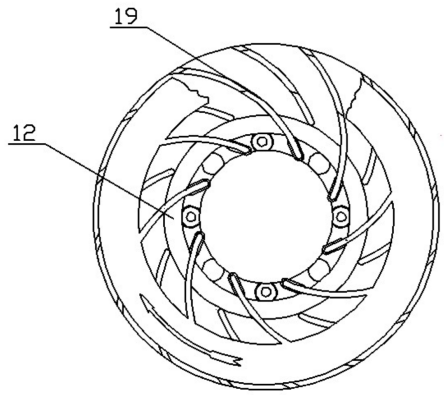

Fig. 3. Structure of electromagnetic brake disc

Figures 4 and 5 show other important parts of the composite brake.

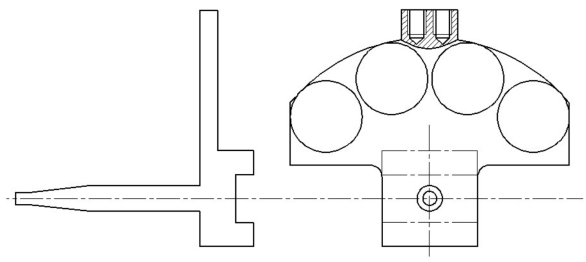

Fig. 4. Steering knuckle of composite brake

The flexible catheter (15) allows the electromagnetic brake disc and field coil to produce relative motion, which does not affect the transmission of the electromagnetic field.

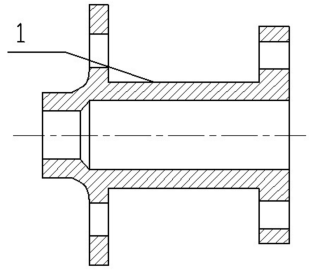

Fig. 5. Wheels hub of composite brake

4. Mathematical model of electromagnetic part of electromagnetic-hydraulic composite brake

Generally, any eddy current disc brake consists of a source of magnetic flux (produced by permanent magnet or, in our case, by electromagnet) and a well electrically conductive disc. The disc driven by an external source of mechanical energy rotates at an angular velocity ω is located very close to the magnetic poles.

The mathematical model of the electromagnetic part of the composite brake may generally be described by the following system of equations:

1. The first equation
$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{1}$$
expresses that electric field strength is produced in the electrically conductive disc moving in a time variable magnetic field of magnetic flux density $\boldsymbol{B}$.

2. The occurrence of electric field $\boldsymbol{E}$ in the rotating disc causes currents in the disc whose density $\boldsymbol{J}$ is described by equation
$$\boldsymbol{J} = \sigma \boldsymbol{E}\,, \tag{2}$$
σ denoting the electric conductivity.

3. Relation between magnetic field strength $\boldsymbol{H}$ and produced current density $\boldsymbol{J}$ is generally given by the equation
$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}\,, \tag{3}$$
but the displacement currents of density $\partial \boldsymbol{D}/\partial t$ do not play any role here and may be neglected.

4. The relation between the field vectors $\boldsymbol{J}=$ and $\boldsymbol{H}$ can be expressed as
$$\boldsymbol{B} = \mu_0 \mu_{\mathrm{r}} \boldsymbol{H}\,, \tag{4}$$

where $\mu_0 = 4\pi 10^{-7}\,\mathrm{H/m}$ is the magnetic permeability of vacuum and μ_r is the relative permeability.

5. The volumetric force $\boldsymbol{f}$ acting in the disc is given (provided that the disc is non-ferromagnetic) by the relation

$$\boldsymbol{f} = \boldsymbol{J} \times \boldsymbol{B}\,. \tag{5}$$

6. Finally, the corresponding volumetric torque $\boldsymbol{t}$ is

$$\boldsymbol{t} = \boldsymbol{r} \times \boldsymbol{f} = \boldsymbol{r} \times (\boldsymbol{J} \times \boldsymbol{B}) \tag{6}$$

and acts against the rotation of the disc, thus decelerating it.

In our case the distribution of magnetic field in the system will be solved in terms of magnetic vector potential defined by the equation

$$\boldsymbol{B} = \operatorname{curl} \boldsymbol{A} \tag{7}$$

normalized by the Coulomb condition

$$\operatorname{div} \boldsymbol{A} = 0\,. \tag{8}$$

and aalso electric potential φ

The electric field strength $\boldsymbol{E}$ in the investigated domain containing the whole system generally consists of three terms

$$\boldsymbol{E} = -\operatorname{grad} \varphi - \frac{\partial \boldsymbol{A}}{\partial t} + \boldsymbol{v} \times \boldsymbol{B} = -\operatorname{grad} \varphi - \frac{\partial \boldsymbol{A}}{\partial t} + \boldsymbol{v} \times \operatorname{curl} \boldsymbol{A}\,, \tag{9}$$

where the first term denotes the component coming from the external source, the second term is produced by time variations of magnetic fields and the third term (existing only in the disc) is produced by its rotation. Symbol $\boldsymbol{v}$ denotes the velocity.

After inserting the above equation into (7), using of simplified equation (3) and (4) provides

$$\operatorname{curl}\, \left(\mu^{-1} \operatorname{curl} \boldsymbol{A}\right) = \gamma \left[-\operatorname{grad} \varphi - \frac{\partial \boldsymbol{A}}{\partial t} + \boldsymbol{v} \times \operatorname{curl} \boldsymbol{A}\right]. \tag{10}$$

which represents the general equation describing the field in the system consisting of the electromagnet (containing the field coil, magnetic circuit, catheter and poles) and disc.

This equation contains, however, two unknowns: $\boldsymbol{A}$ and φ, which means that another, supplementary, equation has to be added. This equation is represented by the zero divergence of the current density at any point of the system, i.e.

$$\operatorname{div} \boldsymbol{J} = \operatorname{div}\, \left(\gamma \left[-\operatorname{grad} \varphi - \frac{\partial \boldsymbol{A}}{\partial t} + \boldsymbol{v} \times \operatorname{curl} \boldsymbol{A}\right]\right) = 0\,. \tag{11}$$

Both equations must be supplemented with correct boundary condition along a sufficiently distant boundary. The condition for the magnetic vector potential is of the Dirichlet type, the condition for the electric vector potential is of the Neumann type.

