

The Use of Least Square, Matrix Decomposition and Compression Techniques for Oil and Gas Pipeline Simulation

Dai-Xi Xie and Da-Cheng Mei

Abstract—Oil and gas pipeline transmission network simulation is important for the petroleum industry. Existing node method monitors node pressure, simulate pipeline flow rate, and identify different risks. However, it may not be efficient enough to meet the real-time requirement. In this paper, we propose to improve the node method with two techniques. Using the node method, Comparing real data, average deviation, goodness of fit and height of matrix should be raised. First, the least square method is employed to replace the fluid distribution method. Second, matrix decomposition and compression are applied to matrix calculation. Experiments are undertaken on synthesis networks with different sizes. Results show that the new method is at least twice faster than the counterpart.

Index Terms—Pipeline network, simulation, sparse matrix.

I. INTRODUCTION

The simulation of gas and oil transmission systems requires increasingly sophisticated computer software as these systems grow in size and complexity [1]-[4]. Because the current price, this represents a significant amount of cost for operating large pipeline network system. For instance, considering the transmission pipeline network system, [5] indicated that a hundredth improvement on the performance of the transmission pipeline network system could result a saving of 48.6 million dollars.

The nodal pressure and pipeline flow rate are essential to realize real time simulation of the transmission pipeline network system. Through monitoring nodal pressure, company should observe whether the pipeline network is unsafety.

In general, a brief review on the previous researches is presented: Model of node equation method base on nodal law of conservation of mass. A Newton-Raphson method has been extended to a system of nonlinear equations given in the parameter form and it leads to a node method [6].

However, it is inefficient. Real time simulation needs to be fast. The fast speed should better realize early warning. Therefore, the increasing speed is needed in natural gas and oil transmission simulation systems.

In this paper, in order to enhance efficiency of node equation method, model and algorithm is improved. The revised model focused on improve originally fluid distribution in order to enhance efficiency of node method. It

should strengthen the effect of real time emulation. In the revised algorithms, methods of matrix compression storage and decomposition based solution procedures were used in order to cut down running time of algorithms. On a typical hard problem, the Cholesky decomposition is about 8000 times faster than the point Gauss-Seidel method, 200 times faster than the alternating direction implicit method, and 30 times faster than the block successive over relaxation method with optimum relaxation factor [7]. LU decomposition also is faster than no decomposition.

Two experiments are used to revise node method. Firstly, the result of the simulation system compare with Aspen Plus software, in order to validate availability of the revised node method. The comparative result with Aspen Plus verify that the revised node method are effective and reliable. Secondly, the total running time of the simulation system in different matrix size are presented. The totally running time demonstrate that the revised node method is high-efficiency.

In simulation system of transmission pipeline network system, real time monitoring is very important, therefore an efficient method are importance to simulation system. This paper make enormous contribution to solve the problem. As a result, the simulation model and algorithm could assist in decisions regarding the design and operations of the transmission pipeline network system.

II. BASIC THEORY

In this section, model and algorithm of node method, methods of matrix compression and calculation are presented.

A. Model of Node Equation Method

In this section, mass balance equations and pressure drop equation and establishment process of model of node equation method are presented.

1) Formulation of mass balance equations

Mass balance provides the basic equations in order to have a complete mathematical formulation for the simulation model of a given transmission pipeline network system. The mass balance equations are obtained based on the principle of conservation of mass at each junction of transmission pipeline network system. At any junction within transmission pipeline network system, the generalized mass balance equation can be summarized as:

$$\sum Q_i - \sum q_j - \sum D_L = 0, \quad (1)$$

Manuscript received December 20, 2015; revised February 20, 2016.

The authors are with school of computer science, Southwest Petroleum University, Chengdu, China (e-mail. xiedaisy@163.com)

where Q_i flow through incoming are pipelines to junction node, q_j are flow through outgoing pipelines from junction node and D_L . are the flow rate from junction node.

2) Pressure drop equation

The calculated branch flows rate and pressures can be utilized in gas and oil network design [8]. Their deviations from expected values are used for corrections of pipe flow rate. Branch characteristic is a nonlinear function, describing quadratic dependence of the pressure drop H_i .

$$H_i = R_i \times q_i^2, \quad (2)$$

where q_i express every pipeline flow rates, R_i being aerodynamic resistance of the i th pipeline, the resistance R_i is defined by the following Equation (3):

$$R_i = 10.67 \times C^{1.852} \times d^{4.87} \times L, \quad (3)$$

where L denote length, d denote diameter and C denote pipeline material resistance (or friction coefficient). NP is nodal pressures, then the pressure drop between adjacent nodes can be written as:

$$H = NP_i - NP_j, \quad (4)$$

where NP_i express the pressure of i th node, NP_j express pressure of next node of i th, H express the pressure drop. In the method, first original flow rate is distributed according to Equation (5)

$$q_i = R_e \times \frac{u \times \Pi \times d}{4 \times \rho}, \quad (5)$$

where R_e express Reynolds number, u express viscosity, ρ and express density.

Second, pipeline flow rates are corrected and nodal pressures are calculated. Instead of Equation (1) we use characteristic

$$A \times q_i + Q_i = 0, \quad (6)$$

where A express matrix of pipeline direction and layout. Its storage and calculation method is showed in section 2.2.1 and section 2.2.1. Q_i express flow rate of node. In order to became nonlinear algebraic equations to linear algebraic equations. Instead of Equation (2) we use characteristic

$$H_j = S_j \times q_i^2, \quad (7)$$

where S_i is defined as matrix, instead of Equation (7) we use characteristic, an example of matrix S_i is showed:

$$H_i = \begin{bmatrix} s_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & s_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & s_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & s_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & s_6 \end{bmatrix} \times q_i^2$$

R_i is defined as

$$R_i = S_i \times q_i. \quad (8)$$

Equation (8) are substituted into Equation (2) thus giving instead of Equation(9) we use characteristic

$$H_i = R_i \times q_i \quad (9)$$

$$q_i = 1 / R_i \times H_i \quad (10)$$

where c is defined as

$$c = 1 / R_i, \quad (11)$$

c is diagonal form.

Equation (10) are substituted into Equation (11) thus giving linearized characteristics Equation(12) are introduced into Equation (6).

$$q_i = c H_i, \quad (12)$$

And together with Equation (9). We obtain Equation (13).

$$AcA^T H + Q = 0, \quad (13)$$

where H express pressure of node. An example about Equation (13) should express as:

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} c_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & c_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & c_6 \end{bmatrix} \times \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} + \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \\ Q_6 \end{bmatrix} = 0$$

The original set of nonlinearity simultaneous equations in known node pressures and layout of pipeline. Pipeline flow rates and node pressure as the unknown is calculated by a set of linear equations. The value of the nodal pressure can be calculated from Equation (13) and value of q is obtained on Equation (12).

B. Algorithm of Node Method

Algorithm 1 : Node method

Input: NQ ; e ; L ; C ; d ; u ; A ; ρ .

Output: H ; PQ .

$$q = 2 \times 10^5 \times \frac{u \times \Pi \times d}{4 \times \rho}$$

$$\max(q^{new} - q^{old}) > e$$

$$q^{new} = q;$$

$$q^{old} = q^{new};$$

$$s = 10.67 \times C^{1.852} \times d^{4.87} * L;$$

$$J = 1 / (s \times q^{new});$$

$$B = A \times J \times A';$$

$$H = B \setminus NQ';$$

$$q^{new} = A' \times H \times c;$$

$$q^{new} = 1 / 2 \times (q^{old} + q^{new});$$

TABLE I: ILLUSTRATION OF PARAMETER

Name of parameter	Account for the name of parameter
NQ	every pipeline node flow rate;
q	pipeline flow rate;
H	nodal pressure;
n	The total number of pipeline;
e	value of relative error;
L	length of pipeline;
c	value of aerodynamic resistance;
d	every pipeline diameter;
A	A matrix to illustrate layout and direction of pipeline network;
ρ	density of fluid;
u	viscosity ;

Algorithm 1 is implemented to calculate pipeline flow rates and nodal pressures. It has 8 input parameters and 2 output parameters. They are all listed in Table I, q express every pipeline flow rates.

Newton Raphson solution technique is used.