In the rotating disc itself, the source currents are zero, so that the field may be described just by the equation containing only the magnetic vector potential in the form

$$\text{curl } \left(\mu^{-1}\text{curl } \boldsymbol{A}\right) = \gamma \left[-\frac{\partial \boldsymbol{A}}{\partial t} + \boldsymbol{v} \times \text{curl } \boldsymbol{A}\right] . \tag{12}$$

and in case that the disc is nonferromagnetic and magnetic field does not vary in time, there holds

$$\text{curl } (\text{curl } \boldsymbol{A}) = \mu_0 \gamma \left[\boldsymbol{v} \times \text{curl } \boldsymbol{A}\right] , \tag{13}$$

and after using identity

$$\text{curl } (\text{curl } \boldsymbol{A}) = \text{grad div } \boldsymbol{A} - \Delta \boldsymbol{A} = -\Delta \boldsymbol{A}$$

with respect to the Coulomb condition (8) we finally have

$$\Delta \boldsymbol{A} + \mu_0 \gamma \left[\boldsymbol{v} \times \text{curl } \boldsymbol{A}\right] = \boldsymbol{0} . \tag{14}$$

Particular steps of the general solution include:

1. Computation of the time-dependent distributions of potential $\boldsymbol{A}$ and φ using (10) and (11).

2. Computation of the time-dependent distribution of magnetic flux density in the disc using (7) and computation of time-dependent distribution of density of induced eddy currents ($\boldsymbol{J} = \gamma \boldsymbol{v} \times \boldsymbol{B}$).

3. Computation of the time-dependent distribution of the volumetric forces $\boldsymbol{f}$ in the radial cross-section in the disc using (5).

4. Computation of the time-dependent distribution of the volumetric torques $\boldsymbol{t}$ in the radial cross-section in the disc using (6).

5. Computation of the total drag torque using the formula

$$\boldsymbol{T} = \int_V \boldsymbol{t} \, \mathrm{d}V , \tag{15}$$

where V is the volume of the rotating disc.

5. Simulation results

As the general 3D model of the braking system is rather complicated to solve, we decided to use a simplified 2D computational arrangement. The model was solved by the code Ansoft Maxwell 2D 10.0.

First, it is necessary to initialize the construct environment, and create a new project name through PROJECTS control in the software. Although the Maxwell software can construct model diagram using its own modeler, it is easier to apply AutoCAD software to build an electromagnetic - hydraulic composite brake model. Through Maxwell module TRANSLATOR controls, the AutoCAD DXF files are transformed into SM document format that Maxwell software can identify.

The principal parameters of the electromagnetic-hydraulic composite brake are:

- The distance r_{ab} from the center of the brake disc to the center of the electromagnetic coil is 0.08,m.
- The thickness of brake disc is 8 mm.
- The distance from the center of the electromagnetic guide block to the center of the electromagnetic brake disc is 2 mm.
- The field coil of the electromagnetic brake is made of copper.
- The brake disc is also of copper.
- The material of the electromagnetic guide block and electromagnetic catheter is NdFe35.
- The properties of soft magnetic material are nonlinear.
- The material of background is air.
- The number of turns of the field coil is 4000.

The simulation results are shown in Figs. 6–9. Figure 6 shows the discretization mesh covering the electromagnetic-hydraulic composite brake. Figure 7 shows the distribution of magnetic field in the system for field current 15 A. In this case, the maximum value of magnetic flux density B is 0.7 T (current density J being $8.78 \times 10^3\,\mathrm{A/cm^2}$). According to the simpler formula (20), the braking torque of the brake $T_{\mathrm{b}} = 1780.61\,\mathrm{N\,m}$, while code Ansoft Maxwell provides the force acting on the brake disc $F_{\mathrm{bs}} = 1780.61\,\mathrm{N}$, so that the simulated torque $T_{\mathrm{bs}} = 142.45\,\mathrm{N\,m}$.

Figure 8 shows an analogous magnetic field distribution for field current 20 A. Now the maximum B is 0.93 T (current density J being $1.24 \times 10^4\,\mathrm{A/cm^2}$), which is more larger than in the previous case. According to (20), the braking torque of the electromagnetic-hydraulic composite brake disc $T_{\mathrm{b}} = 239.72\,\mathrm{N},$. Ansoft Maxwell provides the force acting on the brake disc $F_{\mathrm{bs}} = 3165.5\,\mathrm{N}$ so the simulated torque $T_{\mathrm{bs}} = 253.24\,\mathrm{N\,m}$.

For comparison, Fig. 9 shows the distribution of magnetic flux density also for field current 20 A, but now the brake disc is made of iron. In this case, the maximum value of B is 0.67 T, which is substantially smaller than maximum value of B in Fig. 8. This suggests that the ideal material for the brake disc is cooper.

Figure 10, shows the dependence of brake torque on the field current. From It is clear that the torque grows with increasing torque. Moreover, the torque calculated from simulation is somewhat higher than the torque determined from the simpler formula (20), but the difference does not extend about 13 % (for field current 10 A).