$q^{new} = 1/2 \times (q^{old} + q^{new})$ is to produce next iterative value of pipeline flow rate. $\max(q^{new} - q^{old})$ is to control iterations. The value should be computed for each iteration. Iteratively until the relative errors are less than specified tolerance. The value of q^{new} is solution of equation.

C. Methods of Matrix Storage and Calculation

A method of matrix compression and decomposition are used in this paper. The large size matrix and the condition of many technical result in the need for efficient techniques. So it is necessary to solve systems of equations using process mode of sparse matrix.

1) Storage scheme

Efficient storage schemes for large matrices depend on the sparsity pattern of the matrix. In this section, we apply the three one-dimensional arrays format [9].

The principle of the scheme is illustrated in a matrix A with non-zeros a_{ij} .

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

The non-zeros of matrix A are stored row-wise in three one-dimensional arrays. Value contains the values of the non-zeros, j the corresponding column indices. The elements of i point to the position of the beginning of each row in value and column indices. v express value contains the values of the non-zeros.

2) Calculation scheme of matrix multiplication

In this section, the method of matrix multiplication is discussed. Consider the product $C = A \times B$, where A, B, and C are matrices of size $M \times P$, $P \times N$, and $M \times N$, respectively. For a matrix A, let $A(i, j)$, and $A(j, i)$ denote the

ith row and the ith column of B respectively. In other word, it is to get one element in C, we multiply a row in the A matrix with a column in the B matrix, $C(i, j) = A(i, :) \times B(:, j)$ for $i = 1, 2, \dots, M$ and $j = 1, 2, \dots, N$. This is the standard matrix multiplication approach.

TABLE II: THE THREE ONE-DIMENSIONAL ARRAYS FORMAT

i	j	v
1	2	1
1	3	1
3	1	-1
3	6	-1
4	3	1
5	2	1
6	1	1
6	4	-1

3) Matrix decomposition methods

In this section, matrix decomposition is presented, in order to efficiently dispose matrix multiplication.

A is an arbitrary nonsingular matrix, then $A = K - R$ represents a decomposition of the matrix A .

The more K resembles A . A lower triangular $N \times N$ matrix is denoted by $L = (l_{ij})$, so $l_{ij} = 0$ if $i < j$. And an upper triangular $N \times N$ matrix by $U = (u_{ij})$. A matrix K approximating A has to be constructed such that the L and U belonging to K are sparse. This can be realized by making an LU decomposition of A , during which elements are neglected in the L and U matrices in appropriate places. That is the reason that we shall call $K = LU$ an incomplete \$LU\$ decomposition of A . The decomposition $A = LU - R$ is regular. The factors L and U are unique [10].

III. REVISED NODE METHOD

In order to enhance efficiency of node method. We use a method to improve originally fluid distribution and design an algorithm including originally fluid distribution and processing method of sparse matrix.

A. Method of Originally Fluid Distribution

Originally fluid distribution networks is anew designed in order to enhance the efficiency of node method. A method based on the least square method is presented, by which the optimal originally design of a fluid can be obtained. Number of pipeline flow is q , the flow rate linking node i and j can be expressed as:

$$\Omega = \sum q_{ij}^2,$$

Equation (14) includes all the commercially available pipeline sizes for every pipeline in the network. The value of the incoming and outgoing flow rate is equal to the value of node flow rate. The final solution will include only one size for every pipeline in the network.

Therefore, the following constraint is considered for the zero-unity variables for pipeline:

$$\sum q_{ij} + Q_i = 0,$$

where Q_i express node flow rate.

Now, the equation have some constraints here. It is mass balance of node flow. Equation (14) are substituted into Equation (15) thus giving equation (3).

$$\Omega = \sum q_{ij}^2 + \sum \lambda_i (\sum q_{ij} + Q_i),$$

Ω is a minimum value solving process as the objective function through Equation(16).

It can start taking derivatives of this with respect to Equation (16). Then it equal to zero. If the result is negative value, it express opposite direction and branch i leaving node j . If the result is positive value, it express opposite direction and branch i entering node j .

Flow direction should be judged according to this algorithm.

If branch i not enters node j , it is $+1$. If branch i not enter node j , it is 0 . If branch i leave node j , it is -1 .