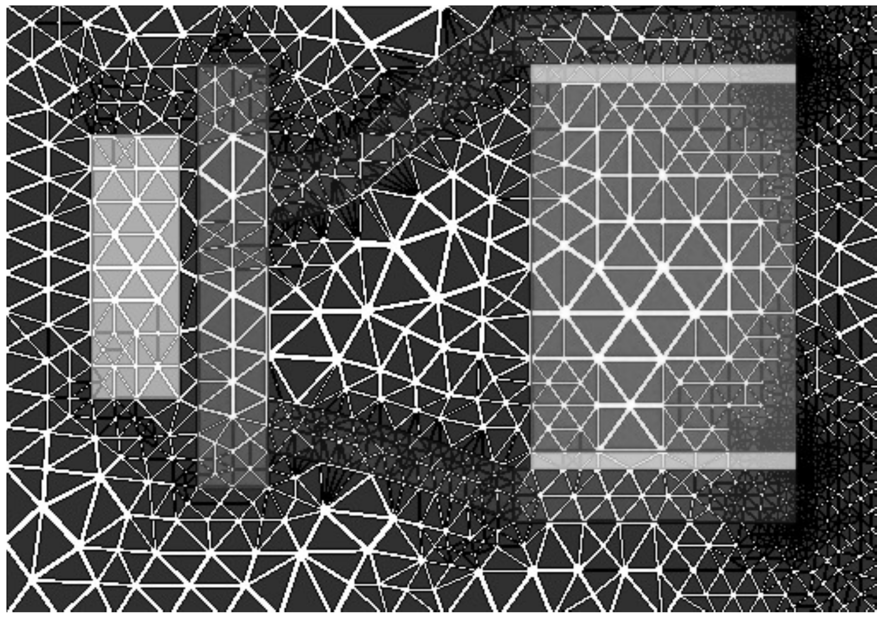

Fig. 6. Discretization mesh

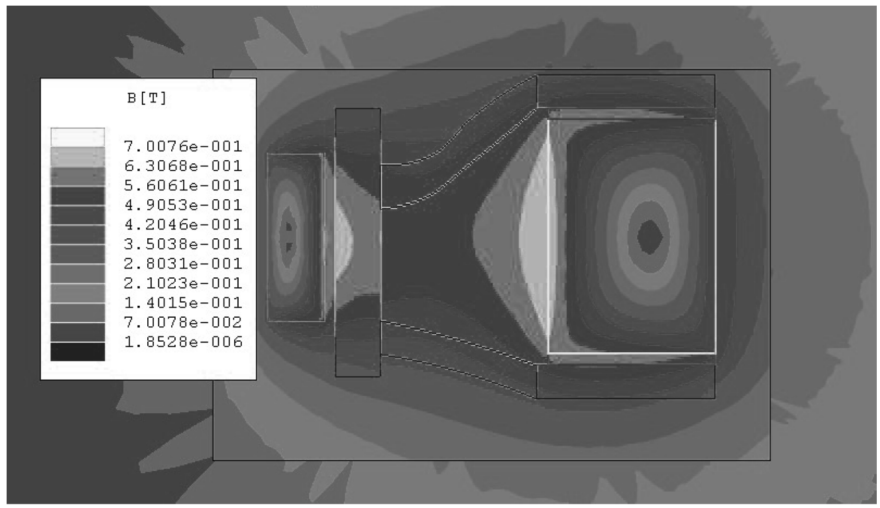

Fig. 7. Magnetic flux density distribution for current $I = 15$ A, disc made of copper

6. Conclusion

As the double-disc brakes use different materials, friction brake and electromagnetic brake can exhibit their particular advantages. The electromagnetic-hydraulic composite brake reduces the wear of the brake system, maintenance costs, danger of

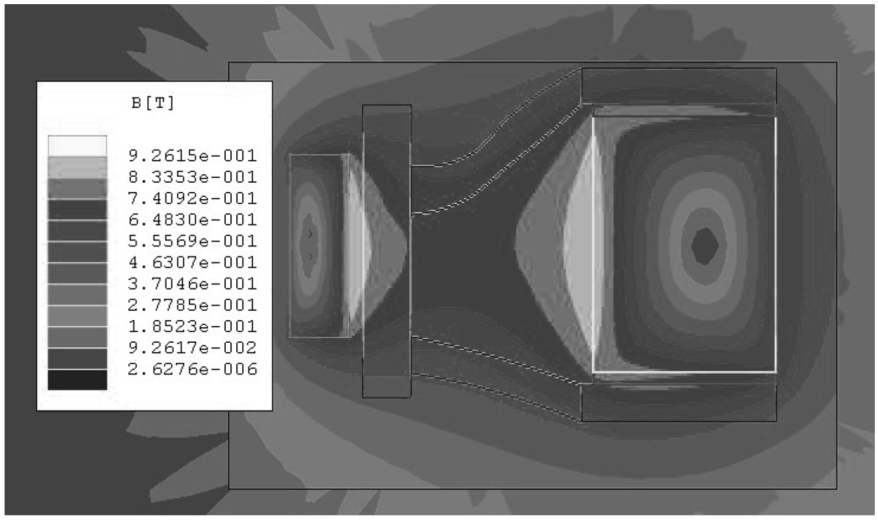

Fig. 8. Magnetic flux density distribution for current $I = 20$ A, disc made of copper

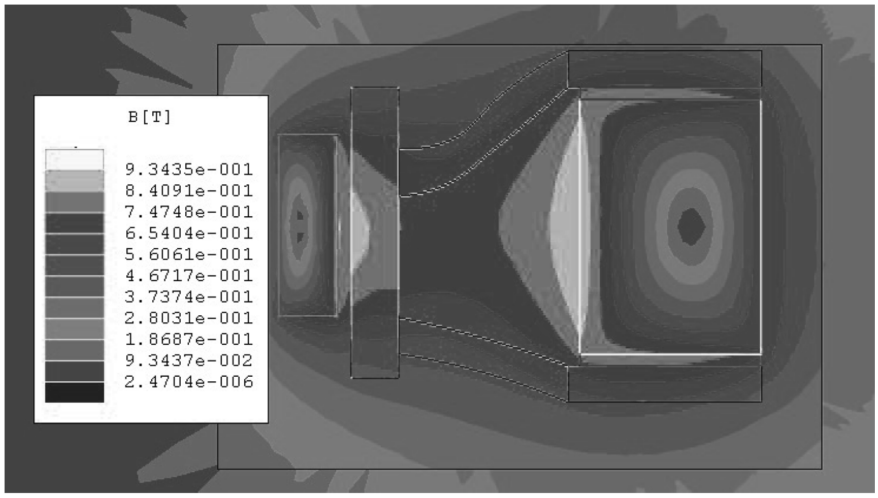

Fig. 9. Magnetic flux density distribution for current $I = 20$ A, disc made of iron

the car's thermal failure, rate of water aging and, moreover, it increases security of the braking system.

Thus, overcoming the shortcomings of the existing hydraulic braking systems using the hydraulic system and non-contact electromagnetic brake system made of composite represents the present direction of development of the braking systems.

Based on the mathematical model, the braking torque of the electromagnetic - hydraulic composite brake is calculated. The computations are carried out by applying the software Ansoft Maxwell 2D/3D. The results show that the torque

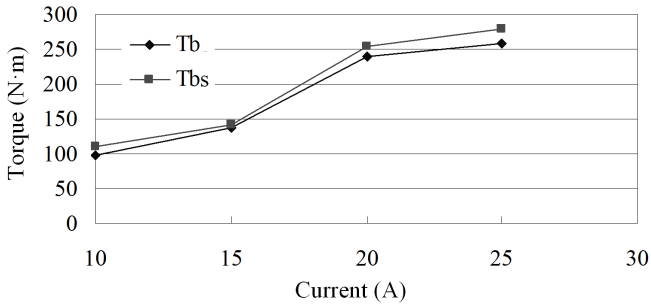

Fig. 10. Torque versus applied field current

grows as the current increases, and the torque value obtained from (20) approaches the value obtained from simulation.

References

[1] S. M. Savaresi, M. Tanelli: *Active braking control systems design for vehicles.* Springer Science & Business Media, Springer London, Dordrecht, Heidelberg, New York, 2010.

[2] A. Sakai, Y. Sakai, A. Otomo, F. Kawahata: *Toyota braking system for hybrid vehicle with regenerative system.* Proc. 14th Int. Electric Vehicle Symposium (EVS-14), 11–15 Dec. 1997, Orlando, FL, USA,

[3] K. Lee, K. Park: *Modeling eddy currents with boundary conditions by using Coulomb's law and the method of images.* IEEE Trans. Magn. *38* (2002), No. 2, 1333–1340.

[4] M. Hara, T. Nihei, N. Matsui, M. Matsuura: *Braking force control apparatus for a vehicle.* US Patent US6957874 B2, 2005.

[5] K. Karakoc, E. J. Park, A. Suleman: *Improved braking torque generation capacity of an eddy current brake with time varying magnetic field: A numerical study.* Finite Elements in Analysis and Design *59* (2012), No. 1, 66–75.

[6] S. M. Jang, J. K. Kwon, S. H. Lee, B. S. Kim, H. J. Cho: *Characteristic analysis of linear eddy-current brakes.* Proc. Sixth IC Electrical Machines and Systems, 9–11 Nov. 2003, Beijing, China, 177–179.

[7] K. J. Lee, J. N. Kang, S. M. Wang, K. H. Park: *Torque analysis and optimization of an eddy current brake system.* Proc. of IMCSD 99, San José, CA, USA, 137–141.

[8] K. Karakoc, E. J. Park, A. Suleman: *Design considerations for an automotive magnetorheological brake.* Mechatronics *18* (2008), No. 8, 434–447.

Received November 16, 2016

Acta Technica *61*, No. 4B/2016, 327–336

Cement concrete composites on the basis of by-passed stone and stone milling wastes[1]

Said-Alvi Yu. Murtazaev[2], Dena K.-S. Batayev[3], Magomed S. Saidumov[4], Madina Sh. Salamanova[4], Salambek A. Aliev[4]

Abstract. The analysis of an environmental situation of the region and results of research of the specified wastes for the purpose of their use in cement concrete production engineering, expansion of a raw-material base and upgrading of road surfacing on their basis is presented. The history of processes of structuring in the course of aging of a low-slump concrete mix with use of the compounded binding materials and a waste of crushing of a demolished concrete and rocks is founded. Character of distribution of pores by the sizes of a cement stone of compounded binding materials on the basis of a waste of sawing stones is revealed. Influence of the modified fine aggregate from screenings of crushing of a demolished concrete and rocks and micro-aggregates on structure formation and the basic physical-mechanical properties of the pressed concretes for construction of cement concrete pavements is also revealed.

Key words. By-products, by-passed stones, secondary fine aggregate, floured filler, compounded binding materials, pressed concretes, concrete strength, environmental safety.

1. Introduction

One of the basic criteria of a country development level is the condition of its street and road network. Now road pavements in Russia by their overall considerably drop behind the world level. Besides, on many roads volume of traffic for this class

[1]The study has been prepared with support of the Ministry of Education and Science of the Russian Federation.

[2]Academy of Sciences of the Chechen Republic, Grozny, M. Esambaeva Avenue 13; e-mail: S.A.yumurtazaev2@gmail.com

[3]Comprehensive Research Institute named after H. I. Ibragimov of the Russian Academy of Sciences, Grozny, Staropromyslovskoye Highway 21a

[4]Groznensky State Oil Technical University named after Academician M.D. Millionshchikov, Grozny, H.A. Isaev Avenue 100

of road has increased almost in 3 times in comparison with the proof load which has been stipulated at their designing [1, 2].

The most promising solution of the problem according to Transport strategy of the Russian Federation for the period till 2020 is construction of road surfacing of a rigid type with use of cement concrete composites [3] that implies a significant increase in the volume of work. For this purpose the road-building industry should have a necessary raw-material base for production of cement concretes for arrangement of such pavements. The basic preconditions for production of effective road cement concretes are decrease in a water demand of designed compositions, raise in reactivity of binding materials with simultaneous reduction of its discharge, and also decrease in cost of produced concrete [4]. Therefore, application of especially low-slump concrete mixes with the reduced charge of cement in comparison with traditional compositions, and wide use of by-products and the compounded binding materials on their basis appear the most expedient.