B. Algorithm Implementation

Algorithm 2 : Pipeline flow rate correction

Input: NQ; e; q; n; L; C; d; A
Output: H; q.

```

x0=ones(1,n);
m=length(NQ);
For t ≤ m;
    beq = 0;
    beq = beq + NQ(t);
End for
Aeq=one(1,n);
For k ≤ m
    f = 0;
    f = f + NQ(k)^2;
End for
[x,fval]= fmincon(f,x0,[],[],Aeq,beq);
qnew = x;
While max(abs(qnew - qold)) > e
    qold = qnew;
    s=10.67×c1.852×d4.87×L;
    J=1/(s×thetanew);
    J1=diag(J);
    J2=sparse(J1);
    J3= ichol(J2); or ilu(J2);
    B=A×J3×A'; or B=A×J2×A'; or B=A×J1×A'
    H=B\Q';
    qnew = A'×H×c;
    qnew = 1/2×(qold + qnew);
End while

```

Algorithm 2 has 8 input parameters and 2 output parameters. They are all listed in Table II. It is implemented to improve Algorithm 2.

$x0$ is a 1-by-n array. It is initial point for result. m express length of \$Q\$. For-loop structure is used to calculate sum of every pipeline flow rate. f express linear objective

function vector. `Aeq` express matrix for linear equality constrains. `beq` express vector for linear equality constrains. Function `fmincon` is called to solve linear programming problems. `fmincon` is method offered by matlab. Initial value of q^{new} equal to value of the processing technic about sparse matrix is presented through way above text.

Matrix J is sparse matrix and it is positive definite symmetric matrices. It is compressed using method of section 2.3.1 and decomposed using method of section 2.3.3. `diag` is a method of matlab. It transform $J1$ to diagonal matrix J . `sparse` is a method of matlab. It is transform $J1$ to sparse matrix $J2$. $J3$ express a result of matrix decomposition. It include two ways, they are incompleteness Cholesky decomposition and incompleteness LU decomposition. LU decomposition and Cholesky decomposition include matrix decomposition and implementation for matrix multiplication after decomposition.

IV. EXPERIMENTAL RESULTS

In this section, we try to answer the following questions through experimentations.

- 1) Are the revised node method effective?
- 2) Do the revised node method enhance efficiency of node method?

A. The Verification of Model and Algorithm

In order to illustrate the methods presented, a carefully selected network, with a representative structure and parameters are given.

The network form is shown in Fig. 1.

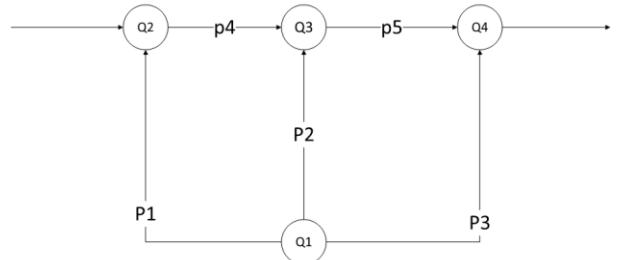


Fig. 1. Layout of pipeline.

The network form consists of all the most common configurations of pipeline network systems, branched, linear and looped configuration.

In order to verify the correctness and effectiveness of the revised node method, the results of example are compared with the Aspen Plus software. The revised method are effective if the value of goodness of fit compare with Aspen Plus is acceptable. The revised method are incorrect if the value of goodness of fit compare with Aspen Plus is unacceptable. Goodness of fit is ratio of regression sum of squares and sum of squares for total. Its range is from 0 to 1. When it approach 1, the value of two sets of data is similar. When it approach 0, the value of two sets of data is unlike.

Aspen Plus software is the market-leading chemical process optimization software used by the bulk, fine, specialty, biochemical industries, as well as the polymers industry for the design, operation, and optimization of safe, profitable manufacturing facilities. The industries that drive

our economies and touch our lives are optimized by Aspen Plus software every day. The world's leading oil gas, chemical, engineering construction, pharmaceutical, food, beverage. Its products represent the industry standard in process optimization, helping more than 150,000 users at over 1,750 companies drive profitability faster and more efficiently than ever before.

Firstly, in order to verify, we set up input parameters. The input parameters come from oil and gas Pipeline Company. Some parameters are defined in Aspen Plus software according to the simulation system.