Effective direction of use of industrial wastes for the solution of a problem of expansion of a raw-material base of the building industry is their complex use with provision of physical and service characteristics required for road surfacing [5–7].

2. Raw materials for cement concrete composites and research procedures

The methods applied in the course of carrying out researches within the frameworks of the study are based on known postulates of the theory of clinker minerals solidification with fillers of various composition, in particular, with by-passed stones and stone millings, mathematical logic, technology of composite materials, the theory of production processes and production automation and control. Researches were carried out taking into account state standards and recommendations.

Both macro and microscopic structures of concrete composites and chemical composition of a feed stock were studied by means of a power dispersion spectrometer (PDS) and raster electron microscope Quanta 3D 200i with integrated microanalysis system Genesis Apex 2 EDS from EDAX. Spectra processing was performed by means of software EDAXTEAM.

3. Secondary filler surface modifying with use of cation-active additive ADMAX

It was noted in proceedings of Pecheny that use of pulverous fraction in concretes reduces hardness of a contact area between filler and a cement stone due to a low adhesive bond between them. At the same time Bazhenov and Shishkin have carried out the researches which have proved that dust particles do not render influence on concrete strength if they are not mechanically connected with a filler surface and even raise it upon introduction of appropriate chemical additives.

Researches of influence of modified cation-active additive ADMAX included in a filler on structure formation processes and properties of the pressed fine concrete

have found that in the course of processing the filler by cation-active surface active agent selective-oriented adsorption of its macro molecules is observed. Upon that, the polar terminal group of molecules is positively charged, and therefore in various media (neutral or alkaline) it is pulled to negatively charged surface of silicon dioxide. Simultaneously, hydrophobic hydrocarbon chains of molecules are aimed to be freed from a fluid phase and to join to each other forming the adsorptive monomolecular layer in the investigated medium whereby the surface becomes hydrophobic. These processes of adsorption provide the raised wetting and the best permeation of cement suspension in a micro relief of filler grains creating upon that the most favorable conditions for formation of the adhesive contact with the increased hardness. It is found that use of a filler modified with ADMAX improves compactness of an agglutinant sand, changes behavior of opened porosity and raises strength characteristics of the concrete (see Table 1).

Table 1. Influence of cation-active additives to properties of pressed cement concretes

Concrete additive name, content (%)	Stripping strength (MPa)	Compaction factor k_{comp}	Medium density (kg/m^3)	Water absorption (%w/w)	Compression strength R_{cs} (MPa)	Opened porosity (%)
ADMAX, 0.1 %	0.96	1.77	2312	2.51	55.8	7.6
Without the additive	0.84	1.71	2283	4.23	42.9	11.2

Besides, formation of an even number of modificator layers on a filler grain surface, as a rule, does not result in reaching the desirable results because the essential gain of compression strength R_{cs} is not observed thus. The maximum positive effect is observed upon introduction of surface active agent on a filler surface by odd quantity of layers, namely, 5 layers especially in researches within the frameworks of the present study (see Table 2).

4. Structure and properties of cement concrete composites with the modified filler

Studying of the pressed fine concrete structure with use of the filler which is superficially processed by the cation-active additive ADMAX was made by a method of radiography analysis and a raster electron microscopy (Figs. 1 and 2). It was found that use of the modified filler in the course of hydration of clinker minerals promotes synthesis of fine-crystalline new growths of low-base calcium hydro silicates.

X-ray analysis researches have shown that the microscopic structure of a cement stone in 28–day age (Fig. 2a, zoom $\times$5000) samples without additives is presented by loosened matrix, there are no new growths with the expressed habitus of crystals. The microscopic structure of samples on the basis of the modified filler (Fig. 2b) is mainly presented in the same age by a dense mass and a fine-grained phase, and

crystal hydrate which is a basic component of its microscopic structure is presented by the needle-shaped form. In such structures crystals of fibrous, needle-shaped and prismatic forms predominate to more extent. Formation of homogeneous structure due to effective intergrowth of hydro silicates in the entire matrix of solidifying system of a composite is observed.

Table 2. Key properties and hardness of the pressed cement concrete depending on a surface treatment type

No.	Quantity of monomolecular layers ADMAX on a filler surface	ADMAX concentration (%)	Water demand (%)	Water-cement ratio at equal place ability parameters	Compression strength R_{cs} (MPa)	
					7 days	28 days
1		10^{-2}, mixing with water	4.4	0.32	25.4	45.8
2	8	$1 \cdot 10^{-1}$	4.1	0.30	22.1	32.2
3	7	$7 \cdot 10^{-2}$	4.0	0.30	25.2	42.9
4	6	$5 \cdot 10^{-2}$	3.8	0.29	25.6	43.7
5	5	$1 \cdot 10^{-2}$	3.5	0.27	29.1	49.8
6	4	$5 \cdot 10^{-3}$	4.2	0.29	25.7	44.1
7	3	$3 \cdot 10^{-3}$	4.5	0.30	25.1	45.8
8	2	$1 \cdot 10^{-3}$	5.2	0.32	22.8	40.1
9	1	$6 \cdot 10^{-4}$	5.2	0.33	21.7	37.9
10			5.2	0.34	21.3	37.4

Structural particles (scales and plates) of low-base calcium hydro silicate penetrate into space between parts of the binding material that provides raise of hardness of a solidifying colloid system.

Besides, one instance of development of new highly effective technology for production of multicomponent compositions for construction of cement concrete pavements is mechanochemical activation of binding materials which allows synergistically to improve properties of components of the complex binding material with the highest filling level: hardness of cement increases by 2–3 grades, and the plasticizing effect of an organic component and the modifying additive increases approximately twice. Therefore the task in view in the study for production of the effective compositions intended for cement concrete pavements with improved properties is to develop the formula of cement concrete mixes with use of a secondary material for a road making.

Thus, filled binding materials (FBM) are received with filler from a waste of sawing of stones (see Table 3). A Portland cement TSEM I 42, 5 H is used in the capacity of a binding material in the course of their production.

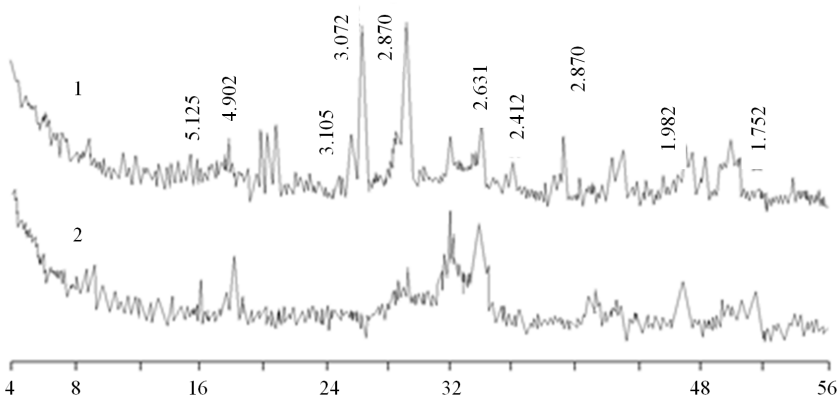

Fig. 1. Roentgenograms of contact areas of cement stone on the basis of modified fillers (1) and usual fillers (2)

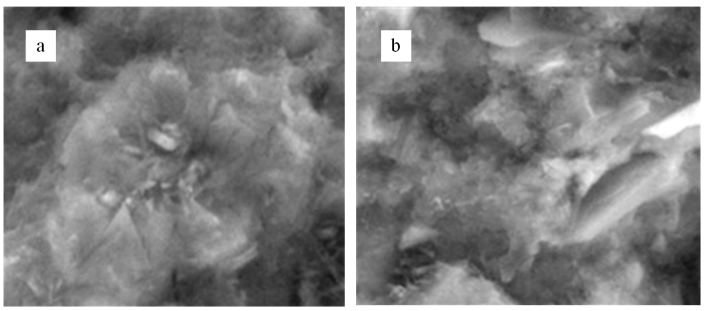

Fig. 2. Photomicrogram of contact area of usual (a) and modified (b) filler and cement stone (zoom ×5000)

Table 3. Formula of filled binding material on the basis of a waste of sawing of stones

Type of the filled binding material	Quantity of cement (%)	Filler level (%)	Quantity of the additive "Bio-NM" (% of FBM mass)	FBM specific surface (m^2/kg)	Compressive strength (MPa)
IA85	85	15	2	558	75.3
IA70	70	30	2	578	69.8
IA55	55	45	2	599	45.2

As is known, the main precondition for production of qualitative concretes is decrease of a water-cement ratio of a concrete mix and raise of binding material reactivity. From this point of view it is considered that the most actual directions use filled binding materials with fineness of grinding to 600 m^2/kg. Such approaches to formula designing allow composites to receive with high density and to upgrade

a cement stone due to raise of cement hydration extent together with the filler and to provide a necessary reserve of anhydrate cement particles for embedment of the defects originating under the influence of various external factors.

Results of the comparative analysis of particle-size compositions of offered filled binding materials HB55 and HB70 with a floured filler produced from the waste of sawing of stones with specific area SSP = 549–599 m^2/kg (see Table 4) have revealed essential difference in their granulometry.

Table 4. Particle diameters D (μm) falling into various fractions

Fraction	1	2	3	4	5	6	7	8	9	10
D (μm) from-to	0.20	0.25	0.30	0.36	0.44	0.49	0.58	0.75	0.90	1.03
	0.25	0.30	0.36	0.44	0.49	0.58	0.75	0.90	1.05	1.31
Fraction	11	12	13	14	15	16	17	18	19	20
D (μm) from-to	1.31	1.50	1.80	2.21	2.61	3.22	3.81	4.53	5.46	6.54
	1.50	1.80	2.21	2.61	3.22	3.81	4.53	5.46	6.54	7.83
Fraction	21	22	23	24	25	26	27	28	29	30
D (μm) from-to	7.83	9.42	11.4	14.1	15.9	20.2	24.8	28.2	34.5	41.5
	9.42	11.4	14.1	15.9	20.2	24.8	28.2	34.5	41.5	49.9
Fraction	31	32	33	34	35	36	37	38	39	40
D (μm) from-to	49.9	58.9	70.2	84.6	100	125	150	175	210	260
	58.9	70.2	84.6	100	125	150	175	210	260	300

Material HB55 has the quantitative content of more fine fractions that increase in the specified interval. The screening curve for HB70 is located in the interval of larger fractions. Thus, content of more fine fractions in HB55 is higher than in HB70. Such difference in granulometric properties upon equal specific surface of particles is caused by the fact that the filler presence promotes finer milling of the binding material.

Hence, it is necessary not only to increase the value of specific surface of the binding material, but also to be aimed to provide rational grain-growing composition by application of fillers of a various origin.

Electron microscope investigations founded more consistent character of the cement stone for samples with HB55 with specific surface of 599 m^2/kg (Fig. 3b) in comparison with a cement stone made from usual Portland cement (Fig. 3a).

The structure of samples with HB55 represents a close packing of clinker grains in a cement stone that is caused by presence of the thinnest films of water between grains and the secured formation in the constrained volume of enough quantity of low-base calcium hydro silicates. Selective point distribution of a surface active agent on cement grains (Fig.3 d) is noted. Therefore, the charge of the additive "Bio-NM" upon production of the filled binding material makes a small share from a total surface of cement grains. Among the factors influencing structure formation and properties of both newly compressed mixes, and matured concretes it is possible to select the main ones: the water content, compacting pressure, charge of the binding material, conditions of preparation and subsequent concreting of a composite. Water content influence in moldable mixes on properties of concretes was studied on the

samples obtained with use of a filler made from screening of crushed materials in the Argunsky quarry and HB70. Pressing of samples was carried out at pressure about 30 MPa which have been tested at the age of 28 days (see Table 5).