TABLE III: INPUT PARAMETERS

Q_i	1	2	3	4
Node Flow Rate	530	250	100	180

TABLE IV: BRANCH PARAMETERS

Parameter	Value
L	500
d	120
c	130

Table III, Table IV are input parameters in simulation system and Aspen Plus software. In Table III, Table IV, number of node is arranged according to Fig. 1. In Table III, Q_i express number of node. In table 4, every pipeline branch parameters is presented. L express length of pipeline. d express diameter of pipeline. C express resistance of pipeline material.

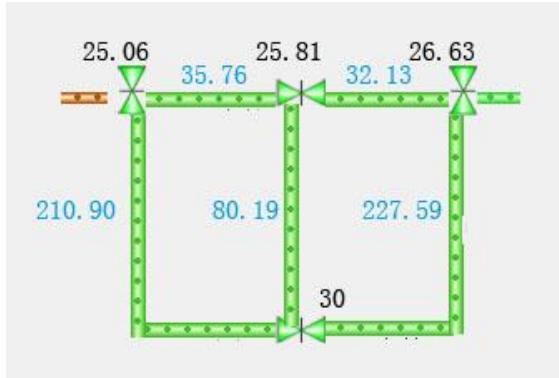


Fig. 2. Result of simulation system.

In Fig. 2, illustration about setting up pipeline layout and layout in Aspen Plus software is showed. The direction and layout equal to the example.

Secondly, the output results of the example and analysis of comparative result with Aspen Plus software are shown in the text below.

The output results of example are presented in Table V and Fig. 3.

In Table V and VI, p_i express number of pipeline. v_i express pipeline flow rate. H_i express node pressure. Fig. 2 show value of every node pressure and pipeline flow rate.

Black text express node pressure, it could be in units of Mbar. Blue text express pipeline flow rate, it could be in units of m^3 .

TABLE V: PIPELINE FLOW RATE

p_i	1	2	3	4	5
$v_i [m/s]$	210.9 027	80.1 9	227. 59	35.7 58	321.296
p_i	1	2	3	4	
$H_i [P_A]$	30	25.06	25.81	26.8 3	13

Through the example, values of pressure and flow rate are visualized.

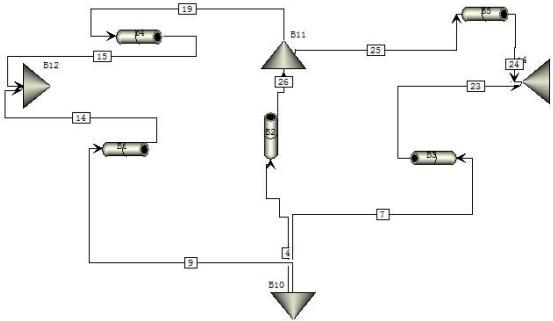


Fig. 3. Layout of aspen plus.

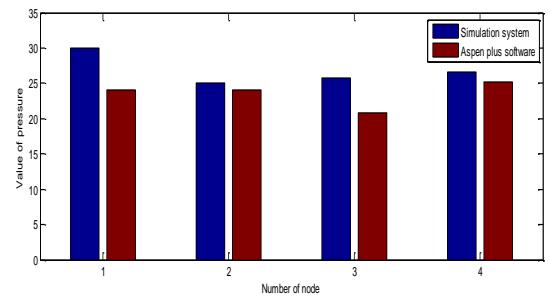
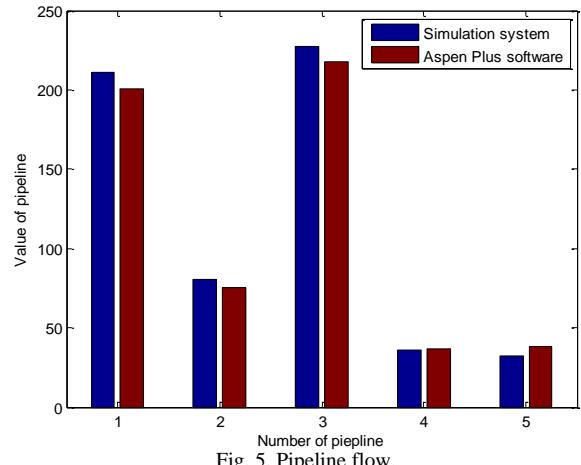


Fig. 4. Nodal pressure.