By the conducted researches it is proved that the concretes made from sand with the water content 6.5 % (Fig. 4) have the greatest strength and therefore, further properties were investigated for concretes with this initial water content.

Studying of influence of the binding materials charge and compacting pressure on properties of investigated cement concretes (Fig. 5) was carried out with use of sands made on the basis of screening of rubble crushing in Argunsky quarry with the charge of binding materials 20–35 % from mass of the prepared mix and the initial water content in a mix of 6.5 %. Samples were pressed within the interval of pressure 20–50 MPa, and then stored in normal conditions for 28 days.

Raise of the cement charge promotes increase in physical-mechanical characteristics of green composites. At the same time the increase in the charge of cement at 25 % and more does not result in a notable gain of concrete strengths (see Table 6).

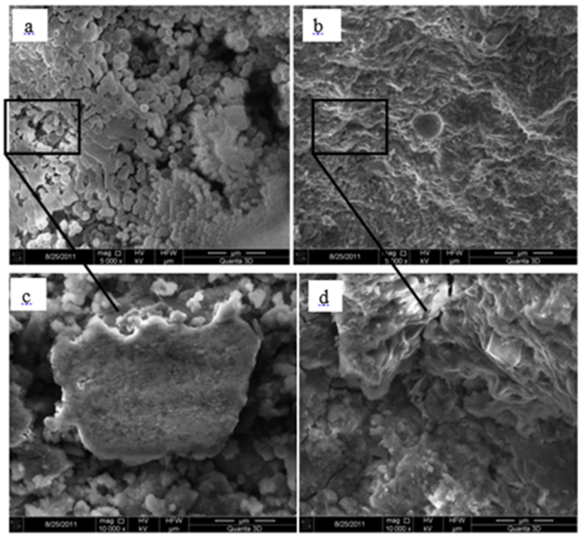

Fig. 3.
Photomicrography of a cement stone made from portland cement
(a, c–zoom ×5000) and the binding material HB55 (b, d–zoom ×10000)

The increase in compacting pressure more than 30 MPa makes insignificant impact on properties both newly compacted, and green composites (see Table 7), therefore, in the further research there was accepted compacting pressure 30 MPa as the most effective (see Figs. 4–5).

So, possibility of cement concrete composites manufacturing on the basis of a secondary material made from by-passed stones is proved. Influence of a type and grain-growing composition of a filler to cement concrete properties is investigated. Influence of filler made from a waste of sawing of stones on processes of structure formation and physical-mechanical properties of a green composite is studied.

Table 5. Properties of cement concretes depending on initial water content in a sand

No.	Initial water content (%) from mass of dry components	Compaction factor k_{comp}	Medium density of a concrete mix (kg/m^3)	Medium density of cement concrete (kg/m^3)	Compressing strength (MPa)
1	5.5	1.52	2302	2276	51.1
2	6.5	1.73	2328	2294	56.0
3	7.5	1.81	2316	2288	55.1
4	8.5	1.84	2294	2264	52.0
5	9.5	1.87	2272	2252	50.7

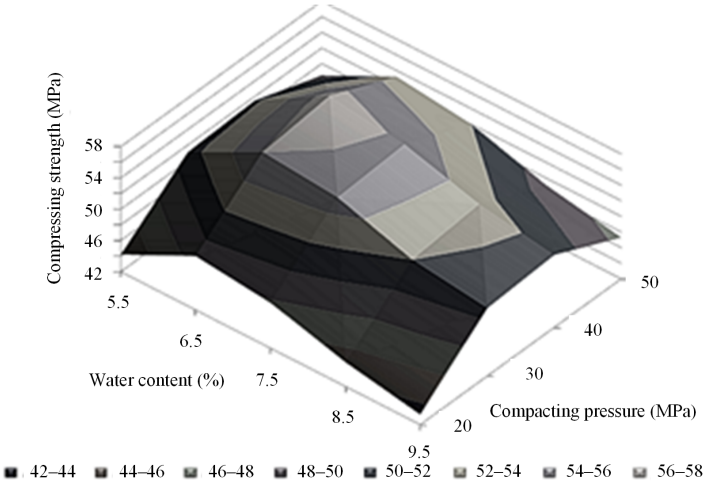

Fig. 4. Dependence of hardness pressed cement concrete from the initial water content and compacting pressure

To forecast strength of cement concretes with use of the modified filler and application of a floured micro filler made from a waste of sawing of stones researches with use of the mathematical experiment design apparatus were carried out. In the capacity of variables there have been selected water-to-binding material relation (W:BM) from 0.20 to 0.40; a type of a compounded binding material and a relationship between quantity of screenings after crushing (CS) and the compounded binding material in concrete (CS:BM) from 2:1 to 4:1.

For the investigated compositions of concretes with the charge of binding materials 435–585 kg/m^3 without additives, compression strength at the age of 28 days varied within the range of 34.2–42.3 MPa. Analogous compositions with application of the filler modified by ADMAX have compression strength 46.6–61.4 MPa and a prism strength 40.8–54.0 MPa. So, use of the filler processed by the cation-active additive ADMAX in the proposed effective compositions allows compression

strength and a prism strength to raise approximately by 20–30 %. Values of static and dynamic elastic moduli have also increased.

Table 6. Influence of the binding material content in a sand on properties of cement concrete composites

No.	Relative content of binding materials in a sand (%)	Medium density of a concrete mix (kg/m^3)	Medium density of the cement concrete (kg/m^3)	Compressing strength (MPa)
1	20	2232	2223	50.2
2	25	2323	2294	56.0
3	30	2345	2312	57.6
4	35	2354	2325	59.2

Table 7. Influence of compacting pressure on properties of cement concretes

No.	Compacting pressure (MPa)	Compaction factor k_{comp}	Medium density of a concrete mix (kg/m^3)	Medium density of cement concrete (kg/m^3)	Compressing strength (MPa)
1	20	1.56	2245	2215	48.9
2	30	1.77	2330	2294	56.0
3	40	1.82	2350	2320	57.0
4	50	1.83	2360	2325	53.2

5. Conclusion

We developed and investigated cement concrete construction composites with complex utilization of by-passed stone and stone milling waste.

It was proved that compounded binding materials with filler from a waste of sawing of stones, which are partially amorphized in the course of milling, in the case of their hydration rate up a formation process of clinker mineral scatting as subcase upon obtaining of new growths and promoting binding of a lime hydrate in insoluble calcium hydro-silicates with fine-crystalline structure and various basicity factors.

The character of distribution of pores depending on the sizes of cement stone of compounded binding materials on the basis of stone sawing wastes is revealed. The grain size curve for particles of the compounded binding material made from stone sawing wastes has more intermittent character due to the polymineral composition which intensifies milling that has a positive impact on formation of a microscopic structure of a cement stone owing to more consistent spatial arrangement of particles and, therefore, production of stronger cement stone. Influence of the modified fine aggregate made from screenings of a demolished concrete and rocks crushing and

micro fillers on structure formation and the basic physical-mechanical properties of the pressed concretes for construction of cement concrete pavements is revealed.

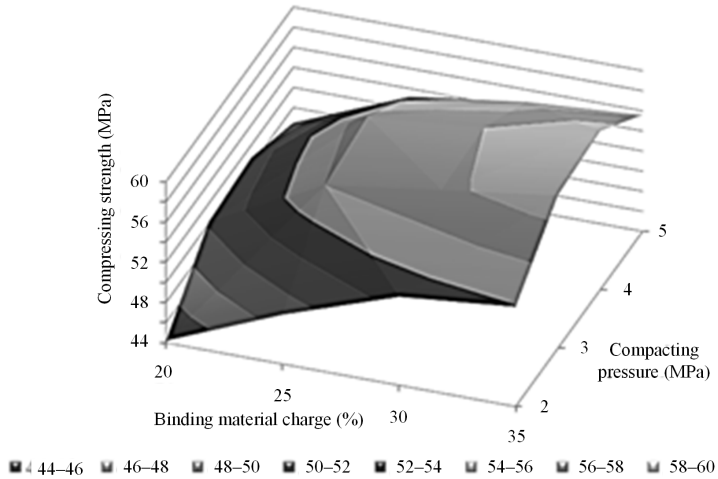

Fig. 5. Dependence of strength of pressed cement concretion the binding material charge

References

[1] C. A. A. Rocha, G. C. Cordeiro, R. D. Toledo Filho: *Influence of stone cutting waste and ground waste clay brick on the hydration and packing density of cement pastes.* Revista IBRACON de Estruturas e Materials *6* (2013), No. 4, 601–680.

[2] S. A. Balwaik, S. P. Raut: *Utilization of waste paper pulp by partial replacement of cement in concrete.* IJ Engineering Research and Applications *1* (2009), No. 2, 300 to 309.

[3] Y. M. Chun, T. R. Naik: *Concrete with paper industry fibrous residuals: Mixture proportioning.* ACI Materials Journal *102* (2005), No. 4, 237–243.

[4] P. Soroushiana, M. Hassan: *Evaluation of cement-bonded strawboard against alternative cement-based siding products.* Construction and Building Materials *34* (2012), 77–82.

[5] R. Slávik, M. Čekon, P. Oravec, R. Kolář: *Operative temperature predicting of a room in summer: an approach for validating of empirical calculation models.* Applied Mechanics and Materials *824* (2016), 519–526.

[6] P. Garcés, E. Zornoza, E. G. Alcocel, Ó. Galao, L. G. Andión: *Mechanical properties and corrosion of CAC mortars with carbon fibers.* Construction and Building Materials *34* (2012) 91–96.

[7] N. Sebaibi, M. Benzerzour, N. E. Abriak, C. Binetruy: *Mechanical and physical properties of a cement matrix through the recycling of thermoset composites.* Construction and Building Materials *34* (2012), 226–235.

[8] X. An, K. Maekawa: *Shear resistance and ductility of RC columns after yield of main reinforcement.* J Materials, Concrete Structures and Pavements, JSCE *38* (1998), No. 585, 233–247.

Received November 16, 2016

Aim and Scope

Acta Technica (formerly Acta Technica CSAV) is a refereed international journal founded in 1956. It is devoted to presenting progress in electrical engineering, electronics, mechatronics, thermodynamics, hydromechanics, materials science and engineering, mechanics of solids and structures, plasma physics, and others, together with corresponding mathematical and computer tools. It also covers interdisciplinary, coupled and newly emerging problems (e.g. nanotechnology, bioelectromagnetics, etc.).

Frequency: Quarterly, four times a year.

Submission of Contributions

Only contributions satisfying the following requests can be accepted:

- original works not yet published and not being under consideration for publication elsewhere,
- relating the aim and scope of the journal,
- written in a sufficiently good English and sent electronically.

The reprint of the published paper together with the complete issue will be provided to each author in electronic form in PDF.

Form of manuscript. Any standard format of TEX (preferable) or MS Word. Use double spacing, number the pages and indent paragraphs.

Structure, style and notation. Detailed instructions are presented at

`http://journal.it.cas.cz/author_guide.pdf`

Read them carefully before elaborating the manuscript.

Figures. Acceptable formats: PDF, PNG (recommended for line art) and JPEG (recommended for photographs). Use black-and-white or greyscale figures. Avoid using hair lines (problems with printing).

Delivering the manuscripts. Via e-mail to the editorial office (see the inside front cover). The Editor will confirm the acceptance to the corresponding author.

Review stage. Manuscripts are usually evaluated by two independent reviews. The referees' identities are not released.

Proof stage. Authors obtain proofs in PDF via e-mail to correct possible errors. Do not return a corrected manuscript but send a list of corrections by an e-mail message (e.g. as an attachment file).

Copyright. Before publishing the paper, the authors are asked by the Editor to sign transfer of the copyright to the Publisher.

ACTA TECHNICA

Volume *61* (2016), Number 4B