In Fig. 4 and Fig. 5, the result of the example compared with Aspen Plus software.

In Fig. 4, firstly, in pipeline 2, 4, the height of blue matrix and red matrix is similar. Secondly, in pipeline 1, 5 there was a deviation. Mean range in pipeline 1, 5 is 1.33, the deviation is acceptability. Finally, in total 4 pipeline, their goodness of fit is 0.72. From the result, we can summarize that there are little changed on node pressure.

The analysis of Fig. 5 is similar to Fig. 4.

Firstly, the height of blue matrix and red matrix is similar.

Secondly, in total 5 pipeline, their goodness of fit is 0.99. From the result, we can summarize that there are little changed on pipeline. We could see that the revised node method is effective.

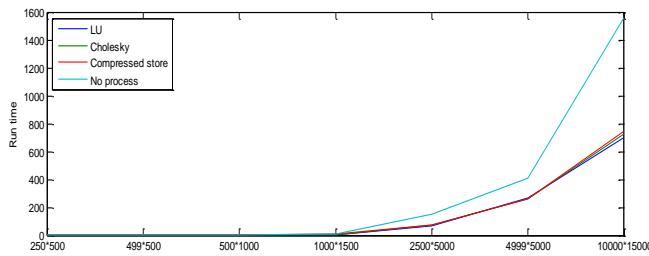
Finally, early warning can be realized based on the example. Our works can be used to safety precaution and explosion prevention. In natural gas pipeline in city, a reasonable pressure range is from 40Mbar to 96Mbar to ensure supply and safety. When nodal pressure is below 40Mbar, natural gas can't transmit to high-rise building. People's life will encounter difficulty. In the experiment, node pressure is at unreasonable level. They are under 40Mbar. In application, the city including the four will encounter severe problem. According to the result, we should inform company. Company should adjust nodal flow rate.

B. The Verification of Algorithm Efficiency

In the experiment, the run time are compared among matrix with LU decomposition, matrix with Cholesky decomposition, using matrix compression, and no process matrix in different matrix size.

Firstly, input matrix are structured. A kind of matrix is structured by 0, -1 and 1. Each column only have two number. It express layout and direction of the pipeline to be similar to matrix A. Other a kind of matrix is one dimensional vector quantity, which express original nodal flow rates. It is to be similar to Q in Table III. Matrix of 250×500 and 499×500 and 500×1000 and 1000×1500 and 2500×5000 and 4999×5000 10000×15000 is structured in order to replace matrix A. One dimensional vector quantity of 250, 499, 500, 1000, 2500, 4999, 10000 is structured to replace Q.

Second, the results of total running time in different size matrix are shown in Table V of matrix decomposition time} and Fig. 6.



- Mathematical and Computer Modelling*, vol. 31, no. 2, pp. 197-220, 2000.
- [7] A. D. Woldeyohannes and M. A. A. Majid, "Simulation model for natural gas transmission pipeline network system," *Simulation Modelling Practice and Theory*, vol. 19, no. 1, pp. 196-212, 2011.
- [8] O. Baker, "Design of pipelines for the simultaneous flow of oil and gas," in *Proc. Fall Meeting of the Petroleum Branch of AIME, Society of Petroleum Engineers*, 1953.
- [9] F. S. Smailbegovic, G. N. Gaydadjiev, and S. Vassiliadis, "Sparse matrix storage format," in *Proc. the 16th Annual Workshop on Circuits, Systems and Signal Processing*, 2005, pp. 445-448.
- [10] J. A. Meijerink and H. A. V. D. Vorst, "An iterative solution method for linear systems of which the coefficient matrix is a symmetric *M*-matrix," *Mathematics of computation*, vol. 31, no. 137, pp. 148-162, 1977.



Dai-Xi Xie is with the School of Computer Science, Southwest Petroleum University, Chengdu, China. She is graduate student. Her major is simulation of oil pipeline.

Da-Cheng Mei is a professor at the School of Computer Science, Southwest Petroleum University. His major is simulation system.

He has been study simulation of oil pipeline two years old in Southwest Petroleum University.